



**Centraal Bureau voor de Statistiek**

Centrum voor Beleidsstatistiek

---

## **Documentatierapport GBA\_Uniciteit-LMR 1995-2011**

*Datum:* 11 december 2013

## **Bronvermelding**

Publicatie van uitkomsten geschiedt door het onderzoeksbureau of de opdrachtgever op eigen titel. Verwijzing naar het CBS betreft uitsluitend het gebruik van de microdatabestanden. Dat wordt als volgt geformuleerd:

“Eigen berekening [*naam onderzoeksbureau, c.q. opdrachtgever*] op basis van door het CBS gegenereerde hulpbestanden van de in de Gemeentelijke Basisadministratie (GBA) ingeschreven personen vanaf 1995, waarmee longitudinale studies kunnen worden uitgevoerd”

## Versiegeschiedenis

### *Versie 1*

131211 GBA\_Uniciteit-LMR 1995-20130116V1

131211 GBA\_Uniciteit\_compleet-LMR 1995-20130116V1

131211 GBA\_Uniciteit\_pj-LMR 1995-2011V1

---

De gebruiker dient rekening te houden met het volgende:

- Uitsluitend de variabele RINPERSOON identificeert een persoon.
- Vanwege een gewijzigd verwerkingsproces bij het CBS van de gegevens van de Gemeentelijke Basisadministratie (waarvan de GBA\_Uniciteitsbestanden afgeleid worden) bevat deze versie van de GBA\_Uniciteitsbestanden voor een klein deel van de personen een wijziging in de uniciteitsstatus in vergelijking met eerdere versies van dit bestand (zie paragraaf 2.2).
- De startdatum van het GBA\_Uniciteitsbestand is vanaf 2012 gewijzigd in 1 oktober 1994 in plaats van 1 januari 1995. Dit is ook een gevolg van de wijzigingen in het verwerkingsproces van de GBA.
- De koppelsleutel voor personen is veranderd van SRTNUM en RIN naar RINPERSOONS en RINPERSOON. Het omzetten van de ene sleutel naar de andere wordt beschreven in het script “koppelvariabelen gelijk zetten spss.sps” voor de Remote Access en On-site gebruikers te vinden in de map \8\_Uilities\Tools\. Let vooral op het omzetten van RIN naar RINPERSOON omdat RINPERSOON wordt weggeschreven mét voorloophnullen.



## Inhoudsopgave

<b><i>Bronvermelding</i></b> .....	<b>2</b>
<b><i>Versiegeschiedenis</i></b> .....	<b>3</b>
<b>1. <i>Inleiding</i></b> .....	<b>7</b>
<b>2. <i>Toelichting op de samenstelling van de bestanden</i></b> .....	<b>8</b>
<b>2.1 Microdatabestanden</b> .....	<b>8</b>
13121 GBA_Uniciteit-LMR en GBA_Uniciteit_compleet-LMR 1995-20130116V1 .....	8
131211 GBA_Uniciteit_pj-LMR 1995-2011V1 .....	9
<b>2.2 Inhoud uniciteitsbestanden</b> .....	<b>9</b>
<b>2.3 Gebruik van de uniciteitsbestanden</b> .....	<b>10</b>
Bepalen van begin- en einddatum follow-up periode.....	11
Selecteren van unieke populaties over hele kalenderjaren .....	13
<b>2.4 Representativiteit van de unieke populatie</b> .....	<b>13</b>
<b>2.5 Missende en niet koppelende records</b> .....	<b>14</b>
<b>2.6 Kwaliteit</b> .....	<b>16</b>
<b>3. <i>Bestandsopbouw en toelichting</i></b> .....	<b>17</b>
<b>3.1 Bestandsopbouw van het microdatabestanden</b> .....	<b>17</b>
<b>3.2 Toelichting op de variabelen</b> .....	<b>19</b>
RINPERSOON.....	19
VOLGNR.....	19
BEGGELY .....	19
BEGGELM .....	19
BEGGELD .....	19
EINDGELY .....	19
EINDGELM.....	19
EINDGELD .....	19
STATUS .....	19
UNIJJ .....	19
<b>3.3 Voorkomende waarden van de variabelen</b> .....	<b>20</b>
131211 GBA_Uniciteit -LMR 1995-20130116V1 .....	20
131211 GBA_Uniciteit_compleet-LMR 1995-20130116V1 .....	20
131211 GBA_Uniciteit_pj-LMR 1995-2011V1 .....	21



