

<p>Informatieveiligheidscomité Kamer sociale zekerheid en gezondheid</p>
--

IVC/KSZG/26/308

BERAADSLAGING NR. 26/154 VAN 6 OKTOBER 2026 BETREFFENDE DE KOPPELING VAN GEPSEUDONIMISEERDE PERSOONSgegevens DIE DE GEZONDHEID BETREFFEN TUSSEN INTEG0 – VLAAMS HUISARTSENREGISTRATIENETWERK EN DE ALGEMENE DIRECTIE STATISTIEK - STATISTICS BELGIUM (STATBEL) AAN DE UNIVERSITEIT ANTWERPEN, DE KU LEUVEN EN DE VRIJE UNIVERSITEIT BRUSSEL IN HET KADER VAN HET PROJECT “POPULATION HEALTH RESEARCH ON MULTI-MORBIDITY AND SOCIAL DETERMINANTS USING PRIMARY CARE AND POPULATION REGISTER DATA”

Het Informatieveiligheidscomité, kamer sociale zekerheid en gezondheid, (hierna “het Comité” genoemd;

Gelet op de Verordening (EU) nr. 2016/679 van het Europees Parlement en de Raad van 27 april 2016 *betreffende de bescherming van natuurlijke personen in verband met de verwerking van persoonsgegevens en betreffende het vrije verkeer van die gegevens en tot intrekking van Richtlijn 95/46/EG* (Algemene Verordening Gegevensbescherming of AVG);

Gelet op de wet van 30 juli 2018 *betreffende de bescherming van natuurlijke personen met betrekking tot de verwerking van persoonsgegevens*;

Gelet op de wet van 13 december 2006 *houdende diverse bepalingen betreffende gezondheid*, in het bijzonder artikel 42, § 2, 3°, gewijzigd bij de wet van 5 september 2018;

Gelet op de wet van 15 januari 1990 *houdende oprichting en organisatie van een Kruispuntbank van de Sociale Zekerheid*, inzonderheid op artikel 15;

Gelet op de aanvraag van Universiteit Antwerpen, KU Leuven en VUB;

Gelet op het auditoraatsrapport van het eHealth-platform van 15 september 2026;

Gelet op het verslag van de heer Michel Deneyer;

Beslist op 6 oktober 2026, na beraadslaging, als volgt:

I. VOORWERP VAN DE AANVRAAG

1. Het Centrum voor Demografie, Familie en Gezondheid (CDFG) en Eerstelijns- en Interdisciplinaire zorg (ELIZA) van de Universiteit Antwerpen, het Academische Centrum Huisartsgeneeskunde (ACHG) van de KU Leuven en Brussels Institute for Social and Population Studies van de Vrije Universiteit Brussel hebben een aanvraag ingediend voor de koppeling van gepseudonimiseerde persoonsgegevens die de gezondheid betreffen tussen het Vlaams huisartsenregistratienetwerk Intego van de KU Leuven en de Algemene Directie Statistiek van Statbel in het kader van het project “Population Health Research on Multi-Morbidity and Social Determinants Using Linked Primary Care and Population Register Data”.
2. Het project richt zich op de clustering van gezondheidscondities, sociale gezondheidsongelijkheden vanuit een intersectioneel perspectief, en individuele levenslooppatronen in relatie tot gezondheid en zorggebruik.
3. Het doel van deze datakoppelingsaanvraag is het koppelen van individuele patiëntgegevens uit de Intego dataset, een Vlaams huisartsenregistratienetwerk, voor de periode 2018 tot en met 2026 met individuele kenmerken van Statbel. De Intego-gegevens betreffen de jaarlijkse contactgroep van personen ouder dan 18 jaar, waarbij de contactgroep gedefinieerd wordt als patiënten die minstens één keer per jaar een deelnemende huisartsenpraktijk hebben bezocht. De koppeling is noodzakelijk om onderzoeksvragen te beantwoorden die niet beantwoord kunnen worden door gebruik te maken van slechts één van beide databanken afzonderlijk.
4. Dit datakoppelingsproject is een samenwerking tussen de volgende partners: KU Leuven (Intego), Universiteit Antwerpen en Vrije Universiteit Brussel. Elk van deze partners brengt unieke en complementaire expertise in het project in en maakt deel uit van het Data4PHM-consortium, dat een proactieve en datagedreven benadering van de gezondheidszorg beoogt te bevorderen. Dit project sluit aan bij het quintuple aim (5AIM) van population health management: (1) het verbeteren van de publieke gezondheid, (2) de kwaliteit van de zorg, (3) het welzijn van het zorgpersoneel en (4) de gezondheidsgelijkheid, (5) zonder verdere kostenstijging.
5. De jaarlijkse contactgroep uit de Intego dataset voor periode 2018-2026 bevat de gegevens van circa 400 000 personen). De jaarlijkse contactgroep bestaat uit personen die een huisartsenpraktijk hebben bezocht in één jaar. Het precieze aantal personen verschilt per jaar. De periode van selectie is gebaseerd op het gebruik van CareConnect software voor de registratie van patiëntgegevens, wat een betrouwbaardere en vollediger datacollectie toeliet vanaf 2018. 2026 vormt tot op heden het meest recente jaar van patiëntgegevens. De reden dat de gegevens van de volledige jaarlijkse contactgroep worden aangevraagd, is dat voor het bestuderen van (1) multi-morbiditeit en haar relatie met de positie van individuen op diverse sociale dimensies en diens intersecties, (2) de rol van omgevingsfactoren in deze relaties, (3) het analyseren van levenslooptrajecten (‘life-courses’) van patiënten en (4) de impact van bepaalde levensgebeurtenissen (‘life events’) op chronische aandoeningen en het management en verloop ervan, omvangrijke

longitudinale datasets nodig zijn. Dit staat de onderzoekers toe om geavanceerdere modellen te schatten met indicatoren op verschillende niveaus (tijdsperiode, individu, directe omgeving), waardoor de bruikbaarheid van de onderzoeksresultaten voor het beheer van volksgezondheid ('public health management') en gezondheidsbeleid wordt geoptimaliseerd. Het doel is om hierbij bijzondere aandacht te besteden aan de rol van sociale gezondheidsdeterminanten.

6. Alle personen vanaf 18 jaar die een deelnemende huisartsenpraktijk uit Intego hebben bezocht in één of meerdere jaren gedurende de referentieperiode (2018-2026) worden geselecteerd, inclusief eerder geregistreerde gegevens aangaande hun medische geschiedenis voor zover deze beschikbaar zijn. Voor deze personen worden vanaf hun 18^{de} levensjaar longitudinale Demobel en Census gegevens van Statbel opgevraagd, met als doel socio-demografische en socio-economische informatie te koppelen aan patiënten gegevens uit Intego.

II. BEVOEGDHEID

7. Krachtens artikel 42, § 2, 3°, van de wet van 13 december 2006 houdende diverse bepalingen betreffende gezondheid vereist iedere mededeling van persoonsgegevens die de gezondheid betreffen, behoudens voorziene uitzonderingen, een principiële beraadslaging van de kamer sociale zekerheid en gezondheid van het Informatieveiligheidscomité.
8. De kamer sociale zekerheid en gezondheid van het Informatieveiligheidscomité acht zich bijgevolg bevoegd om zich uit te spreken over de mededeling van persoonsgegevens die de gezondheid betreffen.

III. BEHANDELING

A. TOELAATBAARHEID

9. De verwerking van persoonsgegevens die de gezondheid betreffen is in principe verboden¹. Dit verbod is echter niet van toepassing onder meer wanneer de verwerking noodzakelijk is met het oog op archivering in het algemeen belang, wetenschappelijk of historisch onderzoek of statistische doeleinden overeenkomstig artikel 89, lid 1, op grond van Unierecht of lidstatelijk recht, waarbij de evenredigheid met het nagestreefde doel wordt gewaarborgd, de wezenlijke inhoud van het recht op bescherming van persoonsgegevens wordt geëerbiedigd en passende en specifieke maatregelen worden getroffen ter bescherming van de grondrechten en de belangen van de betrokkene².
10. Het verbod op verwerking van persoonsgegevens die de gezondheid betreffen is niet van toepassing indien de verwerking noodzakelijk is met het oog op archivering in het algemeen belang, wetenschappelijk of historisch onderzoek of statistische doeleinden overeenkomstig artikel 89, lid 1, op grond van Unierecht of lidstatelijk recht, waarbij de evenredigheid met het nagestreefde doel wordt gewaarborgd, de wezenlijke inhoud van het recht op

¹ Art. 9, §1 van de AVG.

² Art. 9, §2, j) van de AVG.

bescherming van persoonsgegevens wordt geëerbiedigd en passende en specifieke maatregelen worden getroffen ter bescherming van de grondrechten en belangen van de betrokkene.³ Het betreft een secundair gebruik van gegevens die oorspronkelijk voor een ander doel zijn verzameld.

11. Intego staat voor “Integrated Computer Network”. In 1990 is het project gestart in het Academisch Centrum Huisartsengeneeskunde van de KU Leuven. De bedoeling was om een grote database met eerstelijnszorgmorbidity op te bouwen met informatie over de incidentie en prevalentie van ziekten in Vlaanderen. Ook data over diagnostische tests en behandelingen in de huisartsenpraktijk kwamen zo ter beschikking. Vandaag wordt Intego nog steeds gefinancierd door de Vlaamse Gemeenschap en omvat een netwerk van ongeveer 500 huisartsen werkzaam in meer dan 125 huisartsenpraktijken verspreid over heel Vlaanderen. De verwerking van de Intego gegevens is noodzakelijk voor het wetenschappelijk onderzoek en wordt verricht onder de voorwaarden vastgesteld in artikel 9.2, j) van de AVG.
12. De opdrachten van de Algemene Directie Statistiek (Statbel) worden geregeld in de wet van 22 maart 2006 *tot wijziging van de wet van 4 juli 1962 betreffende de openbare statistiek en van de wet van 8 augustus 1983 tot regeling van een Rijksregister van de natuurlijke personen*. Dit wettelijk kader definieert de bevoegdheden inzake inzameling, opmaak en verspreiding van betrouwbare statistieken die relevant zijn voor de Belgische economie, de maatschappij en het grondgebied.
13. Het Comité is bijgevolg van mening dat er een grondslag bestaat voor de beoogde verwerking van persoonsgegevens die de gezondheid betreffen.

