

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

CSI/CSSS/24/280

DÉLIBÉRATION N° 24/134 DU 3 SEPTEMBRE 2024 PORTANT SUR LA COMMUNICATION DE DONNÉES À CARACTÈRE PERSONNEL PSEUDONYMISÉES RELATIVES À LA SANTÉ PAR L'AGENCE INTERMUTUALISTE, SCIENSANO ET LA FONDATION REGISTRE DU CANCER À SCIENSANO DANS LE CADRE DE L'ÉTUDE SCIENTIFIQUE SPADIS

Le Comité de sécurité de l'information ;

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 relatif à la protection des données ou RGPD);

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*, notamment l'article 37 ;

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

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

Vu la demande de Sciensano ;

Vu le rapport d'auditorat de la Plate-forme eHealth du 26 août 2024;

Vu le rapport de monsieur Michel Deneyer.

Émet, après délibération, la décision suivante, le 3 septembre 2024:

I. OBJET DE LA DEMANDE

1. La présente demande s'inscrit dans le cadre du projet SPADIS qui vise à mesurer et à mieux comprendre la participation sociale (en ce compris la participation au marché du travail) de personnes souffrant d'une maladie chronique en Belgique. Dans le cadre de la génération de nouvelles connaissances détaillées sur la participation sociale de personnes souffrant de maladies chroniques, c'est-à-dire la mesure dans laquelle les individus participent à la société ou à la communauté, le projet SPADIS a été développé au sein de Sciensano. SPADIS examine les manières de mesurer la participation sociale, tout en mettant l'accent sur la participation au marché du travail, de personnes souffrant de maladies chroniques ainsi que leurs déterminants. En outre, SPADIS évaluera la faisabilité de l'extension de la capacité de monitoring épidémiologique actuelle de Sciensano à la participation sociale, grâce au couplage de banques de données et à l'implémentation d'instruments de mesure rapportés par les patients.
2. Au début du projet SPADIS, il a été exploré en détail dans quelle mesure des données sont disponibles, au niveau régional et/ou au niveau national, qui permettent de comprendre les aspects sociaux de la santé. De manière plus spécifique, des données ont été recherchées qui montrent dans quelle mesure les malades chroniques savent participer aux activités impliquant des interactions avec d'autres personnes dans la société ou la communauté, à savoir la participation sociale. Cet exercice a montré que des données sont disponibles (pour un couplage) concernant deux domaines essentiels de la vie qui facilitent la participation sociale, à savoir le marché du travail et l'enseignement. La demande de données actuelle concerne un couplage de registres cliniques actuels aux données relatives au marché du travail.
3. Les données de la demande actuelle seront utilisées pour répondre aux (principales) questions de la recherche suivantes: Quelle est la participation au marché du travail de personnes souffrant du VIH, d'un cancer ou de diabète et quels sont les déterminants? Quels sont les éléments moteurs des transitions sur le marché du travail et comment diffèrent-ils en fonction de la pathologie, des caractéristiques socio-démographiques et de la situation socio-économique ?
4. Les informations qui se recoupent et/ou qui divergent et qui sont issues de cet effort peuvent être utilisées pour informer sur des soins sociaux sur mesure dans la pratique clinique et sur des mesures politiques sur mesure, pour identifier les inégalités et pour informer sur le renforcement des capacités au sein de Sciensano afin d'étendre sa capacité actuelle à surveiller de manière exhaustive la santé aux personnes souffrant du VIH, du diabète et du cancer.
5. Motivation générale préalable pour les personnes sélectionnées dont les données à caractère personnel seront traitées: le projet SPADIS vise à mesurer et à comprendre la participation sociale en Belgique de malades chroniques. Le but est d'évaluer des différences éventuelles avec la population générale, mais aussi d'identifier des facteurs déterminants y relatifs et des inégalités sous-jacentes. C'est la raison pour laquelle les critères de sélection pour la population actuelle à étudier sont les suivants:

Les patients qui sont âgés entre 18 et 60 ans au moment où les données les plus récentes sont enregistrées dans le registre Healthdata et qui sont domiciliés en Belgique. Étant donné les objectifs actuels de SPADIS (étude des positions sur le marché du travail et des transitions), les patients âgés de moins de 18 ans et de plus de 60 ans échapperont, dans une large mesure, aux dimensions que les chercheurs souhaitent étudier. Leur enregistrement n'est donc pas pertinent. Pour les patients souffrant d'un cancer, les critères d'âge sont par ailleurs limités à la tranche d'âge 18-55 ans. Il ressort notamment d'une étude antérieure en Belgique (cohorte EMPCAN) que peu de transitions entre des positions sur le marché du travail sont constatées pour les patients âgés de 55 à 60 ans au moment du diagnostic d'un des types de cancer énumérés.

6. Les personnes concernées sont les suivantes:

- Les patients chez lesquels le diabète type 1 ou 2 a été diagnostiqué et qui ont été intégrés à l'âge de 18 à 60 ans dans l'audit 9 (2015-2016), l'audit 10 (2017-2018) ou l'audit 11 (2020-2021) du registre IPQED (initiative pour la promotion de la qualité et de l'épidémiologie du diabète sucré);
- Les patients repris dans la cohorte VIH à l'âge de 18 à 60 ans et ayant reçu leur diagnostic entre le 1^{er} janvier 2003 et le 31 décembre 2021;
- Les patients enregistrés auprès de la Fondation Registre du Cancer (FRC) et ayant reçu (entre le 1^{er} janvier 2016 et le 31 décembre 2021) le diagnostic d'un des types de cancer suivants:
 1. cancer du sein (C50)
 2. cancer colorectal (C18-C20),
 3. cancer tête et cou (C00-C14; C30-C32),
 4. cancer du testicule (C62), ou
 5. malignité hématologique

Cette population présente des maladies chroniques très fréquentes en Belgique qui sont associées avec un lourd fardeau de la maladie à un niveau individuel et au niveau de la population, dont l'âge moyen tombe dans les critères d'âge définis.

7. Au cas où plusieurs cancers ont été constatés chez un individu pendant la période d'étude (2016-2021), la FRC fournit les données demandées séparément pour chaque diagnostic depuis le début de la période d'étude à Healthdata.be, une section de Sciensano. Sur la base du NISS pseudonymisé correspondant dans l'environnement de Healthdata, il est possible de déterminer quel individu souffre de plusieurs cancers. Il est ainsi possible d'en tenir compte dans les analyses envisagées. Par ailleurs, la FRC fournit à l'AIM uniquement les données pour le premier diagnostic du cancer au cours de la période d'étude.

8. Dans les analyses prévues relatives aux transitions sur le marché du travail, il y a, par ailleurs, aussi lieu de tenir compte de diagnostics éventuels qui ont lieu antérieurement à la période d'étude (2016-2021) ou de diagnostics qui ne font pas partie des types de cancer sélectionnés. Ces types de diagnostic peuvent, en effet, avoir un impact individuel majeur sur les transitions sur le marché du travail (p.ex. déclaration pour inactivité existant déjà avant le diagnostic repris dans notre sélection). Pour tous les diagnostics ne faisant pas partie de la sélection ou qui ont eu lieu au cours de la période 2011-2016 (tant les diagnostics tombant

dans qu'en dehors de notre sélection), l'année du diagnostic doit être communiquée à Healthdata, ainsi que le type de cancer lorsque la donnée « autre cancer » est déclarée. Pour l'autre cancer qui fait effectivement partie de la période d'étude et de la sélection, il y a lieu de suivre la procédure normale (l'AIM reçoit l'année et le mois du diagnostic et le NISS pseudonymisé et HealthData reçoit toutes les informations demandées). Sur la base du NISS pseudonymisé correspondant et des données de diagnostic dans l'environnement de Healthdata, il est possible d'identifier la succession des diagnostics.

Cohorte VIH

9. En raison de son caractère d'appui à la politique, l'étude SPADIS vise à offrir une mesure, au niveau de la population, de la participation sociale de patients souffrant du VIH (PLWH) en Belgique. Actuellement, la cohorte VIH a un taux de couverture élevé, estimé à plus de 95 %, ce qui permet de générer, dans le cadre de cette étude, des connaissances qui concernent la quasi-totalité de la population des personnes vivant avec le VIH. Tous les individus diagnostiqués à un âge compris entre 18 et 60 ans et dans la période de 2003 à 2021, sont donc intégrés dans l'étude actuelle.

IPQED

10. Trois audits (9, 10 et 11) ont été choisis pour l'étude actuelle. L'IPQED est le résultat d'échantillons ou d'audits biennaux de patients diabétiques qui ont signé une convention de soins. En l'absence d'un registre au niveau de la population comprenant tous les patients diabétiques, la méthode d'audit de l'IPQED permet de réaliser les études les plus adéquates afin d'extrapoler les connaissances au niveau de la population. Les patients qui ont signé la convention précitée ont, par définition, des stades plus avancés du diabète qui entraînent un fardeau de la maladie plus élevé, tant au niveau individuel qu'au niveau de la population. Ce fardeau de la maladie se manifeste tant au niveau physique, psychologique que social et justifie une étude plus détaillée.

Fondation Registre du Cancer

11. Tous les patients diagnostiqués en Belgique avec un des types de cancer énumérés ci-dessus sont enregistrés par la FRC. En incluant tous les patients diagnostiqués à un âge compris entre 18 et 55 ans et dans la période 2016 - 2021, l'étude décrite ici permet de réaliser des analyses et d'acquérir des connaissances au niveau de la population, ce qui est essentiel pour les finalités du projet SPADIS.

Procédure de sélection

12. La procédure de sélection est différente selon qu'il s'agit de la cohorte VIH, IPQED et BCR¹.

IPQED

¹ Voir annexe.

13. Healthdata applique un query sur la (les) banque(s) de données existante(s) pour l'audit 9, 10 et 11 de l'IPQED afin de sélectionner des patients diabétiques type 1 et 2 pour le projet SPADIS. Ce query est informé par les spécifications techniques des chercheurs de Sciensano et sélectionne les patients dont les données les plus récentes ont été enregistrées dans l'IPQED à un âge compris entre 18 et 55 ans. Le résultat du query (comprenant le PsINSZHD et la date du diagnostic) est fourni à l'AIM par Healthdata.
14. Sur la base de ces informations et après intervention du CIN et de la TTP OA (BCSS), l'AIM consultera les données souhaitées (voir le point 6.2 Ensemble de données demandé, à indiquer par source de données), à commencer un an avant la date du diagnostic et ensuite sur base annuelle jusqu'à l'année 2022 incluse. Par exemple, pour un patient diagnostiqué en mars 2015, l'AIM doit fournir les données à partir de mars 2014 jusqu'à l'année 2022 comprise (à condition qu'elles soient disponibles). Si l'approche mensuelle n'est pas possible, l'année complète sera fournie, ce qui signifie que pour l'exemple précité des données seront fournies pour la période de 2014 à 2022 inclus (à condition qu'elles soient disponibles).

VIH

15. En Belgique, la quasi-totalité de la population vivant avec le VIH reçoit une thérapie antirétrovirale (ART). L'AIM dispose des données de remboursement pour cette thérapie. Par ailleurs, les traitements ART sont principalement réservés aux personnes vivant avec le VIH. L'AIM est ainsi en mesure de sélectionner la population de patients de la cohorte VIH envisagée sans l'intervention de Healthdata, sur la base de critères prévus par les chercheurs de Sciensano qui sont décrits ci-dessus. Cette manière de procéder permet d'éviter le flux de données des PsINSZHD au départ de Healthdata, via la plate-forme eHealth qui déseudonymise en le NISS, vers l'AIM. La date du diagnostic peut également être déduite des données relatives au remboursement auprès de l'AIM et concerne la date du dernier test de diagnostic pour le VIH préalablement à l'initialisation de la thérapie ART.
16. Healthdata.be recevra les données de l'AIM pour la population VIH de la manière décrite à la figure 2. L'AIM consultera les données souhaitées à partir d'un an précédant la date du diagnostic (date du dernier test d'anticorps VIH enregistré préalablement à l'initialisation de la thérapie ART) et ensuite annuellement jusqu'en 2022 inclus. Par exemple, pour un patient avec un test d'anticorps VIH qui a été remboursé en mars 2015 et le premier remboursement de ART en juin 2014, l'AIM doit fournir les données à partir de mars 2014 jusqu'en 2022 inclus. Si l'approche mensuelle n'est pas possible, l'année complète sera communiquée, ce qui signifie que pour l'exemple précité des données sont fournies à partir de 2014 jusqu'en 2022 inclus (si elles sont disponibles).

Fondation Registre du Cancer

17. La BCR sélectionne la population de patients cancéreux sur la base des critères définis par les chercheurs de Sciensano et qui sont indiqués ci-dessus.

18. Healthdata.be recevra les données BCR pour la population EMPCAN 2.0 selon la figure 3, flux 6a/b.
19. Healthdata.be recevra les données de l'AIM souhaitées pour la population EMPCAN 2.0 selon la figure 3, flux 6a/b. L'AIM consultera les données à partir d'une période d'un an avant le mois/l'année de l'incident et ensuite sur base annuelle jusqu'en 2022 inclus. Par exemple, pour un patient avec une date d'incidence en février 2013, l'AIM doit fournir les données à partir de février 2012 jusqu'en 2022 inclus (si elles sont disponibles). Si l'approche mensuelle n'est pas possible, l'année complète sera communiquée, ce qui signifie que pour l'exemple précité des données sont fournies à partir de 2013 jusqu'en 2022 inclus (si elles sont disponibles).