## 1. Inleiding

Het Centrum voor Beleidsstatistiek stelt, onder [bepaalde voorwaarden](#), microdata (geanonimiseerde data op persoons- en bedrijfsniveau) ter beschikking voor onderzoek. Dit rapport beschrijft de inhoud en structuur van de bestanden:

- ‘131211 GBA\_Uniciteit-LMR 1995-20130116V1’;
- ‘131211 GBA\_Uniciteit\_compleet-LMR 1995-20130116V1’;
- ‘131211 GBA\_Uniciteit\_pj-LMR 1995-2011V1’.

Deze bestanden zijn hulpbestanden die in eerste instantie gebruikt worden bij het analyseren van data uit de Landelijke Medische Registratie (LMR; Dutch Hospital Data, Utrecht). Daarom is het belangrijk om eerst het documentatierapport te lezen behorend bij de microdatabestanden van de [LMR](#).

Hoofdstuk 2 beschrijft de microdatabestanden zoals de populatieafbakening, methodologische bijzonderheden, kwaliteit en herkomst van de gegevens en er wordt een toelichting gegeven op het gebruik van de bestanden. In hoofdstuk 3 wordt de bestandsindeling per microdatabestand gegeven. Daarnaast wordt per microdatabestand per variabele meer informatie gegeven. Voor de categoriale variabelen worden alle mogelijke scores en hun betekenis opgesomd.

## 2. Toelichting op de samenstelling van de bestanden

### 2.1 Microdatabestanden

Deze GBA\_Uniciteit-LMR bestanden zijn hulpbestanden die in eerste instantie gebruikt worden bij het analyseren van data uit de Landelijke Medische Registratie (LMR; Dutch Hospital Data, Utrecht). Daarom is het belangrijk om eerst het documentatierapport te lezen behorend bij de microdatabestanden van de [LMR](#).

In deze GBA\_Uniciteitsbestanden wordt voor alle personen die ingeschreven staan in de Gemeentelijke Basisadministratie (GBA) weergegeven in welke periode(n) vanaf 1995 zij aanwezig waren in Nederland en welke tijd zij daarvan uniek identificeerbaar waren op de variabelen geslacht, geboortedatum en 4-cijferige postcode. Het bestand bevat dus gegevens over zowel geboorte/sterfte, immigratie/emigratie en wel/niet unieke koppelbaarheid. De bestanden zijn ontwikkeld omdat voor de koppeling van de LMR met de Gemeentelijke Basisadministratie (GBA) geen volledig identificerende koppelsleutel beschikbaar is, namelijk alleen de eerder genoemde koppelsleutel geslacht, geboortedatum en 4-cijferige postcode (zie [Koppeling van LMR- en GBA-gegevens: methode, resultaten en kwaliteitsonderzoek](#)). Voor analyses waarbij personen gedurende een bepaalde tijd gevolgd dienen te worden op ziekenhuisopnamen is het daarom belangrijk te weten dat alle opnamen die plaatsgevonden hebben tijdens deze periode uniek gekoppeld konden worden aan de GBA. Met behulp van de GBA\_Uniciteitsbestanden kan de onderzoeker alleen die GBA-personen voor het onderzoek selecteren, waarvan bekend is dat ze gedurende de onderzoeksperiode uniek koppelbaar waren op basis van hun koppelsleutel. Daardoor wordt gewaarborgd dat vrijwel alle geregistreerde LMR-opnamen van deze personen gekoppeld konden worden.