B. FINALITEIT

14. Volgens artikel 5 van de AVG moeten de persoonsgegevens worden verwerkt op een wijze die ten aanzien van de betrokkene rechtmatig, behoorlijk en transparant is. Ze moeten voor welbepaalde, uitdrukkelijk omschreven en gerechtvaardigde doeleinden worden ingezameld en mogen vervolgens niet verder worden verwerkt op een met die doeleinden onverenigbare wijze.
15. Het doel van de koppeling van (1) patiëntgegevens uit de eerstelijnsgezondheidszorg van Intego met (2) bevolkingsregistergegevens van Statbel (Demobel en Census) en (3) de aan Intego gekoppelde omgevingsgegevens is het bestuderen van sociale kwetsbaarheid en multimorbiditeit in Vlaanderen, België. Dit maakt baanbrekend onderzoek mogelijk naar intersecterende sociale ongelijkheden, cruciale levensgebeurtenissen en relevante omgevingsfactoren, teneinde de kennis over de belangrijkste syndemieën in België te vergroten. De resultaten van dit werk beogen een verbetering van het beleid en van interventies die de kwaliteit van de zorg en de gezondheidsgelijkheid bevorderen. Belangrijk is dat deze data-integratie eveneens het publieke gezondheidsmanagement en de kwaliteit van de zorg in de klinische praktijk zal versterken. Sociale risicostratificatie en predictiemodellen zullen bijdragen tot de ontwikkeling van proxy-indicatoren voor

³ Art. 9, §2, j) van de AVG.

patiëntkwetsbaarheden en tot de identificatie van hoogrisicopatiënten, waardoor zorgverleners over bijkomende instrumenten beschikken om patiënttrajecten op te volgen. Ten slotte zal microsimulatie worden ingezet voor langetermijnvoorspellingen van multimorbiditeit, ter onderbouwing van duurzame beleidsopties inzake gezondheid. Het gebruik van grootschalige eerstelijnsgezondheidsgegevens is van cruciaal belang, aangezien de onderzoeksdoelstellingen grote patiëntenaantallen vereisen om variaties in multimorbiditeit te identificeren en te analyseren, in het bijzonder vanuit een intersectioneel en longitudinaal perspectief. Daarnaast is het essentieel om de patiëntinformatie aan te vullen met bevolkingsregistergegevens, aangezien de eerstgenoemde gegevens gedetailleerde sociodemografische kenmerken (bv. burgerlijke staat en migratieachtergrond) en socio-economische kenmerken (bv. opleidingsniveau, tewerkstelling, woonsituatie en inkomen) missen.

1. Fundamentele kennis: Sociale determinanten van chronische aandoeningen en multimorbiditeit vanuit een syndemische benadering

16. Het eerste deel van dit project richt zich op innovatief onderzoek naar de sociale determinanten van chronische aandoeningen en multimorbiditeit vanuit een syndemische benadering, aangevuld met een intersectioneel en levensloopperspectief.

1° Het **syndemisch kader** is een nuttig instrument om de interacties te bestuderen tussen samen voorkomende gezondheidsaandoeningen en sociale factoren op individueel en contextueel niveau. De maatschappelijke waarde van syndemisch onderzoek ligt in het vermogen ervan om doeltreffendere, rechtvaardigere en meer holistische antwoorden vanuit de volksgezondheid te onderbouwen. In plaats van aandoeningen afzonderlijk te benaderen, brengt syndemisch onderzoek in kaart hoe gezondheidsproblemen onderling en met sociale omstandigheden interageren, wat helpt om de onderliggende oorzaken van slechte gezondheid aan te pakken.

Een syndemie verwijst naar de aanwezigheid van twee of meer nadelige gezondheidsaandoeningen binnen kwetsbare bevolkingsgroepen, die synergetisch met elkaar interageren en tegelijkertijd worden aangedreven door contextuele krachten — waardoor het voorkomen en de clustering van deze gezondheidsaandoeningen in bepaalde groepen worden versterkt. De syndemische benadering omvat drie kernelementen:

- (1) Het samen voorkomen van twee of meer ziekten of aandoeningen (bv. depressie en diabetes, of COPD en hart- en vaatziekten) binnen dezelfde populatie of bij dezelfde individuen. Deze aandoeningen bestaan niet louter naast elkaar — zij interageren biologisch, psychologisch of op gedragsmatig vlak.
- (2) De interactie tussen deze aandoeningen is synergetisch, wat betekent dat de aanwezigheid van de ene aandoening de ernst of de uitkomsten van de andere verergert. Zo kan het gelijktijdig hebben van diabetes en depressie het risico op gebrekkig zelfmanagement en op bijkomende complicaties verhogen.
- (3) Deze interagerende aandoeningen worden aangedreven of versterkt door sociale, economische en omgevingsgebonden factoren, zoals structurele ongelijkheden op basis van gender, migratieachtergrond en socio-economische kwetsbaarheid.

Syndemieën zijn dus = (1) het samen voorkomen van aandoeningen + (2) interactie tussen aandoeningen + (3) sociale/structurele drijfveren. Op basis van deze kernelementen worden de volgende onderzoeksvragen (OV) geformuleerd:

OV1.1: Wat zijn de belangrijkste syndemieën in Vlaanderen (België), namelijk (a) twee of meer ziekten die samen clusteren [syndemisch element 1] en (b) synergetisch interageren, waarbij zij elkaars ziekteprogressie, behandeltraject en uitkomsten verslechteren [syndemisch element 2]?

OV1.2: Welke individuele kenmerken versterken de clustering van ziekten (geselecteerd in OV1.1) [syndemisch element 3]?

OV1.3: Wat zijn de contextuele drijfveren van de syndemieën (geselecteerd in OV1.1) die de clustering van en de interactie tussen ziekten versterken [syndemisch element 3]?⁴

- 2° Het hanteren van een intersectionele benadering in kwantitatieve analyses werpt licht op de fijnmazige ongelijkheden in het samen voorkomen van ziekten en in de nadelige interacties tussen ziekten, over sociale strata heen die gedefinieerd worden door de intersectie van meerdere sociale dimensies (bv. leeftijd, gender, socio-economische indicatoren, migratieachtergrond, enz.). De maatschappelijke relevantie van een intersectionele benadering in syndemisch onderzoek ligt in het vermogen ervan om zichtbaar te maken hoe gezondheidsongelijkheden niet louter door één sociale factor, zoals gender of migratieachtergrond, worden gevormd, maar door de intersecties daarvan — bijvoorbeeld hoe gender, migratiestatus en socio-economische status samen unieke gezondheidsrisico's en drempels tot zorg creëren.

De integratie van een intersectionele benadering leidt tot de volgende onderzoeksvragen:

OV2.1: Hoe variëren de geïdentificeerde ziekteclusters over verschillende sociale strata op basis van intersecterende sociale determinanten [syndemische elementen 1 en 3]? Bijvoorbeeld: in welke sociale strata is de clustering van de geïdentificeerde ziekten bijzonder sterk en in welke strata is zij zwak of zelfs afwezig?

OV2.2: Hoe variëren de nadelige interacties tussen clusterende ziekten over verschillende sociale strata [syndemische elementen 2 en 3]? Bijvoorbeeld: voor welke sociale strata is het samen voorkomen van diabetes en depressie het sterkst geassocieerd met toegenomen diabetescomplicaties?⁵

⁴ *Data rationale.* Om deze onderzoeksvragen te beantwoorden is een combinatie van grootschalige administratieve gegevens uit meerdere bronnen noodzakelijk. Syndemisch element 1 (OV1.1.a) vereist uitgebreide patiëntgegevens, met inbegrip van omvattende diagnostische informatie over een reeks gezondheidsaandoeningen, aangezien het gelijktijdig voorkomen van meerdere gezondheidsaandoeningen moeilijk vast te stellen is. Voor syndemisch element 2 (OV1.1.b) is niet alleen diagnostische informatie nodig, maar ook informatie over het zorgtraject, het ziekteverloop en de uitkomsten, om de synergetische interactie tussen de gezondheidsaandoeningen te kunnen beoordelen. Syndemisch element 3 vergt een combinatie van gezondheids- en sociale gegevens om te bestuderen hoe individuele factoren (OV1.2) en contextuele factoren (OV1.3) als drijfveren van de syndemie fungeren. Dit laatste steunt op de opname van de bevolkingsregistergegevens van Statbel en aan Intego gekoppelde omgevingsindicatoren.

⁵ *Data rationale.* De integratie van patiëntgegevens uit de eerstelijnsgezondheidszorg met bevolkingsregistergegevens biedt onderzoekers de unieke mogelijkheid om intersecterende sociale ongelijkheden in gezondheidsaandoeningen omvattend te documenteren en te analyseren. Deze benadering overstijgt de beperkingen van andere gegevensbronnen, zoals surveys, die er vaak niet in slagen de complexiteit van meerdere intersecterende sociale dimensies in voldoende aantallen te vatten (bv. ondervertegenwoordiging van personen met een migratieachtergrond). De Demobel- en Censusgegevens van Statbel laten toe granulaire intersectionele analyses uit te voeren, dankzij hun gedetailleerde indicatoren over sociodemografische en socio-economische kenmerken.

- 3° Het interdisciplinaire Life Course Health Development (LCHD)-kader erkent de complexe, niet-lineaire, multilevel en ontwikkelingsgebonden dimensies van gezondheid. Het LCHD-kader reikt een geheel van principes aan die het wetenschappelijk onderzoek naar gezondheidsontwikkeling over de levensloop sturen. Dit project steunt op het LCHD-kader en gaat uit van de centrale stellingen dat i) de individuele gezondheid zich continu ontwikkelt over de levensloop en gevormd wordt door eerdere ervaringen, ii) gezondheidsontwikkeling voortkomt uit adaptieve en wederkerige relaties tussen individuen en hun sociale omgeving, en iii) gezondheidsontwikkeling afhankelijk is van de timing en de sociale structurering van gebeurtenissen en blootstellingen. Een longitudinale levensloopbenadering laat toe rekening te houden met de temporele ordening tussen levensloopgebeurtenissen enerzijds en zorggebruik en gezondheidsuitkomsten anderzijds, waardoor potentiële endogeniteitsproblemen en selectiebias in de associaties kunnen worden aangepakt.

De volgende onderzoeksvragen staan daarom centraal in dit project:

OV3.1: (a) Hoe verhouden levensloopgebeurtenissen (bv. echtscheiding of het verlies van een partner) zich tot het voorkomen van de geïdentificeerde ziekteclusters [syndemische elementen 1 en 3] en (b) tot de nadelige interacties ervan [syndemische elementen 1 en 2]?

OV3.2: Hoe veranderen ziekteclusters en hun nadelige interacties naar aanleiding van levensloopgebeurtenissen (a) doorheen de tijd en (b) over verschillende sociale strata heen [syndemische elementen 1-3]?⁶

2. Valorisatiekennis: Sociale risicostratificatie en risicopredictie-modellen voor de klinische praktijk

17. Het tweede deel van het project focust op de vertaling van de hierboven beschreven fundamentele kennis naar valorisatiekennis — dit is de vertaling naar evidence-based procedures en beleid met reële toepasbaarheid. Meer bepaald behelst dit de vertaling van de fundamentele kennis naar (1) patiëntrisico-identificatie in de eerstelijnspraktijk, (2) de ontwikkeling en validatie van maten voor sociale kwetsbaarheid, en (3) microsimulatiemodellering ter voorspelling van de toekomstige ziektelast. Samen zullen de resultaten van deze valorisatie-inspanningen leiden tot beter geïnformeerde en gecontextualiseerde beleidsvorming.

- 1° Er bestaat een duidelijke nood aan geïnformeerde patiëntrisico-identificatie in de eerstelijnsgezondheidszorg. De ontwikkeling van performante risicopredictie-algoritmen is essentieel om te komen tot efficiënte en hanteerbare risicopredictie-instrumenten in de klinische praktijk. Een accurate identificatie van hoog- en laagrisicopatiënten (bv. voor de ontwikkeling van ischemisch hartlijden) is een noodzakelijke stap om risicobeperkende interventies te organiseren. De risicopredictiebenadering is waardevol voor zowel preventie als

⁶ *Data rationale*. Naast de vereiste van grootschalige gegevens die omvattende informatie op patiëntniveau bevatten (zoals diagnostische informatie, behandeling en zorgdetails, uitkomsten en complexiteiten) en rijke informatie over sociale kenmerken (zowel sociodemografisch als socio-economisch), vergt een levensloopbenadering bijkomend longitudinale informatie over de timing van levensloopgebeurtenissen, ziekteprogressie en zorgtrajecten op individueel niveau.

gezondheidsmanagement vanuit een syndemische bril, waarbij erkend wordt hoe ziekten clusteren en interageren binnen specifieke sociale contexten. De opname van sociale gezondheidsdeterminanten is cruciaal om deze modellen voor risicostratificatie te verbeteren, doordat een uitgebreidere set relevante risicopredictoren in rekening wordt gebracht. Dit werk bouwt dan ook voort op de onderzoeksresultaten van het eerste projectdeel. Belangrijk is dat een accurate risicostratificatie leidt tot een efficiëntere allocatie van financiële en menselijke middelen, wat helpt om patiënten te prioriteren en te voorspellen hoe hun noden zullen evolueren. Valorisatiedoelstelling VD1: Het ontwikkelen van gezondheidsrisicopredictiemodellen in de eerstelijnsgezondheidszorg.⁷

- 2° De gegevensverzameling in de eerstelijnsgezondheidszorg is niet in de eerste plaats opgezet om de sociale kwetsbaarheid van patiënten op te volgen. Het statuut van verhoogde tegemoetkoming is momenteel een van de weinige proxy-indicatoren voor socio-economische status (SES) in eerstelijnsgezondheidsgegevens (en in de gegevens van de Belgische ziekteverzekering (IMA)). Deze maat wordt vervolgens in nagenoeg alle effectstudies van gezondheidsbeleid gebruikt om de gelijkheidsuitkomsten van gezondheidsinterventies te beoordelen. De validiteit van deze maat als proxy voor SES is echter onbekend: vat deze maat daadwerkelijk de bevolkingsgroep met de laagste SES-niveaus, of zijn andere proxy's of maten vereist? De unieke koppeling tussen gedetailleerde sociodemografische gegevens en longitudinale gezondheidsgegevens stelt datagebruikers in staat om (1) te beoordelen of de huidige dominante proxy (verhoogde tegemoetkoming) personen in kwetsbaarheid correct identificeert, zowel op het vlak van sensitiviteit (het vermogen om personen met een lage SES correct te identificeren) als specificiteit (het vermogen om personen zonder lage SES correct te identificeren), en (2) alternatieve proxy's voor sociale kwetsbaarheid te ontwikkelen. Daarnaast zal de koppeling van deze gecombineerde dataset met omgevingsgegevens ons ook toelaten na te gaan hoe contexten met een lage SES omgevingen vormen waarin het verband tussen SES en gezondheid mogelijk nog sterker is — wat aangeeft hoe het statuut van verhoogde tegemoetkoming mogelijk gecontextualiseerd dient te worden. VD2: Het ontwikkelen van gevalideerde proxy's voor patiëntkwetsbaarheden.⁸
- 3° Vanuit een syndemische benadering kunnen alle voorgaande inzichten in de sociale determinanten van multimorbiditeit, voortkomend uit de integratie van patiëntgegevens uit de eerstelijnsgezondheidszorg (ziekteprogressie en behandel-effectiviteit), bevolkingsregistergegevens (populatiestratificatie en identificatie van risicogroepen) en omgevingsgegevens (omgevingsdeterminanten), worden gecombineerd om

⁷ *Data rationale.* De longitudinale eerstelijnsgezondheidsgegevens van Intego missen gedetailleerde achtergrondkenmerken van patiënten. Daarom is het cruciaal dat de patiëntgegevens worden aangevuld met sociodemografische en socio-economische informatie op individueel niveau van Statbel. Deze combinatie van datasets laat onderzoekers toe risicopredictie-algoritmen over de tijd heen te ontwikkelen, rekening houdend met een geheel van granulaire sociale determinanten om de accuraatheid van de modellen te verbeteren.

⁸ *Data rationale.* Door de koppeling van Intego- en Statbelgegevens kan klinische diagnose-informatie worden gebruikt om na te gaan hoe gezondheidsaandoeningen clusteren naar sociodemografische en socio-economische kenmerken. In een volgende stap kunnen proxy's voor patiëntkwetsbaarheden in de Intego-dataset worden geïdentificeerd. Door de gedetailleerde sociale indicatoren uit de Statbelgegevens te vergelijken met de beperkt beschikbare socio-economische achtergrondinformatie, kan de validiteit van deze proxy's (bv. verhoogde tegemoetkoming) worden beoordeeld.

microsimulatiemodellen te voeden die langetermijntrends in multimorbiditeit — en dus de toekomstige ziektelast (en de bijhorende kosten) — voorspellen. De kracht van dynamische projecties is dat de evolutie van multimorbiditeit kan worden gemodelleerd onder verschillende scenario's doorheen de tijd (bv. onderwijsexpansie, migratietrends, wisselende arbeidsmarktparticipatie, enz.) en de ruimte (bv. omgevingsblootstellingen voor uiteenlopende subpopulaties), wat robuuste schattingen kan opleveren in onzekere contexten. Deze reeks projecties van multimorbiditeit wordt als bijzonder informatief beschouwd voor de vormgeving van toekomstig gezondheidsbeleid, aangezien deze proactieve benadering kan leiden tot een betere voorbereiding en gerichtere interventies, en zo uiteindelijk de totale druk op het gezondheidszorgsysteem kan verminderen. VD3: Het inschatten van langetermijntrends in multimorbiditeit via microsimulatiemodellering.⁹

- 4° Naar schatting is medische zorg verantwoordelijk voor slechts 10 tot 20 procent van de beïnvloedbare factoren die bijdragen tot gezonde uitkomsten voor een populatie. De overige 80 tot 90 procent wordt soms in brede zin de sociale gezondheidsdeterminanten genoemd — wat de koppeling van gezondheidsgegevens aan sociodemografische gegevens noodzakelijk maakt om omvattende interventies te kunnen ontwikkelen. De data-integratie zal de ontwikkeling en implementatie mogelijk maken van evidence-based interventies die zijn afgestemd op de (sociale, gezondheids- en omgevingsgebonden) noden van de personen met het hoogste risico. Meer bepaald zal zij leiden tot: (a) Betere preventieve dienstverlening: aan de hand van gekoppelde gegevens kunnen zorgverleners doeltreffendere preventieve diensten ontwikkelen die zijn afgestemd op de specifieke noden van de bevolking. Door trends en risicofactoren te analyseren, kunnen zij hoogrisicogroepen identificeren en gerichte preventieprogramma's opzetten. Dergelijke programma's kunnen screenings, vaccinaties en gezondheidsvoorlichtingscampagnes omvatten. Verbeterde preventieve dienstverlening kan de incidentie van ziekten verlagen, de zorgkosten drukken en de algemene gezondheid van de gemeenschap versterken. (b) Multicomponent-interventies: gezondheid is afhankelijk van een complex samenspel van factoren (sociale gezondheidsdeterminanten, gezondheidszorgdiensten, omgevingscontext, enz.). Het combineren van gegevens uit uiteenlopende bronnen maakt het ontwerp en de implementatie mogelijk van multicomponent-interventies die deze gelaagdheid van gezondheid en patiëntenzorg gelijktijdig aanpakken, op basis van kennis die deze complexiteit weerspiegelt. Dergelijke interventies kunnen een mix omvatten van medische behandelingen, leefstijlveranderingen en omgevingsaanpassingen. Door meerdere factoren tegelijk te viseren, kunnen multicomponent-interventies leiden tot aanzienlijkere verbeteringen in gezondheidsuitkomsten. Deze holistische benadering kan ook de patiënttevredenheid verhogen en de zorgkosten verlagen door complicaties en heropnames te voorkomen. (c) Betere afstemming tussen gezondheidszorg en welzijnsdiensten: de integratie van gezondheidszorggegevens met bevolkings- en omgevingsgegevens kan de coördinatie tussen welzijnsdiensten en zorgverleners

⁹ *Data rationale.* Om populatiedynamieken in microsimulatiemodellering te vatten, is naast relevante sociodemografische, socio-economische en omgevingsgebonden gezondheidspredictoren een omvattend geheel aan informatie over levensloopgebeurtenissen (bv. mortaliteit en migratie) vereist. Diagnostische informatie over een reeks gezondheidsaandoeningen is noodzakelijk om multimorbiditeit bij patiënten te kunnen vatten.

verbeteren. Deze afstemming zorgt ervoor dat patiënten omvattende zorg ontvangen die zowel hun medische als hun sociale noden adresseert. Zo kan gegevenskoppeling helpen om personen te identificeren van wie het profiel (sociodemografisch, economisch of omgevingsgebonden) hen blootstelt aan een risico op ongunstige gezondheidsuitkomsten (bv. recent weduwnaars of weduwen met een chronische aandoening). Deze geïntegreerde databenadering kan leiden tot betere gezondheidsuitkomsten en een efficiënter gebruik van middelen. VD4: Het mogelijk maken van beter afgestemde interventies door triangulatie van de geïntegreerde datasets.¹⁰

18. Deze studie betreft een verdere verwerking van gegevens die oorspronkelijk voor andere doeleinden werden ingezameld.

- a. INTEGRO

Intego staat voor "Integrated Computer Network". In 1990 is het project gestart in het Academisch Centrum Huisartsengeneeskunde (www.achg.be) van de KU Leuven. De bedoeling was om een grote database met eerstelijnszorgmorbidity op te bouwen met informatie over de incidentie en prevalentie van ziekten in Vlaanderen. Ook data over diagnostische tests en behandelingen in de huisartsenpraktijk kwamen zo ter beschikking. Anno 2025 wordt Intego nog steeds gefinancierd door de Vlaamse Gemeenschap en omvat een netwerk van ongeveer 500 huisartsen werkzaam in meer dan 130 huisartsenpraktijken verspreid over heel Vlaanderen. De data staan ter beschikking van onderzoekers, beleidsmakers, masterstudenten en eenieder die denkt de data nuttig te kunnen gebruiken.

- b. Statbel-gegevens

Het Belgische statistiekbureau Statbel verzamelt, produceert en verspreidt betrouwbare en relevante statistieken over de Belgische economie, samenleving en territorium. De inzameling gebeurt op basis van administratieve gegevensbronnen en enquêtes. De aanmaak van statistieken gebeurt op een kwalitatieve, wetenschappelijke manier. Statbel is krachtens de artikelen 24bis en 24quater van de statistiekwet gemachtigd om administratieve gegevens waaronder het rijksregisternummer te gebruiken voor het opmaken en verspreiden van statistieken.

19. Het Comité is van mening dat de beoogde verwerking van persoonsgegevens wel degelijk welbepaalde, uitdrukkelijk omschreven en gerechtvaardigde doeleinden nastreeft.

C. PROPORTIONALITEIT

20. Volgens artikel 5 van de AVG moeten de persoonsgegevens toereikend, ter zake dienend en beperkt zijn tot wat noodzakelijk is voor de doeleinden waarvoor zij worden verwerkt.

¹⁰ *Data rationale.* De Intego-gegevens missen gedetailleerde achtergrondkenmerken van patiënten. Daarom is het cruciaal dat de patiëntgegevens worden aangevuld met sociodemografische en socio-economische informatie op individueel niveau van Statbel. De koppeling van Intego- en Statbelgegevens zal de ontwikkeling mogelijk maken van evidence-based interventies die gelijktijdig de klinische en de socio-economische drijfveren van gezondheid aanpakken, wat de impact van deze interventies vergroot.

21. Het INSZ (identificatienummer van de sociale zekerheid, d.w.z. hetzij het rijksregisternummer, hetzij het Kruispuntbanknummer) van de persoon wordt gebruikt door de instellingen die betrokken zijn bij de mededeling en koppeling van de gegevens. De onderzoekers zullen enkel toegang hebben tot gepseudonimiseerde identificatienummers.
22. Het verband tussen de gepseudonimiseerde dataset en de identiteit van de personen wordt niet bijgehouden.
23. Verschillende pseudonimiseringsmaatregelen worden geïmplementeerd. (1) Persoonlijke identificatienummers worden in gepseudonimiseerde vorm door een Trusted Third Party (TTP1) bezorgd aan de onderzoekers en hebben louter tot doel om individuele gegevens afkomstig van verschillende gegevensbronnen aan elkaar te koppelen. (2) Variabelen die datums bevatten (in de reeds gekoppelde dataset) worden steeds in gepseudonimiseerde vorm voorzien. De hierna beschreven methode pseudonimiseert datums door deze te verschuiven op basis van een per individu willekeurige fictieve geboortedatum, waarbij alle onderlinge tijdrelaties correct behouden blijven. Stap 1: voor elk individu worden alle relevante datums (bv. diagnose) uitgedrukt als het aantal dagen ten opzichte van de originele geboortedatum. Stap 2: per individu wordt een nieuwe fictieve geboortedatum gegenereerd. Hierbij worden dag en maand willekeurig gekozen, terwijl het geboortjaar correct blijft. Stap 3: de tijdintervallen uit Stap 1 worden toegepast op de fictieve geboortedatum uit Stap 2. Dit levert de uiteindelijke gepseudonimiseerde datums op, waarbij alle onderlinge tijdrelaties exact behouden blijven. De fictieve geboortedatum (dag en maand) is voor de verschillende brondatasets identiek en wordt vastgesteld door een Trusted Third Party (TTP2), (3) Afstammingsvariabelen die de oorspronkelijke nationaliteit of het land van herkomst weergeven als onderdeel van de aangevraagde gegevens bij Statbel worden voor aanlevering gegroepeerd in landengroepen. Dit vermijdt dat afstamming op individueel niveau heel precies achterhaald kan worden, voornamelijk voor lage frequentie landen.
24. Het Comité stelt vast dat het Health Data Agency een instantie binnen de SCRA-pool zal aanduiden om een *small cell risk analysis* te verrichten op de gegevens alvorens de gegevens worden meegedeeld aan de onderzoekers. Alle noodzakelijke pseudonimiseringsmaatregelen zullen worden geïmplementeerd om de re-identificatierisico's van de betrokkenen te beperken.
25. Overeenkomstig artikel 5, 8°, van de wet van 21 augustus 2008 *houdende oprichting en organisatie van het eHealth-platform* komt het eHealth-platform tussen als intermediaire organisatie. conversiesleutels in kader van pseudonimisering van de rijksregisternummers, worden behouden door de TTP1 (eHealth-platform) in kader van eventuele data-aanvragen of updates voor bijkomende onderzoeksvragen of vervolg-onderzoeksprojecten.
26. Het gaat om een eenmalige gegevensmededeling.
27. Het Comité acht dat de verwerking van deze persoonsgegevens in principe toereikend, ter zake dienend en niet overmatig is.

D. OPSLAGBEPERKING

28. Overeenkomstig artikel 5 van de AVG moeten persoonsgegevens worden bewaard in een vorm die het mogelijk maakt de betrokkenen niet langer te identificeren dan voor de doeleinden waarvoor de persoonsgegevens worden verwerkt noodzakelijk is; persoonsgegevens mogen voor langere perioden worden opgeslagen voor zover de persoonsgegevens louter met het oog op archivering in het algemeen belang, wetenschappelijk of historisch onderzoek of statistische doeleinden worden verwerkt overeenkomstig artikel 89, § 1, mits de bij deze verordening vereiste passende technische en organisatorische maatregelen worden getroffen om de rechten en vrijheden van de betrokkene te beschermen (opslagbeperking).
29. De voorgestelde geldigheidsduur van de beraadslaging is 15 jaar of tot december 2041, waarna ze ter hernieuwing kan voorgelegd worden aan het Informatieveiligheidscomité. Dit komt overeen met de voorgestelde bewaartermijn van de gekoppelde gegevens. Deze zullen vijftien jaar na de aanvraag vernietigd worden.
30. De gekoppelde informatie zal bewaard worden voor een periode van 15 jaar. Dit betekent dat alle data ten laatste in december 2041 vernietigd zal worden.
31. Het Comité acht de beoogde bewaartermijn redelijk.

E. TRANSPARANTIE

32. Wanneer de persoonsgegevens niet van de betrokkene zijn verkregen, verstrekt de verwerkingsverantwoordelijke hem/haar alle informatie bedoeld in artikel 14 van de AVG.
33. De betrokkene beschikt over de informatie in de vorm van een affiche die in de wachtzaal van de huisarts opgehangen wordt. De inhoud van de affiche wordt hieronder opgenomen.

F. VEILIGHEIDSMATREGELEN

34. Krachtens artikel 5 van de AVG moeten persoonsgegevens, door het nemen van passende technische of organisatorische maatregelen, op een dusdanige manier worden verwerkt dat een passende beveiliging ervan gewaarborgd is, en dat zij onder meer beschermd zijn tegen ongeoorloofde of onrechtmatige verwerking en tegen onopzettelijk verlies, vernietiging of beschadiging (integriteit en vertrouwelijkheid). Deze maatregelen moeten een passend beveiligingsniveau verzekeren rekening houdend, enerzijds, met de stand van de techniek ter zake en de kosten voor het toepassen van de maatregelen en, anderzijds, met de aard van de te beveiligen gegevens en de potentiële risico's.
35. De verwerking met het oog op archivering in het algemeen belang, wetenschappelijk of historisch onderzoek of statistische doeleinden is onderworpen aan passende waarborgen in overeenstemming met de AVG voor de rechten en vrijheden van de betrokkene. Die waarborgen verzekeren dat er technische en organisatorische maatregelen zijn getroffen om de naleving van het beginsel van minimale gegevensverwerking te garanderen. Deze maatregelen kunnen pseudonimisering omvatten, voor zover de doeleinden in kwestie op die manier kunnen worden verwezenlijkt. Wanneer die doeleinden kunnen worden

verwezenlijkt door verdere verwerking die de identificatie van betrokkenen niet of niet langer toelaat, moeten zij aldus worden verwezenlijkt¹¹.

36. Om de vertrouwelijkheid en de veiligheid van de gegevensverwerking te garanderen, moet iedere instelling die persoonsgegevens bewaart, verwerkt of meedeelt maatregelen nemen in de volgende elf actiedomeinen die betrekking hebben op de informatieveiligheid: veiligheidsbeleid; aanstelling van een functionaris inzake gegevensbescherming; organisatorische en menselijke aspecten van de veiligheid (vertrouwelijkheidsverbintenis van het personeel, regelmatige informatieverstrekking en opleidingen ten behoeve van het personeel inzake bescherming van de privacy en veiligheidsregels); fysieke veiligheid en veiligheid van de omgeving; netwerkbeveiliging; logische toegangs- en netwerkbeveiliging; loggings, opsporing en analyse van de toegangen; toezicht, nazicht en onderhoud; systeem van beheer van de veiligheidsincidenten en de continuïteit (backup-systemen, fault tolerance-systemen, ...) en documentatie.
37. Het Comité stelt vast dat de Universiteit Antwerpen, de KU Leuven en de VUB een functionaris voor gegevensbescherming hebben aangesteld.
38. Het Comité stelt vast dat er een gegevensbeschermingseffectbeoordeling werd verricht en werd meegedeeld aan het Comité.
39. Het Comité stelt vast dat het eHealth-platform en de KSZ tussenkomen als trusted third party (TTP) voor de koppeling en pseudonimisering van de gegevens betreft.
40. Het Comité stelt vast dat er een small cell risk analyse uitgevoerd zal worden door de SCRA pool alvorens de gegevens ter beschikking worden gesteld.
41. Het Comité stelt vast dat de medewerkers van de Universiteit Antwerpen, de KU Leuven en de VUB een contractuele vertrouwelijkheidsplicht hebben ten aanzien van de gegevens die ze verwerken in het kader van hun functie.
42. Het Comité herinnert eraan dat noch de Universiteit Antwerpen, de KU Leuven en de VUB, noch hun medewerkers stappen mogen ondernemen om de betrokkenen te heridentificeren. De resultaten van de studie moeten in een anonieme vorm worden gepubliceerd.
43. Het Comité stelt vast dat zowel de Universiteit Antwerpen, de KU Leuven als de VUB een arts hebben aangesteld die verantwoordelijk is voor de verwerking van de persoonsgegevens die de gezondheid betreffen, alsook een functionaris voor gegevensbescherming.
44. Het Comité herinnert eraan dat krachtens artikel 9 van de wet van 30 juli 2018 betreffende de bescherming van natuurlijke personen met betrekking tot de verwerking van persoonsgegevens de verwerkingsverantwoordelijke de volgende bijkomende maatregelen neemt bij de verwerking van genetische, biometrische of gezondheidsgegevens:

¹¹ Art. 89, §1, van de AVG.

1° hij of, in voorkomend geval, de verwerker wijst de categorieën van personen die toegang hebben tot de persoonsgegevens aan waarbij hun hoedanigheid ten opzichte van de verwerking van de betrokken gegevens nauwkeurig wordt omschreven;

2° hij of, in voorkomend geval, de verwerker houdt de lijst van de aldus aangewezen categorieën van personen ter beschikking van de bevoegde toezichthoudende autoriteit;

3° hij zorgt ervoor dat de aangewezen personen door een wettelijke of statutaire verplichting, of door een evenwaardige contractuele bepaling ertoe gehouden zijn het vertrouwelijk karakter van de betrokken gegevens in acht te nemen.

45. Het Comité herinnert uitdrukkelijk aan de bepalingen van titel 6. Sancties van de wet van 30 juli 2018 *betreffende de bescherming van natuurlijke personen met betrekking tot de verwerking van persoonsgegevens* waarin strenge administratieve en strafsancties zijn voorzien in hoofde van de verwerkingsverantwoordelijke en van de verwerkers voor het niet-naleven van de voorwaarden voorzien in de AVG en de voormelde wet van 30 juli 2018.

Om deze redenen, besluit

de kamer sociale zekerheid en gezondheid van het informatieveiligheidscomité, wat de mededeling betreft van de persoonsgegevens die tot zijn bevoegdheid behoren

dat:

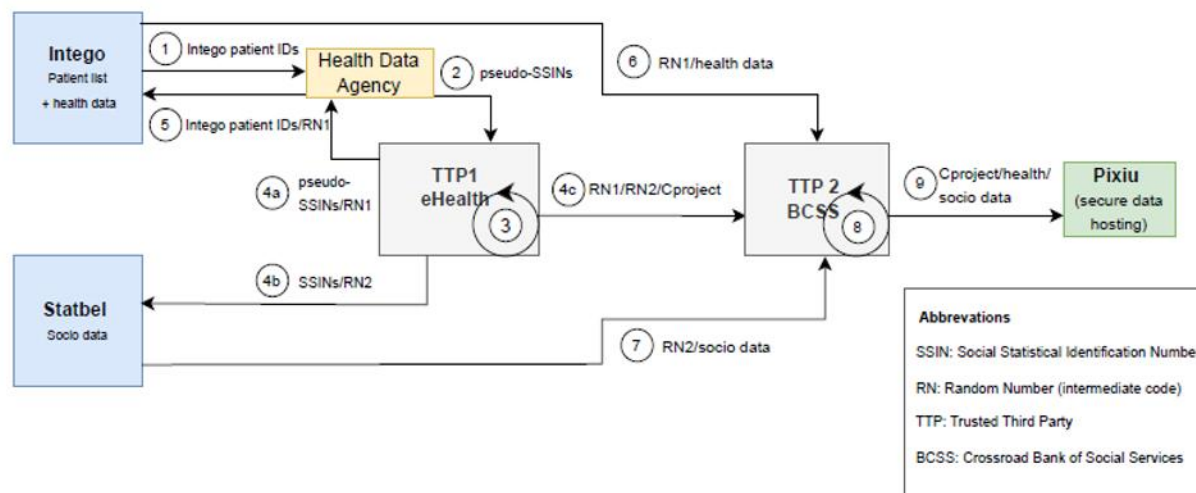
de mededeling van de persoonsgegevens, zoals beschreven in deze beraadslaging, toegestaan is, mits er wordt voldaan aan de in deze beraadslaging vastgestelde maatregelen ter waarborging van de gegevensbescherming, in het bijzonder de maatregelen op het vlak van doelbinding, minimale gegevensverwerking, opslagbeperking en informatieveiligheid.

Deze beraadslaging treedt in werking op 21 oktober 2026.

Michel DENEYER
Voorzitter

De zetel van de kamer sociale zekerheid en gezondheid van het informatieveiligheidscomité is ugevestigd in de kantoren van de Kruispuntbank van de Sociale Zekerheid, op het volgende adres: Willebroekkaai 38 – 1000 Brussel (tel. 32-2-741 83 11).
--

Bijlage 1– Gegevensstroom



Stap 1: Intego identificeert de studiepopulatie binnen zijn register en bezorgt de geselecteerde lijst van Intego-patiënt-ID's aan het Health Data Agency.¹

Stap 2: Het Health Data Agency zet de Intego-patiënt-ID's om in een lijst van pseudo-INSZ's en bezorgt deze aan het eHealth-platform (TTP1), zonder enige gezondheidsgegevens over te maken.¹

Stap 3: eHealth (TTP1) zet de pseudo-INSZ's om in INSZ's ten behoeve van de gegevensselectie bij Statbel.¹ De lijst met pseudo-INSZ's wordt gebruikt in stap 4a (naar het Health Data Agency), terwijl de lijst met INSZ's wordt gebruikt in stap 4b (naar Statbel).

Stap 4a: eHealth (TTP1) genereert een willekeurige, eenmalige tussencode (RN1) voor de geselecteerde Intego-patiënten en bezorgt de lijst pseudo-INSZ/RN1 aan het Health Data Agency. RN1 vervangt de Intego-patiënt-ID's in de verdere verwerking bij Intego.

Stap 4b: eHealth (TTP1) genereert een afzonderlijke willekeurige, eenmalige tussencode (RN2) en bezorgt de lijst INSZ/RN2 aan Statbel. RN2 vervangt het INSZ in de verdere verwerking bij Statbel.

Stap 4c: eHealth (TTP1) bezorgt RN1/RN2/Cproject aan de KSZ (TTP2). Cproject wordt gegenereerd als de finale projectspecifieke pseudonieme identificator. Er wordt geen enkel INSZ overgemaakt.

Stap 5: Het Health Data Agency bezorgt na omzetting de lijst Intego-patiënt-ID/RN1 aan Intego.

Stap 6: Intego bezorgt de gezondheidsgegevens, gelabeld met RN1, aan de KSZ (TTP2).

Stap 7: Statbel bezorgt de sociale gegevens, gelabeld met RN2, aan de KSZ (TTP2).

Stap 8: De KSZ (TTP2) koppelt de datasets aan de hand van RN1 en RN2, verwijdert deze identificatoren en pseudonimiseert datums (bv. diagnose) in de gekoppelde dataset. Koppeling en gepseudonimiseerde identificatie gebeurt door middel van Cproject.

Stap 9: De KSZ (TTP2) bezorgt de volledig gepseudonimiseerde dataset, gelabeld met Cproject, aan de beveiligde analyseomgeving Pixiu.

¹ De omzettingen tussen Intego-patiënt-ID's en pseudo-INSZ's (Intego – Health Data Agency, zie stap 2), evenals tussen pseudo-INSZ's en INSZ's (Health Data Agency – eHealth, zie stap 3), zijn vooraf vastgelegd als onderdeel van de bestaande gegevensstromen tussen deze entiteiten. Alle gegevensoverdrachten zijn versleuteld.

Bijlage 3 – informatie met betrekking tot transparantie

De machtiging voor INTEGEO vereist het gebruik van een affiche die in de wachtzaal opgehangen wordt met volgende melding:

“Uw huisarts registreert voor Intego

Uw huisarts registreert gegevens over u en uw gezondheid in het medisch dossier. Zij worden in de eerste plaats bijgehouden opdat de zorg voor u zo goed mogelijk kan gebeuren. Daarnaast worden sommige gegevens ook gebruikt voor wetenschappelijk onderzoek. Onderzoekers proberen hiermee ziekten in kaart te brengen en de aanpak ervan te verbeteren.

Dit betekent dat uw arts gegevens over u opstuurt naar het Academisch Centrum Huisartsgeneeskunde (ACHG) van de KU Leuven in samenwerking met de dienst healthdata.be, onderdeel van het Wetenschappelijk Instituut Volksgezondheid (Sciensano). Dit gebeurt op een gecodeerde manier waardoor u niet herkenbaar bent, noch voor ACHG, noch voor Sciensano. Met deze gegevens worden verslagen gemaakt voor de overheid en de volksgezondheid, om bv. toenames of afnames van infectieziekten in kaart te brengen. Met de resultaten kunnen dan bv. campagnes opgezet worden ter voorkoming van verdere verspreiding van ziektes. Zo zullen ook de geografische coördinaten van het adres op een beveiligde manier gebruikt worden voor analyses in het kader van de studie “surveillance van milieu-gerelateerde morbiditeit”. U hebt het recht de doorgestuurde gegevens in te kijken en aan te passen. U hebt ook het recht om te weigeren dat het ACHG, in samenwerking met healthdata.be, uw gegevens verzamelt. U kan hiervoor uw huisarts contacteren.

De mededeling van de gecodeerde persoonsgegevens werd gemachtigd door de afdeling gezondheid van het Sectoraal Comité van de Sociale Zekerheid en van de Gezondheid (SCSZG) met referentienr. SCSZG/13/079 bij beraadslaging 13/026 van 13 maart 2013, laatst gewijzigd op 20 maart 2018 en referentienr. 18/176 bij beraadslaging 18/079 van 3 juli 2018, laatst gewijzigd op 7 mei 2019. Het SCSZG maakt onderdeel uit van de Commissie voor de Bescherming van de Persoonlijke Levenssfeer (“Privacycommissie”). Voor elk gebruik van persoonsgegevens dat nog niet beschreven is, moet een nieuwe aanvraag gericht worden aan de Gegevensbeschermingsautoriteit.

De goedkeuringen zijn terug te vinden op de website van de Gegevensbeschermingsautoriteit: www.privacycommission.be. Indien u hieromtrent nog vragen of bemerkingen heeft, kan u dat steeds ter sprake brengen.

Uw huisarts”

Op de website van Intego (<https://www.intego.be/>) staat bovendien de volgende informatie voor patiënten:

“Uw huisarts registreert gegevens over u en uw gezondheid in het medisch dossier. Zij worden in de eerste plaats bijgehouden opdat de zorg voor u zo goed mogelijk kan gebeuren. Als uw huisarts deelneemt aan het INTEGEO-netwerk worden sommige gegevens ook gebruikt voor wetenschappelijk onderzoek. Onderzoekers proberen hiermee ziekten in kaart te brengen en de

aanpak ervan te verbeteren. Het onderzoek wordt verricht in het algemeen belang tot een vermeerdering van kennis en inzicht die de maatschappij (direct of indirect) ten goede komt.

Dit betekent dat uw arts gegevens over u opstuurt naar het Academisch Centrum Huisartsgeneeskunde (ACHG) van de KU Leuven in samenwerking met de dienst healthdata.be, onderdeel van het Wetenschappelijk Instituut Volksgezondheid (Sciensano). Dit gebeurt op een gecodeerde manier waardoor u niet herkenbaar bent, noch voor ACHG, noch voor Sciensano. De gegevens die verzameld worden door INTEGO kan u raadplegen op <https://www.healthdata.be/dcd/#/collections>. Uw gegevens worden tot 30 jaar na uw overlijden bewaard in de INTEGO databank.

De INTEGO data wordt geanalyseerd door onderzoekers van ACHG – KU Leuven en onderzoekers werkzaam op het INTEGO project binnen een interuniversitaire samenwerking. Met deze gegevens worden verslagen gemaakt voor de overheid en de volksgezondheid, om bv. toenames of afnames van infectieziekten in kaart te brengen. Met de resultaten kunnen dan bv. campagnes opgezet worden ter voorkoming van verdere verspreiding van ziektes. Zo zullen ook de geografische coördinaten van uw adres op een beveiligde manier gebruikt worden voor analyses in het kader van de studie “surveillance van milieu-gerelateerde morbiditeit”.

In het algemeen belang kan de INTEGO-databank ook geaggregeerde gegevens ter beschikking stellen van commerciële partners zoals farmaceutische bedrijven (met het oog op optimalisering van diensten en producten). De commerciële partners krijgen geen rechtstreekse toegang tot de data en krijgen evenmin tot personen herleidbare data. De wetenschappelijke rapporten op hun verzoek gemaakt worden naderhand ook publiek gemaakt op de INTEGO website.

U hebt het recht uw doorgestuurde gegevens in te kijken en te laten aanpassen. U hebt ook het recht om te weigeren dat het ACHG, in samenwerking met healthdata.be, uw gegevens verzamelt in de INTEGO-databank. U kan hiervoor uw huisarts contacteren.

De mededeling van de gecodeerde persoonsgegevens werd gemachtigd door de afdeling gezondheid van het Sectoraal Comité van de Sociale Zekerheid en van de Gezondheid (SCSZG) met referentienr. SCSZG/13/079 bij beraadslaging 13/026 van 13 maart 2013, laatst gewijzigd op 20 maart 2018 en referentienr. 18/176 bij beraadslaging 18/079 van 3 juli 2018, laatst gewijzigd op 7 mei 2019. Het SCSZG maakt onderdeel uit van de Commissie voor de Bescherming van de Persoonlijke Levenssfeer (“Privacycommissie”). Voor elk gebruik van persoonsgegevens dat nog niet beschreven is, moet een nieuwe aanvraag gericht worden aan de Gegevensbeschermingsautoriteit.

De goedkeuringen zijn terug te vinden op de website van de Gegevensbeschermingsautoriteit: www.privacycommission.be.

Als u hieromtrent nog vragen of bemerkingen heeft, kan u dat steeds ter sprake brengen bij uw huisarts.”

<https://www.intego.be/patienten>

Bijlage 3: overzicht variabelen met omstandige motivatie

BRON: Statbel			
Dataset: Demobel			
Variabele	Beschrijving	Opmerking	Motivatie
ID_DEMO	Coded identification number – DEMO	Gepseudonimiseerd door TTP	Deze variabele is nodig om Statbel data te linken met patiëntendata van Intego via TTP.
CD_SEX	Sex		Geslacht is een cruciale sociodemografische gezondheidsdeterminant en daarom een individueel kenmerk dat de te bestuderen clustering van gezondheidscondities kan beïnvloeden (RQ1.2). Geslacht is ook een proxy om structurele ongelijkheden te meten op basis van gender en zal dus ook als één van de sociale dimensies van ongelijkheid bestudeerd worden vanuit een intersectionele benadering (RQs 2.1, 2.2 en 3.2). Tot slot is geslacht van belang om mee te nemen in alle valorisatie objectieven (VO1 -VO4).
DT_BTH	Date of birth	Dag en maand van de geboortedatum worden per persoon gerandomiseerd voor vrijgave. Op basis van de tijd tussen de originele geboortedatum en andere datumvariabelen worden deze laatste gepseudonimiseerd	Geboortedatum is nodig om de leeftijd en geboortecohorte van de respondent te berekenen. Leeftijd is een belangrijke sociodemografische gezondheidsdeterminant en daarom een individueel kenmerk dat de te bestuderen clustering van gezondheidscondities kan beïnvloeden (RQ1.2). Leeftijdsverschillen zijn ook belangrijk om mee te nemen bij een intersectionele benadering (RQs 2.1, 2.2 en 3.2) en elementair als tijdschaal voor het in kaart brengen van individuele levenslooppatronen (RQs 3.1 en 3.2). Tot slot is leeftijd van belang om mee te nemen in alle valorisatie objectieven (VO1 -VO4).
DT_DEATH	Date of death	Voor deze variabele wordt een gepseudonimiseerde datum berekend. Dit betreft het aantal jaren, maanden en dagen tussen de originele geboortedatum en de betreffende datum, toegepast op de fictieve geboortedatum.	De datum van overlijden is essentieel om het ultieme gezondheidsrisico vast te stellen en maakt analyses van overleving, mortaliteit en gezondheidsverschillen mogelijk (RQ1.1, RQ1.2). Het is cruciaal voor kosten- en zorganalyses, aangezien overlijden direct de zorggebruik- en kostpatronen beïnvloedt (RQ2.3). Daarnaast ondersteunt het een intersectionele benadering door overlijden te koppelen aan sociodemografische kenmerken zoals leeftijd, geslacht en SES (RQs 2.1, 2.2, 3.2) en vormt het een belangrijke tijdsmarkering voor levensloopperspectieven (RQs 3.1, 3.2). Tot slot draagt het bij aan alle valorisatie-objectieven (VO1–VO4) door inzicht te geven in kritieke gezondheidsuitkomsten.
MS_AGE	Age at reference date		Leeftijd is een belangrijke sociodemografische gezondheidsdeterminant en daarom een individueel kenmerk dat de te bestuderen clustering van gezondheidscondities kan beïnvloeden (RQ1.2). Leeftijdsverschillen zijn ook belangrijk om mee te nemen bij een intersectionele benadering (RQs 2.1, 2.2 en 3.2) en elementair als tijdschaal voor het in kaart brengen van individuele levenslooppatronen (RQs 3.1 en 3.2). Tot slot is leeftijd van belang om mee te nemen in alle valorisatie objectieven (VO1 -VO4).
CD_CNTRY_BTH	Country of birth	Landen en nationaliteiten worden gegroepeerd opgevraagd.	Geboorteland, nationaliteit op de referentiedatum, herkomstnationaliteit en burgerlijke staat zijn belangrijke indicatoren om verblijfsduur, juridische status en culturele identiteit te construeren. Migratie status is een belangrijke sociodemografische gezondheidsdeterminant en daarom een individueel kenmerk dat de te bestuderen clustering van gezondheidscondities kan beïnvloeden
CD_NATLTY	Nationality at reference date		

CD_FST_NA TLTY	First registered nationality		(RQ1.2). Migratie status is ook een proxy om structurele ongelijkheden mee te meten op basis van migratie achtergrond en zal dus ook als één van de sociale dimensies van ongelijkheid bestudeerd worden vanuit een intersectionele benadering (RQs 2.1, 2.2 en 3.2), gezien het verband met gezondheidsongelijkheden en onderdrukkingssystemen. Tal van eerdere studies hebben ook aangetoond dat migratie achtergrond relateert aan zorgtrajecten en erkennen daarnaast het belang van verblijfsduur en herkomstland. Tot slot is migratie achtergrond van belang om mee te nemen in alle valorisatie objectieven (VO1-VO4).
CD_CIV	Civil status at reference date		
ID_HH	Identification number of household		Deze informatie is nodig om de individuele data te linken met andere personen uit het huishouden om clustering in kaart te brengen.
HH_TYPE_LI PRO	Household type according to lipro typology		Huishoudtype volgens LIPRO en de individuele LIPRO positie doorheen de tijd zijn elementair in het identificeren van levenslooptransities en dus onontbeerlijk voor het construeren van individuele levenslooppatronen. Gezondheid bestuderen vanuit een longitudinaal en levensloopperspectief geeft een dieper inzicht over hoe gezondheid en zorgtrajecten temporeel samenhangen met verschillende levenslooptransities en -posities (RQs 3.1 en 3.2). Ook voor het in kaart brengen van multi-morbiditeit trends met behulp van micro-simulatie is het modelleren van levenslooptransities cruciaal (VO3), terwijl informatie over levenslooptransities potentieel ook een rol spelen in het ontwerpen van gerichte gezondheidszorginterventies (VO4).
HH_POS_LIP RO	Position within the household according to lipro typology		
CD_DSCNT	Origin	Landen en nationaliteiten worden gegroepeerd opgevraagd.	Afstammingsinformatie is cruciaal om migratie achtergrond in kaart te brengen. Met behulp van deze variabelen kan onderscheid worden gemaakt tussen generaties van migratie en dus individuele migratiegeschiedenis. Deze gedetailleerde informatie is van belang aangezien migratie status een belangrijke sociodemografische gezondheidsdeterminant is (RQ1.2). Migratie status is ook een proxy om structurele ongelijkheden mee te meten op basis van migratie achtergrond en zal dus ook als één van de sociale dimensies van ongelijkheid bestudeerd worden vanuit een intersectionele benadering (RQs 2.1, 2.2 en 3.2), gezien het verband met gezondheidsongelijkheden en onderdrukkingssystemen. Tal van eerdere studies hebben ook aangetoond dat migratie achtergrond relateert aan zorgtrajecten en erkennen daarnaast het belang van verblijfsduur en herkomstland. Tot slot is migratie achtergrond van belang om mee te nemen in alle valorisatie objectieven (VO1-VO4).
CD_DSCNT_ NATLTY	Origin nationality		
DR_MAR	Start date marriage	Jaar van huwelijk	De startdatum van een huwelijk wordt gebruikt ter aanvulling van LIPRO om levenslooppatronen te construeren. Gezondheid bestuderen vanuit een longitudinaal en levensloopperspectief geeft een dieper inzicht over hoe gezondheid en zorgtrajecten temporeel samenhangen met verschillende levenslooptransities en -posities (RQs 3.1 en 3.2). Ook voor het in kaart brengen van multi-morbiditeit trends met behulp van micro-simulatie is het modelleren van levenslooptransities cruciaal (VO3), terwijl informatie over levenslooptransities potentieel ook een rol spelen in het ontwerpen van gerichte gezondheidszorginterventies (VO4).
CD_STATUS	Indicator on status of individual (year + 1): alive, death, migration or administrative deletion	Apart aangemaakte variabele voor de specifieke noden van dit project.	Deze variabele geeft de status van een individu doorheen de tijd weer op basis van informatie uit het Belgisch rijksregister. Dit is noodzakelijk om te achterhalen waarom een persoon eventueel verdwijnt uit de observatie. CD_STATUS geeft weer of een individu nog in leven is of nog steeds in het land verblijft in het volgend geobserveerde jaar.

CD_STATUS_PTNR	Indicator on status of partner (year + 1): alive, death, migration or administrative deletion		Deze variabele geeft de status van een eventuele partner doorheen de tijd weer op basis van informatie uit het Belgisch rijksregister. Dit is noodzakelijk voor de opvolging van partners van patiënten uit Intego die zelf geen deel uitmaken van de Intego patiëntenpopulatie. CD_STATUS_PTNR geeft weer of een partner nog in leven is of nog steeds in het land verblijft in het volgend geobserveerde jaar. In geval van een verandering in de huishoudsamenstelling en -positie op basis van LIPRO, laat deze informatie toe levenslooptransities accuraat in kaart te brengen door te differentiëren tussen de scheiding of verweduwing van een persoon. Het in kaart brengen van levenslooppatronen is cruciaal met betrekking tot RQs 3.1 en 3.2, en VO3 en VO4.
Dataset: Census 2011/2021			
CAS	Arbeidsmarktsituatie	Voor 1 januari 2011 en 2021. Jaarlijkse updates beschikbaar voor periode 2016-2020.	Arbeidsmarktsituatie is een belangrijk socio-economisch kenmerk, gegeven dat iemands relatie tot de arbeidsmarkt een gezondheidsdeterminant en voorspeller van zorgtrajecten vormt. Toegang tot gezondheidszorg is in belangrijke mate gerelateerd aan arbeidsmarktparticipatie en dus van belang voor het bestuderen clustering van gezondheidcondities (RQ1.2) en een intersectionele analyse van geclusterde ziekten (RQs 2.1, 2.2 en 3.2). Tot slot is arbeidsmarktsituatie van belang om mee te nemen in alle valorisatie objectieven (VO1-VO4).
EDU	Onderwijsniveau volgens ISCED-2011	Voor 1 januari 2011 en 2021. Jaarlijkse updates beschikbaar voor periode 2012-2017.	Onderwijsniveau is een belangrijk socio-economisch en socio-cultureel kenmerk, gegeven dat iemands onderwijsniveau een gezondheidsdeterminant en voorspeller van zorgtrajecten vormt. Gezondheidsgedrag, -geletterdheid en toegang tot gezondheidszorg is in belangrijke mate gerelateerd aan opleiding en dus van belang voor het bestuderen clustering van gezondheidcondities (RQ1.2) en een intersectionele analyse van geclusterde ziekten (RQs 2.1, 2.2 en 3.2). Tot slot is onderwijsniveau van belang om mee te nemen in alle valorisatie objectieven (VO1-VO4).
CD_NACE	Economische sector	NACE-code niveau 1: 21 waarden	Economische sector is als component van werkstatus een belangrijk socio-economisch kenmerk, gegeven dat iemands werksituatie betekenisvol is voor de clustering van gezondheidcondities (RQ1.2) en van belang is voor een intersectionele analyse van geclusterde ziekten (RQs 2.1, 2.2 en 3.2).
CD_SIE_DET	Beroepsstatus		Beroepsstatuut is als component van werkstatus een belangrijk socio-economisch kenmerk, gegeven dat iemands werksituatie betekenisvol is voor de clustering van gezondheidcondities (RQ1.2) en van belang is voor een intersectionele analyse van geclusterde ziekten (RQs 2.1, 2.2 en 3.2).
CD_TAUXPT	Werktijden (per kwartiel)	per kwartiel ingedeeld	Tewerkstellingspercentage is als component van werkstatus een belangrijk socio-economisch kenmerk, gegeven dat iemands werksituatie betekenisvol is voor de clustering van gezondheidcondities (RQ1.2) en van belang is voor een intersectionele analyse van geclusterde ziekten (RQs 2.1, 2.2 en 3.2).
TSH	Eigendomssituatie mbt de wooneenheid van huishoudens		Wooneenheid eigendomssituatie is een belangrijk socio-economisch kenmerk, gegeven dat iemands sociale positie betekenisvol is voor de clustering van gezondheidcondities (RQ1.2) en van belang is voor een intersectionele analyse van geclusterde ziekten (RQs 2.1, 2.2 en 3.2).
YAE	Jaar van aankomst sinds 1980		Jaar van aankomst is noodzakelijk om verblijfsduur als onderdeel van migratie status te berekenen. Migratie status is een belangrijk sociodemografische gezondheidsdeterminant en daarom een
YAT	Jaar van aankomst sinds 2000		individueel kenmerk dat de te bestuderen clustering van gezondheidscondities kan beïnvloeden (RQ1.2). Migratie status is ook een proxy om structurele ongelijkheden mee te meten op basis van migratie achtergrond en zal dus ook als één van de sociale dimensies van ongelijkheid bestudeerd worden vanuit een intersectionele benadering (RQs 2.1, 2.2 en 3.2), gezien het verband met gezondheidsongelijkheden en onderdrukkingssystemen. Tal van eerdere studies hebben ook

			aangetoond dat migratie achtergrond relateert aan zorgtrajecten en erkennen daarnaast het belang van verblijfsduur en herkomstland. Tot slot is migratie achtergrond van belang om mee te nemen in alle valorisatie objectieven (VO1-VO4).
Dataset: IPCAL			
HH_INC_NE T_EQ	Equivalized net household income deciles from 2011 onwards	In decielen op basis van ADI (administratief beschikbaar inkomen van het huishouden)	Inkomen is een belangrijk socio-economisch kenmerk, gegeven dat iemands onderwijsniveau een gezondheidsdeterminant en voorspeller van zorgtrajecten vormt. Toegang tot gezondheidszorg is in belangrijke mate gerelateerd aan inkomen en dus van belang voor het bestuderen van clustering van gezondheidcondities (RQ1.2) en een intersectionele analyse van geclusterde ziekten (RQs 2.1, 2.2 en 3.2).
Bron: Intego			
Dataset: all intego datasets			
Patiënt_id	patient identifier		De gepseudonimiseerde patiënt identificator is noodzakelijk om informatie per patiënt met enerzijds Intego data uit verschillende modules en tabellen te koppelen en anderzijds Statbel data.
Practice_id	practice identifier		De gepseudonimiseerde praktijkniveau identificator is noodzakelijk om de clustering van patiënten in eerstelijnszorg in rekening te brengen.
Dataset: patient_master			
firstreg_date	first intego registration date	Voor deze variabele wordt een gepseudonimiseerde datum berekend. Dit betreft het aantal jaren, maanden en dagen tussen de originele geboortedatum en de betreffende datum, toegepast op de fictieve geboortedatum	De startdatum van patiënten registratie is van belang om de mogelijke observatieperiode per patiënt te definiëren.
lastreg_date	last intego registration date		De einddatum van patiënten registratie is van belang om de mogelijke observatieperiode per patiënt te definiëren.
ycg	list of years with at least one clinical encounter (ycg = yearly contact group)		De actieve jaren van patiënten registratie is van belang om de mogelijke observatieperiode per patiënt te definiëren.
Dataset: patient_dynamic			
increased_com pensation	indicates whether a patient is in the increased reimbursement system		De indicator van het recht op een verhoogde tegemoetkoming in de gezondheidskosten van een patiënt is onontbeerlijk voor validatie objectieven VO1 en VO2. Deze laat namelijk toe om na te gaan in hoeverre het recht op een verhoogde tegemoetkoming gelinkt is aan socio-demografische en socio-economische kenmerken. Daarnaast is deze oefening instrumenteel om proxies voor sociale kwetsbaarheid bij patiënten te construeren.
change_year	year of change in increased reimbursement		De indicator geeft aan wanneer er een verandering plaatsvond in een tijdsvariërend kenmerk van de patiënt. Dit is van belang om een gewijzigde status van verhoogde tegemoetkoming in de tijd te situeren en accuraat te relateren aan tijdsvariërende achtergrondkenmerken uit de nationale registers van Statbel.
Dataset: patient_lifestyle			

lifestyle_date	date of patient lifestyle record / registration	Voor deze variabele wordt een gepseudonimiseerde datum berekend.	Datum van registratie is essentieel om de levensstijlkenmerken van een patiënt accuraat te relateren aan verschillende gezondheidscondities, hun clustering en positie in de levensloop.
physical_activity	physical activity status of a patiënt		Fysieke activiteit is een betekenisvolle voorspeller van gezondheidscondities en daarom belangrijk om in rekening te brengen als controlevariabele in de analyse van de clustering van gezondheidscondities (RQ1.1-RQ1.3 en RQ2.1-RQ2.2) en de link tussen levenslooptransities en gezondheid (RQ3.1-RQ3.2).
smoking_status	smoking status of a patient		Tabaksgebruik is een betekenisvolle voorspeller van gezondheidscondities en daarom belangrijk om in rekening te brengen als controlevariabele in de analyse van de clustering van gezondheidscondities (RQ1.1-RQ1.3 en RQ2.1-RQ2.2) en de link tussen levenslooptransities en gezondheid (RQ3.1-RQ3.2).
alcohol_use	alcohol use of a patient		Alcoholgebruik is een betekenisvolle voorspeller van gezondheidscondities en daarom belangrijk om in rekening te brengen als controlevariabele in de analyse van de clustering van gezondheidscondities (RQ1.1-RQ1.3 en RQ2.1-RQ2.2) en de link tussen levenslooptransities en gezondheid (RQ3.1-RQ3.2).
Dataset: patient_body_measures			
body_measures_date	date of patient body measures record / registration	Voor deze variabele wordt een gepseudonimiseerde datum berekend.	Datum van registratie is essentieel om de levensstijlkenmerken van een patiënt accuraat te relateren aan verschillende gezondheidscondities, hun clustering en positie in de levensloop.
length_cm	length of a patient in cm		Lengte is noodzakelijk om BMI te berekenen als dit ontbreekt. BMI is een relevante risicofactor voor verscheidene chronische (cardiovasculaire en metabole) aandoeningen en daarom belangrijk om in rekening te brengen als controlevariabele in de analyse van de clustering van gezondheidscondities (RQ1.1-RQ1.3 en RQ2.1-RQ2.2) en de link tussen levenslooptransities en gezondheid (RQ3.1-RQ3.2).
weight_kg	weight of a patient in kg		Gewicht is een relevante risicofactor voor verscheidene chronische (cardiovasculaire en metabole) aandoeningen en daarom belangrijk om in rekening te brengen als controlevariabele in de analyse van de clustering van gezondheidscondities (RQ1.1-RQ1.3 en RQ2.1-RQ2.2) en de link tussen levenslooptransities en gezondheid (RQ3.1-RQ3.2).
bmi	bmi of a patient		BMI is een relevante risicofactor voor verscheidene chronische (cardiovasculaire en metabole) aandoeningen en daarom belangrijk om in rekening te brengen als controlevariabele in de analyse van de clustering van gezondheidscondities (RQ1.1-RQ1.3 en RQ2.1-RQ2.2) en de link tussen levenslooptransities en gezondheid (RQ3.1-RQ3.2).
abd_circum	abdominal circumference of a patient		Buikomtrek is een relevante risicofactor voor verscheidene chronische (cardiovasculaire en metabole) aandoeningen en daarom belangrijk om in rekening te brengen als controlevariabele in de analyse van de clustering van gezondheidscondities (RQ1.1-RQ1.3 en RQ2.1-RQ2.2) en de link tussen levenslooptransities en gezondheid (RQ3.1-RQ3.2).
Dataset: history_chronic_diseases			
chronic_disease_date	date of patient chronic disease record / registration	Voor deze variabele wordt een gepseudonimiseerde datum berekend.	Datum van registratie is essentieel om diagnose van chronische aandoeningen van een patiënt accuraat en chronologisch te reconstrueren om verschillende gezondheidscondities, hun clustering en positie in de levensloop in kaart te brengen.
icd10st	ICD 10 code		De ICD-10 code is van belang om de specifieke diagnose van de patiënt zoals het geregistreerd werd door de huisarts te kennen. Deze informatie over chronische aandoeningen waar een patiënt aan lijdt is essentieel voor het beantwoorden van alle onderzoeksvragen (RQ1-RQ3) en het bereiken van elk validatie objectief (VO1-VO4).

icpc2	ICPC2 code		De ICPC-2 code is van belang om de concrete reden van bezoek of diagnose van de patiënt zoals het geregistreerd werd door de huisarts te kennen. Deze informatie over chronische aandoeningen waar een patiënt aan lijdt is essentieel voor het beantwoorden van alle onderzoeksvragen (RQ1-RQ3) en het bereiken van elk validatie objectief (VO1-VO4).
ibui	ibui code		De Belgische Unieke Identificator is van belang om de specifieke diagnose van de patiënt zoals het geregistreerd werd door de huisarts te kennen. Deze informatie over chronische aandoeningen waar een patiënt aan lijdt is essentieel voor het beantwoorden van alle onderzoeksvragen (RQ1-RQ3) en het bereiken van elk validatie objectief (VO1-VO4).
Dataset: history_procedures			
procedure_date	date of patient procedure record / registration	Voor deze variabele wordt een gepseudonimiseerde datum berekend.	Datum van registratie is essentieel om de medische procedures van een patiënt accuraat en chronologisch te reconstrueren om verschillende gezondheidscondities, hun clustering en positie in de levensloop in kaart te brengen.
proced_text	medical procedures in text format		Gegevens over medische procedures van een patiënt bevat mogelijk aanvullende informatie over co-morbiditeiten. Dit is belangrijk om de gezondheidsconditie van een patiënt te construeren en bijgevolg essentieel voor het beantwoorden van alle onderzoeksvragen (RQ1-RQ3) en het bereiken van elk validatie objectief (VO1-VO4).
Dataset: history_vaccines			
vaccination_date	date of patient vaccine record / registration	Voor deze variabele wordt een gepseudonimiseerde datum berekend.	Datum van registratie is essentieel om vaccinatie van een patiënt accuraat en chronologisch te reconstrueren om zorgtrajecten, hun clustering en positie in de levensloop in kaart te brengen.
vac_code	history of vaccination	ATC waar beschikbaar, CNK indien niet beschikbaar.	Informatie over vaccinaties is een belangrijke indicator van zorgkwaliteit. Dit is betekenisvol voor RQ1.1 dat het verloop van zorgtrajecten in functie van de interactie tussen aandoeningen in kaart brengt. Daarnaast is informatie over zorgkwaliteit van belang voor het begrip van hoe levenslooptransities relateren aan de clustering van gezondheidcondities (RQ3.1-RQ3.2).
Dataset: diagnoses			
diagnosis_date	date of patient diagnosis record / registration	Voor deze variabele wordt een gepseudonimiseerde datum berekend.	Datum van registratie is essentieel om diagnoses bij een patiënt accuraat en chronologisch te reconstrueren om verschillende gezondheidscondities, hun clustering en positie in de levensloop in kaart te brengen.
icd10	icd10 code of disease or medical condition diagnosed		De ICD-10 code is van belang om de specifieke diagnose van de patiënt zoals het geregistreerd werd door de huisarts te kennen. Deze informatie over de gezondheidsconditie van een patiënt is essentieel voor het beantwoorden van alle onderzoeksvragen (RQ1-RQ3) en het bereiken van elk validatie objectief (VO1-VO4).
icpc2	icpc2 code of disease or medical condition diagnosed		De ICPC-2 code is van belang om de concrete reden van bezoek of diagnose van de patiënt zoals het geregistreerd werd door de huisarts te kennen. Deze informatie over de gezondheidsconditie van een patiënt is essentieel voor het beantwoorden van alle onderzoeksvragen (RQ1-RQ3) en het bereiken van elk validatie objectief (VO1-VO4).
ibui	ibui code of disease or medical condition diagnosed		De Belgische Unieke Identificator is van belang om de specifieke diagnose van de patiënt zoals het geregistreerd werd door de huisarts te kennen. Deze informatie over de gezondheidsconditie van een

			patiënt is essentieel voor het beantwoorden van alle onderzoeksvragen (RQ1-RQ3) en het bereiken van elk validatie objectief (VO1-VO4).
diag_text	diagnosis information in text format		Aanvullende informatie over de specifieke diagnose van de patiënt is belangrijk om de gezondheidsconditie van een patiënt te construeren en bijgevolg essentieel voor het beantwoorden van alle onderzoeksvragen (RQ1-RQ3) en het bereiken van elk validatie objectief (VO1-VO4).
Dataset: care_approaches			
approach_start_date	start date of care approach	Voor deze variabele wordt een gepseudonimiseerde datum berekend.	De startdatum van een zorgelement is belangrijk om het zorgtraject van een patiënt accuraat en chronologisch te reconstrueren om hun wisselwerking met gezondheidcondities en positie in de levensloop in kaart te brengen.
approach_end_date	end date of care approach	Voor deze variabele wordt een gepseudonimiseerde datum berekend.	De einddatum van een zorgelement is belangrijk om het zorgtraject van een patiënt accuraat en chronologisch te reconstrueren om hun wisselwerking met gezondheidcondities en positie in de levensloop in kaart te brengen.
approach_id	care approach title	Drie mogelijke zorgtrajecten: (1) pre-diabetes, (2) chronische nierinsufficiëntie, (3) diabetes.	De identificatie van zorgelementen is essentieel voor RQ1.1 dat het verloop van zorgtrajecten in functie van de interactie tussen aandoeningen in kaart brengt. Daarnaast is informatie over zorgtrajecten van belang voor het begrip van hoe levenslooptransities relateren aan de clustering van gezondheidcondities (RQ3.1-RQ3.2).
Dataset: prescriptions			
prescription_start_date	start date of prescription	Voor deze variabele wordt een gepseudonimiseerde datum berekend.	De startdatum van een voorschrift is belangrijk om het therapeutische traject van een patiënt accuraat en chronologisch te reconstrueren om hun wisselwerking met gezondheidcondities en positie in de levensloop in kaart te brengen.
prescription_end_date	end date of prescription	Voor deze variabele wordt een gepseudonimiseerde datum berekend.	De einddatum van een voorschrift is belangrijk om het therapeutische traject van een patiënt accuraat en chronologisch te reconstrueren om hun wisselwerking met gezondheidcondities en positie in de levensloop in kaart te brengen.
pres_atc	medication ATC code	ATC waar beschikbaar, CNK indien niet beschikbaar.	De ATC-code is van belang om geneesmiddelen voorschriften van patiënten te kennen. Dit is betekenisvol voor RQ1.1 dat het verloop van therapeutische trajecten in functie van de interactie tussen aandoeningen in kaart brengt. Daarnaast is informatie over therapeutische trajecten van belang voor het begrip van hoe levenslooptransities relateren aan de clustering van gezondheidcondities (RQ3.1-RQ3.2).
pres_status	medication status	Twee mogelijke statussen: (1) actief en (2) passief.	In combinatie met informatie over de start- en einddatum is de status van een voorschrift belangrijk om het therapeutische traject van een patiënt accuraat en chronologisch te reconstrueren om hun wisselwerking met gezondheidcondities en positie in de levensloop in kaart te brengen.
pres_type	medication type	Mogelijke types: (1) eenmalig, (2) acuut, (3) chronisch.	Informatie over de status van voorgeschreven medicatie is van belang om onderscheid te maken tussen een langdurige therapie en beperkt gebruik. Dit is betekenisvol voor RQ1.1 dat het verloop van therapeutische trajecten in functie van de interactie tussen aandoeningen in kaart brengt. Daarnaast is informatie over therapeutische trajecten van belang voor het begrip van hoe levenslooptransities relateren aan de clustering van gezondheidcondities (RQ3.1-RQ3.2).
dosage_text	dosage in text format		De dosering van voorgeschreven geneesmiddelen van belang om eventuele wijzigingen in de therapeutische trajecten van patiënten te kennen. Dit is betekenisvol voor RQ1.1 dat het verloop van therapeutische trajecten in functie van de interactie tussen aandoeningen in kaart brengt. Daarnaast is

			informatie over therapeutische trajecten van belang voor het begrip van hoe levenslooptransities relateren aan de clustering van gezondheidcondities (RQ3.1-RQ3.2).
Dataset: laboratory_results			
laboratory_test_date	date of laboratory test request	Voor deze variabele wordt een gepseudonimiseerde datum berekend.	Datum van registratie is essentieel om labotesten bij een patiënt accuraat en chronologisch te reconstrueren om verschillende gezondheidcondities, hun clustering en positie in de levensloop in kaart te brengen.
medidoc_code	medidoc lab test code		Deze informatie is nodig ter operationalisering van de aangevraagde labotest en de accurate interpretatie ervan.
lab_result	lab test result (clean)		Resultaten van labotesten zijn noodzakelijk om een beter beeld te krijgen van hoe (geclusterde) ziekten verlopen (RQ1.1). Daarnaast verschaffen ze belangrijke informatie over de aanwezige risicofactoren voor verscheidene chronische (bv. cardiovasculaire en metabole) aandoeningen en daarom belangrijk om in rekening te brengen als controlevariabele in de analyse van de clustering van gezondheidcondities (RQ1.1-RQ1.3 en RQ2.1-RQ2.2) en de link tussen levenslooptransities en gezondheid (RQ3.1-RQ3.2). Tot slot is het uitvoeren van labotesten een indicator van zorgkwaliteit (cf. RQ1.1, RQ3.1-RQ3.2).
lab_unit	lab test result unit		Deze informatie is nodig voor een accurate interpretatie van labotest resultaten.
Dataset: measurements			
measure_date	date of patient measurement record / registration	Voor deze variabele wordt een gepseudonimiseerde datum berekend.	Datum van registratie is essentieel om metingen bij een patiënt accuraat en chronologisch te reconstrueren om verschillende gezondheidcondities, hun clustering en positie in de levensloop in kaart te brengen.
measure_code	measurement code	Mogelijke metingen: bloeddruk, hartslag, lichaamstemperatuur, etc.	Resultaten van metingen zijn noodzakelijk om een beter beeld te krijgen van hoe (geclusterde) ziekten verlopen (RQ1.1). Daarnaast verschaffen ze belangrijke informatie over de aanwezige risicofactoren voor verscheidene chronische (bv. cardiovasculaire en metabole) aandoeningen en daarom belangrijk om in rekening te brengen als controlevariabele in de analyse van de clustering van gezondheidcondities (RQ1.1-RQ1.3 en RQ2.1-RQ2.2) en de link tussen levenslooptransities en gezondheid (RQ3.1-RQ3.2). Tot slot is het uitvoeren van deze metingen een indicator van zorgkwaliteit (cf. RQ1.1, RQ3.1-RQ3.2).
measure_value	measurement value		
measure_unit	measurement unit		
Dataset: imaging_orders			
imaging_order_date	date of patient medical imaging request	Voor deze variabele wordt een gepseudonimiseerde datum berekend.	Datum van registratie is essentieel om aanvraag voor medische beeldvorming bij een patiënt accuraat en chronologisch te reconstrueren. Informatie over medische beeldvorming is een belangrijke indicator van zorgkwaliteit. Dit is betekenisvol voor RQ1.1 dat het verloop van zorgtrajecten in functie van de interactie tussen aandoeningen in kaart brengt. Daarnaast is informatie over zorgkwaliteit van belang voor het begrip van hoe levenslooptransities relateren aan de clustering van gezondheidcondities (RQ3.1-RQ3.2).
Dataset: physiotherapy_prescriptions			
physio_prescription_date	date of physiotherapy prescription	Voor deze variabele wordt een gepseudonimiseerde datum berekend.	Datum van registratie van een aanvraag voor fysiotherapie is van belang om het therapeutisch traject van een patiënt accuraat en chronologisch te reconstrueren. Dit is betekenisvol voor RQ1.1 dat het verloop van therapeutische trajecten in functie van de interactie tussen aandoeningen in kaart brengt.

			Daarnaast is informatie over therapeutische trajecten van belang voor het begrip van hoe levenslooptransities relateren aan de clustering van gezondheidcondities (RQ3.1 -RQ3.2).
Dataset: absence_certificates			
absence_start_date	start date of absence certificate	Voor deze variabele wordt een gepseudonimiseerde datum berekend.	Datum van periode van afwezigheid op het werk is van belang voor RQ1.1 om uitkomsten van clustering van gezondheidscondities in kaart te brengen.
absence_end_date	end date of absence certificate		
Dataset: intego environmental			
train_noise	train noise in decibels (annual mean)	Weighted noise level caused by rail transport over 24h (2016)	Omgevingskenmerken vormen een centraal element van een syndemische benadering van gezondheidcondities. Gelinkte omgevingskenmerken zijn van cruciaal belang om de contextuele factoren die samenhangen met de clustering en interactie van gezondheidcondities te onderzoeken (RQ1.3).
road_noise	road noise in decibels (annual mean)	Weighted noise level caused by rail transport over 24h (2018)	
aviation_noise	aviation noise in decibels (annual mean)	Weighted noise level caused by rail transport over 24h (2016)	
access_residential_green_space	access to residential green space	Binaire variabelen voor verschillende buffer zones (2016, 2022)	
coverage_residential_green_space	residential greenspace coverage in percent	Dekking voor verschillende buffer zones (2021)	
view_green_space	view of green from residence	Viewshed Greenness Visibility Index (2021)	
air_quality_pm10	concentration particulate matter in µg/m³ (annual mean)	Mean of daily concentration (2009-2022)	
air_quality_pm25	concentration particulate matter in µg/m (annual mean)		
air_quality_no2	nitrogen dioxide concentration in µg/m (annual mean)		
air_quality_o3	ozone concentration in µg/m³ (annual mean)		

air_quality_bc	black carbon concentration in $\mu\text{g}/\text{m}^3$ (annual mean)	Mean of daily concentration (2010-2022)	
n_heatwave_days	number of days of official heatwaves (yearly)	2000-2021	
heatwave_degree_days	summary measure of heatwave days and average temperature (yearly)		
n_heatalarm_days	number of days of official heatwave alarm (yearly)		
t_mean_daily	daily mean temperature in $^{\circ}\text{C}$		
relative_humidity_max	daily maximum relative humidity in percentage		
relative_humidity_min	daily minimum relative humidity in percentage		
relative_humidity_mean	daily mean relative humidity in percentage		
distance_to_road	distance in to a major road (in meter) (yearly)	2019	
traffic_load	number of vehicles multiplied by length of roads within 200m around the residence (yearly)	2022	