II. COMPETENCE

20. En vertu de l'article 42, § 2, 3° de la loi du 13 décembre 2006 portant dispositions diverses en matière de santé, la chambre sécurité sociale et santé du Comité de sécurité de l'information est compétente pour rendre une autorisation de principe concernant toute communication de données à caractère personnel relatives à la santé.
21. Compte tenu de ce qui précède, le Comité estime qu'il peut se prononcer sur la communication de données à caractère personnel relatives à la santé, telle que décrite dans la demande d'autorisation.

III. EXAMEN

A. ADMISSIBILITÉ

22. Le traitement de données à caractère personnel est uniquement autorisé pour des finalités déterminées, explicites et légitimes et le traitement de données à caractère personnel relatives à la santé est en principe interdit.²
23. 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 la protection des données et prévoir des mesures appropriées et spécifiques pour la sauvegarde des droits fondamentaux et des intérêts de la personne concernée.³

Fondation Registre du Cancer

² Art. 9, point 1 RGPD.

³ Art. 9, point 2, j), du RGPD.

24. Ce traitement de données à caractère personnel trouve son fondement dans l'article 9, j), du RGPD et l'article 138 de la loi coordonnée du 10 mai 2015 *relative à l'exercice des professions des soins de santé*⁴.

Sciensano

25. Ce traitement des données à caractère personnel trouve son fondement dans l'article 4, § 4, de la loi du 25 février 2018 *portant création de Sciensano*.
26. En ce qui concerne le rôle du service Healthdata.be de Sciensano, il est fait référence à l'article 22, 20°, de la loi du 14 juillet 1994 *relative à l'assurance obligatoire soins de santé et indemnités* qui a créé un cadre permettant à l'INAMI de renforcer et de systématiser sa collaboration avec Sciensano, en particulier pour la coordination et le soutien des collectes de données censées renforcer les connaissances sur la santé de la population.

AIM

27. L'Agence intermutualiste (AIM) est une asbl qui a été créée en octobre 2002 par les unions nationales des organismes assureurs. L'objectif de l'AIM consiste à analyser les données recueillies par les organismes assureurs dans le cadre de leurs missions et de fournir les informations à ce propos. Les données sont recueillies pour la réalisation de leur mission légale en vertu des articles 278 à 281 de la loi-programme du 24 décembre 2002. Aussi pour la réalisation de finalités scientifiques pertinentes pour la politique. Pour les finalités scientifiques, l'AIM traite les données sur la base des délibérations rendues par la chambre sécurité sociale et santé du Comité de sécurité de l'information.
28. Par ailleurs, cette étude scientifique peut contribuer aux missions légales des OA telles qu'elles sont définies dans la loi du 6 août 1990 (art. 3a et b), à savoir participer à l'exécution de l'assurance obligatoire soins de santé et indemnités, régie par la loi coordonnée du 14 juillet 1994.
29. 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 pseudonymisées relatives à la santé envisagé.

B. FINALITÉ

30. Conformément à l'art. 5, b) du RGPD, le traitement de données à caractère personnel est uniquement autorisé pour des finalités déterminées, explicites et légitimes.

⁴ 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 la protection des données et prévoir des mesures appropriées et spécifiques pour la sauvegarde des droits fondamentaux et des intérêts de la personne concernée.

- 31.** Les trois principaux objectifs du projet SPADIS sont les suivants: (1) mesurer l'association entre fonctionnalité et participation sociale; (2) développer et mettre en place un ou des outils pour la collecte électronique de résultats et d'expériences rapportés par les patients (PROM et PREM) et (3) estimer l'impact de facteurs contextuels (par ex. systèmes de soins de santé et de la sécurité sociale) sur l'association entre fonctionnalité et participation sociale pour les patients diagnostiqués VIH, souffrant d'un cancer ou du diabète en Belgique.
- 32.** Pour atteindre ces objectifs, le projet est divisé en quatre volets distincts mais interdépendants. La présente demande concerne le volet 2 (WP2) qui a trait à l'objectif 1: mesurer l'association entre fonctionnalité et participation sociale sur la base du couplage de banques de données existantes.
- 33.** Le WP2-SPADIS est axé sur l'épidémiologie du fonctionnement, des activités et de la participation de patients vivant avec un cancer, le diabète ou le VIH. SPADIS définit ces concepts sur la base de la Classification internationale du fonctionnement, du handicap et de la santé (CIF).
- 34.** Les données utilisées pour satisfaire aux objectifs sont recueillies en couplant des données issues de registres existants à un niveau individuel, tant au sein qu'en dehors de Sciensano, à savoir l'Agence intermutualiste (AIM), la Fondation Registre du Cancer (BCR), la cohorte VIH et l'Initiative pour la promotion de la qualité et de l'épidémiologie du diabète sucré (IPQED).
- 35.** En couplant les données de l'AIM annuelles relatives aux positions sur le marché du travail de patients vivant avec un cancer, le diabète ou le VIH en Belgique, la participation sociale sur le marché du travail des populations d'étude décrites sera étudiée. Les analyses envisagées sont des analyses à caractère longitudinal et permettent de réaliser tant des analyses pronostiques qu'étiologiques des positions et transitions sur le marché du travail (p.ex. transition d'un travail vers une incapacité de travail ou une invalidité et à nouveau vers un travail après une incapacité de travail ou une invalidité). Par ailleurs, des données socio-démographiques et socio-économiques ainsi que des données cliniques sont également analysées (informations diagnostiques, observations cliniques, antécédents médicaux, etc.) en tant que déterminants potentiels des transitions sur le marché du travail.
- 36.** Il sera donc analysé s'il existe des différences (significatives) au niveau de (l'évolution de) la participation sociale (sur le marché du travail) entre les patients souffrant de différentes maladies chroniques en Belgique. Les chercheurs étudieront aussi les différences entre les patients diagnostiqués avec une même maladie afin d'identifier des inégalités socio-économiques et socio-démographiques éventuelles. Aussi, en raison de leur enregistrement dans le datawarehouse de l'AIM, le projet actuel permet d'évaluer l'impact d'avantages sociaux (p.ex. reconnaissance du handicap, BIM/OMNIO, statut malade chronique, etc.) sur la participation sociale de la population étudiée.
- 37.** De cette manière, le projet SPADIS permettra de développer des connaissances précieuses afin de mieux comprendre les déterminants de l'exclusion sociale (p.ex. inactivité sur le marché du travail de longue durée) de malades chroniques en Belgique.

38. Au vu des objectifs, le Comité considère que le traitement des données à caractère personnel envisagé poursuit bien des finalités déterminées, explicites et légitimes.

C. PROPORTIONNALITÉ

39. Conformément à l’art. 5, b) et c) du RGPD, les données à caractère personnel doivent être adéquates, pertinentes et non excessives au regard des finalités pour lesquelles elles sont obtenues et pour lesquelles elles sont traitées ultérieurement.
40. Le Comité fait observer que le NISS pseudonymisé est utilisé.
41. Pour chacun des fournisseurs de données et les données respectives, l’annexe contient une justification (scientifique) détaillée (en ce compris des références) de la sélection et de la proportionnalité. En ce qui concerne la proportionnalité, les chercheurs concernés de Sciensano prennent déjà des décisions concernant la transformation de variables pour lesquelles la forme originale est inutilement détaillée.
42. En vue de la justification des données demandées provenant de l’IPQED, de la cohorte VIH et de la BCR, la littérature a été étudiée en détail. De manière spécifique, il a été étudié pour chacune des variables contenues dans ces registres, dans quelle mesure elle peut être utilisée comme variable de substitution ou comme indication d’une fonctionnalité réduite. Sont ainsi retenus des variables et des résultats y liés qui, dans la littérature scientifique, ont déjà été mis en rapport avec le fardeau de la maladie, le handicap de certaines parties ou structures du corps et avec la mesure dans laquelle des individus peuvent exécuter des activités. Il est important de noter que les données décrites ici ne représentent qu’une sélection limitée de variables issues des registres concernés. En ce qui concerne les données de l’AIM demandées, ont été sélectionnées les données qui contribuent à une définition aussi précise que possible des positions sur le marché du travail et des déterminants socio-démographiques et socio-économiques éventuels.

Informations diagnostiques (BCR, IPQED, cohorte VIH)

43. Il existe tant pour le cancer, le diabète que le VIH des différences importantes au niveau de l’impact sur la fonctionnalité, la rapidité de progression de la maladie, la symptomatologie, la pathophysiologie, etc. en fonction du type de cancer (p.ex. cancer du sein versus cancer hématologique), du type de diabète (type 1 et 2) et du type de VIH (type 1 et 2). Par ailleurs, la date du diagnostic est importante afin de tenir compte du temps écoulé depuis le diagnostic, étant donné qu’il s’agit de maladies chroniques, souvent progressives. Les données relatives au stade de la maladie doivent également être prises en compte, pour la même raison évidente. En ce qui concerne la stadification du cancer, la variable Stade du cancer donne la stadification Ann-Arbor pour les hémopathies malignes et le « stade combiné du cancer », plus courant, pour les tumeurs solides.

Informations thérapeutiques (BCR, IPQED, cohorte VIH, AIM)

44. L'inclusion d'informations thérapeutiques dans la demande de données actuelle poursuit deux objectifs principaux. D'une part, la nécessité d'une thérapie déterminée permet de donner une indication plus détaillée de la rapidité de progression ou du stade de la maladie (par exemple, en ce qui concerne les antidiabétiques oraux, l'insuline est utilisée dans les cas de DT2 plus avancés, le traitement antirétroviral du VIH permet de retarder la progression de la maladie ou l'hospitalisation pour une exacerbation, la transplantation pour les patients atteints d'un cancer avancé chez qui les thérapies conventionnelles n'ont pas d'effet, etc.). D'autre part, il y a lieu de tenir compte des effets secondaires (imprévus) éventuels des thérapies mêmes comme la radio- et la chimiothérapie en cas de cancer. Par ailleurs, les hospitalisations éventuelles sont aussi consultées (AIM, cohorte VIH), étant donné qu'il s'agit aussi d'une indication du stade avancé de la maladie. En outre, les hospitalisations peuvent aussi avoir un impact important sur la fonctionnalité des patients.

Observations cliniques en cas de visite à un hôpital ou lors de consultations (BCR, cohorte VIH, IPQED)

45. Les observations formulées pendant une consultation peuvent constituer une indication pour la progression de la maladie. Ceci est notamment le cas pour des IMC faibles ou des analyses de détection d'un déficit immunitaire (CD4) et de la charge virale (ARN-VIH) dans le cas du VIH, et des valeurs de l'hémoglobine glyquée (HbA1c) pour évaluer le contrôle glycémique dans le cas du diabète. Le score de l'OMS au moment du diagnostic représente le score de performance d'un patient au moment du diagnostic du cancer et fournit une mesure du fardeau de la maladie individuel et des limitations de fonctionnement.

Informations relatives aux antécédents médicaux (IKED) et aux maladies opportunistes (cohorte VIH)

46. Une grande partie de l'impact d'une maladie sur la fonctionnalité d'un patient provient des complications à court et à long terme associées à cette maladie. Par exemple, un épisode d'hypoglycémie a un impact significatif sur la vie d'un patient diabétique. D'autres complications à long terme du diabète et ses traitements ont également un impact significatif et évident. Il s'agit par exemple de la rétinopathie ou de la maculopathie diabétique, des amputations des membres inférieurs, des accidents ischémiques transitoires, etc. Pour le VIH, la consultation de données relatives aux maladies opportunistes, en raison de la suppression immunitaire qui le caractérise, se fonde sur un raisonnement similaire.

Informations socio-démographiques (AIM)

47. D'importantes informations socio-démographiques sont enregistrées dans les banques de données de l'AIM. Ces données sont essentielles pour constater au sein de SPADIS des inégalités au niveau de la participation sociale en fonction de certains sous-groupes socio-démographiques. Font notamment partie de ces données, l'âge d'un individu, le sexe et la région du domicile. Sur la base de différentes données de l'AIM, il est aussi possible de tenir compte de la composition et de la taille du ménage.

Informations socio-économiques (AIM) et avantages / statuts sociaux (AIM)

48. Des données relatives aux revenus au niveau du secteur statistique sont utilisées afin d'identifier des inégalités au niveau du revenu ainsi que d'examiner leur rôle au niveau de la prédiction des transitions sur le marché du travail. Le type de revenu des personnes inactives est également étudié afin de mener les recherches nécessaires concernant l'impact des différentes sources de revenus sur la probabilité des transitions sur le marché du travail. Pour cartographier de manière plus détaillée les formes de soutien financier accordées, des données relatives au bénéfice de différents types d'avantages ou de statuts sociaux sont demandées à l'IMA. Par exemple, des données relatives aux interventions majorées, au bénéfice d'un forfait pour maladies chroniques, d'un statut de malade chronique, d'une aide d'un CPAS ou relatives à la catégorie de ménage MAF, etc. Ces données permettent aussi d'obtenir des informations supplémentaires concernant le revenu du ménage de l'individu, étant donné que divers avantages sociaux y sont liés (p.ex. catégorie MAF du ménage).

Position sur le marché du travail

49. Des données longitudinales (annuelles) relatives aux positions sur le marché du travail permettent de définir des transitions sur le marché du travail et de les intégrer comme résultats principaux dans les analyses prévues. Différentes données de l'AIM relatives à l'incapacité de travail, à l'invalidité, à la pension et au décès permettent aux chercheurs du projet SPADIS de définir ces positions de manière précise.
50. Le Comité fait observer qu'une analyse de risque 'small cell' (SCRA) par P95 est prévue.
51. Le Comité fait observer que la Plate-forme eHealth intervient comme tierce partie de confiance (TTP) en vue du couplage et de la pseudonymisation des données⁵.
52. Il est nécessaire que la Plate-forme eHealth conserve le lien entre l'ensemble de données pseudonymisées et l'identité de la personne concernée. Le couplage des données décrit concerne un couplage unique. Pour cette raison, il n'est pas nécessaire de conserver un lien entre l'ensemble de données pseudonymisées et l'identité des personnes. Cependant, le lien doit être conservé jusqu'après l'évaluation finale des données reçues. Si les données obtenues ne correspondent pas aux données demandées, une correction doit avoir lieu auprès des fournisseurs de données (BCR, AIM). Dès que les données obtenues sont considérées comme complètes, le lien entre les données pseudonymisées et l'identité des personnes peut être supprimé.
53. La Plate-forme eHealth est chargée de la pseudonymisation des numéros NISS. La Plate-forme eHealth est ensuite autorisée à conserver le lien entre l'ensemble de données pseudonymisées et le numéro d'identification réel, et ce pendant la durée de l'étude (01/01/2030). Le Comité autorise la Plate-forme eHealth à conserver les clés de codage, et ce pendant la durée complète de l'étude. Ces clés seront ensuite détruites.

⁵ En vertu de l'article 5, 8°, de la loi du 21 août 2008 relative à l'institution et à l'organisation de la Plate-forme eHealth, la Plate-forme eHealth est autorisée à conserver les clés de codage utilisées pendant toute la durée de l'étude.

54. Le Comité précise que les données obtenues sur la base de la présente délibération ne seront pas couplées à d'autres données qui ont été obtenues sur la base de délibérations antérieures.

D. LIMITATION DE LA CONSERVATION

55. Conformément à l'article 5, 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^{er}, 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).
56. Les données demandées seront utilisées pour réaliser les finalités d'étude du projet SPADIS. Pour réaliser ces objectifs, les données doivent être mises à la disposition des chercheurs du projet SPADIS pendant une période de 60 mois ou 5 ans. Les données doivent donc être disponibles pour les chercheurs jusqu'en janvier 2030.
57. Ce délai offre la marge nécessaire dans le cadre des publications qui sont particulièrement longues. Pendant cette période, des comparaisons peuvent être réalisées entre les résultats SPADIS et une étude externe, tant au niveau national qu'international. Par ailleurs, le délai permet aux chercheurs du projet SPADIS de répondre à des questions de suivi concernant les résultats SPADIS qui sont formulées par les autorités, les administrations et autres institutions. Les analyses peuvent par conséquent être adaptées aux besoins de ces instances et autres organisations.

E. TRANSPARENCE

58. Conformément à l'art. 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.

Les articles 13 et 14 du RGPD fixent les conditions auxquelles le responsable du traitement doit satisfaire lorsque des données à caractère personnel sont collectées concernant la personne concernée. Ainsi, les informations suivantes doivent notamment être communiquées: les coordonnées du responsable du traitement et du délégué à la protection des données, les catégories de données à caractère personnel si les données ne sont pas obtenues auprès des personnes concernées, les finalités du traitement et le fondement du traitement, les catégories de destinataires et, si le responsable du traitement a l'intention de transmettre les données à caractère personnel à un destinataire dans un pays tiers, les garanties appropriées.

59. Le responsable du traitement est dispensé de l'obligation de communication d'informations aux personnes dont les données à caractère personnel sont traitées dans la mesure où cette communication s'avère impossible ou demande des efforts disproportionnés. C'est le cas en l'espèce. Le fait de contacter tous les individus de la population étudiée exigerait de gros efforts disproportionnés et une dépseudonymisation de l'identité des individus des populations étudiées qui n'est pas souhaitable. Une déclaration vie privée spécifique ("*privacy notice*") a été rédigée et sera publiée sur la page internet de Sciensano dans le but de communiquer en toute transparence concernant la vie privée (article 14, 5, b) RGPD).
60. Le Comité est par conséquent d'avis que la demande répond aux exigences de transparence.

F. MESURES DE SÉCURITÉ

61. Conformément à l'article 5, f) du RGPD, le demandeur doit prendre toutes les mesures techniques et organisationnelles appropriées pour protéger les données à caractère personnel. Ces mesures doivent garantir un niveau de protection adéquat, compte tenu, d'une part, de l'état de la technique en la matière et des frais qu'entraîne l'application de ces mesures et, d'autre part, de la nature des données à protéger et des risques potentiels.
62. 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.
63. Le Comité fait observer que tant l'Agence intermutualiste, la Fondation Registre du Cancer que l'AIM ont désigné un délégué à la protection des données et un professionnel des soins de santé responsable.
64. Le Comité constate ensuite qu'une analyse d'impact relative à la protection des données a été réalisée par Sciensano.
65. Conformément à l'article 9, point 3, du RGPD, le traitement de données à caractère personnel relatives à la santé peut uniquement être effectué sous la surveillance et la responsabilité d'un professionnel des soins de santé. C'est le cas en l'espèce. Le Comité rappelle que lors du traitement de données à caractère personnel, le professionnel des soins de santé ainsi que ses préposés ou mandataires sont soumis au secret, conformément à l'article 458 du Code pénal.
66. Les collaborateurs sont tenus par un devoir de confidentialité sur base de l'article 458 du Code pénal, de l'article 138, § 4, alinéa 2, de la loi coordonnée du 10 mai 2015, de l'article

17, 3°, a) de la loi du 3 juillet 1978 relative aux contrats de travail et d'une obligation contractuelle.

67. Le Comité fait observer que la Plate-forme eHealth intervient comme tierce partie de confiance (TTP) en vue du couplage et de la pseudonymisation des données.
68. Le Comité attire explicitement l'attention sur les dispositions du Titre 6 de la loi du 30 juillet 2018 relative à la protection des personnes physiques à l'égard des traitements de données à caractère personnel, qui prévoit des sanctions administratives et pénales sévères dans le chef du responsable du traitement et des sous-traitants pour la violation des conditions prévues dans le RGPD et la loi du 30 juillet 2018 précitée.

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 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.

La présente délibération entre en vigueur le 18 septembre 2024.

Michel DENEYER
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
--

Bijlage 1 - Schematisch overzicht van de gegevensstromen + toelichting

We onderscheiden **drie verschillende gegevensstromen**:

1. Het verkrijgen van de gegevens van het IMA voor een deel van de bestaande IKED populatie (**Proces A**).
2. Het verkrijgen van de gegevens van IMA voor een deel van de HIV-populatie (**Proces B**).
3. Het verkrijgen van gegevens van IMA en van BCR voor een deel van de populatie van personen die gediagnosticeerd werden met één van de kankertypes beschreven in deze aanvraag (**Proces C**).

Deze afzonderlijke gegevensstromen resulteren in **drie losstaande SPADIS-registers** waarbij **subsets van gegevens of variabelen** uit het IKED, de HIV cohort, en BCR worden gekoppeld met de gewenste IMA gegevens: respectievelijk SPADIS-IKED (*SPADIS-IKED* in figuur 1), SPADIS-HIV (figuur 2) en SPADIS-EMPCAN2.0 (BCR, figuur 3).

Belangrijk:

De gegevens waarover IMA beschikt, kunnen gebruikt worden om op een indirecte manier individuen te identificeren die gekenmerkt worden door een lagere socio-economische status (SES). Zo kan informatie omtrent het al dan niet krijgen van OCMW steun een indicatie zijn voor een lage SES op basis van inkomen. Echter, dergelijke sociale statuten bieden geen informatie voor zij met een hogere SES en laten dus niet toe om verschillen te identificeren tussen zij met een gemiddelde en hoge SES.

Om de volledige socio-economische gradiënt te onderzoeken en diens effect op *sociale participatie* wordt er gebruik gemaakt van de variabele 'statistische sector' en een open dataset ontwikkeld door **Statbel**. Gebaseerd op de Belgische Censusgegevens uit 2021, verzameld en geanalyseerd door Statbel, kunnen inkomensdecielen worden berekend voor de Belgische bevolking op basis van de belastingaangiftes van elke wettelijke inwoner van België in 2021.

De statistische sector is een maatstaf die de kleinste geografische eenheid vertegenwoordigt waarvoor Statbel sociaaleconomische statistische analyses uitvoert. De IMA Populatie Database registreert ook de statistische sector voor elke individuele persoon (PP0055).

Statbel ontwikkelt een open dataset die de verhouding van individuen in elke inkomensdeciël beschrijft voor elke statistische sector. Deze dataset wordt aan IMA verstrekt, dat de 9 variabelen koppelt aan de andere gewenste IMA-gegevens op basis van de geregistreerde statistische sector voor elk individu van de SPADIS-populaties. Vervolgens bezorgt IMA-gezondheidsgegevens met de variabelen van belang (inclusief inkomensdeciëlvariabelen) aan healthdata.be zonder de statistische sectorvariabele (per flow 2b in Figuur 1, flow 1b in Figuur 2 en flow 3b in Figuur 3).

Op deze manier wordt vermeden dat de statistische sector in de datawarehouse van healthdata.be wordt gekoppeld aan de andere gegevens die beschreven worden in deze aanvraag. Statistische sector is een gevoelige variabele die de identificatie van personen zou kunnen faciliteren en verantwoordelijk zou zijn voor een aanzienlijke *small cell risk* wanneer gekoppeld aan de andere gegevens in de huidige aanvraag.

Toelichting:**PROCES A – IKED**

De IMA-variabelen worden opgevraagd bij IMA volgens **Figuur 1**. In **bijlage 9** kunnen de figuren voor de gegevensstromen in hogere resolutie teruggevonden worden.

De patiëntenpopulatie (Gepseudonimiseerde INSZ (PsINSZ_{HD}) samen met de datum van diagnose ('date_of_diagnosis_of_diabetes') worden door healthdata.be naar NIC/CIN gestuurd via eHealth, voor decodering van de INSZ (**flow 1**). Deze gegevens uit het healthdata.be datawarehouse (DWH) hebben een gepseudonimiseerde patientID dat werd gecreëerd door TTP eHealth door middel van een specifiek healthdata.be pseudonimisatie algoritme.

De ontvangen gedepseudonimiseerde INSZ nummers via eHealth worden vervolgens gepseudonimiseerd door NIC/CIN door middel van diens eigen specifieke pseudonimisatie mechanisme. Deze worden vervolgens doorgestuurd, met de 'date of diagnosis' en technische identificatie codes (*technical ID*, gemaakt door NIC/CIN) naar de TTP van de verzekeringsinstellingen (TTP OA/VI (BCSS/KSZ)) – **flow 2**)).

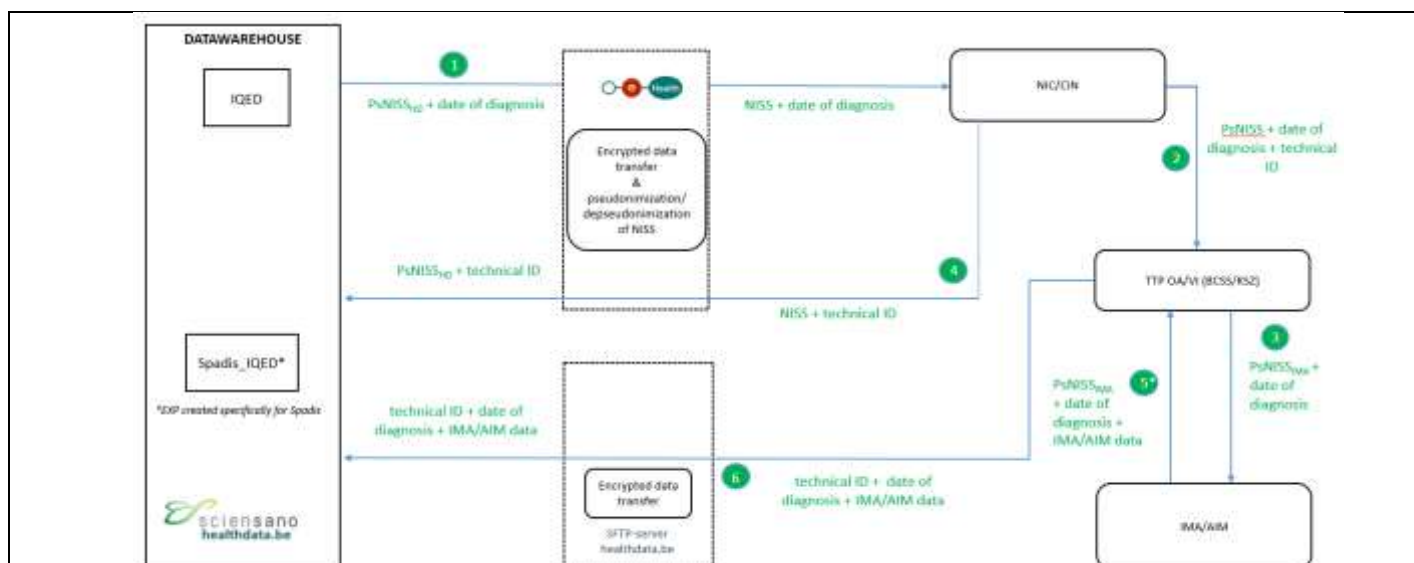
De TTP van de verzekeringsinstellingen zal dan een tweede pseudonimisatie uitvoeren van de gepseudonimiseerde identificatie nummers (PsINSZ_{IMA}) en ze doorsturen, samen met de variabele 'date of diagnosis' maar zonder de technische identificatiecodes, naar het IMA/AIM waar de selectie van de gegevens plaats kan vinden (**flow 3**).

De NIC/CIN stuurt de INSZ met de technische identificatie codes (technical ID, gemaakt door NIC/CIN) naar eHealth. EHealth pseudonimiseert de INSZ naar het healthdata.be pseudonimisatie algoritme en stuurt een correspondentietabel "PsINSZ_{HD}/technical ID" naar het healthdata.be DWH – **flow 4**.

Op basis van de INSZ (PsINSZ_{IMA}) en de 'date of diagnosis' verzamelt IMA/AIM de gegevens en stuurt ze door naar de TTP van de verzekeringsinstellingen (TTP OA/VI (BCSS/KSZ)) – (**flow 5***) die de PsINSZ_{IMA} vervangt met de technische identificatiecode (technical ID).

De TTP van de verzekeringsinstellingen stuurt de data samen met de technische identificatie codes naar healthdata.be via SFTP (**flow 6**).

De *technical ID* wordt versleuteld door de verzender, terwijl de INSZ wordt gecodeerd door eHealth, als TTP. Bij ontvangst consolideert het platform van healthdata.be de afzonderlijke communicaties op basis van de *technical ID*. Na consolidatie en technische kwaliteitscontrole wordt de *technical ID* onmiddellijk en permanent verwijderd uit de infrastructuur van healthdata.be. Het platform van healthdata.be zal een logboek bijhouden van deze technische processen



Figuur 1. Gegevensstromen voor het beoogde SPADIS-IKED register (Proces A)

PROCES B – HIV

De IMA-variabelen worden aangevraagd bij IMA volgens **Figuur 2**. In **bijlage 9** kunnen de figuren voor de gegevensstromen in hogere resolutie teruggevonden worden.

De patiëntenpopulatie wordt geselecteerd door IMA en wordt gedefinieerd door de individuen voor wie IMA de vergoeding voor antiretrovirale therapie (ART) en de leeftijd bij de diagnose (tussen 18 en 60 jaar voor diagnoses tussen 1 januari 2003 en 31 december 2021) heeft geregistreerd: de diagnose-datum wordt gedefinieerd door de laatst geregistreerde vergoeding van een HIV-antilichaamtest (zie **bijlage 3** voor lijst met ATC codes voor ART ter selectie van studiepopulatie) voorafgaand aan de start van ART.

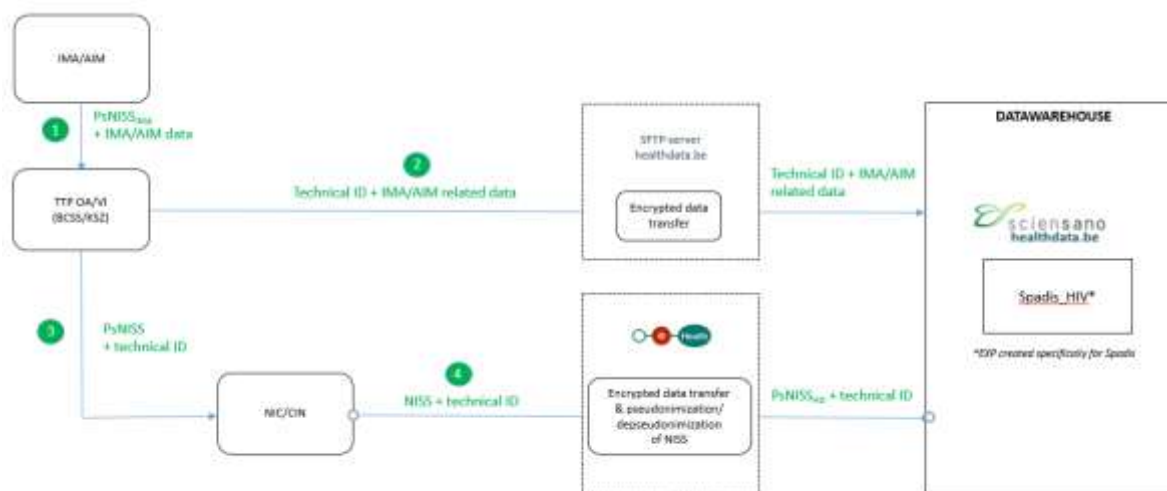
Op basis van de vergoeding voor ART en de diagnose-datum haalt IMA de gegevens op en stuurt ze naar de TTP van de verzekeringsinstellingen (TTP OA/VI (BCSS/KSZ)) – **flow 1**.

De TTP van de verzekeringsinstellingen (TTP OA/VI (BCSS/KSZ)) vervangt de PsINSZ_{IMA} met een technische identificatie code (technical ID, gecreëerd door TTP OA/VI) en verstuurt de gegevens samen met de *technical ID* naar healthdata.be via SFTP (**flow 2**).

De TTP van de verzekeringsinstellingen (TTP OA/VI (BCSS/KSZ)) bezorgt de PsINSZ en dezelfde *technical ID* aan NIC/CIN (**flow 3**).

De NIC/CIN creëert en verstuurt de INSZ van de patiënt samen met de *technical ID* (gegenereerd door TTP OA/VI), via eHealth (die de INSZ vervangt door de PsINSZ_{HD} en de correspondentietabel (PsINSZ_{HD}/technical ID) transfereert), naar het healthdata.be DWH – **flow 4**.

De *technical ID* wordt versleuteld door de afzender, terwijl de INSZ wordt gecodeerd door eHealth, als TTP. Bij ontvangst consolideert het platform van healthdata.be de afzonderlijke communicaties op basis van de *technical ID*. Na consolidatie en technische kwaliteitscontrole wordt de *technical ID* onmiddellijk en permanent verwijderd uit de infrastructuur van healthdata.be. Het platform van healthdata.be zal een logboek bijhouden van deze technische processen.



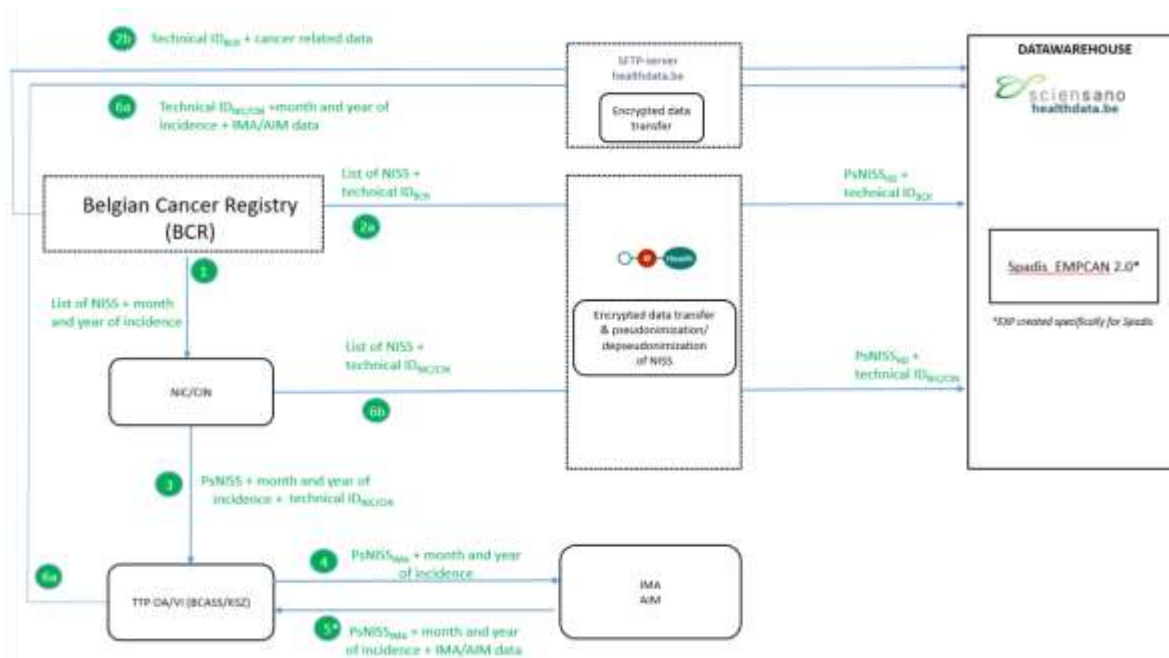
Figuur 2. Gegevensstromen voor het beoogde SPADIS-HIV register (Proces B).

PROCES C – BCR

Data van IMA en van BCR worden aangevraagd volgens **Figuur 3**. In **bijlage 9** kunnen de figuren voor de gegevensstromen in hogere resolutie teruggevonden worden.

1. Stichting Kankerregister (BCR) levert de INSZ (de populatie wordt geselecteerd door BCR op basis van de selectiecriteria die zijn verstrekt door de Sciensano onderzoekers - zie 2.2 *Personen van wie de persoonsgegevens zullen worden verwerkt*) en de maand en het jaar van incidentie aan NIC/CIN (**flow 1**).
2. BCR verstuurt, voor deze populatie, de gegevens naar healthdata.be in 2 afzonderlijke stromen of *flows*:
 - **Flow 2a** bevat de INSZ van de patiënt met een *technical ID* (die wordt gegenereerd door BCR). De gegevens worden overgedragen via eHealth (die de INSZ vervangt door PsINSZ_{HD} en de correspondentietabel (PsINSZ_{HD} / *technical ID*) overdraagt naar het healthdata.be DWH.
 - **Flow 2b** bevat de gegevens gerelateerd aan kanker met dezelfde *technical ID*. Omdat dit tweede bestand geen gegevens bevat die moeten worden gepseudonimiseerd, wordt het bestand rechtstreeks vanuit BCR naar het HD-DWH van healthdata.be verzonden via SFTP of een andere methode voor bestandsoverdracht.
3. NIC/CIN, nadat zij de lijst van INSZ en de maand/jaar van incidentie van BCR hebben ontvangen, pseudonimiseert de INSZ en verstuurt ze, samen met de gegevens voor de maand en het jaar van incidentie en de *technical ID* (gecreëerd door NIC/CIN) naar de TTP van de verzekeringsinstellingen (TTP OA/VI (BCSS/KSZ)) – **flow 3**.
4. De TTP van de verzekeringsinstellingen voorziet een tweede pseudonimisatie van de gepseudonimiseerde identificatienummers (PsINSZ_{IMA}) en verzendt deze, samen met de gegevens voor de maand en het jaar van incidentie maar zonder de technische identificatiecodes (*technical ID*) naar IMA/AIM waar de selectie van de gevraagde gegevens kan plaatsvinden (**flow 4**).
5. Gebaseerd op de INSZ (PsINSZ_{IMA}) en de maand en het jaar van incidentie, verzamelt IMA/AIM de gevraagde gegevens en bezorgt deze aan de TTP van de verzekeringsinstellingen (TTP OA/VI (BCSS/KSZ)) – **flow 5**, die de PsINSZ_{IMA} vervangt met de *technical ID*.

6. De TTP van de verzekeringsinstellingen stuurt de gegevens, samen met de *technical ID*, naar healthdata.be via SFTP (**flow 6a**). NIC/CIN verstuurt de INSZ van de patiënt en de *technical ID* (gecreëerd door NIC/CIN en verschillend aan de *technical ID* genereerd door BCR (flow 2a)), via eHealth (die de INSZ vervangt met de PsINSZ_{HD} en de correspondentie tabel (PsINSZ_{HD} / *technical ID*) doorstuurt) naar het healthdata.be DWH – **flow 6b**. De *technical ID* wordt versleuteld door de verzender, terwijl de INSZ wordt gecodeerd door eHealth, als Trusted Third Party (TTP). Bij ontvangst consolideert het platform van healthdata.be de afzonderlijke communicaties op basis van de technische ID. Na consolidatie en technische kwaliteitscontrole wordt de *technical ID* onmiddellijk en permanent verwijderd uit de infrastructuur van healthdata.be. Het platform van healthdata.be zal een logboek bijhouden van deze technische processen.



Figuur 3. Gegevensstromen voor het beoogde SPADIS-EMPCAN 2.0 (BCR) (Proces C)

Bijlage 2 . Overzicht gevraagde gegevens per databron, bewijs van proportionaliteit, gebruikte referenties

IMA: Overzicht gevraagde gegevens, bewijs van proportionaliteit, gebruikte referenties

Dataset	Naam	Beschrijving	Bewijs van proportionaliteit
Populatie database	PP0015A	Geboortejaar	Het aanvragen van sociaaldemografische gegevens is essentieel om binnen SPADIS mogelijke ongelijkheden vast te stellen inzake sociale participatie. Op basis van het geboortejaar en de geboortemaand kunnen de SPADIS onderzoekers de leeftijd van de studiepopulatie in rekening brengen bij de voorziene analyses.
Populatie database	PP0015B	Geboortemaand	
Populatie database	PP0020	Geslacht	Op basis van de variabele 'Geslacht' kunnen de SPADIS onderzoekers de mogelijke verschillen met betrekking tot sociale participatie en diens determinanten in kaart brengen tussen mannen en vrouwen in België.
Populatie database	PROVINCE	Provincie (hercoderen naar Regio*)	Om na te gaan in welke mate er ongelijkheden bestaan met betrekking tot sociale participatie en diens determinanten tussen individuen die leven in Vlaanderen, Wallonië en het Brussels Hoofdstedelijk Gewest wordt de variabele 'Provincie' aangevraagd in een getransformeerde vorm (transformeren van <i>provincie</i> naar <i>regio</i>)
Populatie database	FAM_SIZE	Gezinsgrootte	Bij het bekijken van transities op de arbeidsmarkt toont de literatuur dat financiële incentives gepaard aan het terugkeren naar of behouden van actieve tewerkstelling van belang zijn ^{1,2,3} . Naast een drijfveer om actieve tewerkstelling te behouden of bekomen, heeft financiële stress ook een impact op de (mentale) gezondheid ⁴ en dient het ook mee in kaart gebracht te worden om gezondheid op een brede manier te benaderen. De SPADIS studiepopulatie betreft een volwassen populatie. Om naast de financiële middelen ook in kaart te brengen hoeveel personen er in het gezin ten laste zijn van een individu in de SPADIS populatie wordt de variabele 'gezinsgrootte' aangevraagd. Als de persoon zelf ten laste is van een andere volwassene in het gezin, kan dit bevestigd worden door naar variabel PP1002 te kijken.
Populatie database	PP4008	Eénoudergezin	Inzake de gezinsgrootte en -samenstelling kan een onderscheid gemaakt worden tussen één- en tweoudergezinnen. Voor éénoudergezinnen tracht SPADIS de hypothese te bevestigen dat het dragen van de volledige financiële verantwoordelijkheid (financiële stress) voor het hele gezin een invloed kan uitoefenen op de onderzochte arbeidsmarktposities en -transities.
Populatie database	PP1002	Gerechtigde of persoon ten laste	Deze variabele geeft aan of de persoon in kwestie bij het ziekenfonds is aangesloten als 'gerechtigde' of 'persoon ten laste'. Deze laatste categorie verwijst naar personen die geen eigen inkomen hebben (als echtgenoot of samenwonende, als descendent van de gerechtigde of als ascendent van de gerechtigde) en hun recht op ziekteverzekering ontlenen aan een gezinslid dat eigen rechten heeft. Deze variabele helpt aldus om een meer gedetailleerd beeld te schetsen van de financiële mogelijkheden van het gezin en de financiële verantwoordelijkheden die rusten op de schouders van het individu in de SPADIS studiepopulatie.

Populatie database	PP1008	Aard inkomen	Deze variabele beschrijft de aard van het inkomen van niet-actieve personen en betreft dus niet het bedrag van het inkomen. Deze variabele draagt bij aan het in kaart brengen van de financiële beperkingen die door het individu ondervonden worden. De categorieën van deze variabele beschrijven of een individu geniet van een (equivalent) leefloon, een inkomen heeft kleiner of gelijk aan het leefloon, een volledige persoonlijke bijdrage betaalt voor de ziekteverzekering of een bruto-(gezins)inkomen heeft dat lager ligt dan bepaalde grenswaarden.
Populatie database	PP1009	Oorsprong erkenning als persoon met handicap	Om de financiële middelen die ter beschikking staan van het individu verder in kaart te brengen dienen de sociale voordelen of statuten die werden toegekend aan het individu meegenomen te worden in de voorziene SPADIS analyses. Om een handicap vast te stellen wordt o.a. gebruik gemaakt van het ICF-model om langdurige en belangrijke participatieproblemen in kaart te brengen, net zoals in het SPADIS project. Deze variabele geeft aan welk traject werd gevolgd om een erkenning te krijgen als persoon met een handicap.
Populatie database	PP1010	Voorwaarde recht verhoogde tegemoetkoming	In lijn met de hierboven vermelde variabelen geeft deze variabele weer of het individu recht heeft op een verhoogde tegemoetkoming voor gezondheidszorgkosten, zij het op basis van inkomensvoorwaarden of op basis van het reeds ontvangen van een sociale uitkering. Deze variabele helpt om meer inzichten te verkrijgen in de financiële druk die individuen ondervinden. Aangezien de SPADIS populatie volwassenen betreft die leven met een chronische aandoening en bijgevolg geconfronteerd worden met relatief hogere gezondheidszorgkosten, is het belangrijk om inzichten te verwerven over de mate waarin zij financieel ondersteund worden om deze kosten te dekken.
Populatie database	PP3015	Statuut Chronische aandoening – financieel criterium	Verband houdend met de variabele <i>PP1010: Voorwaarde recht verhoogde tegemoetkoming</i> , schetst de huidige variabele of een individu gedurende twee jaar uitgaven had voor geneeskundige verzorging van minstens €300 per kwartaal. Door deze variabele op te nemen kan worden bekeken in welke mate het individu hoge uitgaven heeft inzake gezondheidszorg en of het individu toegang heeft tot bijkomende financiële ondersteuning om hiermee om te gaan.
Populatie database	PP3002	Categorie MAF Gezin	De MAF-categorie geeft aan of het gezin van de persoon een terugbetaling kan verkrijgen in het kader van de Maximumfactuur (MAF) op basis van sociale MAF of de inkomens MAF. De sociale MAF wordt toegekend aan het gezin van individuen met een verhoogde tegemoetkoming. De inkomens-MAF bekijkt de gezamenlijk inkomens van alle gezinsleden om 5 inkomenscategorieën te koppelen aan 5 verschillende remgeldplafonds. Deze variabele laat aldus toe om inzichten te verwerven met betrekking tot de financiële middelen die ter beschikking staan van het gezin en of het/een individu in het gezin een verhoogde tegemoetkoming geniet.

Patiënten-databank	Chronical_YN	Chronische ziekte	De variabele 'chronische ziekte' geeft aan of het individu voldoet aan één van de afhankelijkheidssituaties die de voorwaarden vormen voor de toekenning van het forfait voor chronische ziekten of het recht op een toelage voor handicaptten. In combinatie met de variabele 'zorgforfait chronische ziekte' kan een onderscheid gemaakt worden tussen zij met recht op het zorgforfait en zij met recht op een toelage voor gehandicaptten.
Patiënten-databank	ZORGFORFAIT	Zorgforfait chronische ziekte	Deze variabele geeft weer of een individu een zorgforfait voor chronische ziekten krijgt. Personen hebben recht op deze tegemoetkoming voor een bepaald jaar wanneer ze in dat jaar en het jaar ervoor hoge zorguitgaven (remgelden) hebben en sterk afhankelijk zijn van anderen. Het zorgforfait vertegenwoordigt een belangrijke financiële ondersteuning die binnen het SPADIS project in rekening wordt gebracht als mogelijke determinant van arbeidsmarkttransities.
Populatie database	PP3013	OCMW steun	Financiële ondersteuning van het OCMW of CPAS wordt toegekend aan zij die behoeftig zijn en zelf niet over voldoende middelen beschikt om een menswaardig leven te leiden. Deze variabele stelt het SPADIS onderzoeksteam in staat om verschillen naargelang socio-economische status in kaart te brengen voor individuen met lage inkomens (recht op OCMW steun).
Populatie database	PP0030	Code gerechtigde	Deze variabele geeft weer volgens welke voorwaarde een persoon (of de gerechtigde waarvan hij/zij ten laste is) verzekerd is. Met behulp van de verschillende categorieën kan onder andere een onderscheid gemaakt worden tussen zelfstandigen en niet-zelfstandigen bij het onderzoeken van (determinanten van) arbeidsmarkttransities, waarvoor de literatuur significante verschillen aantoon ⁶ . Verder kunnen categorieën van deze variabele gebruikt worden om belangrijke arbeidsmarktposities en -transities (uitkomsten in voorziene SPADIS analyses) te definiëren: actief op de arbeidsmarkt (antwoordcategorie 1: actief), werkzoekend inactief op de arbeidsmarkt (antwoordcategorie 2: invaliden of mindervaliden), gepensioneerd (antwoordcategorie 3: gepensioneerden). Het verkrijgen van deze informatie, op jaarlijkse basis, is daarom essentieel voor het SPADIS project om de voorziene analyses uit te voeren.
Populatie database	PP1003	Sociaal statuut van de gerechtigde	Deze variabele geeft het sociaal statuut weer waaraan de gerechtigde zijn/haar recht op de verplichte ziekteverzekering ontleent. De variabele vertegenwoordigt het (laatste) statuut van tewerkstelling voor werkenden, werklozen en mindervaliden. Deze variabele laat toe om meer detail te schetsen in de hoedanigheid van tewerkstelling of inactiviteit op arbeidsmarkt.
Populatie database	PP1004	Type werkloosheidsuitkering	Deze variabele stelt het SPADIS onderzoeksteam in staat om een onderscheid te maken tussen werkenden en werklozen voor de personen die actief zijn op de arbeidsmarkt. Verder kunnen de antwoordcategorieën met betrekking tot Brugpensioenen (9 en 10) gebruikt worden om personen te identificeren die reeds voor de officiële pensioensleeftijd de actieve arbeidsmarkt verlaten. Deze informatie is essentieel om pensioen op een complete manier als <i>competing event</i> te introduceren in de voorziene SPADIS <i>time-to-event</i> analyses.

Populatie database	PP4002	Aantal dagen arbeidsongeschiktheid	<p>Op basis van hierboven vermelde variabelen kunnen jaarlijkse transities naar arbeidsongeschiktheid en invaliditeit in kaart gebracht worden. De huidige variabelen laten toe om een meer gedetailleerde tijdslijn op te stellen door rekening te houden met hoeveel dagen in het jaar de persoon arbeidsongeschikt of invalide was. Deze informatie is essentieel aangezien ze gebruikt zal worden om de voornaamste <i>events-of-interest</i> te definiëren in de voorziene <i>time-to-event</i> analyses.</p> <p>In de voorziene analyses van arbeidsmarkttransities dient rekening gehouden te worden met het feit dat individuen kunnen overlijden gedurende de retrospectieve follow-up. Door het jaar en de maand van overlijden op te nemen in de gegevensaanvraag, kan overlijden dus worden opgenomen in de SPADIS analyses als <i>competing event</i>. Rekening houdend met de proportionaliteit van de gegevens in de huidige aanvraag werd de dag van de overlijdensdatum niet aangevraagd aangezien dit niveau van detail niet vereist is.</p> <p>Een ziekenhuisopname kan een aanzienlijke impact hebben op de functionaliteit van een individu. Hospitalisaties worden in verband gebracht met een significante reductie van <i>health-related quality of life</i>. Deze variabele wordt in de huidige gegevensaanvraag opgenomen om de relatie met mogelijke arbeidsmarkttransities (sociale participatie) te onderzoeken⁷. De variabelen FIRST_PREST en LAST_PREST laten toe om op basis van, respectievelijk, de datum van de eerste en laatste zorgverstrekking binnen een ziekenhuisverblijf de duur van dit verblijf in kaart te brengen.</p>
Populatie database	PP4003	Aantal dagen invaliditeit	
Populatie database	PP0040A	Overlijdensdatum - jaar	
Populatie database	PP0040B	Overlijdensdatum - maand	
Hospitalisatie-databank	FIRST_PREST	Datum 1ste <i>service</i> bij ziekenhuisverblijf in jaar X	
Hospitalisatie-databank	LAST_PREST	Datum laatste <i>service</i> bij ziekenhuisverblijf in jaar X	

Open dataset	Income_D0D01	<p>Op basis van de Belgische Censusgegevens uit 2021, verzameld en geanalyseerd door Statbel, kunnen inkomensdeciles worden berekend voor elke wettelijke inwoner in België die belastingaangifte heeft gedaan in 2021. De statistische sector is een maat die de kleinste geografische eenheid vertegenwoordigt waarvoor Statbel sociaaleconomische statistische analyses uitvoert. De IMA Populatie Database registreert ook de statistische sector voor elke individuele persoon (PP0055). Statbel zal een open dataset ontwikkelen die de proportie van individuen in elke inkomensdecile</p>	<p>De hierboven vermelde variabelen die meer inzichten bieden omtrent de financiële middelen, verantwoordelijkheden en <i>stress</i> stellen de SPADIS onderzoekers in staat om individuen met een lage SES of met hoge zorguitgaven te identificeren. Ze laten echter niet toe om alle niveaus van SES in kaart te brengen of om te zien hoe sociale participatie en diens determinanten variëren langs het volledige SES spectrum. Zo kunnen er geen verschillen in sociale participatie onderzocht worden tussen individuen met een 'gemiddelde' en 'hoge SES'. Om een dergelijke gradiënt alsnog in kaart te brengen wordt gebruik gemaakt van inkomensgegevens op het niveau van de statistische sector van het individu⁸. Deze inkomensgegevens zijn afkomstig van Statbel en worden via de statistische sector gelinkt aan de personen in de SPADIS populatie. Deze koppeling vindt plaats binnen de IMA-omgeving.</p>
Open dataset	Income_D1D02		
Open dataset	Income_D2D03		
Open dataset	Income_D3D04		
Open dataset	Income_D4D05		
Open dataset	Income_D5D06		
Open dataset	Income_D6D07		
Open dataset	Income_D7D08		

Open dataset	Income D8D09	beschrijft voor elke statistische sector. Deze dataset wordt verstrekt aan IMA, dat de <u>9 variabelen</u> koppelt aan de SPADIS-variabelen van belang op basis van de statistische sector die is geregistreerd voor elke persoon in de SPADIS-populaties. IMA verstrekt vervolgens gezondheidsgegevens met de variabelen van belang (inclusief inkomensdecile-variabelen) zonder de variabele voor statistische sector.	
--------------	------------------------------	--	--

#	IMA gegevens – gebruikte referenties
1	Sabariego C, Coenen M, Ito E, et al. Effectiveness of Integration and Re-Integration into Work Strategies for Persons with Chronic Conditions: A Systematic Review of European Strategies. <i>Int J Environ Res Public Health</i> . 2018;15(3):552. doi:10.3390/ijerph15030552
2	Sheehan LR, Lane TJ, Collie A. The Impact of Income Sources on Financial Stress in Workers' Compensation Claimants. <i>J Occup Rehabil</i> . 2020 Dec;30(4):679-688. doi: 10.1007/s10926-020-09883-1. PMID: 32109310.
3	Weerdesteijn, K.H.N., Schaafsma, F., Bonefaas-Groenewoud, K. et al. Predicting return to work after long-term sickness absence with subjective health complaints: a prospective cohort study. <i>BMC Public Health</i> 20, 1095 (2020). https://doi.org/10.1186/s12889-020-09203-5
4	Guan N, Guariglia A, Moore P, Xu F, Al-Janabi H. Financial stress and depression in adults: A systematic review. <i>PLoS One</i> . 2022 Feb 22;17(2):e0264041. doi: 10.1371/journal.pone.0264041. PMID: 35192652; PMCID: PMC8863240.
5	https://www.vaph.be/voorwaarden/handicap
6	Spierdijk L, van Lomwel G, Peppelman W. The determinants of sick leave durations of Dutch self-employed. <i>J Health Econ</i> . 2009 Dec;28(6):1185-96. doi: 10.1016/j.jhealeco.2009.07.003. Epub 2009 Aug 22. PMID: 19733927.
7	Feemster LC, Cooke CR, Rubenfeld GD, Hough CL, Ehlenbach WJ, Au DH, Fan VS. The influence of hospitalization or intensive care unit admission on declines in health-related quality of life. <i>Ann Am Thorac Soc</i> . 2015 Jan;12(1):35-45. doi: 10.1513/AnnalsATS.201404-172OC. PMID: 25493656; PMCID: PMC4342801.
8	Otavova M, Masquelier B, Faes C, et al. Measuring small-area level deprivation in Belgium: The Belgian Index of Multiple Deprivation. <i>Spatial and Spatio-temporal Epidemiology</i> . 2023;45:100587. doi: https://doi.org/10.1016/j.sste.2023.100587

HIV Cohorte: Overzicht gevraagde gegevens, bewijs van proportionaliteit, gebruikte referenties

Dataset	Naam	Beschrijving	Bewijs van proportionaliteit
HIV_Incidence	RNA_DATE	Datum van HIV-confirmatietest	Een bevestigde diagnose is een belangrijk criterium voor opname in het SPADIS-project, evenals de registratie binnen de HIV-cohorte. De tijd sinds diagnose, in combinatie met waarden voor CD4 en HIV-RNA metingen, draagt bij aan het in kaart brengen van de snelheid van progressie van de aandoening ^{1,2,3} . Om de datum van diagnose te berekenen op basis van de gegevens in de HIV cohorte wordt per <i>default</i> gekeken naar de datum van de HIV-confirmatietest in de HIV_Incidence dataset.
HIV_Cohort_BAS_BREACH	POS_DATE	Datum eerste positieve HIV serologie	Overheen de verschillende datasets die behoren tot de HIV cohorte, worden verschillende diagnostische testen geregistreerd per type test en per bron (Aids referentie laboratoria vs. HIV referentie centra).
HIV_Incidence	NEG_DATE	Datum van laatste negatieve serologie (AIDS referentie laboratorium)	Om rekening te houden met het feit dat de 'datum van HIV-confirmatietest' (RNA_DATE uit HIV_Incidence) niet steeds accuraat is (bv. personen uit het buitenland die reeds een geconfirmeerde diagnose hebben ondergaan deze test niet steeds), wordt er door de onderzoekers van de HIV cohorte gebruik gemaakt van een <i>mapped diagnosis</i> . Hierbij wordt gekeken naar verschillende variabelen die data registreren met betrekking tot diagnostische testen en HIV-gerelateerde (zorg)verstrekkingen. De datum van diagnose wordt dan berekend door te kijken naar de vroegst voorkomende datum waarop een diagnostische test / HIV-gerelateerde zorgverstrekking werd geregistreerd. Deze <i>mapped diagnosis date</i> zal ook binnen SPADIS gebruikt worden om rekening te houden met de tijd sinds diagnose in de voorziene analyses. Daarnaast wordt deze <i>mapped diagnosis date</i> gebruikt om al dan niet te bevestigen of de door IMA geselecteerde gegevens (zie 2.2) de correcte periodes omschrijven voor elk individu.
HIV_Cohort_BAS	FIRST_VIS_DATE	Datum eerste bezoek aan HIV referentiecentrum	
HIV_Cohort_VIS	VIS_DATE	Datum bezoek HIV referentiecentrum	
HIV_VL	RNA_DATE	Datum meting van virale lading in Aids Referentie laboratorium	Gegevens die bijdragen tot de <i>mapped diagnosis date</i> zijn (1) datum HIV-confirmatie test, (2) datum van eerste positieve HIV serologie, (3) datum van laatste negatieve serologie (AIDS referentie laboratorium), (4) datum van meting virale lading in AIDS referentie laboratorium, (5) datum van eerste bezoek aan HIV referentiecentrum.

HIV_Cohort_BAS_BREACH	NEG_DATE	Datum laatste negatieve HIV serologie	Deze variabele geeft weer bij welke individuen de virale lading voldoende onderdrukt is om als ondetecteerbaar te worden aanschouwd. Een ondetecteerbare lading toont niet enkel aan dat de ziekte niet kan overgedragen worden maar indiceert ook dat de verstrekte behandeling effectief is ter consistente suppressie van het virus en ter handhaving van een gezonde immuunrespons.
HIV_Cohort_BAS_BREACH	SEROCO_CLIN_YES	Klinische symptomen van seroconversieziekte	Klinische symptomen van seroconversieziekte en het moment waarop deze zich initieel manifesteerden kunnen inzichten verschaffen in de mate waarop HIV de dagelijkse activiteiten van personen die leven met HIV beïnvloedt. Het optreden van symptomen geeft tevens meer inzicht in de mate en snelheid van ziekteprogressie ⁴ .
HIV_Cohort_BAS_BREACH	SERO_CLIN_DATE	Datum van eerste symptomen	
HIV_Cohort_BAS_BREACH	HIVTYPE	Type HIV	Er bestaan duidelijke verschillen inzake behandelingsstrategieën en virale pathogenese of ziekteprogressie naargelang het HIV subtype. Om hiermee rekening te houden in de voorziene analyses dient de variabele 'type HIV' opgenomen te worden in de huidige aanvraag ^{1,3} .
HIV_Incidence	CLIN_STAGE	Klinisch stadium	De variabele klinisch stadium geeft aan in welke stadium van de ziekte het individu zich bevindt: 'primaire/recente infectie' (<6 maanden), 'asymptotisch' (stadium 1), 'symptomatisch non-AIDS' (stadium 2 en 3), 'AIDS' (stadium 4) en 'niet gekend'. Aldus faciliteert deze variabele het maken van een onderscheid tussen individuen die leven met HIV naargelang de ziekteprogressie en hoe deze zich symptomatisch uit ⁵ .
HIV_Cohort_BAS	HEIGHT	Lengte van patiënt bij bezoek / meest recent (cm)	Op basis van de gemeten lengte en gewicht kan de BMI berekend worden van individuen die werden opgenomen in de HIV cohorte. De BMI is een eenvoudige en vaak gebruikte meting die de nutritionele status vertegenwoordigt. De relatie met een HIV infectie kan onder andere teruggevonden worden bij het <i>wasting syndrome</i> kenmerkend voor AIDS en verantwoordelijk voor meer dan 10% onvrijwillig gewichtsverlies. Een lage BMI wordt bij personen die leven met HIV erkend als een voorspeller van de ziekteprogressie ³ .
HIV_Cohort_VIS	WEIGHT	Gewicht van patiënt bij bezoek (kg)	
HIV_Cohort_LAB_CD4	CD4_DATE	Datum CD4 meting	Er bestaat een omgekeerd evenredig verband tussen CD4+ T cel metingen en geobserveerde overleving in ongeveer 80% van de personen die leven met HIV. CD4+ T cel metingen worden frequent gebruikt om de ziekteprogressie te beoordelen ^{6,7} .
HIV_Cohort_LAB_CD4	CD4_VALUE	Waarde van CD4 meting in cellen/ μ l	
HIV_Cohort_LAB_CD4	CD4_PERCENT	Waarde van CD4 meting in %	
HIV_Cohort_LAB_RNA	RNA_DATE	Datum van HIV – RNA meting	De waarde van de HIV-RNA meting als prognostische <i>marker</i> wordt al lange tijd erkend en kent een recht evenredige relatie met de ziekteprogressie en een omgekeerd evenredige relatie met CD4+ T-cel metingen. Een hogere virale lading is gerelateerd aan hogere kans op AIDS. Hogere <i>baseline</i> HIV-RNA metingen in een vroeg infectiestadium worden geassocieerd met een snellere CD4+ T-cell daling in de eerste twee jaar na infectie ³ .
HIV_Cohort_LAB_RNA	RNA_VALUE	Waarde van HIV – RNA meting (kopieën/ml)	
HIV_Cohort_ART	ART_ID	Antiretrovirale therapie (ART)	Efficiënt gebruik van ART leidt tot een meer efficiënte en duurzame onderdrukking van de virale lading en aldus een vertraging van de ziekteprogressie (predictor) en
HIV_Cohort_ART	ART_START_DATE	Datum start ART behandeling	

HIV_Cohort_ART	ART_END_DATE	Datum van stopzetting therapie	een vermindering van het risico op comorbiditeiten ⁸ . Het aanvragen van de variabele inzake de redenen voor stopzetting van de therapie laat toe een onderscheid te maken tussen zij die om een gerechtvaardigde reden de therapie stopzetten en zij die niet behandelingstrouw zijn.
HIV_Cohort_ART	ART_STOP_REASON	(Reden ter stopzetting ART?)	
HIV_Cohort_HOS	HOSP_START_DATE	Datum start hospitalisatie	Wereldwijd blijven bacteriële en met AIDS gerelateerde infecties de belangrijkste oorzaken van ziekenhuisopnames en intramurale sterfte ⁹ . Een ziekenhuisopname vormt een indicatie voor de ernst van de symptomatologie bij het individu dat leeft met HIV ⁹ . Een ziekenhuisopname wordt gezien als hebbende een belangrijke invloed op de functionaliteit van personen die leven met HIV.
HIV_Cohort_HOS	HOSP_END_DATE	Datum einde hospitalisatie	
HIV_Cohort_DIS	DIS_ID	Aandoening: Identificatie van aandoening / voorval	Opportunistische aandoeningen (OA) of klinische <i>events</i> , veroorzaakt door de immuun suppressie ten gevolge van een HIV-infectie, blijven ondanks vooruitgangen in ART, één van de belangrijkste oorzaken van morbiditeit en mortaliteit bij HIV/AIDS-patiënten ¹⁰ . Informatie omtrent het voorkomen en de aard van OA's en klinische <i>events</i> is essentieel bij het beoordelen van de impact van de aandoening op de functionaliteit van HIV-patiënten.
HIV_Cohort_CEP	CLIN_EVENT_ID	Klinisch <i>event</i> : Identificatie van type <i>event</i> of procedure	
HIV_Cohort_CEP	CLIN_EVENT_SPECIFICATION	Verdere specificatie van klinisch <i>event</i>	
HIV_Cohort_DIS	DIS_DATE	Datum van voorval	
HIV_cohort_CEP	CLIN_EVENT_DATE	Datum aanvang van klinisch <i>event</i>	

#	HIV cohort gegevens - gebruikte Referenties
1	The Stages of HIV Infection NIH. Accessed September 14, 2022. https://hivinfo.nih.gov/understanding-hiv/fact-sheets/stages-hiv-infection
2	Nyamweya S, Hegedus A, Jaye A, Rowland-Jones S, Flanagan KL, Macallan DC. Comparing HIV-1 and HIV-2 infection: Lessons for viral immunopathogenesis. <i>Rev Med Virol.</i> 2013;23(4):221-240. doi:10.1002/rmv.1739
3	Langford SE, Ananworanich J, Cooper DA. Predictors of disease progression in HIV infection: a review. <i>AIDS Res Ther.</i> 2007;4(1):11. doi:10.1186/1742-6405-4-11
4	Simon V, Ho DD, Abdool Karim Q. HIV/AIDS epidemiology, pathogenesis, prevention, and treatment. <i>Lancet.</i> 2006 Aug 5;368(9534):489-504. doi: 10.1016/S0140-6736(06)69157-5. PMID: 16890836; PMCID: PMC2913538.
5	Jennifer L. Weinberg and Carrie L. Kovarik. The WHO Clinical Staging System for HIV/AIDS <i>AMA Journal of Ethics.</i> Accessed November 23, 2023. https://journalofethics.ama-assn.org/article/who-clinical-staging-system-hiv/2010-03
6	Thiébaud R, Pellegrin I, Chêne G, Viallard JF, Fleury H, Moreau JF, Pellegrin JL, Blanco P. Immunological markers after long-term treatment interruption in chronically HIV-1 infected patients with CD4 cell count above 400 x 10 ⁶ cells/l. <i>AIDS.</i> 2005 Jan 3;19(1):53-61. doi: 10.1097/00002030-200501030-00006. PMID: 15627033.
7	Arnaout RA, Lloyd AL, O'Brien TR, Goedert JJ, Leonard JM, Nowak MA. A simple relationship between viral load and survival time in HIV-1 infection. <i>Proc Natl Acad Sci U S A.</i> 1999 Sep 28;96(20):11549-53. doi: 10.1073/pnas.96.20.11549. PMID: 10500214; PMCID: PMC18071.

8	Myezwa H, Hanass-Hancock J, Ajidahun AT, Carpenter B. Disability and health outcomes - from a cohort of people on long-term anti-retroviral therapy. SAHARA J. 2018 Dec;15(1):50-59. doi: 10.1080/17290376.2018.1459813. PMID: 29635976; PMCID: PMC5917329.
9	Ford N, Shubber Z, Meintjes G, et al. Causes of hospital admission among people living with HIV worldwide: a systematic review and meta-analysis. Lancet HIV. 2015;2(10):e438-e444. doi:10.1016/S2352-3018(15)00137-X
10	Kim YJ, Woo JH, Kim MJ, et al. Opportunistic diseases among HIV-infected patients: a multicenter-nationwide Korean HIV/AIDS cohort study, 2006 to 2013. Korean J Intern Med. 2016;31(5):953-960. doi:10.3904/kjim.2014.322

IKED: Overzicht gevraagde gegevens, bewijs van proportionaliteit, gebruikte referenties

Dataset	Naam	Beschrijving	Bewijs van proportionaliteit
IQED	Start_date_audit_period	Startdatum van audit waartoe patiënt in IQED behoort	Op basis van de variabele die de startdatum van de audit weergeeft, kan de de audit waartoe de patiënt behoort afgeleid worden. Daarnaast wordt de variabele die de geboortedatum weergeeft gebruikt om de leeftijd op het moment van de audit te berekenen, die wordt omgevormd naar een categorische variabele alvorens ze ter beschikking gesteld wordt aan de SPADIS onderzoekers. De wijze waarop de nieuwe variabele 'Age_class' wordt geconstrueerd wordt beschreven in bijlage 2.
IQED	Patient_id date_of_birth	Geboortedatum patiënt	Hoewel patiënten die gediagnosticeerd zijn met type 1 en type 2 diabetes mellitus (T1DM en T2DM) vaak vergelijkbare symptomen vertonen, zijn er enkele belangrijke verschillen in de pathofysiologie. Het meest opvallende is dat de aanvang en progressie van symptomen sneller verlopen bij T1DM-patiënten ¹ . Om de heterogeniteit in onze onderzoekspopulatie nauwkeurig in kaart te brengen, is deze diagnostische informatie cruciaal.
IQED	diabetes_type	Type diabetes	Bovendien kan de snelheid van de ziekteprogressie in kaart gebracht worden door de datum van de diagnose op te nemen in onze analyses aangezien onderzoek heeft aangetoond dat het verhoogde risico op invaliditeit door diabetes gemodereerd kan worden door de duur van diabetes ² .
IQED	date_of_diagnosis_of_diabetes	Datum van diagnose (dd/mm/yyyy)	De analyse van geglyceerd hemoglobine (HbA1c) in het bloed biedt informatie met betrekking tot de gemiddelde bloedglucosewaarden van een individu gedurende de afgelopen twee tot drie maanden, wat overeenkomt met de voorspelde halveringstijd van rode bloedcellen ³ . HbA1c wordt beschouwd als de gouden standaard om de glycemische controle te beoordelen en wordt frequent gebruikt om de diabetesstatus te beoordelen ⁴ . Bovendien wordt het ook erkend als een goede voorspeller van lipidenprofielen die het risico op cardiovasculaire complicaties beïnvloeden ⁴ . Per definitie hebben de patiënten die geregistreerd zijn in het IKED, en dus een zorgovereenkomst hebben afgesloten, vrij ernstige vormen van diabetes en suboptimale glycemische controle. In lijn met andere studies zal HbA1c als een controlevariabele worden gebruikt in onze analyses.
IQED	was_hba1c_determined	Werd HbA1c gemeten?	
IQED	was_hba1c_determined hba1c_	HbA1c (%)	
IQED	was_hba1c_determined number_of_hba1c_determinations	Aantal HbA1c metingen	

IQED	body_mass_index_calculated_kgm	Body Mass Index, calculated (kg/m ²)	BMI zal gebruikt worden als een controlevariabele in de beoogde analyses aangezien obesitas op zichzelf een directe invloed kan hebben op het functioneren (bijvoorbeeld via cardiovasculaire effecten) en ook één van de drijvende factoren van T2DM kan zijn ⁵ .
IQED	start_date_treatment_with_oral_antidiabetics	Start datum orale diabetes behandeling	Voor patiënten met T1DM is insuline de enige behandelingsmodaliteit. Mensen met T2DM kunnen daarentegen kiezen voor andere behandelingsmodaliteiten, zoals orale antidiabetica. Veel mensen met T2DM die beginnen met orale antidiabetica hebben echter binnen 6 jaar ⁶ aanvullende insuliner therapie nodig vanwege het progressieve karakter van de ziekte. Door middel van deze variabele is het mogelijk om de ziekteprogressie in kaart te brengen op basis van de toegepaste behandeling en de respectievelijke startdatum.
IQED	start_date_of_insulin_treatment	Start datum van insuline behandeling	Voor patiënten met T1DM is insuline de enige behandelingsmodaliteit. Mensen met T2DM kunnen daarentegen kiezen voor andere behandelingsmodaliteiten, zoals orale antidiabetica. Veel mensen met T2DM die beginnen met orale antidiabetica hebben echter binnen 6 jaar ⁶ aanvullende insuliner therapie nodig vanwege het progressieve karakter van de ziekte. Verdere behandelingsstappen omvatten een verhoogde insulinedosering. Observatoire studies suggereren dat onvoldoende verhoging van de dosis leidt tot slechtere glycemische controle en een verhoogd risico op complicaties ⁶ . Door middel van deze variabelen is het mogelijk om de ziekteprogressie in kaart te brengen op basis van de toegepaste behandeling en de respectievelijke startdatum.
IQED	insulin_regimen*	Insuline regime	
IQED	insulin_regimen if_other_insulin_schedule_specify	Als "ander" insuline regime, specificeer:""	
IQED	did_the_patient_have_episodes_of_severe_hypoglycemia	Had de patiënt ernstige hypoglycemie episodes?	Hypoglycemie heeft een grote impact op het leven van patiënten, zowel op fysiek, mentaal als sociaal vlak; alle aspecten van het leven kunnen worden beïnvloed, waaronder werk, autorijden, reizen en vrijetijdsactiviteiten ⁷ . Hoewel zeldzaam en voornamelijk onderzocht in oudere populaties, werden ziekenhuisopnames ten gevolge van een hypoglycemie geassocieerd met een ernstige prognose, waarbij maar liefst 60% van degenen die een episode meemaakten, overleed overheen een follow-up van 4 jaar ⁸ .
IQED	was_the_patient_admitted_for_ketoacidosis_or_for_hyperosmolar_hyperglycemic_state_with_or_without_ketosis	Werd de patiënt opgenomen voor een ketoacidose of een hyperosmolaire hyperglycemische status (met of zonder ketose)?	De meest ernstige impact van diabetes op het functioneren van patiënten komt voort uit de ontwikkeling van korte- en langetermijncomplicaties en de behandelingen daarvan ^{9,10} , zoals een TIA, een myocardinfarct, een beroerte, een (behandeling voor) diabetische retinopathie of maculopathie, het ondergaan van een amputatie van de onderste ledematen, het ondergaan van een niertransplant of krijgen van hemodialyse, het hebben van een diabetische voetulcus, etc ⁹ . Binnen het IKED bevinden zich verschillende geregistreerde variabelen die verband houden met het optreden van deze complicaties en of patiënten de juiste behandeling hebben ontvangen. Daarenboven beschrijft het IKED of deze voorvallen plaatsvonden in de auditperiode of niet. Deze informatie is essentieel om beperkingen op vlak van fysieke gezondheid in kaart te brengen
IQED	has_the_patient_already_received_a_kidney_transplant_or_is_he_receiving_hemodialysis_or_peritoneal_dialysis	Kreeg de patiënt reeds een niertransplant of krijgt de patiënt hemodialyse of peritoneale dialyse?	

IQED	has_the_patient_ever_received_treatment_laser_photocoagulation_andor_intravitreal_injection_for_diabetic_retinopathy	Kreeg de patiënt ooit een behandeling (laser fotocoagulatie en/of intravitreale injectie) voor diabetische retinopathie?	(inclusief of deze complicaties/behandelingen recentelijk voorvielen) en in verband te brengen met mogelijke transities op de arbeidsmarkt.
IQED	has_the_patient_ever_received_treatment_laser_photocoagulation_andor_intravitreal_injection_for_diabetic_retinopathy if_yes_during_the_audit_period	Zo ja, in de audit periode?	
IQED	has_the_patient_ever_received_treatment_laser_photocoagulation_andor_intravitreal_injection_for_diabetic_maculopathy	Heeft de patiënt ooit behandeling ontvangen (laserfotocoagulatie en/of intravitreale injectie) voor diabetische maculopathie?	
IQED	has_the_patient_ever_received_treatment_laser_photocoagulation_andor_intravitreal_injection_for_diabetic_maculopathy if_yes_during_the_audit_period	Zo ja, in de audit periode?	
IQED	does_the_patient_have_diabetic_retinopathy	Heeft de patiënt diabetische retinopathie	
IQED	is_the_patient_blind	Is de patiënt blind?	
IQED	has_the_patient_ever_suffered_a_diabetic_foot_ulcer	Heeft de patiënt ooit een diabetische voetulcus gehad?	
IQED	has_the_patient_ever_suffered_a_diabetic_foot_ulcer if_yes_during_the_audit_period	Zo ja, in de audit periode?	
IQED	has_the_patient_ever_received_revascularization_of_the_lower_limbs	Heeft de patiënt ooit een revascularisatie van de onderste ledematen ondergaan?	
IQED	has_the_patient_ever_received_revascularization_of_the_lower_limbs if_yes_during_the_audit_period	Zo ja, in de audit periode?	
IQED	has_the_patient_ever_received_a_minor_amputation_	Heeft de patiënt ooit een mineure amputatie van	

	of_the_lower_limbs	de onderste ledematen ondergaan?
IQED	has_the_patient_ever_suffered_a_myocardial_infarction	Heeft de patiënt ooit een myocardinfarct gehad?
IQED	has_the_patient_ever_suffered_a_myocardial_infarction if_yes_during_the_audit_period	Zo ja, in de audit periode?
IQED	has_the_patient_ever_undergone_a_percutaneous_coronary_intervention_pci	Onderging de patiënt ooit een percutane coronaire interventie (PCI?)
IQED	has_the_patient_ever_undergone_a_percutaneous_coronary_intervention_pci if_yes_during_the_audit_period	Zo ja, in de audit periode?
IQED	has_the_patient_ever_received_coronary_bypass_surgery	Onderging de patiënt ooit een bypassoperatie van de kransslagaders?
IQED	has_the_patient_ever_received_coronary_bypass_surgery if_yes_during_the_audit_period	Zo ja, in de audit periode?
IQED	has_the_patient_ever_suffered_a_stroke	Heeft de patiënt ooit een beroerte gehad?
IQED	has_the_patient_ever_suffered_a_stroke if_yes_during_the_audit_period	Zo ja, in de audit periode?
IQED	has_the_patient_ever_suffered_a_transient_ischemic_attack	Heeft de patiënt ooit een transiënte ischemische aanval (TIA) gehad?
IQED	has_the_patient_ever_suffered_a_transient_ischemic_attack if_yes_during_the_audit_period	Zo ja, in de audit periode?

#	IKED gegevens – Gebruikte referenties
1	Zaccardi F, Webb DR, Yates T, Davies MJ. Pathophysiology of type 1 and type 2 diabetes mellitus: a 90-year perspective. Postgrad Med J. 2016;92(1084):63-69. doi:10.1136/postgradmedj-2015-133281
2	Wong E, Backholer K, Gearon E, et al. Diabetes and risk of physical disability in adults: a systematic review and meta-analysis. Lancet Diabetes Endocrinol. 2013;1(2):106-114. doi:10.1016/S2213-8587(13)70046-9
3	Sherwani SI, Khan HA, Ekhezaimy A, Masood A, Sakharkar MK. Significance of HbA1c Test in Diagnosis and Prognosis of Diabetic Patients. Biomark Insights. 2016;11:95-104. doi:10.4137/BMI.S38440

4	Khan HA, Sobki SH, Khan SA. Association between glycaemic control and serum lipids profile in type 2 diabetic patients: HbA1c predicts dyslipidaemia. Clin Exp Med. 2007;7(1):24-29. doi:10.1007/s10238-007-0121-3
5	Nitecki, M., Gerstein, H.C., Balmakov, Y. et al. High BMI and the risk for incident type 1 Diabetes Mellitus: a systematic review and meta-analysis of aggregated cohort studies. Cardiovasc Diabetol 22, 300 (2023). https://doi.org/10.1186/s12933-023-02007-y
6	Cahn A, Miccoli R, Dardano A, Del Prato S. New forms of insulin and insulin therapies for the treatment of type 2 diabetes. Lancet Diabetes Endocrinol. 2015;3(8):638-652. doi:10.1016/S2213-8587(15)00097-2
7	Fidler C, Elmelund Christensen T, Gillard S. Hypoglycemia: An overview of fear of hypoglycemia, quality-of-life, and impact on costs. J Med Econ. 2011;14(5):646-655. doi:10.3111/13696998.2011.610852
8	Majumdar SR, Hemmelgarn BR, Lin M, McBrien K, Manns BJ, Tonelli M. Hypoglycemia Associated With Hospitalization and Adverse Events in Older People: Population-based cohort study. Diabetes Care. 2013;36(11):3585-3590. doi:10.2337/dc13-0523
9	Deshpande AD, Harris-Hayes M, Schootman M. Epidemiology of diabetes and diabetes-related complications. Phys Ther. 2008 Nov;88(11):1254-64. doi: 10.2522/ptj.20080020. Epub 2008 Sep 18. PMID: 18801858; PMCID: PMC3870323.
10	Ling, W., Huang, Y., Huang, YM. et al. Global trend of diabetes mortality attributed to vascular complications, 2000–2016. Cardiovasc Diabetol 19, 182 (2020). https://doi.org/10.1186/s12933-020-01159-5

BCR: Overzicht gevraagde gegevens, bewijs van proportionaliteit, gebruikte referenties

Naam	Beschrijving	Bewijs van proportionaliteit
Cancer_type	Type kanker	Er bestaan duidelijke verschillen inzake de impact op de functionaliteit en de individuele ziektelast wanneer verschillende soorten kanker en de respectievelijke (gecombineerde) stadia met elkaar worden vergeleken. Bovendien is de impact van het type kanker en de stadia afhankelijk van de tijd sinds de diagnose (pre-behandelingsfase, behandelingsfase, post-behandelingsfase, overleving). Hierbij wordt ook rekening gehouden met de impact en secundaire effecten van therapeutische modaliteiten. Belangrijk is om te vermelden dat de stadiëring bij solide tumoren en hematologische maligniteiten verschillend is. De variabele 'Combined_stage' zal voor solide tumoren het meer courante gecombineerd kanker stadium weergeven. Voor hematologische maligniteiten zal de variabele 'Combined_stage' de Ann-Arbor stadiëring rapporteren wanneer deze voor handen is.
Combined_stage	Kanker stadium (gecombineerd)	
Incidence_month	Maand en jaar van incidentie	
WHO	WHO – score	De WHO-score bij de diagnose vertegenwoordigt de performantiescore van een patiënt op het moment van de diagnose en biedt een maatstaf voor de individuele ziektebelasting en beperkingen op het gebied van functioneren ¹ . Het gebruik ervan binnen de BCR wordt beschreven op de website van Stichting Kankerregister ²
Chemo12m	Chemotherapie in 12 maanden na diagnose (ja/neen)	Het opnemen van variabelen met betrekking tot oncologische behandelingen is cruciaal omdat deze behandelingen aanzienlijke korte- en langetermijneffecten kunnen hebben op het niveau van functioneren van een patiënt en op diens algehele morbiditeit. Deze effecten zijn niet enkel afhankelijk van de therapeutische modaliteiten maar zijn ook onderhevig aan het type kanker ³ . De diverse
RT12m	Radiatie therapie in 12 maanden na diagnose (ja/neen)	

Immuno12m	Immunotherapie in 12 maanden na diagnose (ja/nee)	behandelingsmethoden zoals bijvoorbeeld chirurgie, chemotherapie, bestralingstherapie en immunotherapie kunnen fysieke bijwerkingen, psychologische stress en veranderingen in dagelijkse activiteiten veroorzaken ⁴ . Het genereren van inzichten omtrent de wijze waarop therapeutische modaliteiten korte- en langetermijn consequenties hebben op de sociale participatie (overgangen tussen arbeidsmarktposities) is waardevol op verschillende vlakken. Het stelt ons in staat om risicoprofielen op te stellen voor bepaalde overgangen op de arbeidsmarkt (bv. terugkeer naar werk) naargelang de toegepaste behandelingen. Daarnaast kunnen de door SPADIS ontwikkelde inzichten gebruikt worden in de klinische praktijk door rekening te houden met de impact op sociale participatie bij keuzes inzake de behandelingsstrategie.
Hormono12m	Hormonale therapie in 12 maanden na diagnose (ja/nee)	
Surgery12m	Chirurgie in 12 maanden na diagnose (ja/nee)	
Targeted12m	Doelgerichte therapie in 12 maanden na diagnose (ja/nee)	
Transplantation12m	Transplantatie in 12 maanden na diagnose (ja/nee)	

#	BCR gegevens – Gebruikte referenties
1	Mol L, Ottevanger PB, Koopman M, Punt CJA. The prognostic value of WHO performance status in relation to quality of life in advanced colorectal cancer patients. <i>European Journal of Cancer</i> . 2016;66:138-143. doi:10.1016/j.ejca.2016.07.013
2	WHO performantiescore: conversietabel Lansky/Karnofsky Stichting Kankerregister. Accessed November 23rd 2023. https://kankerregister.org/media/docs/downloads/zorgprogramma/ConversietabelLansky-KarnofskyNL.pdf
3	Adam R, Nair R, Duncan LF, et al. Treatment burden in individuals living with and beyond cancer: A systematic review of qualitative literature. <i>PLOS ONE</i> . 2023;18(5):e0286308. doi:10.1371/journal.pone.0286308
4	Long-term Side Effects of Cancer. (n.d.). American Cancer Society. https://www.cancer.org/cancer/survivorship/long-term-health-concerns/long-term-side-effects-of-cancer.html .

Bijlage 3. Transformaties van gegevens van het IKED en de HIV cohorte (engelstalig)

IQED transformations (English)

New variable	Based on...	Technical specifications
Audit_number	start_date_audit_period	If <i>start_date_audit_period</i> equals... - ...'01/10/2015' then <i>Audit_number</i> becomes '9', - ...'01/10/2017' then <i>Audit_number</i> becomes '10' - ...'01/03/2020' then <i>Audit_number</i> becomes '11'
Age_class	patient_id date_of_birth	Step 1. Create a <u>new continuous variable</u> called <i>Age_at_audit</i> . - For individuals for whom <i>Audit_number</i> is '9', <i>Age_at_audit</i> equals the difference in time (in years with no decimals) between the date given by <i>patient_id date_of_birth</i> and 01/04/2016 - For individuals for whom <i>Audit_number</i> is '10', <i>Age_at_audit</i> equals the difference in time (in years with no decimals) between the date given by <i>patient_id date_of_birth</i> and 01/04/2018 - For individuals for whom <i>Audit_number</i> is '11', <i>Age_at_audit</i> equals the difference in time (in years with no decimals) between the date given by <i>patient_id date_of_birth</i> and 01/09/2020 Step 2. Based on the continuous values of the new variable <i>Age_at_audit</i> , a categorical variable <i>Age_class</i> is made to distinguish between the following age classes: <ul style="list-style-type: none"> ○ 18-20 ○ 21-25 ○ 25-30 ○ 31-35 ○ 36-40 ○ 41-45 ○ 46-50 ○ 51-55 ○ 56-60
Date_of_diagnosis	date_of_diagnosis_of_diabetes	A new variable is created that only reports the year and month of diagnosis (MM/YYYY) based on <i>date_of_diagnosis_of_diabetes</i> (DD/MM/YYYY).
BMI	Body_mass_index_calculated_kgm	The variable <i>Body_mass_index_calculated_kgm</i> is a continuous variable that needs to be transformed to a categorical variable based on the following ranges: <ul style="list-style-type: none"> ○ If <i>Body_mass_index_calculated_kgm</i> < 18.5 then <i>BMI</i> becomes 'underweight' ○ If $18.5 \leq \textit{Body_mass_index_calculated_kgm} < 25$ then <i>BMI</i> becomes 'healthy' ○ If $25 \leq \textit{Body_mass_index_calculated_kgm} < 30$ then <i>BMI</i> becomes 'overweight' ○ If <i>Body_mass_index_calculated_kgm</i> ≥ 30 then <i>BMI</i> becomes 'Obese' ○ If <i>Body_mass_index_calculated_kgm</i> is Unknown then <i>BMI</i> becomes 'Unknown'

HIV Cohort transformations (English)

New variable	Based on...	Technical specifications
Month_and_Year_of_first_positive_serology	<i>POS_DATE</i> in <i>BAS_BREACH</i>	Based on the variable <i>POS_DATE</i> (DD/MM/YYYY) a new variable is created that only reports the corresponding months and years of the dates. This new variable is called <i>Month_and_Year_of_first_positive_serology</i> (MM/YYYY).
Month_and_Year_of_first_symptoms	<i>SERO_CLIN_DATE</i> in <i>BAS_BREACH</i>	The new variable <i>Month_and_year_of_first_symptoms</i> only reports the month and year (MM/YYYY) of the date given by <i>SERO_CLIN_DATE</i> (DD/MM/YYYY).
RNA_DATE_Incidence	<i>RNA_DATE</i> in <i>HIV_Incidence</i>	The new variable <i>RNA_DATE_Incidence</i> only reports the month and year (MM/YYYY) of the date given by <i>RNA_DATE</i> (DD/MM/YYYY).
NEG_DATE_Incidence	<i>NEG_DATE</i> in <i>HIV_Incidence</i>	The new variable <i>NEG_DATE_Incidence</i> only reports the month and year (MM/YYYY) of the date given by <i>NEG_DATE</i> (DD/MM/YYYY).
RNA_DATE_VL	<i>RNA_DATE</i> in <i>HIV_VL</i>	The new variable <i>RNA_DATE_VL</i> only reports the month and year (MM/YYYY) of the date given by <i>RNA_DATE</i> (DD/MM/YYYY).
Month_and_year_of_first_visit	<i>FIRST_VIS_DATE</i> in <i>HIV_VIS</i>	The new variable <i>Month_and_year_of_first_visit</i> only reports the month and year (MM/YYYY) of the date given by <i>FIRST_VIS_DATE</i> (DD/MM/YYYY).
<i>Month_and_year_of_DIS</i>	<i>DIS_DATE</i> in <i>HIV_Cohort_DIS</i>	The new variable <i>Month_and_year_of_DIS</i> only reports the month and year (MM/YYYY) of the date given by <i>DIS_DATE</i> (DD/MM/YYYY).
<i>CEP_EVENT_DATE</i>	<i>CLIN_EVENT_DATE</i> in <i>HIV_Cohort_CEP</i>	The new variable <i>CEP_EVENT_DATE</i> only reports the month and year (MM/YYYY) of the date given by <i>CLIN_EVENT_DATE</i> (DD/MM/YYYY).
BMI	<i>HEIGHT</i> in <i>BAS</i> and <i>WEIGHT</i> in <i>VIS</i>	<p>Step 1. Transform <i>HEIGHT</i> variable to report the height of the patient in metres instead of centimetres (in case of multiple measurements, create a variable that reports the average height over all measurements).</p> <p>Step 2. Create variable <i>BMI_con</i> by dividing the individuals' <i>WEIGHT</i> (kg) by the square of the <i>HEIGHT</i> (in metres)</p> <p>Step 3. Transform the new variable <i>BMI_con</i> from a continuous variable to a categorical variable <i>BMI</i> based on the following ranges:</p> <ul style="list-style-type: none"> ○ If <i>BMI_con</i> < 18.5 then <i>BMI</i> becomes 'underweight' ○ If $18.5 \leq BMI_con < 25$ then <i>BMI</i> becomes 'healthy' ○ If $25 \leq BMI_con < 30$ then <i>BMI</i> becomes 'overweight' ○ If <i>BMI_con</i> ≥ 30 then <i>BMI</i> becomes 'obese' ○ If <i>BMI_con</i> is unknown then <i>BMI</i> becomes 'unknown'

CD4_value	<i>CD4_VALUE</i> in <i>LAB_CD4</i>	<p>The variable <i>CD4_VALUE</i> (cells/μl) from <i>LAB_CD4</i> is a continuous variable that needs to be transformed into a new categorical variable named <i>CD4_value</i> based on the following ranges:</p> <ul style="list-style-type: none"> ○ If <i>CD4_VALUE</i> < 200, then <i>CD4_value</i> becomes '< 200 cells/mm³' ○ If <i>CD4_VALUE</i> \geq 200, then <i>CD4_value</i> becomes '\geq 200 cells/mm³' ○ If <i>CD4_VALUE</i> has missing values then <i>CD4_value</i> becomes 'Missing'
RNA_value	<i>RNA_VALUE</i> in <i>LAB_RNA</i>	<p>The variable <i>RNA_VALUE</i> (copies/ml) from <i>LAB_RNA</i> is a continuous variable that needs to be transformed into a new categorical variable named <i>RNA_value</i> based on the following ranges:</p> <ul style="list-style-type: none"> ○ If <i>RNA_VALUE</i> \leq 1.7, then <i>RNA_value</i> becomes '\leq1.7 log copies/mL' ○ If <i>RNA_VALUE</i> > 1.7, then <i>RNA_value</i> becomes '> 1.7 log copies/mL' ○ If <i>RNA_VALUE</i> has missing values then <i>RNA_value</i> becomes 'Missing'
Hospitalized	<i>HOSP_START_DATE</i> in <i>HOS</i>	<p>The new variable <i>Hospitalized</i> needs to report whether or not a hospitalization has ever taken place.</p> <ul style="list-style-type: none"> ○ If <i>HOSP_START_DATE</i> is empty/missing then <i>Hospitalized</i> becomes 'No' ○ If <i>HOSP_START_DATE</i> is <u>not empty/missing</u> (in other words, if there is a record of hospitalization) then <i>Hospitalized</i> becomes 'Yes'

Bijlage 4. ATC codes ter selectie van de HIV studie populatie door IMA

The following ATC codes in the reimbursement category (SS000020) represent people living with HIV based on the use of ART:

J05AE01	J05AR03
J05AE02	J05AR04
J05AE03	J05AR05
J05AE04	J05AR06
J05AE05	J05AR07
J05AE06	J05AR08
J05AE07	J05AR09
J05AE08	J05AR10
J05AE09	
J05AE10	
J05AF01	
J05AF02	
J05AF03	
J05AF04	
J05AF05	not used for the selection of individuals on ART for HIV. ¹
J05AF06	
J05AF07	not used for the selection of individuals on ART for HIV. ^a
J05AF09	
J05AF13	not used for the selection of individuals on ART for HIV. ^a
J05AG01	
J05AG02	
J05AG03	
J05AG04	
J05AG05	
J05AG06	
J05AJ00	

J05AJ01
J05AJ02
J05AJ03
J05AJ04
J05AX07
J05AX09
J05AX29
J05AX31
J05AR01
J05AR02

¹ Lamivudine (J05AF05), Tenofovir disoproxil (J05AF07), Tenofovir alafenamide (J05AF13) are only used combined with other ART for HIV (in monotherapy they can be used for hepatitis B treatment). For this reason those ATC codes are not used for selecting the included PWHIV, but the code is reported in the table "medicine"