#### ***13121 GBA\_Uniciteit-LMR en GBA\_Uniciteit\_compleet-LMR 1995-20130116V1***

Een record in dit microdatabestand geeft van één persoon een periode weer waarin de uniciteit niet is veranderd. De uniciteit geeft aan of de persoon in Nederland aanwezig was en de perioden dat deze persoon op zijn/haar koppelsleutel (geslacht, geboortedatum en 4-cijferige postcode) wel of niet uniek identificeerbaar was. Elke persoon heeft een record in dit bestand dat begint op 1 oktober 1994 (begindatum van het bestand) en dat loopt tot de datum waarop de uniciteit verandert. Indien de uniciteit verandert, wordt in een volgend record de nieuwe status van de uniciteit aangegeven en de datum van het begin en einde van deze status (begin geldigheid en eind geldigheid). Eén persoon kan dus meerdere records hebben in dit microdatabestand. Daarom is de variabele volgnummer (VOLGNR) opgenomen in het bestand om per persoon de volgorde van de records in de tijd aan te geven. Het eerste record van een persoon in dit bestand heeft altijd volgnummer 1. Bij elk volgend record van een persoon wordt het volgnummer met 1 opgehoogd. Van iedere persoon begint het eerste record op 1 oktober 1994 en eindigt het laatste record op 16 januari 2013. Records van één persoon sluiten altijd op elkaar aan in de tijd. Indien de status van de uniciteit gedurende de hele periode niet verandert, heeft deze persoon in dit bestand 1 record dat loopt vanaf 1 oktober 1994 tot en met 16 januari



2013. Personen die emigreren of overlijden krijgen een record beginnend op de dag na emigratie of overlijden tot en met januari 2013 (einddatum bestand) met respectievelijk de status ‘niet aanwezig’ of ‘al overleden’. Personen die geboren worden of immigreren tussen 1995 en 2012 hebben vanaf oktober 1994 tot de dag voor geboorte of immigratie een record met de status ‘nog niet geboren’ of ‘niet aanwezig’. Een record wordt geïdentificeerd door de variabelen RINPERSOON en VOLGNR. Deze variabelen zijn uniek voor een record in dit microdatabestand.

Het bestand ‘131211 GBA\_Uniciteit-LMR 1995-20130116V1’ kan ook gebruikt worden om de overlijdensdatums (of datums van geboorte, immigratie en emigratie) van GBA-personen aan een analysebestand toe te voegen.

### ***131211 GBA\_Uniciteit\_pj-LMR 1995-2011V1***

Een record in dit microdatabestand geeft van één persoon weer welke jaren tussen 1 januari 1995 en 31 december 2011 deze aanwezig was in Nederland en welke kalenderjaren zijn/haar koppelsleutel (geslacht, geboortedatum en 4-cijferige postcode) uniek identificeerbaar was.

De status van een persoon die emigreert wordt over het jaar van emigratie vermeld als zijnde het ‘hele jaar of een deel van het jaar niet aanwezig’. Ook personen die immigreren worden over het jaar van immigratie vermeld met de status ‘gehele jaar of deel van het jaar niet aanwezig’. Er is immers niet zeker of er in de tijd na emigratie/voor immigratie ziekenhuisopnamen plaatsgevonden hebben. Indien een persoon voor zijn overlijden uniek koppelbaar is dan wordt deze persoon als zijnde het volledige jaar uniek vermeld in dit bestand omdat na het overlijden geen ziekenhuisopnamen kunnen hebben plaatsgevonden. Bij een geboorte wordt dezelfde regel aangehouden; als een baby na geboorte de rest van het jaar uniek koppelbaar is op de koppelsleutel, wordt deze over het hele jaar als uniek koppelbaar gezien.

Elke persoon die is ingeschreven in de GBA heeft maar 1 record in dit bestand. Dit bestand geeft een samenvatting van de status van de uniciteit per kalenderjaar. Een record wordt geïdentificeerd door de variabele RINPERSOON. Deze variabele is uniek voor een record in dit microdatabestand.

## **2.2 Inhoud uniciteitsbestanden**

Door de koppeling van de LMR met de GBA ([\*Koppeling van LMR- en GBA-gegevens: methode, resultaten en kwaliteitsonderzoek\*](#)) is het mogelijk om met de LMR analyses op persoonsniveau uit te voeren. Voorbeelden van statistieken op persoonsniveau zijn het aantal personen dat één of meer keer in een jaar in het ziekenhuis is opgenomen geweest, het aantal personen dat voor het eerst in 5 jaar in het ziekenhuis is opgenomen geweest of de sterfte na eerste ziekenhuisopname. Bij analyses op persoonsniveau dienen personen in de tijd gevolgd te worden op ziekenhuisopnamen. Omdat voor de koppeling van de LMR met de GBA geen volledig identificerende koppelsleutel aanwezig was, is het belangrijk te weten dat alle opnamen van een persoon die plaatsgevonden hebben tijdens de onderzoeksperiode uniek gekoppeld konden worden aan de GBA. Met behulp van de hier beschreven GBA\_Uniciteitsbestanden kunnen de uniek koppelbare personen worden

geselecteerd en gevolgd worden in de tijd. In de GBA\_Uniciteitsbestanden zijn de gegevens opgenomen van alle personen die op 1 oktober 1994 ingeschreven stonden in de GBA of na deze datum in de GBA zijn ingeschreven. Dit is een wijziging ten opzichte van de GBA\_Uniciteitsbestanden die eerder ter beschikking zijn gesteld. Hierin was de begindatum van het bestand 1 januari 1995. Door een gewijzigde verwerking van de gegevens van de Gemeentelijke Basisadministratie is de startdatum nu 1 oktober 1994.

De GBA bevat gegevens over alle in Nederland in het bevolkingsregister ingeschreven personen. Op basis van de veranderingen die optreden in de waarden van de GBA-LMR koppelsleutel (geslacht, geboortedatum en 4-cijferige postcode) in de GBA wordt bepaald of de uniciteit van de persoon verandert. Het gaat hierbij met name om een verandering in de postcode. De uniciteit van een persoon kan veranderen als hij verhuist of als een andere persoon van of naar zijn postcodegebied verhuist. Het kan echter ook voorkomen dat in beide situaties de uniciteit gelijk blijft doordat er bijvoorbeeld nog een derde persoon met dezelfde koppelsleutel in het gebied woont. Ook als een persoon emigreert, immigreert, geboren wordt of overlijdt, zal de uniciteit veranderen.

Voor beide hierboven beschreven GBA\_Uniciteitsbestanden geldt dat deze versie voor een klein deel van de personen een wijziging in de uniciteitsstatus bevat in vergelijking met eerdere versies van dit bestand. Dit komt niet alleen in het meest recent toegevoegde jaar (2011) voor maar ook in alle eerdere jaren vanaf 1995. Het gaat hierbij per jaar om tussen de 20.000-30.000 personen waar de uniciteitsstatus is gewijzigd. Dit betreft per jaar grofweg 0,1% van alle personen die opgenomen zijn in de bestanden. Dit wordt veroorzaakt door een gewijzigd verwerkingsproces bij het CBS van de gegevens van de Gemeentelijke Basisadministratie, waarvan de GBA\_Uniciteitsbestanden afgeleid worden. Analyses uitgevoerd met een eerdere versie van de GBA\_Uniciteitsbestanden zullen bij gebruik van dit bestand wellicht licht gewijzigde cijfers laten zien. Indien analyses zijn uitgevoerd met een eerdere versie van de GBA\_Uniciteitsbestanden wordt aanbevolen deze te blijven gebruiken en alleen voor cijfers over 2011 het meest recente GBA\_Uniciteitsbestand te gebruiken. Hiermee wordt voorkomen dat uitkomsten van eerder uitgevoerde analyses wijzigen. Bij het starten van nieuw onderzoek kan gebruik gemaakt worden van het meest recente bestand van de GBA\_Uniciteit.

### **2.3 Gebruik van de uniciteitsbestanden**

De uniciteitsbestanden zijn nodig bij analyses van de LMR op persoonsniveau, waarbij individuele personen gedurende een bepaalde tijdsperiode worden gevolgd op ziekenhuisopnamen. Om er zeker van te zijn dat vrijwel alle in de LMR geregistreerde opnamen (voor uitzonderingen, zie Missende en niet koppelende records) van een persoon die plaatsgevonden hebben in de onderzoeksperiode uniek gekoppeld konden worden aan deze persoon in de GBA, moeten alleen de GBA-personen geselecteerd worden die gedurende de onderzoeksperiode uniek koppelbaar waren met de LMR. Hiervoor kan gebruik gemaakt worden van een van de twee hierna beschreven uniciteitsbestanden. Het bestand '131211 GBA\_Uniciteit\_compleet-LMR 1995-20130116V1' kan gebruikt worden om de exacte follow-up periode te bepalen waarin een persoon uniek koppelbaar was. Met het bestand '131211 GBA\_Uniciteit\_pj-LMR 1995-

2011V1' kan op eenvoudige wijze de uniciteit van personen gedurende hele kalenderjaren worden afgeleid. Voor het volgen van een specifiek onderzoekscohort op ziekenhuisopnamen wordt vaak gebruik gemaakt van het eerstgenoemde bestand. Bij longitudinale analyses op de totale LMR wordt vaak het tweede bestand gebruikt. Maar in principe zijn beide bestanden bruikbaar voor de verschillende soorten analyses op persoonsniveau. Het eerste bestand is alleen meer gedetailleerd en als men overlijdensdatums (of jaar en maand van overlijden) nodig heeft voor het onderzoek, dan kunnen deze alleen met het eerste bestand worden afgeleid.

### ***Bepalen van begin- en einddatum follow-up periode***

Met het '131211 GBA\_Uniciteit\_compleet-LMR 1995-20130116V1' kan de exacte periode waarin een persoon uniek koppelbaar is worden afgeleid. In dit bestand wordt de status van de uniciteit weergegeven in de variabele STATUS. Met deze variabele kan de onderzoeker iedereen die aan het begin van de onderzoeksperiode uniek koppelbaar is in de onderzoeksgroep includeren tot aan het moment dat de persoon niet meer uniek identificeerbaar is (STATUS=2), overlijdt (STATUS=-2) of emigreert (STATUS=9). Als een persoon zijn uniciteit verliest kunnen immers eventuele ziekenhuisopnamen van deze persoon niet meer uniek gekoppeld worden aan de GBA. Na het overlijden vinden er geen opnamen meer plaats en bij emigratie zijn de ziekenhuisopnamen die plaats vinden in het buitenland niet bekend. Vanaf deze momenten worden deze personen uitgesloten van de analyses. Zo kan bijvoorbeeld bij survival-analyses een persoon meegenomen worden tot de datum van verlies uniciteit, overlijden of emigratie of retrospectief tot immigratie (STATUS=9) of geboorte (STATUS=-1). Niet meer uniek zijn wordt dan behandeld op dezelfde manier als 'lost-to-follow-up' in een cohort onderzoek.

Ook is het door een koppeling met dit GBA\_Uniciteitsbestand mogelijk de geboortedatum, overlijdensdatum, immigratiedatum en emigratiedatum aan een onderzoekscohort toe te voegen.

Om de periode te bepalen waarin een persoon uniek koppelbaar was dient gebruik gemaakt te worden van de begin- en eind geldigheid van het desbetreffende record (beschikbaar in de variabelen BEGGELY, BEGGELM, BEGGELD en EINDGELY, EINDGELM en EINDGELD).

In veel gevallen is het onnodig de uniciteit tot op de dag nauwkeurig te bepalen. Daarom werd in eerdere jaren de dag van de begin geldigheid (BEGGELD) en de eind geldigheid (EINDGELD) van een record niet voor alle onderzoeken ter beschikking gesteld. Omdat gebleken is dat hiermee een koppeling met het LMR bestand niet meer goed mogelijk is, zijn deze variabelen wel opgenomen in het bestand '131211 GBA\_Uniciteit\_compleet-LMR 1995-20130116V1'. Voor de volledigheid wordt hieronder nog wel beschreven hoe de juiste uniciteitsperiode bepaald kan worden als de dag van de begin geldigheid en eind geldigheid niet beschikbaar zijn.

Indien de dag van de begin geldigheid en de eind geldigheid niet ter beschikking is gesteld, moet er soms een maand opgeteld of afgetrokken worden van de begin- en/of eindmaand van de geldigheid van het record om op een juiste manier de periode in volledige maanden te bepalen waarin de betreffende status geldt. Dit kan gebeuren

volgens onderstaande richtlijnen (zie tabel 1). Bij het gebruik van deze richtlijnen dient het bestand gesorteerd te zijn op respectievelijk RINPERSOON en VOLG NR. De beginmaand en eindmaand van de geldigheid van een record wordt weergegeven in de variabele BEGGELM en EINDGELM.

Tabel 1. Richtlijnen voor het selecteren van begin- en eindmaand van een status als de variabelen BEGGELD en EINDGELD (de ‘dag’ in de datum) niet beschikbaar zijn.

<i>Status</i>	<i>Beginmaand</i>	<i>Eindmaand</i>
Status 1 na status -1 (uniek na geboorte)	BEGGELM	EINDGELM-1 maand
Status 1 voor status -2 (uniek voor overlijden)	BEGGELM+1 maand	EINDGELM
Status 1 overig (uniek overig)	BEGGELM+1 maand	EINDGELM-1 maand
Status 2 (niet uniek)	BEGGELM	EINDGELM
Status 9 (niet aanwezig; voor immigratie/ na emigratie)	BEGGELM	EINDGELM
Status -1 (nog niet geboren)	BEGGELM	EINDGELM-1 maand
Status -2 (al overleden)	BEGGELM+1 maand	EINDGELM

Om de periode te bepalen waarin de persoon uniek koppelbaar was, dient bij de maand genoemd in begin geldigheid van het record met status 1, één maand opgeteld te worden. Het kan namelijk voorkomen dat de eerste weken van een maand de persoon niet uniek koppelbaar is geweest dus deze periode moet niet meegenomen worden. Indien de eindmaand van een unieke periode bepaald wordt, dient om zeker te zijn dat alleen de unieke periode opgenomen wordt in het onderzoek één maand afgetrokken worden van de maand genoemd bij de eind geldigheid van het unieke record. Als status 1 volgt op status -1 dan hoeft er geen maand bij de begin geldigheid opgeteld te worden. Het is namelijk zeker dat een persoon voor zijn geboorte niet in het ziekenhuis opgenomen is geweest. Indien het record met status 1 gevolgd wordt door een record met status -2 dan hoeft er geen maand van de eind geldigheid af getrokken te worden. Het is ook zeker is dat de persoon na overlijden geen ziekenhuisopnamen meer gehad heeft. Om diezelfde redenen wordt er bij een record met status -1 van de eindmaand 1 maand afgetrokken en bij een record met status -2 bij de beginmaand een maand bij opgeteld.

In tabel 2 is een overzicht gegeven van de te gebruiken regels voor het afleiden van de uniciteitsperiode en voor het afleiden van de geboortedatum, overlijdensdatum, immigratiedatum en emigratiedatum. Daarin is zowel aangegeven hoe dit dient te gebeuren als de ‘dag’ van de begin- en eind geldigheid niet bekend is en wanneer dit wel het geval is. Ook bij het gebruik van de richtlijnen in onderstaande tabel dient het bestand gesorteerd te zijn op respectievelijk RINPERSOON en VOLG NR.

Tabel 2. Regels om uniciteitsperioden en geboorte- overlijdens-, immigratie- en emigratiedatums af te leiden uit ‘131211 GBA\_Uniciteit\_compleet-LMR 1995-20130116V1’.

	<i>Datum beschikbaar exclusief dag</i>	<i>Datum beschikbaar inclusief dag</i>
Begin uniciteitsperiode	Status 1 na status -1: begin geldigheid Status 1 overig: begin geldigheid + 1 maand	Status 1: begin geldigheid
Eind uniciteitsperiode	Status 1 voor status -2: eind geldigheid Status 1 overig: eind geldigheid – 1 maand	Status 1: eind geldigheid
Geboortedatum	Eerstvolgende record na record met status -1 (volnummer [record status –1] + 1): begin geldigheid	Eerstvolgende record na record met status -1 (volnummer [record status –1] + 1): begin geldigheid
Overlijdensdatum	Laatste record voor record met status -2 (volnummer [record status –2] -1): eind geldigheid	Laatste record voor record met status -2 (volnummer [record status –2] -1): eind geldigheid
Immigratiedatum	Eerstvolgende record na record met status 9 (volnummer [record status 9] + 1): begin geldigheid	Eerstvolgende record na record met status 9 (volnummer [record status 9] + 1): begin geldigheid
Emigratiedatum	Laatste record voor record met status 9 (volnummer [record status 9] -1): eind geldigheid	Laatste record voor record met status 9 (volnummer [record status 9] -1): eind geldigheid

### ***Selecteren van unieke populaties over hele kalenderjaren***

Het bestand ