

Informatieveiligheidscomité  
Kamer sociale zekerheid en gezondheid

IVC/KSZG/21/054

**BERAADSLAGING NR. 21/032 VAN 7 SEPTEMBER 2021 BETREFFENDE DE MEDEDELING VAN GEPSEUDONIMISEERDE PERSOONSGEGEVENS DIE DE GEZONDHEID BETREFFEN DOOR DE STICHTING KANKERREGISTER, HET INTERMUTUALISTISCH AGENTSCHAP, DE FOD VOLKSGEZONDHEID EN HET DATAWAREHOUSE VAN DE KRUISPUNTBANK VAN SOCIALE ZEKERHEID AAN DE VRIJE UNIVERSITEIT BRUSSEL, ONDERZOEKSGROEP INTERUNIVERSITY CENTRE FOR HEALTH ECONOMICS RESEARCH (I-CHER) IN HET KADER VAN HUN ONDERZOEKSPROJECT**

Het Informatieveiligheidscomité;

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 15 januari 1990 *houdende oprichting en organisatie van een Kruispuntbank van de Sociale Zekerheid*, inzonderheid op artikel 37;

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 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 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 21 augustus 2008 *houdende oprichting en organisatie van het eHealth-platform en diverse bepalingen*;

Gelet op de aanvraag vanwege de Vrije Universiteit Brussel, onderzoeksgroep Interuniversity Centre for Health Economics Research (I-CHER);

Gelet op het auditoraatsrapport van het eHealth-platform van 12 augustus 2021;

Gelet op het verslag van de heer Bart Viaene.

Beslist op 7 september 2021, na beraadslaging, als volgt:

## **I. ONDERWERP VAN DE AANVRAAG**

1. De Vrije Universiteit Brussel, onderzoeksgroep Interuniversity Centre for Health Economics Research (I-CHER) dient een aanvraag in teneinde een machtiging van het informatieveiligheidscomité te verkrijgen voor de mededeling van persoonsgegevens die de gezondheid betreffen door de Stichting Kankerregister, het Intermutualistisch Agentschap, de FOD Volksgezondheid en het datawarehouse van de Kruispuntbank van Sociale Zekerheid in het kader van hun onderzoeksproject.
2. Verschillende organisaties in België beschikken elk afzonderlijk over een belangrijke bron van informatie betreffende de diagnoses, het zorggebruik en de kosten van verleende zorg, de oorzaak-specifieke sterfte, en de sociale situatie van vrouwen met borstkanker. Het linken van deze data biedt een meerwaarde vanuit wetenschappelijk oogpunt gezien het de mogelijkheid biedt om onderzoek te doen op het niveau van de Belgische bevolking.
3. Het onderzoeksproject van I-CHER tracht het traject van vrouwen met borstkanker in kaart te brengen over een periode van twintig jaar na de diagnose. Binnen dit project wordt er gefocust op de vraag welke interventies kunnen leiden tot een verlaging van de ziektelast van een borstkankerbehandeling bij vrouwen in België.
4. Alle Belgische vrouwen die tijdens het incidentiejaar 2010 een diagnose van borstkanker kregen, worden geïncludeerd in de studie. Meer specifiek wordt gezocht naar vrouwen met ICD-10-codes C50 (malignant neoplasms of breast) en D05 (carcinoma in situ of breast). De datum van de eerste microscopische bevestiging van de tumor of, bij afwezigheid hiervan, de datum van het klinisch onderzoek dat geleid heeft tot de diagnose van de tumor (variabele: “incidentiedatum”) wordt beschouwd als de ‘indexdatum’. De dataset wordt geraamd op 13.000 personen.
5. Enkele subdoelstellingen worden vooropgesteld:
  - 1) *Onderzoek naar de maatschappelijke kost van kankerbehandeling in België tot 1,5 jaar na de diagnose van borstkanker.*

Aan de hand van de diagnoses van borstkanker worden de vrouwen die tot de cohort behoren, geïdentificeerd. Er wordt getracht om het zorggebruik van deze vrouwen in kaart te brengen opdat inzicht verworven kan worden in het traject dat deze vrouwen afleggen tot 1,5 jaar na de diagnose.

De prestatiedata van de Belgische ziekenfondsen vormen een belangrijke bron van informatie om de directe medische kosten te berekenen. Zowel de kosten ten laste van de ziekteverzekering als het persoonlijk aandeel van de patiënt worden in detail bestudeerd. De directe medische kosten na de diagnose van borstkanker dienen gecorrigeerd te worden voor de reguliere zorg voor de patiënt. Dit gebeurt aan de hand van de berekening van de kosten in de periode voorafgaand aan de indexdatum. Hierdoor fungeert elke patiënt als haar eigen controlepersoon.

Mits rekening te houden met verhoogd zorggebruik in de pre-diagnostische fase als gevolg van pre-diagnostische consulten en onderzoeken, kan het verschil in de kosten voor en na indexdatum toegewezen worden aan het zorggebruik gerelateerd aan de borstkankerbehandeling. Informatie over de tewerkstellingssituatie voor en na indexdatum geeft een indicatie over de indirecte kosten (werkabsenteïsme). Door zowel de directe medische kosten te berekenen, alsook een inschatting te maken van de indirecte kosten, kan de totale maatschappelijke kost van de behandeling van borstkanker in België berekend worden.

*2) Onderzoek naar de incidentie en sterfte van cardiovasculaire aandoeningen en longtumoren in een cohort van Belgische vrouwen met borstkanker tot 20 jaar na de diagnose van borstkanker*

Het traject dat de patiënten in de cohort afleggen zal gevolgd worden tot 20 jaar na indexdatum. Data uit de ziekenhuizen (MZG en IMA) en vanuit Stichting Kankerregister zullen gebruikt worden om de incidentie van cardiovasculaire aandoeningen en (radiatiegeïnduceerde) longtumoren bij de cohort te bepalen. Hierbij kunnen opnieuw de kosten van de verleende zorg in kader van deze nieuwe diagnoses in kaart gebracht worden. Daarnaast beoogt het project ook de oorzakspecifieke sterfte voor deze vrouwen te onderzoeken tot 20 jaar na indexdatum. Op basis van deze data kan de maatschappelijke kost van alle trajecten die vrouwen met borstkanker afleggen tot 20 jaar na de diagnose berekend worden.

*3) Onderzoek naar de ontwikkeling van nieuwe strategieën die de ziektelast gerelateerd aan borstkanker kunnen verminderen.*

De inzichten uit dit onderzoek kunnen aanleiding geven tot het ontwikkelen van methodieken om de verwachte kosten en effecten van nieuwe, preventieve interventies te berekenen. Deze nieuwe strategieën kunnen leiden tot een vermindering van de ziektelast overheen het volledige ziekte-traject van de populatie.

6. Vanuit **de Stichting Kankerregister** worden de volgende gepseudonimiseerde gegevens opgevraagd: gepseudonimiseerd INSZ-nummer, geslacht, leeftijd bij diagnose (jaar), incidentiedatum omgezet in aantal dagen t.o.v. de indexdatum, ICD10 (numeriek), cTNM, pTNM, combined stage, histologisch subtype, ER, PR, HER2, aantal tumoren, multiple, lateraliteit, vitale status, status datum omgezet in aantal dagen t.o.v. de indexdatum, onmiddellijke doodsoorzaak, intermediaire doodsoorzaak (1 t.e.m. 2), oorspronkelijke doodsoorzaak, bijkomende doodsoorzaak (1 t.e.m. 3).
7. Het **Intermutualistisch Agentschap** zal volgende gegevens vanuit *Farmanet* meedelen: gepseudonimiseerde identificatie rechthebbende (SS00010), datum van aflevering

omgezet in aantal dagen t.o.v. de indexdatum (gebaseerd op SS00015), code categorie geneesmiddel (SS00020), hoeveelheid (SS00050), ZIV-tegemoetkoming 1 (SS00060), code gerechtigde 1 {CG1} (SS00095), code gerechtigde 2 {CG2} (SS00100)), nummer product (SS00135), persoonlijk aandeel (SS00160), vermindering verzekeringstegemoetkoming/ bijdrage van de apothekers (SS00165), ZIV tegemoetkoming 2 (SS00195).

Vanuit de databank *Gezondheidszorgen* worden de volgende gegevens meegedeeld: gepseudonimiseerde identificatie rechthebbende (SS00010), begindatum verstrekking omgezet in aantal dagen t.o.v. de indexdatum (gebaseerd op SS00015), nomenclatuurcode (SS00020), aantal gevallen (SS00050), aantal dagen (SS00055), bedrag terugbetaling (SS00060), zorgverstrekker-identificatiecode gepseudonimiseerd (SS00065A), zorgverstrekker-bekwaming (SS00065B), identificatie instelling gepseudonimiseerd (SS00075), dienstcode (SS00080), plaats van verstrekking (SS00085), code gerechtigde 1 {CG1} (SS00095), code gerechtigde 2 {CG2} (SS00100), nummer derde gepseudonimiseerd (SS00105), datum opname omgezet in aantal dagen t.o.v. de indexdatum (gebaseerd op SS00110), datum ontslag omgezet in aantal dagen t.o.v. de indexdatum (gebaseerd op SS00115), nummer product (SS00135), gefactureerde nomenclatuurcodes (SS00150), persoonlijke tussenkomst (SS00160), bedrag supplement of bedrag voor niet-vergoedbare producten, verstrekkingen of diensten (SS00165), code implant (SS00170), derdebetalersregeling (SS00175).

Vanuit de databank *Populatie* worden de volgende gegevens meegedeeld: gepseudonimiseerde identificatie rechthebbende (PP0010), aard VT/OMNIO (PP1010), forfait B geneeskundige zorgen (PP2001), forfait C geneeskundige zorgen (PP2002), kinesitherapie E of fysiotherapie (PP2003), verhoogde kinderbijlagen (PP2004), toelage voor integratie van gehandicapten (PP2005), toelage voor hulp aan ouderen {cat III, IV, V} (PP2006), uitkering hulp aan derden (PP2007), verhoogde invaliditeitsuitkering wegens hulp aan derden (PP2008), forfaitaire tegemoetkoming voor hulp aan derden (PP2009), criterium hospitalisatie {120 dagen} (PP2010), criterium hospitalisaties {6 hospitalisaties} (PP2011), recht op toelage voor gehandicapten (PP3011), code gerechtigde 1 {CG1} en 2 {CG2} (PP0030 en PP0035), recht MAF-gezin (PP3001), terugbetaling en plafond-gezin (PP3004), terugbetaling en plafond-individu of gedeelte van gezin (PP3005), aantal dagen werkloosheid (PP4001) aantal dagen invaliditeit (PP4002), CHRONICAL\_YN, MAJOR\_COVERAGE\_YN, ATC\_ANAT/THER/PHAR/CHEM/PROD\_L.

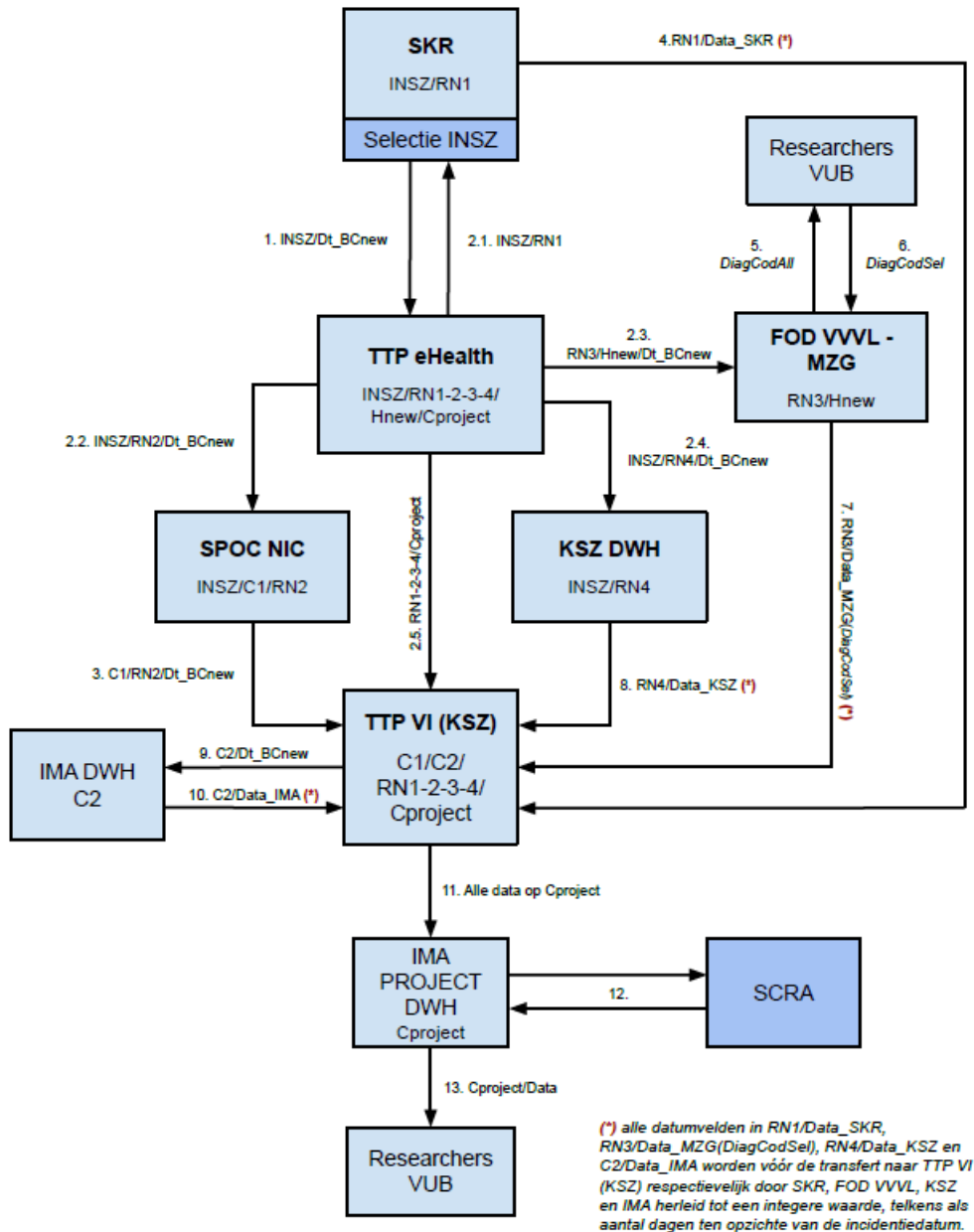
8. Vanuit de **FOD Volksgezondheid** worden de volgende gegevens opgevraagd: anoniem patiëntnummer (PATNUM), erkenningsnummer ziekenhuis gepseudonimiseerd (CODE\_AGR), verblijfsnummer (STAYNUM), datum van opname en ontslag omgezet in aantal dagen t.o.v. de indexdatum (gebaseerd op A2\_YEAR/MONTH/DAY\_HOSP\_IN/OUT), categorie van ziekenhuisverblijf (A2\_HOSPTYPE\_CAT), categorie ziekenhuisverblijf gebaseerd op facturatie (A2\_HOSPTYPE\_FAC), aantal te factureren ligdagen (A2\_TOTAL\_NUMBER\_DAY\_FAC), code heropname (A2\_CODE\_READMISSION), code verzekeringsstatus tijdens verblijf (A2\_CODE\_STAT\_INSURANCE), plaats van opname (A2\_CODE\_PLACE\_BEFORE\_ADM), type opname (A2\_CODE\_ADM), bestemming (A2\_CODE\_DESTINATE), type ontslag (A2\_CODE\_DISCHARGE), volgnummer bedindex (ORDER\_BEDINDEX), code van de bedindex gebaseerd op

facturatie (A4\_CODE\_BEDINDEX\_FAC), opnamedatum bedindex omgezet in aantal dagen t.o.v. de indexdatum (gebaseerd op A4\_YEAR/MONTH/DAY\_BEDINDEX\_IN), Aantal te factureren ligdagen bedindex in het huidige registratiejaar (A4\_NUMBER\_DAY\_FAC), aantal te factureren ligdagen in het vorige registratiejaar (A4\_NUMBER\_DAY\_FAC\_PREV), volgnummer specialisme (ORDER\_SPEC), type diagnose {hoofd- of nevendiagnose} (TYPE\_DIAGNOSE), codeersysteem voor diagnoses (CODE\_SYS\_DIAGNOSE), ICD-code diagnose (CODE\_DIAGNOSE), codeersysteem voor uitgevoerde ingrepen (M2\_CODE\_SYS\_PROCEDURE), ICD-code van de procedure (M2\_CODE\_PROCEDURE), datum van uitvoering van de procedure omgezet in aantal dagen t.o.v. de indexdatum (gebaseerd op M2\_YEAR/MONTH/DAY\_PROCEDURE), aantal keer dat dezelfde procedure op dezelfde dag voorkomt (M2\_NUMBER\_PROC\_1\_DAY).

9. De **datawarehouse van de Kruispuntbank van de Sociale Zekerheid** deelt de volgende gegevens mee: gepseudonimiseerd INSZ-nummer, nomenclatuur van socio-economische positie en bijkomende socio-economische positie omgezet naar een geclusterde variabele {werkend en gedeeltelijke loopbaanonderbreking/deeltijds tijdscrediet, werkend en halftijds brugpensioen, werkend en voltijds brugpensioen, werkend en volledige loopbaanonderbreking/ volledig tijdscrediet, werkend en gedeeltelijk in terbeschikkingstelling voorafgaand aan pensioen, werkend en volledig in terbeschikkingstelling voorafgaand aan pensioen, werkend en halftijds in stelsel van werkloosheid met bedrijfstoelage, werkend en volledig in stelsel van werkloosheid met bedrijfstoelage} (*bron: KSZ*), gedetailleerd onderwijsniveau (*bron: STATBEL*) opleidingsniveau ISCED (*bron: KSZ*); studiecategorie niet-universitair onderwijs (cat\_shu), studiecategorie (cat\_études) (*bron: CREF*), code voor diploma (CO\_DIPLOME) (*bron: Saturn*), arbeidsregime (T\_Prest), tewerkstellingspercentage (Eqtpsp), tewerkstellingspercentage (Tauxaa) (*bron: RSZ*); begin- en einddatum arbeidsongeschiktheid omgezet in aantal dagen t.o.v. de indexdatum (Begindatum, Einddatum), aantal dagen arbeidsongeschiktheid (aantal\_dagen\_ao), samenstelling gezin (samenst\_gezin), (*bron: CIN*); codeersysteem voor ziekte verantwoordelijk voor invaliditeit (Type\_ziekte), oorzaak invaliditeit (Gr\_ziekte), ICD-code voor oorzaak invaliditeit (Diag\_ziekte), begindatum voor primaire arbeidsongeschiktheid omgezet in aantal dagen t.o.v. de indexdatum (B\_ziekte), begin- en einddatum invaliditeit omgezet in aantal dagen t.o.v. de indexdatum (B\_inval, E\_inval), correctievariabele voor samenvoeging kwartalen (R\_exclus) (*bron: RIZIV*).

10. De gegevensstromen zullen volgens onderstaand schema verlopen:

**Data Traject Koppeling SKR, TCT-MZG, IMA en KSZ data:  
Studie Maatschappelijke Kost Borstkanker (P\_CVAL01)**



De Stichting kankerregister (SKR) stelt in het kader van deze koppeling de lijst samen van unieke INSZ-nummer en diagnosedatums (Dt\_BCnew) van de studiepopulatie (alle vrouwen met een diagnose van invasief en/of in situ borstkanker in het jaar 2010), en zendt de lijst INSZ/Dt\_BCnew (\*) door naar de TTP eHealth (stap 1). In geval van meer dan één incidentiedatum voor de selectie van borstkanker binnen indexjaar 2010, wordt voor deze patiënt de indexdatum vastgelegd op de eerste incidentiedatum. De aangeleverde lijst bevat slechts één record per patiënt.

Na ontvangst van de lijst zal TTP eHealth:

- Een RN1 toekennen aan iedere unieke INSZ uit de SKR lijst en de conversielijst INSZ/RN1 naar de SKR doorzenden (2.1.)
- Een RN2 toekennen aan iedere unieke INSZ uit de SKR lijst en de conversielijst INSZ/RN2 inclusief diagnosedatums (INSZ/RN2/Dt\_BCnew) naar de veiligheidsconsulent van het Nationaal Intermutualistisch College (SPOC NIC) doorzenden (2.2).
- Een RN3 toekennen aan iedere unieke INSZ uit de SKR lijst. De TTP eHealth pseudonimiseert eveneens elke unieke INSZ naar Hnew, de identifier die eHealth standaard gebruikt in de MZG flow sinds 2012. De conversielijst RN3/Hnew wordt dan inclusief diagnosedatums (RN3/Hnew/Dt\_BCnew) naar de verantwoordelijke FOD VVVL-MZG verzonden (2.3.).
- Een RN4 toekennen aan iedere unieke INSZ uit de SKR lijst en de conversielijst INSZ/RN4 inclusief diagnosedatums (INSZ/RN4/Dt\_BCnew) naar de DWH (datawarehouse) verantwoordelijke van de Kruispuntbank van Sociale Zekerheid (KSZ) doorzenden (stap 2.4).
- Een Cproject toekennen aan iedere unieke INSZ uit de SKR lijst en de conversielijst RN1/RN2/RN3/RN4/Cproject naar de TTP VI (KSZ) doorzenden (2.5).

De NIC SPOC vervangt iedere unieke INSZ uit de lijst INSZ/RN2/Dt\_BCnew door de C1 en zendt de nieuwe conversielijst C1/RN2/Dt\_BCnew door naar de TTP VI (KSZ) (stap 3).

De SKR selecteert de data (Data\_SKR) van de gevraagde selectieperiode (\*) op INSZ. De SKR herleidt alle datumvelden tot een integere waarde ten opzichte van de incidentiedatum en vervangt tevens elke INSZ door de corresponderende RN1. De SKR zendt hierna de RN1/Data\_SKR door naar de TTP VI (KSZ) (stap 4).

De verantwoordelijke FOD VVVL – MZG stelt op basis van de Hnew een lijst samen met alle codes<sup>1</sup> geassocieerd aan de hoofd- en nevendiaagnoses van de opnames van deze patiënten binnen de gevraagde selectieperiode (\*), en zendt de lijst (DiagCodAll) door naar de VUB-onderzoekers (stap 5).

De VUB-onderzoekers selecteren enkel deze codes die mogelijk verband houden met de doelstellingen van de studie (codes cardiovasculaire aandoeningen en radiatiegeïnduceerde longtumoren) en zenden deze lijst (DiagCodSel) door naar de verantwoordelijke FOD VVVL – MZG (stap 6).

De verantwoordelijke FOD VVVL – MZG selecteert op basis van de Hnew enkel deze opnamegegevens binnen de gevraagde selectieperiode (\*) die geassocieerd zijn aan een code uit de DiagCodeSel lijst van de VUB-onderzoekers. De verantwoordelijke FOD VVVL – MZG herleidt alle datumvelden tot een integere waarde ten opzichte van de

---

<sup>1</sup> De MZG- variabelen die in de DiagCodAll lijst worden opgenomen zijn:

- *CODE\_DIAGNOSE*: ICD code voor de hoofd- en nevendiaagnoses tijdens de hospitalisatie.
- *TYPE\_DIAGNOSE*: om hoofd- en nevendiaagnoses tijdens de hospitalisatie te differentiëren.
- *CODE\_SYS\_DIAGNOSE*: soort coderingssysteem (ICD-9-CM of ICD-10-BE) voor de diagnose.
- *M2\_CODE\_PROCEDURE*: code van de uitgevoerde ingrepen tijdens de hospitalisatie.

incidentiedatum en brengt tevens de data op RN3. De verantwoordelijke FOD VVVL - MZG zendt de data RN3/Data\_MZG(*DiagCodSel*) door naar de TTP VI (KSZ) (stap 7).

De verantwoordelijke KSZ DWH selecteert op basis van de INSZ de KSZ-gegevens binnen de gevraagde selectieperiode (\*), herleidt alle datumvelden tot een integrale waarde ten opzichte van de incidentiedatum, en brengt de data op RN4. De verantwoordelijke KSZ DWH zendt hierna de data RN4/KSZ\_Data door naar de TTP VI (KSZ) (stap 8).

De TTP VI (KSZ) kent de C2 toe aan iedere unieke C1 uit de lijst C1/RN2/Dt\_BCnew (zie stap 3) en zet C2/Dt\_BCnew op de IMA DWH (stap 9).

Het IMA selecteert op basis van de C2 en de diagnosedatums Dt\_BCnew de data binnen de gevraagde selectieperiode (\*), herleidt alle datumvelden tot een integrale waarde ten opzichte van de incidentiedatum, en zendt de data C2/Data\_IMA door naar de TTP VI (KSZ) (stap 10).

De TTP VI (KSZ) brengt alle data op Cproject en zet de data op IMA PROJECT DWH (stap 11).

Een Small Cells Risk Analysis wordt uitgevoerd door KCE (Stap 12).

De datasets Cproject/Data\_SKR, Cproject/Data\_IMA, Cproject/Data\_KSZ en Cproject/Data\_MZG (*DiagCodSol*) worden ter beschikking gesteld aan de onderzoekers op Cproject (stap 13).

*(\*) het traject dient 4 keer doorlopen te worden voor de selectie van gegevens uit de volgende periodes:*

	<i>Selectieperiode 1</i>	<i>Selectieperiode 2</i>	<i>Selectieperiode 3</i>	<i>Selectieperiode 4</i>
<i>SKR</i>	<i>01/07/2008-31/06/2012</i>	<i>01/07/2012-31/12/2020</i>	<i>01/01/2021- 31/12/2025</i>	<i>01/01/2026-31/12/2030</i>
<i>KSZ</i>	<i>01/07/2008-31/06/2012</i>	<i>01/07/2012-31/12/2020</i>	<i>01/01/2021- 31/12/2025</i>	<i>01/01/2026-31/12/2030</i>
<i>MZG</i>	-	<b><i>01/01/2012-31/12/2020</i></b>	<i>01/01/2021- 31/12/2025</i>	<i>01/01/2026-31/12/2030</i>
<i>IMA</i>	<i>01/07/2008-31/06/2012</i>	<i>01/07/2012-31/12/2020</i>	<i>01/01/2021- 31/12/2025</i>	<i>01/01/2026-31/12/2030</i>

*De lijst INSZ/Dt\_BCnew wordt voor elke selectieperiode opnieuw verzonden aan de TTP eHealth (stap 1). De volgende stappen (2 t.e.m. 13) worden voorzien zoals omschreven.*

*In stap 4, stap 7, stap 8 en stap 10 herleiden de dataleveranciers de datumvelden tot een integrale waarde ten opzicht van de incidentiedatum op basis van de aangeleverde incidentiedatum. De verschillende dataleveranciers kunnen de lijst Dt\_BCnew verwijderen, aangezien deze bij elke data-extractie opnieuw aangeleverd wordt door SKR en TTP eHealth.*

11. Het comité stelt vast dat het ethisch comité UZ Brussel een positief advies gegeven heeft omtrent dit onderzoek.

## II. BEVOEGDHEID

12. Ingevolge artikel 42, § 2, 3° van de wet van 13 december 2006 houdende diverse bepalingen betreffende gezondheid is de kamer sociale zekerheid en gezondheid van het



informatieveiligheidscomité bevoegd voor het verlenen van een principiële machtiging met betrekking tot elke mededeling van persoonsgegevens die de gezondheid betreffen.

13. Overeenkomstig artikel 42, § 2, 2°, a), van de wet van 13 december 2006 *houdende diverse bepalingen betreffende gezondheid* is de kamer sociale zekerheid en gezondheid van het informatieveiligheidscomité eveneens bevoegd voor het verlenen van een principiële machtiging voor de koppeling van persoonsgegevens uit het Kankerregister met andere externe gegevens.
14. Krachtens 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 zekerheid (artikel 5, § 1). Een dergelijke mededeling van persoonsgegevens vergt een principiële beraadslaging van de kamer sociale zekerheid en gezondheid van het informatieveiligheidscomité (artikel 15, § 1). Vermits de onderzoekers in het voorliggend geval gebruik willen maken van sociaaleconomische persoonsgegevens uit het datawarehouse arbeidsmarkt en sociale bescherming is de kamer sociale zekerheid en gezondheid bevoegd om zich uit te spreken.
15. Overeenkomstig artikel 5, 8°, van de wet van 21 augustus 2008 houdende oprichting en organisatie van het eHealth-platform is voor de tussenkomst van het eHealth-platform als intermediaire organisatie voor de koppeling en de pseudonimisering van persoonsgegevens de beraadslaging van het informatieveiligheidscomité vereist.
16. Het Informatieveiligheidscomité oordeelt bijgevolg dat het bevoegd is.

### **III. BEHANDELING**

#### **A. TOELAATBAARHEID**

17. De verwerking van persoonsgegevens is enkel toegelaten voor welbepaalde, uitdrukkelijk omschreven en gerechtvaardigde doeleinden en is de verwerking van persoonsgegevens die de gezondheid betreffen in principe verboden.<sup>2</sup>
18. 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, 1ste lid, 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.
19. Het comité oordeelt dat de verwerking van persoonsgegevens toelaatbaar is.

---

<sup>2</sup> Art. 9, punt 1 GDPR

## B. FINALITEIT

20. Overeenkomstig art. 5, b) van de GDPR is de verwerking van persoonsgegevens enkel toegelaten voor welbepaalde, uitdrukkelijk omschreven en gerechtvaardigde doeleinden.
21. Het doel van deze studie is het traject van vrouwen met borstkanker in kaart te brengen over een periode van twintig jaar na de diagnose. Tijdens dit project ligt de focus op twee aspecten: (1) het zorggebruik van deze vrouwen in België en (2) onderzoek naar de langetermijneffecten (met name cardiotoxiciteit en radiatie-geïnduceerde longtumoren) van de behandeling in deze cohort.
22. 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.

## C. PROPORTIONALITEIT

23. Overeenkomstig art. 5, b) en c) van de GDPR dienen de persoonsgegevens toereikend, ter zake dienend en niet overmatig te zijn, uitgaande van de doeleinden waarvoor zij worden verkregen of waarvoor zij verder worden verwerkt.
24. De gegevens afkomstig van **de Stichting Kankerregister** betreffen:
  - *Diagnoses van borstkanker tijdens incidentiejaar 2010*: Deze zijn noodzakelijk voor de identificatie van de cohort.
  - *Diagnoses van longkanker bij vrouwen uit de cohort*: Het bestuderen van het ontstaan van (radiatie-geïnduceerde) longtumoren in de cohort van vrouwen met borstkanker vormt een belangrijk onderdeel van de studie (zie subdoelstelling 2);
  - *Oorzaak-specifieke overlijdens aan de hand van overlijdenscertificaten bij vrouwen uit de cohort*: Het bestuderen van de sterfte in de cohort van vrouwen met borstkanker is een belangrijk onderdeel van de studie (zie subdoelstellingen 2 en 3);
25. Het analyseren van **IMA-data** laat toe om de directe kost van de primaire borstkankerbehandeling in kaart te brengen. Dit vormt een belangrijk onderdeel in de berekening van de maatschappelijke kost van de borstkankerbehandeling (zie subdoelstellingen 1 en 3). Vervolgens maakt de IMA-data het mogelijk om verwachte langetermijneffecten in de cohort te identificeren en de directe kosten ervan te berekenen. Dit vormt een belangrijk onderdeel in de berekening van de maatschappelijke kost ten gevolge van cardiotoxiciteit en radiatie-geïnduceerde longtumoren bij borstkanker (zie subdoelstellingen 2 en 3).
26. De analyse van **MZG-data** (*Bron: FOD VVVL*) maakt het mogelijk om verwachte langetermijneffecten in de cohort te identificeren en de directe kosten ervan te berekenen.

Dit vormt een belangrijk onderdeel in de berekening van de maatschappelijke kost ten gevolge van cardiotoxiciteit en radiatie-geïnduceerde longtumoren bij borstkanker (zie subdoelstellingen 2 en 3).

27. Het analyseren van de gegevens betreffende de socio-economische situatie en tewerkstelling afkomstig uit **datawarehouse van de KSZ** maken het mogelijk om de indirecte kosten van een borstkankerbehandeling in te schatten (zie subdoelstellingen 1, 2 en 3).
28. Het INSZ-nummer/uniiek identificatienummer zal enkel gebruikt worden voor selectie- en koppelingsdoeleinden. Deze variabele wordt omgezet naar een gepseudonimiseerd identificatienummer (variabele: Cproject) dat ter beschikking gesteld wordt voor de onderzoekers.
29. In dit onderzoek worden heel wat tijdreeksanalyses alsook Cox-regressie analyses uitgevoerd die de onderzoekers toelaten predictiemodellen op te stellen naar ‘time-to-event’. Dergelijke statistische analyses refereren naar het analyseren van de duurtijd tot een bepaalde gebeurtenis plaatsvindt (bijv. opname, cardiovasculair event, diagnose longtumor). Vereiste hierbij is dat de data van het beginpunt en gebeurtenis gekend zijn zodoende niet enkel predicties te berekenen of een event al dan niet zal plaatsvinden, maar ook naar predicties wanneer een event zou plaatsvinden. Om te vermijden dat de onderzoeker toegang krijgt tot exacte datums, zullen de dataleveranciers deze datums herleiden tot een integere waarde ten opzichte van de incidentiedatum op basis van de aangeleverde incidentiedatum.

Concreet betreft het volgende variabelen:

	<i>Variabele</i>	<i>Opmerking</i>
<b>SKR</b>	<i>Incidentiedatum (borstkanker)</i>	<i>Gedefinieerd als ‘aantal dagen t.o.v. incidentiedatum’</i>
	<i>Incidentiedatum (longtumor)</i>	<i>Gedefinieerd als ‘aantal dagen t.o.v. incidentiedatum’</i>
	<i>Status datum</i>	<i>Gedefinieerd als ‘aantal dagen t.o.v. incidentiedatum’</i>
<b>KSZ</b>	<i>DWH_CIN : Begindatum</i>	<i>Gedefinieerd als ‘aantal dagen t.o.v. incidentiedatum’</i>
	<i>DWH_CIN : Einddatum</i>	<i>Gedefinieerd als ‘aantal dagen t.o.v. incidentiedatum’</i>
	<i>DWH_IMANI_CMI : B_ziekte</i>	<i>Gedefinieerd als ‘aantal dagen t.o.v. incidentiedatum’</i>
	<i>DWH_INAMI_CMI : B_inval</i>	<i>Gedefinieerd als ‘aantal dagen t.o.v. incidentiedatum’</i>
	<i>DWH_INAMI_CMI : E_inval</i>	<i>Gedefinieerd als ‘aantal dagen t.o.v. incidentiedatum’</i>
<b>MZG</b>	<i>A2_YEAR_HOSP_IN</i>	<i>Gedefinieerd als ‘aantal dagen t.o.v. incidentiedatum’</i>
	<i>A2_MONTH_HOSP_IN</i>	
	<i>A2_DAY_HOSP_IN</i>	
	<i>A2_YEAR_HOSP_OUT</i>	<i>Gedefinieerd als ‘aantal dagen t.o.v. incidentiedatum’</i>
	<i>A2_MONTH_HOSP_OUT</i>	
	<i>A2_DAY_HOSP_OUT</i>	
	<i>A4_YEAR_BEDINDEX_IN</i>	<i>Gedefinieerd als ‘aantal dagen t.o.v. incidentiedatum’</i>
	<i>A4_MONTH_BEDINDEX_IN</i>	
	<i>A4_DAY_BEDINDEX_IN</i>	
	<i>M2_YEAR_PROCEDURE</i>	<i>Gedefinieerd als ‘aantal dagen t.o.v. incidentiedatum’</i>
	<i>M2_MONTH_PROCEDURE</i>	
<i>M2_DAY_PROCEDURE</i>		
<b>IMA</b>	<i>SS00015 – Begindatum verstrekking / Datum van aflevering</i>	<i>Gedefinieerd als ‘aantal dagen t.o.v. incidentiedatum’</i>
	<i>SS00110 – Datum opname</i>	<i>(1) Gedefinieerd als ‘aantal dagen t.o.v. incidentiedatum’</i>

		(2) Beperkt tot data-extractie 1 (ter vervanging van variabelen vanuit FOD VVVL-MZG)
	SS00115 – Datum ontslag	(1) Gedefinieerd als 'aantal dagen t.o.v. incidentiedatum' (2) Beperkt tot data-extractie 1 (ter vervanging van variabelen vanuit FOD VVVL-MZG)

- 30.** Een overzicht van al de gevraagde gegevens samen met een omstandige motivatie per variabele wordt opgenomen in bijlage.
- 31.** Het proportionaliteitsprincipe veronderstelt dat de verwerking in principe verricht wordt aan de hand van anonieme gegevens. Indien het doeleinde echter niet verwezenlijkt kan worden aan de hand van anonieme gegevens, kunnen gepseudonimiseerde persoonsgegevens worden verwerkt. Gelet op de noodzaak om zeer gedetailleerde analyses uit te voeren op basis van deze gegevens, hebben de onderzoekers behoefte aan toegang tot gepseudonimiseerde gegevens om deze analyses te kunnen uitvoeren aangezien ze niet kunnen worden verricht aan de hand van anonieme gegevens. Dit doeleinde rechtvaardigt aldus de verwerking van gepseudonimiseerde persoonsgegevens
- 32.** Persoonsgegevens moeten 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.
- 33.** De onderzochte periode beslaat een termijn van 20 jaar na het vooropgestelde incidentiejaar 2010 voor diagnose van borstkanker (tot 31/12/2030). Er zijn 4 dataleveringen voorzien, gespreid in de tijd. Het gaat over de gegevens voor de volgende periodes: 1,5 pre-index tot 1,5j post index, 1,5j post index tot 10j post index, 10j post index tot 15j post index en 15j post index tot 20 jaar post index.
- 34.** Om analyses mogelijk te maken voor de volledige onderzochte periode, wordt bij elke nieuwe datalevering tevens de data uit de voorgaande dataleveringen beschikbaar gesteld. In het kader van de wetenschappelijke opvolging van het project<sup>3</sup>, is de duur van beschikbaarheid van de data, met name de nieuwe datalevering plus alle data uit alle voorafgaande dataleveringen telkens gesteld op 4 jaar, te rekenen vanaf het moment dat alle dataleveranciers hun data beschikbaar stelden aan de onderzoekers. Bijgevolg blijft de volledige dataset bewaard tot 4 jaar na de laatste datalevering.
- 35.** Opdat het koppelen van de data tussen de verschillende leveranciers mogelijk blijft, is het van belang dat het verband tussen de gepseudonimiseerde dataset en de identiteit van de personen niet verloren gaat. De gefaseerde data-extractie vermijdt het oplopen van vertraging tijdens de voorziene projecttermijn. Het comité oordeelt bijgevolg dat het eHealth-platform het verband tussen de gepseudonimiseerde gegevens en de identiteit van de personen voor dit doeleinde mag bijhouden.
- 36.** Rekening houdend met de doeleinden van de studie is het comité van oordeel dat de beoogde mededeling toereikend, ter zake dienend en niet overmatig is.

<sup>3</sup> Bijvoorbeeld voor het nazicht van statistische gegevens naar aanleiding van peer-reviewed wetenschappelijke publicaties.

## **D. TRANSPARANTIE**

37. De verantwoordelijke van de verwerking van persoonsgegevens verzameld voor bepaalde, uitdrukkelijk omschreven en gerechtvaardigde doeleinden of de intermediaire organisatie, voorafgaand aan de pseudonimisering van de gegevens, moet in principe bepaalde informatie meedelen aan de betrokken persoon.
38. De verantwoordelijke voor de verwerking is vrijgesteld van de verplichting tot informatiemededing aan de personen wiens gegevens worden verwerkt, wanneer het verstrekken van die informatie onmogelijk blijkt of onevenredig veel inspanning zou vergen, in het bijzonder bij verwerking met het oog op archivering in het algemeen belang, wetenschappelijk of historisch onderzoek of statistische doeleinden, behoudens de in artikel 89, lid 1, bedoelde voorwaarden en waarborgen.<sup>4</sup>
39. de gegevens worden niet bekomen bij de betrokkenen zelf en de gegevens worden gepseudonimiseerd waardoor de onderzoekers geen toegang hebben tot de persoonsgegevens van de betrokkenen. Vervolgens is de gevraagde dataset zodanig groot (naar schatting 13.000 patiënten) dat de kennisgeving aan alle betrokkenen zowel praktisch<sup>5</sup> als financieel niet realiseerbaar is.
40. Het comité oordeelt bijgevolg dat de aanvraag voldoet aan de transparantie-eisen.

## **E. VEILIGHEIDSMATREGELEN**

41. De aanvrager moet, overeenkomstig art. 5, f) van de GDPR, alle gepaste technische en organisatorische maatregelen treffen die nodig zijn voor de bescherming van de persoonsgegevens. Deze maatregelen moeten een passend beveiligingsniveau verzekeren, rekening houdend, enerzijds, met de stand van de techniek terzake en de kosten voor het toepassen van de maatregelen en, anderzijds, met de aard van de te beveiligen gegevens en de potentiële risico's.
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 informatieveiligheidsconsulent; 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.

---

<sup>4</sup> Art. 14 lid 5 b) GDPR

<sup>5</sup> Het is onrealistisch om een geactualiseerde lijst met contactgegevens van alle betrokkenen te bekomen, gezien de omvang van de dataset.

43. Overeenkomstig art. 9, punt 3 van de GDPR mogen persoonsgegevens betreffende de gezondheid enkel worden verwerkt onder het toezicht en de verantwoordelijkheid van een beroepsbeoefenaar in de gezondheidszorg. Dit is in casu het geval.
44. Het comité herinnert eraan dat de beroepsbeoefenaar in de gezondheidszorg en zijn aangestelden of gemachtigden bij de verwerking van persoonsgegevens tot geheimhouding verplicht zijn.
45. Het is verboden om handelingen te stellen die ertoe strekken de meegedeelde gepseudonimiseerde persoonsgegevens om te zetten in niet-gepseudonimiseerde persoonsgegevens.
46. De toegang tot de beveiligde server waarop de data worden bewaard, wordt gelogd door het IMA. Een lijst met gemachtigde onderzoekers wordt ter beschikking gesteld aan het IMA door de diensthoofd van I-CHER.
47. Het comité neemt akte van het feit dat er een gegevensbeschermingseffectbeoordeling werd uitgevoerd.
48. Ten slotte stelt het comité vast dat er voorzien wordt in een "small cell risico-analyse" (SCRA) en dat deze zal worden uitgevoerd door het KCE.

Om deze redenen, besluit

#### **de kamer sociale zekerheid en gezondheid van het informatieveiligheidscomité**

dat de mededeling van de persoonsgegevens zoals beschreven in deze beraadslaging toegestaan is mits 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, op voorwaarde dat een "small cell risico-analyse" uitgevoerd zal worden door KCE om de mogelijkheid tot heridentificatie van de betrokkenen uit te sluiten

Dat het eHealth-platform wordt gemachtigd om het verband tussen de gepseudonimiseerde dataset en het reële identificatienummer te bewaren zodat het koppelen van de data tussen de verschillende leveranciers mogelijk blijft.

Bart VIAENE  
Voorzitter

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.
---

## Bijlage: Selectie van variabelen

### A. Toelichting en motivering

#### i. Algemeen

In deze studie zitten vier data-extracties vervat. Dit laat toe om de verschillende doelstellingen trapsgewijs in de tijd te realiseren in plaats van alles op het einde van de follow-up periode te analyseren. Dit maakt het mogelijk om de looptijd van het project efficiënt te gebruiken. Deze stapsgewijze benadering heeft als bijkomend voordeel dat de data incrementeel toeneemt in plaats van dat alles in een keer beschikbaar wordt gesteld. Dit alles zorgt ervoor dat, inzake datamanagement, alle procedures en protocollen beter beheersbaar en controleerbaar zijn.

De studie ambiert het opstellen van predictiemodellen naar ‘time to event’ via tijdreeksanalyses en Cox regressie analyse. Deze statistische modellen analyseren de duurtijd vanaf indexdatum (diagnose borstkanker) tot een bepaald event (bijvoorbeeld: opnamedatum, cardiovasculair event en diagnose longkanker). Een vereiste hierbij is dat het beginpunt (diagnose borstkanker) en de gebeurtenis/event gekend zijn zodoende niet enkel predicties te berekenen naar het voorkomen van events maar ook naar het tijdstip (vanaf indexdatum) waarop het event plaatsvindt. Echter, om te vermijden dat de onderzoeker toegang krijgt tot exacte datums, zullen de dataleveranciers de exacte datums herleiden tot een integere waarde ten opzichte van de incidentiedatum op basis van de aangeleverde incidentiedatum.

#### ii. Met betrekking tot aanvraag van socio-economische variabelen

Deze aanvraag kadert in een onderzoek naar de maatschappelijke kost van borstkanker. Het inschatten van zowel directe als indirecte kosten zijn inherent bij dergelijke economische evaluaties vanuit het maatschappelijk perspectief. De sociaaleconomische positie (bijvoorbeeld: arbeidsmarktpositie) laten toe om de indirecte kosten van zorg in te schatten. Daarnaast heeft de sociaaleconomische positie -en de evolutie hiervan in de tijd- een belangrijk effect op de verschillende sociale (financiële) beschermingsmaatregelen in België. De uitgebreidheid van socio-economische variabelen binnen dit onderzoek is noodzakelijk om die sociaaleconomische posities en sociale beschermingsmaatregelen correct te bepalen zodoende de indirecte kosten correct in te schatten en op te volgen in de tijd.

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

Data-leverancier	Variabele	Bewijs van proportionaliteit	Selectieperiode			
			1	2	3	4
SKR	INSZ	Het INSZ-nummer/uniiek identificatienummer zal enkel gebruikt worden voor selectie- en koppelingsdoeleinden. Deze variabele wordt omgezet naar een gepseudonimiseerd identificatienummer (variabele: C <sub>project</sub> ) dat ter beschikking gesteld wordt voor de onderzoekers.	X	X	X	X
IMA	SS00010 / PP0010 – Identificatie rechthebbende		X	X	X	X
KSZ	INSZ – Identificatienummer van de sociale zekerheid		X	X	X	X
FOD VVVL (MZG)	PATNUM		X	X	X	X

SKR	Geslacht	Deze variabele is noodzakelijk als selectievariabele voor het samenstellen van de cohort. Op basis van de combinatie van variabelen “Geslacht”, “Incidentiedatum” en “ICD10 (numeriek)” kunnen alle vrouwelijke patiënten met een diagnose van borstkanker tijdens incidentiejaar 2010 geïdentificeerd worden.	X	X	X	X
	Incidentiedatum (borstkanker) = Indexdatum	Deze variabele is noodzakelijk als selectievariabele voor het samenstellen van de cohort. In kader van het dataminimalisatieprincipe wordt geopteerd voor een methode waarbij geen exacte datums ter beschikking gesteld worden van de onderzoekers in de aangeleverde dataset. De exacte incidentiedatum (borstkanker) wordt door SKR aangeleverd aan alle dataleveranciers opdat zij op hun beurt alle datumvariabelen kunnen omzetten naar ‘aantal dagen t.o.v. indexdatum’. De lijst met incidentiedatums en patient-ID’s worden niet bewaard door de dataleveranciers, maar worden bij elke nieuwe datalevering opnieuw aangeleverd door SKR waarbij eHealth codeersleutels kan voorzien. In geval van meer dan één diagnose van borstkanker tijdens het indexjaar, geldt de eerste diagnose (incidentiedatum) als indexdatum. De aangeleverde lijst met incidentiedatums omvat slechts één record per patiënt. De indexdatum wordt gefixeerd bij de eerste data-extractie. Bij laattijdige correcties binnen indexjaar 2010 (cfr. wijziging in maand en/of dag), blijft de initiële indexdatum behouden tijdens data-extracties 2-4. Laattijdige correcties buiten het indexjaar (cfr. wijziging in jaar) resulteren in het verwijderen van de patiënt uit de dataset. Er worden geen nieuwe diagnoses (laattijdige correcties m.b.t. type tumor) t.o.v. de eerste data-extractie opgenomen in de dataset. Concreet betekent dit dat er ten opzichte van de eerste datalevering geen nieuwe patiënten toegevoegd kunnen worden aan de initiële selectie, het verwijderen van patiënten n.a.v. correcties is wel mogelijk. De conversiesleutel INSZ/Cproject blijft steeds dezelfde over het volledige verloop van het project. Deze sleutels dienen -evenals de aangeleverde data beschikbaar te blijven tot vier jaar na de aanlevering van de laatste dataset. Dit is noodzakelijk voor het voorkomend geval dat patiënten verwijderd worden uit de dataset.	X	X	X	X
	ICD10 (numeriek)	Deze variabele is noodzakelijk als selectievariabele voor het samenstellen van de cohort. Op basis van de combinatie van variabelen “Geslacht”, “Incidentiedatum” en “ICD10 (numeriek)” kunnen alle vrouwelijke patiënten met een diagnose van borstkanker tijdens incidentiejaar 2010 geïdentificeerd worden. Concreet betreft het ICD-10 codes C50.x en D05.x. <i>(!) Deze variabele is daarnaast ook noodzakelijk voor het identificeren van vrouwen in de cohort die</i>	X	X	X	X



		<i>gediagnosticeerd worden met een longtumor tot 20 jaar na incidentiedatum (cfr. infra).</i>				
--	--	---	--	--	--	--

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

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

<b>Dataleverancier</b>	<b>: SKR</b>
Aantal gevraagde variabelen (excl. variabelen voor koppeling en selectie)	: 22

Variabele	Bewijs van proportionaliteit	Selectieperiode			
		1	2	3	4
Leeftijd (bij diagnose)	Deze variabele is van belang voor het omschrijven van de cohort, en eventueel voor het corrigeren voor leeftijd (jaar) als risicofactor voor langetermijneffecten (zoals cardiovasculaire aandoeningen).	X	X	X	X
Incidentiedatum (longtumor) Incidentiedatum (borstkanker)	Incidentiedatum (longtumor of borstkanker) wordt omgezet naar een gecodeerde datum ('aantal dagen t.o.v. indexdatum'). Deze variabele is noodzakelijk om secundaire tumoren te kunnen identificeren in de cohort.	X	X	X	X
ICD10 (numeriek)	Naast het belang van deze variabele als selectievariabele (cfr. supra), dient deze meegenomen te worden over de volledige onderzoekstermijn voor de identificatie van de diagnose van longtumoren in de cohort. Concreet betreft het ICD-10 code C34.x, alsook ICD-10 codes C50.x en D05.x.	X	X	X	X
cTNM (cT + cN + cM)	Deze variabelen laten toe om de stadiëring te bepalen van de tumoren waardoor het mogelijk wordt om de verschillende behandelingstrajecten uit te zetten in functie van de stadia van ziekte.	X	X	X	X
pTNM (pT + pN + pM)		X	X	X	X
Combined stage		X	X	X	X
Histologisch subtype	Deze variabele laat subanalyses toe volgens histologisch subtype wat van invloed kan zijn op het behandelingstraject van de patiënt.	X	X	X	X
ER	Deze variabele laat toe om de ER-receptorstatus te bepalen wat van invloed kan zijn op het behandelingstraject van de patiënt.	X			
PR	Deze variabele laat toe om de PR-receptorstatus te bepalen wat van invloed kan zijn op het behandelingstraject van de patiënt.	X			
HER2	Deze variabele laat toe om de HER2-receptorstatus te bepalen wat van invloed kan zijn op het behandelingstraject van de patiënt.	X			
Aantal tumoren	Deze variabele geeft het aantal geregistreerde tumoren weer bij de patiënt en maakt het (in combinatie met variabele "Multiple") mogelijk om na te gaan of het borstcarcinoom de eerste gediagnosticeerde tumor is bij de patiënt.	X	X	X	X
Multiple	Deze variabele geeft inzicht in het chronologisch voorkomen van tumoren waardoor het borstcarcinoom met zekerheid als voorafgaande tumor aangeduid kan worden. Indien andere tumortypes gediagnosticeerd worden na het borstcarcinoom, zullen deze ook mee opgenomen worden in de telling, zonder dat de records van deze tumoren zullen worden meegedeeld.	X	X	X	X
Lateraliteit	Deze variabele geeft inzicht in de lateraliteit van de tumor. Op deze manier kan een onderscheid gemaakt worden tussen	X	X	X	X

	linker- en rechterborstcarcinomen, wat van belang kan zijn voor het inschatten van de langetermijneffecten.				
Vitale status	Deze variabelen maken het mogelijk om te bepalen of (en tot welke datum) patiënten uit de cohort nog in leven, dan wel overleden, of <i>lost to follow-up</i> zijn. Status datum wordt geconverteerd naar een gecodeerde datum ('aantal dagen t.o.v. indexdatum').	X	X	X	X
Status datum		X	X	X	X
Onmiddellijke doodsoorzaak	Deze variabelen zijn noodzakelijk om de sterfte en de oorzaak-specifieke doodsoorzaak in de cohort te bepalen.	X	X	X	X
Intermediaire doodsoorzaak 1		X	X	X	X
Intermediaire doodsoorzaak 2		X	X	X	X
Oorspronkelijke doodsoorzaak		X	X	X	X
Bijkomende doodsoorzaak 1		X	X	X	X
Bijkomende doodsoorzaak 2		X	X	X	X
Bijkomende doodsoorzaak 3		X	X	X	X

Tabel 2. Selectie van variabelen vanuit de databank van Stichting Kankerregister

#### D. Selectie van variabelen vanuit de databank van IMA<sup>6</sup>

<b>Dataleverancier</b>	<b>: IMA</b>
Aantal gevraagde variabelen (excl. variabelen voor koppeling en selectie)	: 21 (Gezondheidszorgen), 10 (Farmanet), 23 (Populatiedata)

Bestand(en)	Variabele	Bewijs van proportionaliteit	Selectieperiode			
			1	2	3	4
Gezondheidszorgen en Farmanet	SS00015 – Begindatum verstrekking / Datum van aflevering	Deze variabele laat toe na te gaan wanneer een bepaalde verstrekking (voor de eerste keer) gefactureerd werd. Het betreft verpleegdagen, (toediening van) farmaceutische producten en prestaties.  Deze datum wordt omgezet naar een gecodeerde datum ('aantal dagen t.o.v. indexdatum'). In combinatie met variabelen SS00020, SS00050 en SS00055 kunnen de directe zorgkosten op verschillende tijdsintervallen bepaald worden.	X	X	X	X
Gezondheidszorgen en Farmanet	SS00020 – Nomenclatuur / Code categorie geneesmiddel	De nomenclatuurcode of de categorie van het geneesmiddel (zoals voorgeschreven door het RIZIV) laten toe om na te gaan voor welke	X	X	X	X

		verstrekkingen de patiënt een vergoeding ontvangen heeft.				
Gezondheidszorgen en Farmanet	SS00050 – Aantal gevallen / Hoeveelheid	Deze variabele geeft het aantal keer weer dat een verstrekking verleend is, het aantal leveringen/afgeleverde eenheden of hoeveelheden of het aantal keer dat een eenheidstarief aangerekend werd op een bepaalde dag/periode. Deze variabele laat toe om de totale kost voor een specifieke verstrekking te berekenen.	X	X	X	X
Gezondheidszorgen	SS00055 – Aantal dagen	Deze variabele geeft - vergelijkbaar met SS00050- het aantal gefactureerde dagen of forfaits weer in een bepaalde periode. Het IMA adviseert om beide variabelen op te nemen omdat de informatie niet altijd op correcte wijze geregistreerd wordt.	X	X	X	X
Gezondheidszorgen en Farmanet	SS00060 – Bedrag terugbetaling / ZIV-tegemoetkoming 1	Deze variabele geeft het bedrag weer dat terugbetaald werd door de verzekeringsinstelling aan de patiënt. Deze variabele vormt, in combinatie met variabelen SS00160 en SS00165, de belangrijkste bron van informatie over de directe medische kosten van het zorggebruik.	X	X	X	X
Gezondheidszorgen	SS00065(A) – Zorgverstrekkers-identificatiecode	Deze variabelen laten toe om de zorgverstrekker die verantwoordelijk is voor het uitvoeren van de prestatie te identificeren. Dit vormt een belangrijke bron van informatie omtrent het voorkomen van ziektes/aandoeningen. Variabele SS00065(A) wordt in gecodeerde vorm aangeleverd.	X	X	X	X
Gezondheidszorgen	SS00065(B) – Zorgverstrekkers-bekwaming		X	X	X	X
Gezondheidszorgen	SS00075 – Identificatie instelling	Deze variabele is noodzakelijk om te traceren in welke instelling de rechthebbende van de verleende prestaties opgenomen is (of verblijft). Deze variabele wordt voorzien in een unieke gepseudonimiseerde code per voorziening.	X	X	X	X

Gezondheidszorgen	SS00080 – Dienstcode	Deze variabele laat toe om in detail de kostenplaats/dienstcode voor de verleende zorg binnen een bepaalde zorgvoorziening te bepalen.	X	X	X	X
Gezondheidszorgen	SS00085 – Plaats van verstrekking	Deze variabele laat toe om - op gecodeerde wijze- de instelling te identificeren waar de prestatie werkelijk uitgevoerd werd.	X	X	X	X
Gezondheidszorgen en Farmanet	SS00095 – Code gerechtigde 1 (CG1)	Deze variabelen laten toe om het sociaal statuut van de patiënt te bepalen. Deze variabelen laten toe om het sociaal statuut gelinkt aan de primaire borstkankerbehandeling in detail te bestuderen. In functie van doelstelling 2 is dergelijk detail minder noodzakelijk waardoor deze variabelen bij de volgende data-extracties (2 t.e.m. 4) worden vervangen door PP0030 en PP0035.	X			
	SS00100 – Code gerechtigde 2 (CG2)		X			
Gezondheidszorgen	SS00105 – Nummer derde	Deze variabele laat toe om - op gecodeerde wijze- de instelling of verstrekker die de prestatie gefactureerd heeft te identificeren.	X	X	X	X
Gezondheidszorgen	SS00110 – Datum opname	Deze variabelen geven aan wanneer een patiënt opgenomen/ontslagen is uit een zorgvoorziening. Beide variabelen worden geconverteerd naar gecodeerde variabele ('aantal dagen t.o.v. indexdatum'). Deze variabelen worden enkel opgevraagd tijdens de eerste data-extractie aangezien MZG-variabelen niet beschikbaar zijn tijdens deze extractie.  Deze variabele wordt enkel opgevraagd tijdens de eerste data-extractie aangezien de MZG-variabelen niet beschikbaar zijn tijdens deze extractie.	X			
	SS00115 – Datum ontslag		X			
Gezondheidszorgen en Farmanet	SS00135 – Nummer product	Deze code staat voor de code van het gefactureerde farmaceutisch product. Hierdoor kan exact bepaald worden welke geneesmiddelen de patiënt	X	X	X	X

		gebruikt. Hieruit kan o.a. afgeleid worden aan welke aandoeningen de patiënt lijdt.				
Gezondheidszorgen	SS00150 – Gefactureerde nomenclatuurcodes	De gefactureerde nomenclatuurcode geeft de (pseudo-) nomenclatuurcode weer voor de verstrekking die de zorgverstrekker aangerekend heeft aan de patiënt. Het geeft informatie over pseudocodes waardoor de specifieke aard van de verstrekking bepaald kan worden.	X	X	X	X
Gezondheidszorgen en Farmanet	SS00160 – Persoonlijke tussenkomst / Persoonlijk aandeel	Deze variabele geeft het reëel persoonlijk aandeel (wettelijk remgeld) weer dat de patiënt heeft betaald voor de verstrekking. In combinatie met variabelen SS00060 en SS00165 kan hierdoor de directe medische kost berekend worden.	X	X	X	X
Gezondheidszorgen	SS00165 – Bedrag supplement of bedrag voor niet-vergoedbare producten, verstrekkingen of diensten	Deze variabele geeft informatie over supplementen die aangerekend werden aan de patiënt bovenop het persoonlijk aandeel. Deze supplementen vormen, in combinatie met variabelen SS00060 en SS00160, de basis voor de berekening van de directe medische kosten.	X	X	X	X
Farmanet	SS00165 – Vermindering verzekeringsstegemoetkoming/Bijdrage van de apothekers	Deze variabele geeft de vermindering op de verzekeringsstegemoetkoming van geneesmiddelen weer. Naar analogie van variabele SS00165 in het bestand Gezondheidszorgen, is deze variabele noodzakelijk om de kost van het geneesmiddelengebruik te berekenen.	X	X	X	X
Gezondheidszorgen	SS00170 – Code implant	Deze variabele geeft de identificatiecode voor implantaten weer (zoals gefactureerd in variabele SS00150). Dit kan van belang zijn om het zorggebruik in kaart te brengen bij vrouwen die reconstructieve heelkunde ondergaan na een borstamputatie.	X	X	X	X
Gezondheidszorgen	SS00175 – Derdebetalersregeling	Deze variabele laat toe om te bepalen of een patiënt al dan niet geniet van een	X	X	X	X

		voorkeursregeling volgens het derdebetalerssysteem. Dit kan relevant zijn voor patiënten die recht hebben op een verhoogde tegemoetkoming (identificeerbaar in variabele SS00160).				
Farmanet	SS00195 – ZIV-tegemoetkoming 2	Deze variabele geeft de som weer van de aangerekende honoraria en laat een gedetailleerde berekening van de kost van het geneesmiddelengebruik toe.	X	X	X	X
Populatiedata	PP1010 – Aard VT/OMNIO	Deze variabele geeft weer of de patiënt recht heeft op een verhoogde tegemoetkoming (voorheen OMNIO-statuut). Deze variabele laat toe om het sociaal statuut gelinkt aan de primaire borstkankerbehandeling in detail te bestuderen. In functie van doelstelling 2 is dergelijk detail minder noodzakelijk waardoor deze variabelen bij de volgende data-extracties worden vervangen door CHRONICAL_YN.	X			
Populatiedata	PP2001 – Forfait B geneeskundige zorgen	Deze variabelen laten toe om te bepalen of de patiënt recht heeft op de omschreven tegemoetkomingen/verstrekkingen. Deze variabele laat toe om het sociaal statuut, gelinkt aan de primaire borstkankerbehandeling, in detail te bestuderen. In kader van de berekening van de incrementele kost (doelstelling 1) is het van belang om exact te bepalen welke sociale maatregelen een impact hebben op de vermindering in de directe en indirecte kosten van zorg (tot 1,5 jaar pre- en post-indexdatum). In functie van doelstelling 2 is dergelijk detail minder noodzakelijk waardoor deze variabelen bij de volgende data-extracties worden vervangen door CHRONICAL_YN.	X			
	PP2002 – Forfait C geneeskundige zorgen					
	PP2003 – Kinesithérapie E of fysiotherapie					
	PP2004 – Verhoogde kinderbijslagen					
	PP2005 – Toelage voor integratie van gehandicapten					
	PP2006 – Toelage voor hulp aan ouderen (cat III, IV of V)					

	PP2007 – Uitkering hulp aan derden					
	PP2008 – Verhoogde invaliditeitsuitkering wegens hulp aan derden					
	PP2009 – Forfaitaire tegemoetkoming voor hulp aan derden					
	PP2010 – Criterium hospitalisatie (120 dagen)					
	PP2011 – Criterium hospitalisaties (6 hospitalisaties)					
	PP3011 – Recht op toelage voor gehandicapten					
Populatiedata	PP0030 – Code gerechtigde 1 (CG1)	Deze variabelen laten toe om het sociaal statuut van de patiënt te bepalen. Dit kan een verklarende factor zijn voor (verhoogde) tegemoetkomingen voor verleende zorg.	X	X	X	X
	PP0035 – Code Gerechtigde 2 (CG2)		X	X	X	X
Populatiedata	PP3001 – Recht MAF – gezin	Deze variabelen laat toe te bepalen welk type MAF toegekend kan worden aan de patiënt en/of het gezin. Deze variabelen dienen samen geïnterpreteerd te worden.	X	X	X	X
Populatiedata	PP3004 – Terugbetaling en plafond - gezin	Deze variabelen omschrijft de plafonnering die de basis vormt voor de hoogte van de terugbetaling.	X	X	X	X
Populatiedata	PP3005 – Terugbetaling en plafond - individu of gedeelte van het gezin	Deze variabele omschrijft het recht, en de plafonnering op basis waarvan de hoogte van de terugbetaling bepaald wordt.	X	X	X	X
Populatiedata	PP4001 – Aantal dagen werkloosheid	Deze variabelen laten toe een inschatting te maken van het productiviteitsverlies. Deze variabele is noodzakelijk voor de berekening van de indirecte kosten.	X	X	X	X
	PP4002 – Aantal dagen invaliditeit		X	X	X	X
Populatiedata	CHRONICAL_YN	Deze flag laat toe om te bepalen of de patiënt in het afgelopen jaar recht had op	X	X	X	X

		een attest/statuut Chronische aandoening en/of een toelage voor gehandicapten. Deze variabele kan verklarend zijn voor verschillen in de kosten van zorg (als gevolg van de toekenning van verhoogde tegemoetkoming). Deze variabele wordt enkel opgevraagd tijdens data-extracties 2, 3 en 4 aangezien voor de eerste data-extractie een selectie van variabelen opgevraagd wordt die meer detail toelaat.				
Populatiedata	MAJOR_COVERAGE_YN	Deze flag laat toe om te bepalen of een patiënt al dan niet geniet van een voorkeursregeling. Dit kan relevant zijn voor patiënten die recht hebben op een verhoogde tegemoetkoming (identificeerbaar in variabele SS00160).	X	X	X	X
Populatiedata	ATC_ANAT_L + ATC_THER_L + ATC_PHAR_L + ATC_CHEM_L + ATC_PROD_L	Deze flag is een afgeleide van variabele SS00135 (nummer product) volgens de referentietabellen van de ATC codes zoals voorgesteld door het KCE.	X	X	X	X

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

#### E. Selectie van variabelen vanuit de databank van FOD VVVL (MZG databank)<sup>7</sup>

<b>Dataleverancier</b>	<b>: MZG</b>
Aantal gevraagde variabelen (excl. variabelen voor koppeling en selectie)	: 26 (totaal 34 variabelen waarvan 4 gehercodeerde exacte datums)

Bestand(en)	Variabele	Bewijs van proportionaliteit	Selectieperiode			
			1	2	3	4
DIAGNOSE, PROCEDUR, PROCRIZI en TESTRESU	CODE_AGR	Deze variabele geeft het erkenningsnummer van het ziekenhuis weer waardoor het mogelijk wordt om te bepalen waar (ziekenhuis) de zorg verleend werd. Deze variabele wordt voorzien in een unieke gepseudonimiseerde code per voorziening.		X	X	X
STAYHOSP, STAYINDX, DIAGNOSE, PROCEDUR,	STAYNUM	Een nieuw verblijfsnummer wordt toegekend bij iedere ziekenhuisopname waardoor differentiatie tussen verleende		X	X	X

<sup>7</sup> De kwaliteit van gekoppelde data vanuit de MZG databank kan niet gegarandeerd worden voor data voor 2012. Om de kwaliteit van analyses niet te compromitteren, wordt geopteerd om enkel gebruik te maken van MZG data voor dataleveringen 2 tot en met 4 (doelstellingen 2 en 3).



PROCRIZI en TESTRESU		zorg tijdens verschillende ziekenhuisopnames mogelijk is.				
STAYHOSP	A2_YEAR_HOSP_IN	Deze variabelen maken het mogelijk om de opnamedatum voor de gespecificeerde hospitalisatie (STAYNUM) te bepalen. De combinatie van deze drie variabelen wordt omgezet naar één gecodeerde variabele: ‘aantal dagen t.o.v. indexdatum’.		X	X	X
	A2_MONTH_HOSP_IN			X	X	X
	A2_DAY_HOSP_IN			X	X	X
STAYHOSP	A2_YEAR_HOSP_OUT	Deze variabelen maken het mogelijk om de ontslagdatum voor de gespecificeerde hospitalisatie (STAYNUM) te bepalen. De combinatie van deze drie variabelen wordt omgezet naar één gecodeerde variabele: ‘aantal dagen t.o.v. indexdatum’.		X	X	X
	A2_MONTH_HOSP_OUT			X	X	X
	A2_DAY_HOSP_OUT			X	X	X
STAYHOSP	A2_HOSP_TYPE_CAT	Deze variabele geeft het verblijfstype van de hospitalisatie weer met o.a. een onderscheid tussen “alle soorten verblijf” en “oncologisch daghospitaal”. Dit kan van belang zijn om de zorgen verleend aan vrouwen uit het cohort te bepalen.		X	X	X
STAYHOSP	A2_HOSP_TYPE_FAC	Deze variabele geeft het verblijfstype weer op basis van facturatiegegevens met o.a. een onderscheid tussen “klassieke hospitalisatie” en “niet-chirurgische daghospitalisatie”. Dit kan van belang zijn om de zorgen verleend aan vrouwen uit het cohort te bepalen.		X	X	X
STAYHOSP	A2_TOTAL_NUMBER_DAY_FAC	In deze variabele wordt de som van het totaal aantal gefactureerde ligdagen tijdens de registratieperiode weergegeven. Het aantal gefactureerde ligdagen is van belang om de kosten van de verleende zorg te kunnen berekenen.		X	X	X
STAYHOSP	A2_CODE_READMISSION	Deze variabele codeert heropnames binnen hetzelfde jaar, binnen hetzelfde ziekenhuis. Het is een indicator voor toegenomen zorggebruik.		X	X	X
STAYHOSP	A2_CODE_STAT_INSURANCE	Deze code geeft de verzekeringsstatus van de patiënt weer tijdens het verblijf. Deze variabele kan verklarend zijn voor het toekennen van verhoogde tegemoetkomingen.		X	X	X
STAYHOSP	A2_CODE_PLACE_BEFORE_ADM	Deze variabele geeft weer waar de patiënt zich bevond vlak voor de opname. Onder de codes die toegekend worden o.a. “Ander ziekenhuis” en “Ongepland		X	X	X

		klassiek verblijf onmiddellijk aansluitend op een verblijf in daghospitalisatie". Deze codes kunnen van belang zijn om de trajecten van de cohort in kaart te brengen.			
STAYHOSP	A2_CODE_ADM	Deze variabele geeft het type opname weer (o.a. "Geplande hospitalisatie" en "Dringende opname ..."). Deze codes kunnen van belang zijn om de urgentie van de behandelingen (en daaraan gerelateerde kosten) in te schatten.		X	X
STAYHOSP	A2_CODE_DESTINATE	Deze variabele geeft weer waarnaar de patiënt verwezen wordt na ontslag. Onder de codes die toegekend worden o.a. "Overleden". Deze code kan van belang zijn om te bepalen of de patiënt overleden is tijdens het verblijf in het ziekenhuis.		X	X
STAYHOSP	A2_CODE_DISCHARGE	Deze variabele geeft of de patiënt ontslagen is uit het ziekenhuis (evt. op verwijzing). Onder de codes die toegekend worden o.a. "Overleden". Deze code kan van belang zijn om te bepalen of de patiënt overleden is tijdens het verblijf in het ziekenhuis.		X	X
STAYINDX en PROCRIZI	ORDER_BEDINDEX	Deze variabele geeft in chronologische volgorde de bedindexen (dienstcodes/kostenplaatsen) weer waar de patiënt verbleven heeft tijdens de gespecificeerde hospitalisatie. Deze codes kunnen van belang zijn om de trajecten van de cohort in kaart te brengen.		X	X
STAYINDX	A4_CODE_BEDINDEX_FAC	Deze variabele geeft de code weer van de dienst waarop een bepaalde verstrekking gefactureerd werd. Deze codes kunnen van belang zijn om de trajecten van de cohort in kaart te brengen.		X	X
STAYINDX	A4_YEAR_BEDINDEX_IN	Deze variabelen maken het mogelijk om de opnamedag op een bepaalde bedindex tijdens een gespecificeerde hospitalisatie te bepalen. De combinatie van deze drie variabelen wordt omgezet naar één gecodeerde variabele: 'aantal dagen t.o.v. indexdatum'		X	X
	A4_MONTH_BEDINDEX_IN			X	X
	A4_DAY_BEDINDEX_IN			X	X
STAYINDX	A4_NUMBER_DAY_FAC	Deze variabele geeft het aantal aaneensluitende dagen op een bepaalde bedindex weer en is van belang om de trajecten van de cohort in kaart te brengen.		X	X

STAYINDEX	A4_NUMBER_DAY_FAC_PREV	Deze variabele laat toe om te corrigeren voor gefactureerde dagen in het voorgaande registratiejaar. Deze variabele dient samen geïnterpreteerd te worden met variabele "A4_NUMBER_DAY_FAC".		X	X	X
DIAGNOSE en PROCEDUR	ORDER_SPEC	Deze variabele geeft het volgnummer van de verschillende specialisten weer die verantwoordelijk zijn voor de hoofd- en nevendiaagnoses tijdens een gespecificeerde hospitalisaties.		X	X	X
DIAGNOSE en PROCEDUR	TYPE_DIAGNOSE <sup>8</sup>	Deze variabele maakt het mogelijk om hoofd- en nevendiaagnoses (variabele "CODE_DIAGNOSE") tijdens een gespecificeerde hospitalisatie te differentiëren.		X	X	X
DIAGNOSE en PROCEDUR	CODE_SYS_DIAGNOSE <sup>8</sup>	Deze variabele geeft het soort coderingssysteem (ICD-9-CM of ICD-10-BE) weer voor de diagnose. Deze variabele is noodzakelijk om variabele "CODE_DIAGNOSE" te interpreteren.		X	X	X
DIAGNOSE	CODE_DIAGNOSE <sup>8</sup>	Deze variabele vermeldt de ICD code voor de hoofd- en nevendiaagnoses tijdens een gespecificeerde hospitalisatie. Deze variabele is noodzakelijk om het zorggebruik te linken aan ziekte.		X	X	X
PROCEDUR	M2_CODE_SYS_PROCEDURE <sup>8</sup>	Deze variabele geeft het soort coderingssysteem (ICD-9-CM of ICD-10-BE) weer voor de uitgevoerde ingreep. Deze variabele is noodzakelijk om variabele "M2_CODE_PROCEDURE" te interpreteren.		X	X	X
PROCEDUR	M2_CODE_PROCEDURE	Deze variabele omschrijft de code van de uitgevoerde ingrepen tijdens een gespecificeerde hospitalisatie. Deze variabele is noodzakelijk om uitgevoerde ingrepen (verleende zorgen) te bepalen.		X	X	X
PROCEDUR	M2_YEAR_PROCEDURE	Deze variabelen laten toe om de datum van de uitgevoerde ingreep te bepalen.		X	X	X
	M2_MONTH_PROCEDURE			X	X	X
	M2_DAY_PROCEDURE	De combinatie van deze drie variabele wordt omgezet naar één		X	X	X

<sup>8</sup> Dit project focust op cardiotoxiciteit en radiatie-geïnduceerde langetermijneffecten van borstkankerbehandeling. Opnames in kader van niet-gerelateerde diagnoses zijn niet relevant. Voorafgaand aan de datalevering zal, in samenspraak met FOD VVVL, een lijst opgesteld worden van relevante ICD codes. Enkel opnames als gevolg van deze ICD codes (hoofd- en/of nevendiaagnoses) worden geëxtraheerd.

		gecodeerde variabele: 'aantal dagen t.o.v. indexdatum'.  Deze procedures stellen ons in staat om de radiotherapeutische bestralingsprogramma's te reconstrueren voor de vrouwen in het cohort. Op deze manier kan nagegaan worden welke bestralingstechnieken en -schema's een impact hebben op de langetermijneffecten van de behandeling.				
PROCEDUR	M2_NUMBER_PROC_1_DAY	Deze variabele geeft aan of de ingreep meerdere keren uitgevoerd werd op dezelfde dag. Deze variabele is noodzakelijk om kosten die voortvloeien uit een bepaalde ingreep te kunnen verklaren.		X	X	X

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

#### F. Selectie van variabelen vanuit de databank van KSZ

<b>Dataverancier</b>	<b>: IKSZ</b>
Aantal gevraagde variabelen (excl. variabelen voor koppeling en selectie)	: 21

Bron	Variabele	Bewijs van proportionaliteit	Selectieperiode			
			1	2	3	4
DWH_BCSS_NomenclatureVarDer	Nomenclatuur van socio-economische positie (Nomenc)	Deze variabele laat toe om de socio-economische positie (tewerkstellingsvorm) te bepalen. Via deze variabele kan een inschatting gemaakt worden van het percentage en de snelheid van werkhervatting na ziekte.  Dit is noodzakelijk om een inschatting te maken van de indirecte kosten van zorg (in functie van doelstellingen 1 en 2). Deze variabele wordt geïnterpreteerd in combinatie met de variabelen van "bijkomende socio-economische positie" (zie hieronder).  Om de heridentificatie van personen in de dataset te verhinderen, wordt het detail van de variabelen beginnend met 1 (werkend) beperkt tot dit eerste niveau.	X	X	X	X

	Bijkomende positie van socio-economische positie (hercodering)	<p>Om de heridentificatie van personen in de dataset te verhinderen, werden een aantal variabelen gehercodeerd waarbij het plausibel geacht wordt dat deze personen administratief werkend zijn, doch feitelijk hun beroepsactiviteit (tijdelijk) stopgezet hebben. Concreet betreft het volgende variabelen:</p> <ul style="list-style-type: none"> <li>- Werkend en gedeeltelijk in terbeschikkingstelling voorafgaand aan pensioen (Werkend_en_gedeeltelijk_TBS)</li> <li>- Werkend en gedeeltelijke loopbaanonderbreking/deel tijds tijds krediet (interr_carriere_temps_partiel)</li> <li>- Werkend en halftijds brugpensioen (prepension_mi_temps)</li> <li>- Werkend en halftijds in een stelstel van werkloosheid met bedrijfstoelag (prepension_mi_temps)</li> <li>- Werkend en volledig in terbeschikkingstelling voorafgaand aan pensioen (Werkend_en_volledig_TB S)</li> <li>- Werkend en volledige loopbaanonderbreking/volledig tijds krediet (interr_carriere_complete)</li> <li>- Werkend en volledig brugpensioen (prepension_complete)</li> <li>- Werkend en volledig in stelstel van werkloosheid met bedrijfstoelag (prepension_complete)</li> </ul>	X	X	X	X
DWH_STATBEL_CENSUS2011	Gedetailleerd onderwijsniveau (CD_ISCED_DET)	Deze variabele geeft een momentopname van het hoogst behaalde opleidingsniveau op 01/01/2011. Om echter na te gaan of de personen in het cohort een hoger onderwijsniveau behaalden in de jaren volgend op deze momentopname, dienen onderstaande vier variabelen (Opleidingsniveau ISCED, cat_shu, cat_études, CO_DIPLOME) eveneens opgevraagd te worden.	X	X	X	X

DWH_VDAB_FOREM_ACTIRIS_ADG	Opleidingsniveau ISCED	Deze variabele geeft het opleidingsniveau weer (Vlaamse gemeenschap), en kan verklarend zijn voor de snelheid van werkhervatting na een periode van ziekte.	X	X	X	X
DWH_CREF_diplomes	Studie categorie (niet-universitair onderwijs) (cat_shu)	Deze variabele geeft het opleidingsniveau weer (Franstalige gemeenschap), en kan verklarend zijn voor de snelheid van werkhervatting na een periode van ziekte.	X	X	X	X
	Studie categorie (cat_études)					
	Code van diploma (CO_DIPLOME)					
DWH_ONSS_UniStatbaseDMFA	Prestatietype (T_Prest)	Deze variabele geeft het arbeidsregime (voltijds, deeltijds, speciaal, onbepaald, afwezigheid wegens ziekte) op de laatste dag van het kwartaal weer. Deze variabele laat toe om te bepalen of (en in welk regime) de patiënt werkzaam is.	X	X	X	X
	Voltijdsequivalent exclusief gelijkgestelde dagen (Eqtpsp)	Deze variabele geeft het tewerkstellingspercentage per arbeidsplaats weer gedurende het kwartaal.	X	X	X	X
	Voltijdsequivalent inclusief gelijkgestelde dagen (Tauxaa)	Deze variabele geeft het tewerkstellingspercentage per arbeidsplaats weer gedurende het kwartaal.	X	X	X	X
DWH_CIN	Begindatum arbeidsongeschiktheid (Begindatum)	Deze variabele geeft de begin- en einddatum van arbeidsongeschiktheid weer. Deze variabelen worden geconverteerd naar twee gecodeerde variabelen ('aantal dagen t.o.v. indexdatum'). Deze variabele laat toe om het productiviteitsverlies te bepalen. Deze variabele is noodzakelijk voor de berekening van de indirecte kosten. De exacte berekening van indirecte kosten zijn inherent bij het bepalen van de maatschappelijke kost.	X	X	X	X
	Einddatum arbeidsongeschiktheid (Einddatum)					
	Aantal dagen arbeidsongeschiktheid (Aantal_dagen_ao)	Deze variabele wordt beschouwd als een controlevariabele op de variabelen "Begindatum" en "Einddatum". Deze variabele maakt het mogelijk om op eenvoudige wijze berekeningen te maken omtrent productiviteitsverlies en indirecte kosten.	X	X	X	X

	Samenstelling gezin (Samenst_gezin)	Deze variabele geeft de samenstelling van het gezin weer en kan verklarend zijn voor het toekennen van verhoogde tegemoetkomingen.	X	X	X	X
DWH_INAMI_CMI	Type ziekte (Type_ziekte)	Deze variabele is noodzakelijk om variabele "Diag_ziekte" te interpreteren. Deze variabele geeft het coderingssysteem (ICD-9 of ICD-10) weer waarin de ziekte gecodeerd werd.	X	X	X	X
	Medische code (Gr_ziekte)	Deze variabele laat toe om de oorzaak van invaliditeit (erkend door Geneeskundige Raad voor Invaliditeit) te bepalen. Dit laat toe om de invaliditeit te linken aan de borstkankerbehandeling.	X	X	X	X
	Diag ziekte (Diag_ziekte)	Deze variabele vermeldt de ICD code op basis waarvan de invaliditeit (variabele "Gr_ziekte") toegekend werd.	X	X	X	X
	Begin (B_ziekte)	Deze variabele geeft de begindatum van de primaire arbeidsongeschiktheid weer. Deze variabele wordt omgezet naar een geconverteerde variabele ('aantal dagen t.o.v. indexdatum'). Deze variabele wordt beschouwd als een controlevariabele op de variabele "Begindatum".	X	X	X	X
	Begindatum [invaliditeit] (B_inval)	Deze variabele geeft de begin- en einddatum van invaliditeit weer. Deze variabelen worden geconverteerd naar twee gecodeerde variabelen ('aantal dagen t.o.v. indexdatum'). Deze variabele laat toe om het productiviteitsverlies te bepalen. Deze variabele is noodzakelijk voor de berekening van de indirecte kosten. De exacte berekening van indirecte kosten zijn inherent bij het bepalen van de maatschappelijke kost.	X	X	X	X
	Einddatum [invaliditeit] (E_inval)					
	R_exclus	Deze variabele is noodzakelijk om te corrigeren bij het samenvoegen van data uit verschillende kwartalen	X	X	X	X

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