

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

IVC/KSZG/21/032

BERAADSLAGING NR. 21/038 VAN 1 JUNI 2021 MET BETREKKING TOT DE VERWERKING VAN PERSOONSgegevens DOOR DE KRUIPUNTBANK VAN DE SOCIALE ZEKERHEID EN DE VITO EN DE MEDEDELING VAN PERSOONSgegevens DIE DE GEZONDHEID BETREFFEN AFKOMSTIG VAN DE FOD VOLKSgezONDHEID EN HET RIZIV AAN DE VITO VIA HET HEALTHDATA.BE-PLATFORM IN HET KADER VAN EEN EPIDEMIOLOGISCHE STUDIE OVER HET VOORKOMEN VAN MYOCARDINFARCT - EMIR STUDY

Het Informatieveiligheidscomité, kamer sociale zekerheid en gezondheid,

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);

Gelet op de wet van 3 december 2017 *tot oprichting van de Gegevensbeschermingsautoriteit*, in het bijzonder artikel 114, gewijzigd bij de wet van 25 mei 2018;

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 5 september 2018 *tot oprichting van het informatieveiligheidscomité en tot wijziging van diverse wetten betreffende de uitvoering van verordening (EU) 2016/679 van 27 april 2016 van het Europees Parlement en de Raad 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*, in het bijzonder artikel 97;

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 de artikelen 5, 15 en 42;

Gelet op de wet van 21 augustus 2008 *houdende oprichting en organisatie van het eHealth-platform*;

Gelet op de machtigingsaanvraag van Healthdata;

Gelet op het auditoraatsrapport van het eHealth-platform van 26 mei 2021;

Gelet op het verslag van de heer Michel Deneyer;

Beslist op 1 juni 2021, na beraadslaging, als volgt:

I. VOORWERP VAN DE AANVRAAG

1. Het vroegere Sectoraal Comité van de Sociale Zekerheid en van de Gezondheid heeft zich in het verleden uitgesproken over een epidemiologische studie met betrekking tot het Tako-Tsubo-syndroom¹. Deze studie heeft de link aangetoond tussen variaties in de luchtkwaliteit en het voorkomen binnen de 24 uur van een bepaald type van hartinfarct. De onderzoekers wensen deze resultaten nu te verfijnen door middel van koppeling van verschillende gegevensbanken met persoonsgegevens die de gezondheid betreffen en met milieugegevens aangereikt door de VITO² en GPS-coördinaten van de verblijfplaats en de werkplaats van de betrokken patiënten.
2. De betrokkenen zijn patiënten die in de Belgische ziekenhuizen worden opgenomen wegens een myocardinfarct van het type STEMI (ST elevation myocardial infarction), NSTEMI (Non ST elevation myocardial infarction) of Tako-Tsubo (stressgerelateerde cardiomyopathie).

Er worden twee aparte cohorten van patiënten geanalyseerd:

- 1) patiënten opgenomen tussen 1/01/2012 en 31/12/2018, ofwel in totaal 81297 patiënten,
 - 2) patiënten opgenomen tussen 1/01/2019 en 31/12/2021 (gegevens beschikbaar begin 2023).
3. De selectie van de patiënten gebeurt op het niveau van de Minimale Ziekenhuisgegevens (MZG) op basis van de ICD-codes die met de bestudeerde ziektebeelden overeenstemmen³. De ICD-codes in nomenclatuur ICD 9 (2012,2013,2014) en hun overeenstemming in nomenclatuur ICD 10 (2016,2017,2018).
 4. Deze studie zal worden verricht aan de hand van persoonsgegevens die de gezondheid betreffen die door de technische cel voor het beheer van de Minimale Ziekenhuisgegevens van het RIZIV worden meegedeeld en aan de hand van persoonsgegevens die de gezondheid betreffen afkomstig uit het register Qermid TUCO van het RIZIV. Deze gegevens worden aan de KSZ en aan Healthdata meegedeeld op basis van het gepseudonimiseerde INSZ van de patiënt.

A) Gegevens uit de Minimale Ziekenhuisgegevens (MZG)

¹ Beraadslaging nr. 18/ 090 van 3 juli 2018, gewijzigd op 4 september 2018, met betrekking tot de mededeling van anonieme persoonsgegevens die de gezondheid betreffen door de FOD Volksgezondheid aan het UMC Sint-Pieter in het kader van een epidemiologisch onderzoek over het Tako-Tsubo-syndroom

² Vlaamse Instelling voor Technologisch Onderzoek.

³ De lijst van de noodzakelijke ICD-codes voor de studie gaat als bijlage.

- 1) Administratieve MZG: datum van ziekenhuisopname, geboortjaar van de patiënt, leeftijd van de patiënt bij de opname, geslacht, postcode van de verblijfplaats van de patiënt, het erkenningsnummer van het ziekenhuis, het type opname, de aanduiding van de persoon die de persoon heeft doorverwezen, het type ontslag, de datum van ontslag uit het ziekenhuis, de verblijfsduur, de bestemming, de categorie van hoofddiagnose (mdc), de voor de groep gebruikte hoofddiagnose, de diagnose gerelateerd aan de groep, de graad van ernst en het mortaliteitsrisico.
- 2) MZG hoofddiagnose: STEMI, NSTEMI, Tako-Tsubo.
- 3) MZG nevend diagnoses: diabetes, nierinsufficiëntie, arteriële hypertensie, perifere vasculaire insufficiëntie, cerebrovasculaire aandoening, roken, dyslipidemie.
- 4) MZG therapeutische procedure: procedure PCI, procedure CABG

B) Gegevens afkomstig van het register Oermid TUCO (RIZIV)

- 1) Administratieve gegevens: opnamedatum, datum PCI, erkenningsnummer ziekenhuis, leeftijd op het ogenblik van de opname in het ziekenhuis, geslacht van de patiënt, postcode van de verblijfplaats.
- 2) Indicatie van de behandeling (hoofddiagnose): hoofdindicatie, bijkomende indicatie
- 3) Cardiovasculaire risicofactor (nevend diagnose): grootte, gewicht, BMI, diabetes, nierinsufficiëntie (<30 ml/min), cerebrovasculair accident, perifere arteriopathie
- 4) Klinische ernst: cardiogene shock bij opname, aorto-coronaire bypass gepland/dringend/nee, toestand na ontslag (levend/overleden), datum van ontslag indien levend, oorzaak van het overlijden.
- 5) Anatomische ernst: ziekte 1-2-3 vaten, gemeenschappelijke stam, rechts/codominantie/links, aantal segmenten behandeld in de procedure, stent-trombose, laesie, type materiaal, segment (PL 1-2-3-4)

C) Gegevens meegedeeld door de Kruispuntbank van de Sociale Zekerheid (KSZ)

5. De KSZ deelt een selectie van gepseudonimiseerde persoonsgegevens uit het datawarehouse arbeidsmarkt en sociale bescherming en de GPS-coördinaten van het adres van de hoofdverblijfplaats en van het beroepsadres van de patiënt mee. Deze omzetting zal door VITO worden gerealiseerd in onderaanneming van de KSZ in de lokalen van de KSZ.

- 1) Adres van de hoofdverblijfplaats op het ogenblik van het infarct, gegeocodeerd in GPS-gegevens
- 2) Beroepsadres op het ogenblik van het infarct, gegeocodeerd in GPS-gegevens
- 3) Socio-economische gegevens uit het datawarehouse arbeidsmarkt en sociale bescherming: burgerlijke staat, socioprofessionele categorie (loontrekkende bediende, loontrekkende arbeider, ambtenaar, zelfstandige, gepensioneerde, werkloze), categorie van het jaarinkomen (categorie 1: 1^{ste} deciel, categorie 2: 2^{de} deciel, categorie 3: 3^{de} deciel, categorie 4: 4^{de} deciel, categorie 5: 5^{de} deciel, categorie 6: 6^{de} deciel, categorie 7: 7^{de} deciel, categorie 8: 8^{de} deciel, categorie 9: 9^{de} deciel categorie 10: 10^{de} deciel)
- 4) Status levende of overleden, jaar van het overlijden.

D) Gegevens afkomstig van de BIWAC-gegevensbank

6. De studie vereist eveneens de mededeling van gedepersonaliseerde persoonsgegevens afkomstig van de Belgian Interdisciplinary Working Group on Acute Cardiology (BIWAC).
- 1) BIWAC administratieve gegevens: opnamedatum, erkenningsnummer ziekenhuis, geslacht, leeftijd bij opname
 - 2) BIWAC indicatie van de behandeling (hoofddiagnose): type STEMI-infarct (anterior, niet-anterior)
 - 3) BIWAC cardiovasculaire risicofactoren (nevendiagnose): roken, diabetes, hoge bloeddruk, dyslipidemie, nierinsufficiëntie, cerebrovasculaire aandoening, gewicht en grootte
 - 4) BIWAC klinische ernst: Killip-score (1 tot 4), cardiogene shock of hartstilstand bij opname, tijdstip van begin van de pijn, tijdstip van reperfusie (stent of trombolysie), overlijden tijdens het ziekenhuisverblijf
 - 5) BIWAC anatomische ernst: datum van PCI, soort PCI (culprit vessel)

E) Anonieme gegevens uit het ILI-register

7. Gedepersonaliseerde gegevens uit het ILI-register (influenza-like illness) zullen door Sciensano worden meegedeeld.

Het betreft 3 types gegevens over de incidentie van luchtweginfecties met influenza (seizoensgriep):

- wekelijkse incidentie van de raadplegingen wegens griepsymptomen per leeftijdscategorie en per gewest;
- positiviteitsratio voor influenza per leeftijdscategorie en per gewest;
- dagelijks percentage van positieve serologische tests in het laboratoriumnetwerk dat ongeveer 60% van de Belgische microbiologische laboratoria beslaat.

F) Milieugegevens - VITO

VITO levert 6 groepen van milieugegevens met geocodes die overeenstemmen met de door de KSZ geleverde verblijfs- en beroepsadressen:

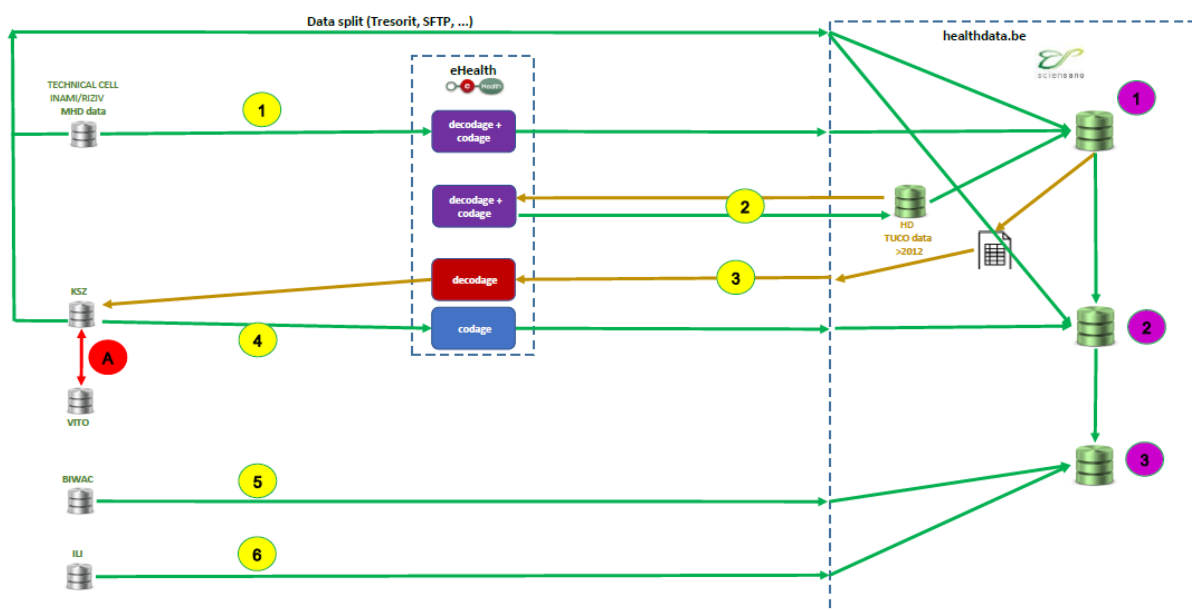
- 1) luchtkwaliteit: daggemiddelde van deeltjesverontreiniging (PM 10, PM2,5), gasvormige verontreiniging (NO₂, O₃) met een resolutie van 4 km²;
- 2) luchttemperatuur: daggemiddelde met een resolutie van 9km²;
- 3) blootstelling aan lawaai van wegen en luchthavens: jaargemiddelde met GPS-resolutie;
- 4) vegetatie-index: jaargemiddelde met GPS-resolutie ;
- 5) afstand tot de dichtstbijzijnde hoofdweg: jaargemiddelde met GPS-resolutie ;
- 6) de code van de cel van de RIO-cartografie met de geocodes van de verblijfsadressen en de beroepsadressen. De RIO-cartografie is een kaart van België waarop een rooster met cellen van 4x4km² is toegepast. Elke GPS-coördinaat wordt gelinkt met de code van de overeenstemmende cel.

VITO levert ook niet-milieugebonden gegevens afkomstig van cartografieën beschikbaar bij StatBel en Censat en met geocodes die overeenkomen met de verblijfs- en beroepsadressen. Deze niet-milieugebonden gegevens betreffen 2 categorieën:

7) bevolkingsdichtheid;

8) socio-economische indexen van het niveau van opleiding en van inkomen

8. De gepseudonimiseerde persoonsgegevens die de gezondheid betreffen zullen via 6 verschillende gegevensstromen en drie gegevensverwerkingspunten worden ingezameld.



Toelichting van de verschillende gegevensstromen (gele cirkels):

- 1) De minimale ziekenhuisgegevens worden aan het RIZIV gevraagd. Enkel de registraties voor een specifieke lijst van ICD-codes worden gevraagd. De MZG-gegevens zullen worden overgemaakt van de technische cel van het RIZIV naar healthdata.be in 2 aparte bestanden.
 - Een dossier met het INSZ van de patiënt en een technische ID. Na codering van het INSZ door de dienst codage eHBox worden de gegevens overgemaakt naar het gegevensdepot van healthdata.be.
 - Het andere bestand bestaat uit MZG-gegevens met dezelfde technische ID. Aangezien dit tweede bestand geen te coderen gegevens bevat, wordt het bestand rechtstreeks verstuurd door de technische cel aan het gegevensdepot van healthdata.be met de hulp van Tresorit, SFTP of een andere methode van gecrypteerde bestandstransfer.
- 2) Healthdata.be zamelt de TUCO-gegevens in sedert 2012. Wat de TUCO-gegevens betreft, is op dit ogenblik enkel een patiëntidentificatienummer QERMID aanwezig in het HD-gegevensdepot. Een hercodering via het QERMID-team en vervolgens via de eHBox decodering/codering is noodzakelijk om een gecodeerd INSZ patiënt 'healthdata.be' te

krijgen dat aan de MZG-gegevens kan worden gelinkt. De gelinkte dataset van TUCO- en MZG-gegevens wordt als basis gebruikt voor het EMIR-project.

- 3) Na de MZG- en de TUCO-gegevens te hebben gekoppeld, wordt een lijst van patiënten opgeroepen voor wie bijkomende gegevens aan de KSZ worden gevraagd. De INSZ van de patiënten die gecodeerd zijn volgens de 'healthdata.be'-sleutel en de datum van de gebeurtenis (infarct) worden voor decoding aan de eHBox overgemaakt. De KSZ gebruikt het niet-gecodeerd INSZ om de adressen en andere gegevens van de patiënten op te halen.
- 4) De KSZ staat in voor het **coderen van de adressen in GPS-coördinaten** door gebruik te maken van de gegevensbank "Belgian Street and Adresses (BeSt)" van de federale overheidsdienst Beleid en Ondersteuning (FOD BOSA). VITO zal de koppeling realiseren met **de milieugegevens en de niet-milieugegevens** van het geografisch informatiesysteem (GIS) waarbij het zich enkel zal baseren op de geocode, de datum van het infarct en de technische ID. Wat de milieugegevens met dagelijkse variatie betreft (luchttemperatuur, PM 10, PM2.5; ozon, NO2), VITO zal de gegevenskoppeling beperken tot de 45 voorafgaande dagen en tot 45 dagen volgend op het infarct. Deze koppelingsprocedure van de milieu- en niet-milieugegevens met de GPS-coördinaten en de datum van het infarct zal geschieden door VITO.

De koppeling van de milieu- en de niet-milieugegevens zal gebeuren in de lokalen van de KSZ onder haar toezicht. De milieugegevens zullen door VITO worden meegebracht naar de KSZ om de koppeling met de geolocalisatiegegevens te verrichten door gebruik te maken van python-scripts. De gekoppelde gegevens zullen enkel op de server van de KSZ worden opgeslagen. De KSZ verwijdert vervolgens de GPS-coördinaten en stuurt de gegevens op naar healthdata.be in 2 aparte bestanden op dezelfde manier als voor de MZG-gegevens (zie gegevensstroom 1).

- 5) De **BIWAC-gegevens** worden rechtstreeks opgestuurd naar healthdata.be. Deze gegevens bevatten geen identificatiegegevens van de patiënt en moeten niet door het eHealth-platform worden gecodeerd.
- 6) De **gegevens van het ILI-register** worden als bijkomende covariante in de studie ingezameld. De ILI-gegevens bevatten geografische informatie over de griep in België. Het gaat om **geaggregeerde gegevens**; er is geen codering door eHealth vereist.

Verwerkingspunten (paarse cirkels):

- 1) De 2 bestanden met MZG-gegevens worden samengevoegd volgens de technische ID die in de twee bestanden aanwezig is. De MZG- en de TUCO-gegevens worden vervolgens samengevoegd op basis van het gecodeerd INSZ patiënt.
- 2) De 2 gegevensbestanden van de KSZ / VITO worden samengevoegd volgens de technische ID die in beide bestanden aanwezig is. De KSZ- / VITO-gegevens worden vervolgens samengevoegd met de MZG / TUCO gegevens op basis van het gecodeerd INSZ patiënt.
- 3) De BIWAC-gegevens worden met de MZG / TUCO / KSZ / VITO samengevoegd op basis van verschillende attributen (opnamedatum, code ziekenhuisakkoord, geslacht,

opnamedatum). De ILI-gegevens worden samengevoegd op basis van het geografisch mandaat (postcode, district, ...).

9. De samengevoegde gegevens als gevolg van deze koppeling zullen door de onderzoekers van het EMIR-team worden geanalyseerd. Dit team is samengesteld uit onderzoekers van de Universiteit van Hasselt, de VUB en de ULB. De onderzoekers krijgen geen toegang tot de gepseudonimiseerde gegevens die voor de koppeling zijn gebruikt.

II. BEVOEGDHEID

10. Ingevolge artikel 42, § 2, 3^o van de wet van 13 december 2006 *houdende diverse bepalingen betreffende gezondheid* is de kamer sociale zekerheid en gezondheid van het Informatieveiligheidscomité in beginsel bevoegd voor het verlenen van een beraadslaging met betrekking tot elke mededeling van persoonsgegevens die de gezondheid betreffen.
11. Krachtens artikel 5, § 1, van de wet van 15 januari 1990 *houdende oprichting en organisatie van een Kruispuntbank van de Sociale Zekerheid* verzamelt de Kruispuntbank van de Sociale Zekerheid persoonsgegevens bij de instellingen van sociale zekerheid, slaat ze op, voegt ze samen en deelt ze mee aan de personen die ze nodig hebben voor het verrichten van onderzoeken die nuttig zijn voor de kennis, de conceptie en het beheer van de sociale bescherming.
12. Wat de mededeling van persoonsgegevens uit het datawarehouse arbeidsmarkt en sociale bescherming betreft, het gaat voorts om een mededeling van persoonsgegevens die volgens artikel 15, § 1, van dezelfde wet van 15 januari 1990 een voorafgaande beraadslaging van de kamer sociale zekerheid en gezondheid van het Informatieveiligheidscomité vereist.
13. Het comité oordeelt bijgevolg dat het bevoegd is om zich over deze machtigingsaanvraag uit te spreken.

III. BEHANDELING

A. TOELAATBAARHEID

14. De verwerking van persoonsgegevens die de gezondheid betreffen is in principe verboden, overeenkomstig artikel 9, § 1, van de Verordening (EU) 2016/679 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 (AVG).
15. Dit verbod is echter niet van toepassing wanneer de verwerking noodzakelijk is met het oog op wetenschappelijk onderzoek of statistische doeleinden overeenkomstig artikel 89, lid 1⁴

⁴ “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 deze Verordening voor de rechten en vrijheden van de betrokkene. Die waarborgen zorgen ervoor dat er technische en organisatorische maatregelen zijn getroffen om de inachtneming van het beginsel van minimale gegevensverwerking te garanderen.

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 betrokkene.

16. Wat de verwerking van persoonsgegevens betreft, is de verwerking enkel rechtmatig indien en voor zover minstens één van de voorwaarden van artikel 6 van de AVG vervuld is.
17. Krachtens artikel 5, § 1, van de wet van 15 januari 1990 *houdende oprichting en organisatie van een Kruispuntbank van de Sociale Zekerheid* verzamelt de Kruispuntbank van de Sociale Zekerheid persoonsgegevens bij de instellingen van sociale zekerheid, slaat ze op, voegt ze samen en deelt ze mee aan de personen die ze nodig hebben voor het verrichten van onderzoeken die nuttig zijn voor de kennis, de conceptie en het beheer van de sociale bescherming.
18. Wat de mededeling van persoonsgegevens uit het datawarehouse arbeidsmarkt en sociale bescherming betreft, het gaat voorts om een mededeling van persoonsgegevens die volgens artikel 15, § 1, van dezelfde wet van 15 januari 1990 een voorafgaande beraadslaging van de kamer sociale zekerheid en gezondheid van het Informatieveiligheidscomité vereist.
19. In het licht van het voorgaande is het comité van oordeel dat er een toelaatbare grond bestaat voor de beoogde verwerking van gepseudonimiseerde persoonsgegevens die de gezondheid betreffen.

B. PRINCIPES MET BETREKKING TOT DE VERWERKING VAN PERSOONSGEGEVENS

1. DOELEINDEN

20. Volgens artikel 5 van de AVG moeten de persoonsgegevens op een wijze die ten aanzien van de betrokkene rechtmatig, behoorlijk en transparant is worden verwerkt. Ze moeten voor welbepaalde, uitdrukkelijk omschreven en gerechtvaardigde doeleinden worden ingezameld en mogen vervolgens niet verder op een met die doeleinden onverenigbare wijze worden verwerkt. De latere verwerking met het oog op archivering in het algemeen belang, wetenschappelijk of historisch onderzoek of statistische doeleinden wordt krachtens artikel 89, paragraaf 1, niet beschouwd als niet verenigbaar met de oorspronkelijke doeleinden (doelbinding).
21. Het betreft een latere verwerking voor wetenschappelijke doeleinden in het kader van een epidemiologische studie over het voorkomen van myocardinfarct in relatie met milieumomstandigheden.

Deze maatregelen kunnen pseudonimisering omvatten, mits aldus die doeleinden in kwestie 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 verwezenlijkt.”

22. Gelet op de doelstellingen van de verwerking zoals hierboven beschreven, oordeelt het Comité dat de verwerking van de voormelde persoonsgegevens wel degelijk welbepaalde, uitdrukkelijk omschreven en gerechtvaardigde doeleinden nastreeft.

2. MINIMALE GEGEVENSVERWERKING

23. 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.
24. Het Comité stelt vast dat de volgende gepseudonimiseerde persoonsgegevens zullen worden verwerkt in het kader van deze studie. De gegevens worden als gepseudonimiseerd beschouwd aangezien de gegevensinzameling ofwel op basis van het INSZ (dat niet aan de onderzoeker wordt overgemaakt), ofwel op basis van de datum van het infarct geschiedt. De identiteit van de betrokken patiënten wordt niet aan de onderzoeker meegedeeld.

a. Gegevens afkomstig van de MZG-gegevensbank - Minimale Ziekenhuisgegevens

1) administratieve MZG

Opnamedatum: Het gaat om een hoofdgegeven voor de EMIR-studie. Dit gegeven wordt gerechtvaardigd door de noodzaak om de milieugegevens op een welbepaald tijdstip te projecteren. Zo zal er een analyse worden uitgevoerd van de lag-structuur (relatie tussen de gebeurtenissen van het infarct en de milieukeurmerken van de vorige dagen) tot 7 dagen voorafgaand aan de opname (lag -7).

Geboortjaar / leeftijd bij opname: Vanuit medisch oogpunt is de leeftijd een cardiovasculaire risicofactor die de invloed van het milieu kan wijzigen. Ouderen personen zijn meer gevoelig voor fijnstof terwijl jongeren een infarct ontwikkelen in relatie met luchtvervuiling door stikstofdioxide. Het geboortjaar zal eveneens dienen voor de kruising van de MZG met gedepersonaliseerde gegevens uit het BIWAC-register. De leeftijd en het geslacht van de patiënt gekoppeld aan de opnamedatum en de erkenningscode van het ziekenhuis van opname zullen het immers mogelijk maken om de gegevens te matchen teneinde een meer nauwkeurige evaluatie te krijgen van het cardiovasculair risicoprofiel van de patiënten.

Geslacht: Vanuit medisch oogpunt is het geslacht een cardiovasculaire risicofactor die de invloed van het milieu kan wijzigen. Cardiovasculaire aandoeningen komen vaker voor bij mannen maar Tako-Tsubo treft vooral vrouwen. Wat de milieu-impact betreft, is de invloed van het geslacht nog niet gekend. Het geslacht zal ook dienen voor het kruisen van de MZG met gedepersonaliseerde gegevens uit het BIWAC-register.

Postcode van de patiënt: De postcode is in de MZG-gegevens beschikbaar en is samen met de code van de cel van de RIO-kaart het enige geografisch gegeven dat nog aanwezig is in de uiteindelijke database die voor de onderzoekers toegankelijk is. De gegevens met betrekking tot het fysiek en het beroepsadres die door de KSZ aan VITO worden overgemaakt, worden immers verwijderd zodra de milieugegevens zijn ingevuld. De

postcode dient dus voor het opstellen van locoregionale distributiestatistieken van de myocardinfarcten maar bij gebrek aan voldoende spatiale resolutie zal hij niet kunnen dienen voor het modelleren van de impact van milieuverontreiniging.

Erkenningsnummer ziekenhuis: de code van het ziekenhuis van opname zal samen met de opnamedatum, de leeftijd van de patiënt, zijn geslacht dienen voor het kruisen van de MZG-gegevens met de gedepersonaliseerde gegevens van het BIWAC-register.

Soort opname: De soort opname van de patiënten uit de EMIR-studie moet hoofdzakelijk uit dringende opnames bestaan. Aan de hand van deze parameter kan worden nagegaan of de gegevens niet besmet zijn door fouten bij het inbrengen van electief toegelaten patiënten.

Doorverwezen door / bestemming: vele STEMI-patiënten worden in een eerste centrum behandeld en vervolgens doorgestuurd naar een derde centrum waar er een angioplastie wordt verricht alvorens ze teruggaan naar het doorverwijzend centrum. Op basis van deze parameters kunnen de patiënten die bijkomend werden doorverwezen, worden geïdentificeerd. In het kader van de STEMI kan de transfertijd een invloed hebben op de timing van de reperfusie en dus de ernst van het klinische ziektebeeld. Het is dus belangrijk om te anticiperen op de impact van deze cofactor in de evaluatie van de milieu-impact op de ernst van het oorspronkelijke ziektebeeld en de mortaliteit van de infarcten.

Type ontslag / datum ontslag uit het ziekenhuis / verblijfsduur (datum van ontslag - opnamedatum) : het type ontslag uit het ziekenhuis naar huis of een revalidatiecentrum en de totale verblijfsduur zijn ook onrechtstreekse indicaties van de ernst van het ziektebeeld die in de EMIR-studie moeten worden bestudeerd.

Major diagnosis category/ Diagnosis related group/ mate van ernst / mortaliteitsrisico : De MZG-gegevens bevatten naast de ICD-nomenclatuur van de hoofddiagnose een code die overeenstemt met een categorie van de hoofddiagnose en een code diagnosegroep. Deze codes worden in verband gebracht met een berekening van de ernst van het ziektebeeld en het verwachte sterfterisico. Deze berekeningen van a priori risico's zijn gebaseerd op klinische gegevens en aangezien de EMIR-studie een reële evaluatie van de waargenomen mortaliteit inhoudt, is het belangrijk om te bepalen of de impact van het milieu de mogelijke verschillen tussen de verwachte sterfte en de werkelijke sterfte kan verklaren.

2) MZG hoofddiagnose

Het betreft gegevens met betrekking tot de verschillende ICD-codes als hoofddiagnose en op basis waarvan de subtypes van het acuut coronair syndroom zoals STEMI, NSTEMI en Tako-Tsubo kunnen worden geïdentificeerd. Dit hoofdgegeven voor de EMIR-studie wordt verklaard door de noodzaak om een differentiële analyse te verrichten van de milieu-impact op deze 3 subklassen van acuut coronair syndroom. De fysiopathologie ervan is immers zeer verschillend. STEMI is te wijten aan het loskomen van een cholesterolplaque, NSTEMI aan een onevenwicht tussen zuurstoftoevoer en -verbruik in de hartspier, en Tako-Tsubo aan een probleem van coronaire microvasculaire spasmen, waarschijnlijk secundair aan een onevenwicht in het autonome zenuwstelsel. Uit de preliminaire resultaten blijkt dat het milieu een verschillende impact heeft op deze infarct-subtypes.

3) MZG nevendiagnose

Deze gegevenscategorie omvat een reeks **ICD-codes** uit nevendiaagnoses die betrekking hebben op traditionele risicofactoren voor cardiovasculaire aandoeningen. Het is de bedoeling om het cardiovasculair risicoprofiel van de patiënt te bepalen aan de hand van gegevens die aanwezig zijn in de MZG en die werden samengevoegd met de gegevens uit de QERMID-TUCO en de gegevens uit het BIWAC-register. Enkel op basis van een kruising van deze verschillende gegevensbronnen is immers een kwaliteitsvolle evaluatie van het cardiovasculair risico mogelijk. Deze gegevenskruising wordt gerechtvaardigd door het feit dat de nevendiaagnoses van de MZG niet als hoofddoel hebben om de cardiovasculaire risicofactoren weer te geven maar als voordeel hebben dat ze betrekking hebben op de hele studiepopulatie, terwijl de QERMID- en BIWAC-gegevens tot doel hebben om dit risicoprofiel te bepalen maar daarentegen enkel betrekking hebben op de patiënten die met een stent (QERMID) zijn behandeld of die een STEMI (BIWAC) hebben gehad. De evaluatie van het cardiovasculair risicoprofiel van een patiënt is gebaseerd op de identificatie van de aanwezigheid van een tabaksprobleem, diabetes, hypertensie, dyslipidemie en een chronische nierinsufficiëntie. De patiënten die reeds een perifere vasculaire aandoening, een cerebrovasculaire aandoening of een chronisch coronair probleem hebben gehad, worden automatisch als patiënt met een hoog risico beschouwd.

4) MZG therapeutische aanpak

Het grote merendeel van de STEMI- en de NSTEMI-patiënten wordt met een coronaire revascularisatie behandeld, ofwel een percutane revascularisatie door stentimplantatie (PCI) ofwel een revascularisatie met overbrugging (CABG). Het is belangrijk om een statistiek te geven over deze revascularisatiemodaliteiten in de beschrijving van de populatie van de EMIR-studie. De mortaliteit van de infarcten in verband met het milieu kan enkel worden geïnterpreteerd indien alle gekende factoren die op deze mortaliteit een invloed hebben, worden bestudeerd (cardiovasculair risicoprofiel, oorspronkelijke ernst van het ziektebeeld en therapeutische aanpak). Daar de PCI-gegevens bovendien het voorwerp uitmaken van de QERMID-gegevensbank, is het belangrijk om de coherentie van de gegevens na te gaan aan de hand van verificatiepunten. Alle patiënten in de QERMID zouden dus als PCI in het MZG-register moeten zijn opgenomen. Deze coherentie tussen de gegevensbanken zal deel uitmaken van het evaluatieproces van de kwaliteit van de finale database.

b. Gegevens uit het register Qermid TUCO

1) TUCO administratief

De rechtvaardigingen van de administratieve gegevens van het Qermid TUCO register zoals het erkenningsnummer van het ziekenhuis, de opnamedatum, de leeftijd bij opname, het geslacht van de patiënt, het jaar van overlijden van de patiënt, de postcode zijn dezelfde als voor de MZG-gegevens.

De Qermid TUCO gegevensbank bevat ook de datum van de PCI die verschillend kan zijn van de opnamedatum, vooral voor de infarcten van het type NSTEMI. Volgens de aanbevelingen voor de STEMI moet de PCI immers binnen de 90 minuten na de opname gebeuren. De tijd tussen de opname en de PCI voor het NSTEMI-infarct is ook een element

waarmee de klinische ernst van het infarct kan worden geëvalueerd. De EMIR-studie zal deze parameter moeten gebruiken in de modellering van de impact van het milieu op de ernst van het oorspronkelijke ziektebeeld van het infarct.

2) TUCO indicatie van de behandeling

De Qermid TUCO gegevensbank identificeert duidelijk de hoofdindicatie van de stent. Dit gegeven is essentieel voor de EMIR-studie omdat enkel de indicaties STEMI en NSTEMI worden weerhouden. Alle gegevens betreffende de in stabiele omstandigheden verrichte PCI worden niet in deze studie opgenomen. Soms vermeldt de chirurg toekomstige informatie die nuttig kan zijn voor een betere beschrijving van de gevallen vervat in de EMIR-studie. Het gaat hoofdzakelijk om het type PCI in de STEMI en de termijn voor de PCI in de NSTEMI.

3) TUCO cardiovasculaire risicofactoren

De Qermid TUCO gegevensbank bevat een redelijk nauwkeurige evaluatie van het cardiovasculair risico van de patiënt (diabetes, nierinsufficiëntie, cerebrovasculaire aandoening (ACV), perifere vasculaire aandoening). Het Qermid-register bevat bovendien ook gegevens met betrekking tot het gewicht en de grootte en dus de body mass index (BMI) die niet in de MZG-gegevens zijn opgenomen. Het tabaksgebruik is daarentegen niet opgenomen in de Qermid TUCO-gegevens. Deze situatie illustreert waarom de onderzoekers voor een transversale evaluatie van het cardiovasculair risico opteren door de MZG-gegevens en de gegevens uit Qermid TUCO en het BIWAC-register met elkaar te kruisen. Enkel door de complementariteit van deze 3 gegevensbanken krijgt men een betrouwbare evaluatie van het cardiovasculair risico van de patiënten.

4) TUCO klinische ernst

De EMIR-studie heeft als doel om na te gaan of het milieu een invloed heeft op het voorkomen van infarcten, maar ook wat de ernst ervan is op het vlak van sterfte. De evaluatie van de impact van het milieu op de oorspronkelijke ernst van het ziektebeeld is enkel mogelijk indien een reeks klinische factoren worden bestudeerd. Het TUCO-register bevat duidelijk informatie over de aanwezigheid van een cardiogene shock bij opname en indien de patiënt een dringende overbrugging heeft gekregen, geven 2 factoren de ernst aan van het klinisch ziektebeeld. Het Qermid TUCO-register bevat ten slotte de mortaliteitsgegevens waarbij wordt aangegeven of de mortaliteit cardiovasculair was of er een andere interfererende factor was (bloeding, infectie, ...). Dit is belangrijk voor de EMIR-studie die enkel betrekking heeft op de cardiovasculaire mortaliteit die rechtstreeks te wijten is aan een infarct.

5) TUCO anatomische ernst

De mortaliteit moet eveneens worden bestudeerd in functie van de coronaire aantasting en het Qermid TUCO-register bevat een betrouwbare evaluatie ervan. De invloed van het milieu is immers niet dezelfde bij een patiënt met een aandoening die slechts één van de 3 kransslagaders treft dan bij patiënten met een aandoening op alle 3 aders of de gemeenschappelijke stam. De Qermid-gegevens geven een beschrijving van de ernst van de

ziekte (aantasting van de gemeenschappelijke stam, aantasting van 1, 2 of 3 aders), de proximale of distale positie van het behandeld letsel (segmentnummer), een bifurcatieletsel (aantal stents), een aantasting van de dominante ader (rechtse, linkse dominantie, co-dominantie). Bovendien kunnen in het Qermid-register de stent-tromboses worden geïdentificeerd die een zeer specifieke categorie van STEMI-infarcten zijn die eerder het gevolg zijn van een technische complicatie en die dus volgens ons geïdentificeerd moeten worden omdat ze de invloed van het milieu op de STEMI kunnen vertekenen. Bovendien kan het aantal ingrepen tijdens de ziekenhuisopname worden geïdentificeerd in het Qermid-register. Dit is belangrijk omdat enkel de primaire ingrepen van belang zijn voor de EMIR-studie en dus niet de eventuele nieuwe ingrepen na technische complicaties. Dit register bevat ten slotte ook belangrijke informatie die kan worden overgenomen in de beschrijving van de infarcten zoals het type gebruikte stent.

c. Gegevens meegedeeld door de Kruispuntbank van de Sociale Zekerheid (KSZ)

Het adres van de hoofdverblijfplaats van de patiënt op het ogenblik van het infarct, gegeocodeerd in GPS-gegevens: dit gegeven is van groot belang omdat de EMIR-studie tot doel heeft om de biotoop van de patiënt te analyseren, met andere woorden de omgeving van zijn woonst op het vlak van temperatuur, luchtkwaliteit, blootstelling aan lawaai en vegetatiepercentage. Er wordt opgemerkt dat het adres van de hoofdverblijfplaats wordt gevraagd op het ogenblik van de gebeurtenis. Dat adres kan dus verschillend zijn van het laatst gekende adres van de patiënt in geval van verhuis tijdens de onderzoeksperiode. Andere studies op de schaal van een ziekenhuis of een stad hebben reeds gefocust op de verblijfsadressen van de patiënt maar de EMIR-studie is de eerste die de luchtverontreiniging modelleert op het leefadres dankzij een interpolatieproces. In de vorige studies werd het dichtstbijzijnde meetstation van de luchtverontreiniging die echter vaak een tiental kilometers verder ligt immers als een betrouwbare meting van de luchtverontreiniging van de woning beschouwd.

Deze geolocalisatiegegevens van de adressen zullen niet meer aanwezig zijn in de gegevensbank die voor de onderzoekers toegankelijk is (niveau 2) en in de wetenschappelijke publicaties (niveau 3). Deze GPS-gegevens zullen door de VITO binnen de KSZ worden verwerkt en onder haar toezicht om ze te koppelen aan de overeenstemmende milieu- en niet-milieuparameters. Zodra deze koppeling is verricht, zullen de geolocalisatiegegevens door de KSZ worden verwijderd.

Het beroepsadres op het ogenblik van het infarct dat in GPS-gegevens is gegeocodeerd: Geen enkele milieustudie heeft tot nu toe rekening gehouden met de tijd doorgebracht buiten de woning. De EMIR-studie heeft tot bijkomende doel om het blootstellingsniveau te temperen tussen de tijd die thuis wordt doorgebracht en de tijd op de werkplek. Het betreft een mathematische modellering om op verschillende manieren de tijd thuis, op de werkplek en de woon-werkverplaatsing te wegen aangezien de reële blootstellingstijd thuis/op het werk niet toegankelijk is in deze retrospectieve studie. Zoals de geolocalisatiegegevens van de verblijfsadressen, zullen deze gegevens van niveau 1 door de KSZ worden verwijderd en zullen ze dus niet toegankelijk zijn op niveau 2 en 3.

Socio-economische gegevens: de milieu-impact kan pas worden onderzocht indien de cardiovasculaire risicofactoren op een betrouwbare manier worden geëvalueerd. Naast medische factoren zoals roken, diabetes, enz. zijn socio-economische factoren belangrijke determinanten van cardiovasculaire aandoeningen. Voor die reden vereist de EMIR-studie een evaluatie van verschillende factoren zoals de burgerlijke stand, de socioprofessionele categorie en de categorie van het jaarinkomen. Uit verschillende studies blijkt immers dat alleenstaande personen en/of zonder werk en/of met een laag inkomen vatbaarder zijn om cardiovasculaire aandoeningen te ontwikkelen. Deze bevolkingsgroepen leven ook vaker in achtergestelde, meer stedelijke wijken waar de luchtverontreiniging potentieel groter is. Enkel de studie van een groot aantal patiënten waarbij zowel gezondheidsgegevens als socio-economische en milieugegevens worden onderzocht, zal het mogelijk maken om de impact eigen aan het milieu te identificeren alsook de bevolkingsgroepen die het meest risico lopen om schadelijke gevolgen op te lopen.

Status levend of overleden, jaar van overlijden: het overlijden van de patiënt tijdens zijn ziekenhuisverblijf naar aanleiding van zijn infarct is een van de voornaamste markers waarop de EMIR-studie betrekking zal hebben. Behalve het voorkomen van een infarct en een ziekenhuisopname wenst de EMIR-studie na te gaan of het milieu de ernst van het ziektebeeld van het infarct verergert en dus tot de dood leidt. Sommige patiënten overlijden in de weken en maanden die volgen op hun ontslag uit het ziekenhuis. De EMIR-studie zal dus niet alleen betrekking hebben op de ziekenhuismortaliteit maar ook op de latere mortaliteit (meer dan 1 jaar na het infarct). Het is dus noodzakelijk om enkel toegang te krijgen tot het jaar van overlijden voor de overlijdens buiten het ziekenhuis.

Het Comité stelt vast dat de onderzoeker enkel het jaar van overlijden van de betrokken patiënt vraagt.

d. Gegevens afkomstig uit het BIWAC-register

De gegevens uit het BIWAC-register zijn gedepersonaliseerde persoonsgegevens die door de werkgroep acute cardiologie van de Belgische cardiologievereniging worden beheerd en die enkel betrekking hebben op patiënten met een STEMI-infarct. Het gaat om de volgende gegevens: administratieve gegevens, indicatie van de behandeling met PCI, cardiovasculaire risicofactoren, factoren van klinische en anatomische ernst die gelijkaardig zijn in de rechtvaardigingen ervan met die van de Qermid TUCO gegevensbank. Het gaat om zeer betrouwbare maar beperkte gegevens omdat ze enkel afkomstig zijn van de ziekenhuizen die vrijwillig aan het register deelnemen. We wensen het cardiovasculair risicoprofiel van de patiënten en de klinische en anatomische ernst ervan beter in kaart te brengen door de BIWAC-gegevens te koppelen aan die uit het Qermid-register en aan de MZG. Aangezien deze koppeling niet kan gebeuren op basis van het INSZ dat in de BIWAC-gegevensbank ontbreekt, wensen de onderzoekers deze gegevens te kruisen met de gegevens van de 2 andere gegevensbanken op basis van de opnamedatum, de leeftijd, het geslacht van de patiënt en het erkenningsnummer van het ziekenhuis van opname.

Naast de Qermid-gegevens die eerder werden gerechtvaardigd, evalueert het BIWAC-register het roken; de Killip-score is een evaluatie van de ernst van het ziektebeeld (ventrikeldysfunctie als gevolg van het infarct); het voorkomen van een hartstilstand tijdens

de behandeling die een impact heeft op de klinische ernst van het infarct; het tijdstip van begin van de symptomen en het tijdstip van de reperfusie door het plaatsen van een stent. Op basis hiervan kan de totale ischemie worden berekend die de ernst van het infarct en de mortaliteit ervan bepaalt.

e. Gegevens met betrekking tot de luchtweginfecties ILI

Het betreft in dit geval samengevoegde gegevens die de evolutie van de griep epidemieën en andere luchtweginfecties op regionaal niveau weergeven. De 3 beschikbare parameters zullen het mogelijk maken om de intensiteit van de epidemieën van seizoensgriep te evalueren. De seizoensgriep is immers een uitlokkende factor van het myocardinfarct en moet als een mogelijke confounder van de impact van fijnstofverontreiniging worden beschouwd. Deze verontreiniging is erger in de wintermaanden.

f. Milieugegevens die door de VITO worden meegedeeld

Het betreft gegevens die betrekking hebben op persoonsgegevens (verblijfsadres en beroepsadres die op voorhand in GPS-coördinaten zijn gegeocodeerd). De aan VITO gevraagde gegevens hebben betrekking op milieufactoren op de schaal van de bevolking: luchtverontreiniging, luchttemperatuur, blootstelling aan geluid en vegetatiegraad van de leefomgeving. VITO heeft toegang tot de GPS-coördinaten die door de KSZ zijn gecodeerd en die overeenstemmen met het verblijfsadres en het beroepsadres op het ogenblik van het infarct. Voor elke geocode deelt VITO de volgende gegevens mee: dagelijks gemiddeld van de vorige 45 dagen en van de 45 dagen volgend op het infarct voor de luchttemperatuur en de concentratie aan vervuilende stoffen en de jaargemiddelden van de overige parameters voor de 8 onderzochte jaren. De milieugegevens na het voorkomen van het infarct kunnen overbodig blijken maar zijn noodzakelijk voor de bidirectionele aanpak van de statistische analyse in case cross-over studies. Een andere statistische aanpak via machine learning zal de complexe dynamiek van de tijdelijke fluctuering van milieufactoren en van de fluctuering van de regionale en subregionale infarctpercentages bestuderen.

Voor dit soort analyse zullen de gegevens over het infarct worden geaggregeerd op een spatiale resolutie van 16 km² die overeenstemt met de RIO-cartografie die is gebaseerd op een rooster van cellen van 4*4 km dat heel België beslaat. VITO zal dus de code van de cel van de RIO-cartografie met inbegrip van de geocode van het adres van de hoofdverblijfplaats en van het beroepsadres meedelen. Het zal niet mogelijk zijn om een specifieke patiënt te identificeren op basis van deze code van de RIO-cel aangezien de cartografie van België uit ongeveer 1900 cellen van 16km² bestaat en dat op basis van de gemiddelde bevolkingsdichtheid ongeveer 6000 personen woonachtig zijn in elke cel. De statistici zullen dus de geaggregeerde gegevens kunnen analyseren op basis van deze code van de RIO-cel (gegevens van niveau 2). De gegevens die toegankelijk zijn voor het publiek (niveau 3) zullen daarentegen enkel als een illustratie van de RIO-cartografie worden meegedeeld met een kleurcode die overeenstemt met de dagelijkse incidentie van de infarcten per cel.

25. Het Comité stelt vast dat de gegevens toegankelijk zijn in functie van 3 niveaus:
- Niveau 1: basisgegevens die gecombineerd en geaggregeerd zullen worden door vertrouwensderden (KSZ/eHealth) ;

- Niveau 2: de dataset die voor de onderzoekers beschikbaar is (gepseudonimiseerd, geaggregeerd op een eerste niveau) ;
- Niveau 3: wetenschappelijke publicatie (volledig geanonimiseerd).

Toegang tot de gegevens

i. De **gegevens van niveau 2** zullen door de op de server van Healthdata beschikbare statistische tools worden opgeslagen en verwerkt. Geen enkel gegeven zal kunnen worden gedownload of geëxporteerd in welk formaat ook.

ii. De **gegevens van niveau 2** zullen toegankelijk zijn voor de 3 onderzoekers elk met hun eigen toegangscode. Deze onderzoekers zullen de “non disclose form” ondertekenen.

iii. Betreffende de **tijdelijke gegevens**: De gegevens van niveau 2 bevatten de datum van de gebeurtenis (infarct) die noodzakelijk is voor een epidemiologische analyse. Alle andere tijdelijke gegevens werden beperkt. Enkel de leeftijd van de patiënt en het jaar van het eventuele overlijden zullen toegankelijk zijn in niveau 2. Er zal in niveau 2 worden vermeld of het overlijden al dan niet tijdens het ziekenhuisverblijf heeft plaatsgevonden.

iv. Betreffende de **geolocalisatiegegevens**: De gegevens van niveau 2 bevatten geen geolocalisatiegegevens van het verblijfsadres of het beroepsadres meer.

Dit punt is essentieel om te garanderen dat een patiënt of een groep patiënten niet kan worden geïdentificeerd. Het centraal punt is dat de gegevens van niveau 1 waarbij de patiënten worden geolocaliseerd door een vertrouwensderde (VITO) op de KSZ zullen worden verwerkt. VITO zal de milieu- en de niet-milieugegevens aanvullen met de GPS-coördinaten. Voor de analyse in een case-crossover studie, zullen de milieugegevens vastgesteld op de respectievelijke geocodes die betrekking op 45 dagen voor en 45 dagen na de gebeurtenis worden samengesteld. Zodra het adres is ingevuld, zullen de geolocalisatiegegevens van de adressen door de KSZ worden verwijderd vooraleer ze in niveau 2 worden overgemaakt. De onderzoekers van niveau 2 hebben dus geen toegang tot de geolocalisatiegegevens van de adressen.

v. Wat de **keuze van de grootte van de statistische cellen** betreft voor de analyse van de gegevens in niveau 2, verduidelijken we voortaan onze strategie in het aanvraagformulier. Voor de “machine learning” analyse zullen de geolocalisatiegegevens op een resolutie van 16 km² worden samengevoegd. Deze resolutie is die van de RIO-kaart. Het betreft een kaart van België waarop een rooster van cellen van 4x4 km is toegepast.

De onderzoekers zullen een analyse van het dagelijks percentage van infarcten per cel verrichten in functie van de gelijktijdige veranderingen van de milieukeurmerken van de cel. Rekening houdende met de grootte van 16 km² van de statistische RIO-cellen en de gemiddelde dichtheid van de Belgische bevolking, zullen de onderzoekers van niveau 2 onmogelijk een patiënt of een specifieke groep patiënten kunnen identificeren. De kaart van België is immers samengesteld uit meer dan 1900 RIO-cellen en gemiddeld 6000 personen zijn woonachtig in elke cel.

vi. Betreffende **de gegevens van niveau 3** : Het gaat hoofdzakelijk om beschrijvende statistiek, inferentiële statistiek (odds ration) en random forest diagrammen opgemaakt door “machine learning”. In de wetenschappelijke communicaties en in de communicaties voor het grote publiek zullen ter illustratie de RIO-kaarten worden gepubliceerd met de dagelijkse incidenties van de infarcten in functie van variaties in de luchtkwaliteit en de temperatuur. Voor deze kaarten zal een kleurschaal per cel van 16 km² worden gebruikt.

26. Het Comité eist dat een “small cell risks”-analyse zou worden uitgevoerd op reële gegevens om het heridentificatierisico van de betrokken patiënten en onder meer de grootte van de in deze studie beoogde geografische cellen te evalueren.
27. Het Comité is van mening dat enkel de datum van het voorkomen van het infarct aan de onderzoeker mag worden meegedeeld om de risico’s op heridentificatie van de betrokken persoon te beperken. Wat de exacte data van het ziekenhuisverblijf betreft, stelt het Comité voor om ofwel een statistische methode toe te passen waarbij gebruik wordt gemaakt van een delta, ofwel de verblijfsduur mee te delen.
28. Hiertoe herinnert het Comité eraan dat alle noodzakelijke maatregelen moeten worden getroffen om het risico op heridentificatie van de betrokken patiënten te beperken. Bij de publicatie van de resultaten mogen er dus geen kaarten worden gepubliceerd met daarop de woonplaats van de betrokken persoon.
29. Het Comité is van oordeel dat de gegevens toereikend, ter zake dienend en strikt beperkt worden gelet op de beoogde studie.

3. OPSLAGBEPERKING

30. Volgens artikel 5, § 1, e) van de AVG mogen de persoonsgegevens niet langer worden bewaard in een vorm die het mogelijk maakt de betrokkenen te identificeren, dan noodzakelijk is voor de verwezenlijking van de doeleinden waarvoor zij worden verkregen. De 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 de verordening vereiste passende technische en organisatorische maatregelen worden getroffen om de rechten en vrijheden van de betrokkene te beschermen ("opslagbeperking").
31. De gepseudonimiseerde persoonsgegevens zullen gedurende 10 jaar worden bewaard vanaf de datum van ontvangst van de gegevens voor het jaar 2021 (die pas mogelijk zal zijn in 2023). De gegevens worden in gepseudonimiseerde vorm bewaard binnen het healthdata.be-platform. De gepseudonimiseerde gegevens zullen daarna worden vernietigd.
32. Het Comité stelt vast dat deze duur van 10 jaar noodzakelijk is om op bijkomende vragen van de wetenschappelijke wereld te antwoorden naar aanleiding van de publicatie van de resultaten van de studie in internationale wetenschappelijke tijdschriften.

33. Het Comité meent dat de voorgestelde bewaarduur redelijk is rekening houdend met de beoogde studie. Het Comité verleent bijgevolg een machtiging voor de bewaring van de gegevens tot 31 december 2033. De gegevens mogen daarna enkel nog in een strikt anonieme vorm worden bewaard.

4. TRANSPARANTIE

34. Wanneer de persoonsgegevens niet bij de betrokkene werden verkregen, verstrekt de verwerkingsverantwoordelijke hem/haar alle informatie bedoeld in artikel 14, § 1 en § 2, van de AVG.
35. De beoogde verwerking van de persoonsgegevens is een latere bewerking van de persoonsgegevens die werden ingezameld in het kader van een eerste doeleinde ander dan de wetenschappelijke EMIR-studie. Er is bijgevolg geen rechtstreeks contact tussen de analist van de geselecteerde gepseudonimiseerde gegevens en de patiënten, dit in tegenstelling tot een klinische studie op basis van een gegevensbank eigen aan een ziekenhuis.
36. De aanvrager is dus van mening dat het informeren van deze personen onevenredig veel inspanning zou vergen. Een duidelijke informatie en het verkrijgen van een individuele toestemming houdt een zeer grote werklast in rekening houdende met de 81297 verwachte patiënten en de vroegere gebeurtenissen (bestudeerde periode loopt van 1/01/2012 tot 31/12/2018). Bovendien zijn sommige patiënten overleden tijdens hun infarct of als gevolg van andere aandoeningen. De mortaliteit binnen het ziekenhuis van STEMI-infarcten bedraagt aldus ongeveer 6% in België. Er wordt geschat dat 5000 à 8000 patiënten zullen zijn overleden op het ogenblik van de gegevensexport.
37. Vanuit praktisch oogpunt kan enkel een schriftelijke communicatie op basis van de adressen van de patiënten worden overwogen, wat de beveiliging van de verblijfsadressen des te complexer maakt. De respons op deze soort schriftelijke informatie ligt bovendien vaak heel laag. De EMIR-studie beoogt een evaluatie van het absolute risico op een infarct, wat een volledige dekking van de bevolking inhoudt en niet een analyse van een subgroep van vrijwillige deelnemers.
38. Ter bevordering van de transparantie werd er over de doelstellingen en de methodologie van de EMIR-studie reeds breed gecommuniceerd bij de overheidsinstanties, de nationale wetenschappelijke centra voor cardiologie (Belgische Cardiologische Liga en de Belgian Society Of Cardiology) en in de medische pers. Wat de transparantie naar het grote publiek toe betreft, zullen de studie en de resultaten ervan worden beschreven op een website.
39. Het Comité meent dat alle nodige maatregelen moeten worden genomen om de rechten van de betrokkenen te beschermen, ook door de informatie voor het publiek beschikbaar te stellen op een website.
40. Het Comité is van oordeel dat er voldoende transparantie is omtrent de beoogde gegevensverwerking.

5. INFORMATIEVEILIGHEID

41. Volgens artikel 5, § 1, f) van de GDPR 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).
42. 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.
43. Het Comité stelt vast dat de KSZ, het eHealth-platform, Healthdata en het UZ Brussel een functionaris voor gegevensbescherming hebben aangesteld.
44. Het Comité stelt vast dat een gegevensbeschermingseffectbeoordeling werd uitgevoerd. De resultaten van deze analyse werden meegedeeld.
45. Het Comité stelt vast dat de geïmplementeerde gegevensuitwisseling een beroep doet op de infrastructuur van het healthdata.be-platform. De uitwisselingen geschieden in overeenstemming met de bepalingen van beraadslaging nr. 15/009 van 17 februari 2015, laatst gewijzigd op 3 maart 2020, betreffende de generieke methode voor de uitwisseling van al dan niet gecodeerde persoonsgegevens die de gezondheid betreffen in het kader van healthdata.be en healthstat.be.
46. Het Comité stelt vast dat de KSZ de omzetting van de privé- en beroepsadressen van de patiënten naar GPS-coördinaten aan VITO zal toevertrouwen. De KSZ beschikt hiervoor immers niet over de nodige tools.
47. Het Comité eist dat de koppeling van de GPS-coördinaten met de milieugegevens plaatsvindt in de lokalen van de KSZ.
48. Het Comité stelt vast dat alle onderzoekers die bij de studie zijn betrokken, contractueel ertoe gehouden zijn de vertrouwelijkheid van de verwerkte gegevens in acht te nemen.
49. Het Comité stelt vast dat het EMIR-onderzoeksteam voorziet dat een onderzoeker van de universiteit van Harvard als consultant in de machine learning-analyse zal worden betrokken. Het Comité herinnert er in dat verband aan dat er met toepassing van artikel 45

van de AVG geen mededeling van gepseudonimiseerde persoonsgegevens naar de VS kan plaatsvinden indien er geen adequaatheidsbesluit voorhanden is. De studieverantwoordelijke bevestigt dat die onderzoeker geen gegevens zal krijgen, ook niet een andere onderzoeker van een Belgische of buitenlandse universiteit. Er zal bovendien geen gegevenstransfer buiten de Healthdata-server mogelijk zijn en alle analyses zullen moeten worden verricht aan de hand van de softwarepakketten die op de server beschikbaar zijn.

- 50.** Hoewel dit strikt genomen niet wordt vereist in de AVG, verdient het volgens het Comité de voorkeur dat dergelijke gegevens worden verwerkt onder de verantwoordelijkheid van een arts⁵. Het Comité herinnert eraan dat de beroepsbeoefenaar in de gezondheidszorg en zijn aangestelden of gemachtigden bij de verwerking van persoonsgegevens tot geheimhouding verplicht zijn.
- 51.** Het Comité stelt vast dat de verwerking van persoonsgegevens zal plaatsvinden onder het toezicht en de verantwoordelijkheid van de voor de studie verantwoordelijke arts en van de verantwoordelijke arts voor healthdata.be.
- 52.** Het Comité herinnert eraan dat er krachtens artikel 35 van de AVG en artikel 23 van de wet van 30 juli 2018 *betreffende de bescherming van natuurlijke personen met betrekking tot de verwerking van persoonsgegevens in voorkomend geval een gegevensbeschermingseffectbeoordeling moet worden gerealiseerd*⁶⁷. Het Comité stelt vast dat de resultaten van deze analyse hem werden meegedeeld.
- 53.** 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:
- 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.
- 54.** Het Comité vestigt de aandacht op 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 strafsancities zijn voorzien in hoofde

⁵Het Sectoraal Comité heeft deze voorkeur opgesteld in beraadslaging nr. 07/034 van 4 september 2007 m.b.t. de mededeling van persoonsgegevens aan het Federaal Kenniscentrum voor de Gezondheidszorg met het oog op het onderzoek 2007-16-HSR "Onderzoek naar mogelijke financieringsmechanismen voor het geriatrisch dagziekenhuis".

⁶ <https://www.gegevensbeschermingsautoriteit.be/publications/aanbeveling-nr.-01-2018.pdf>

⁷ <https://www.gegevensbeschermingsautoriteit.be/publications/beslissing-nr.-01-2019-van-16-januari-2019.pdf>

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

dat:

de mededeling van 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.

Michel DENEYER

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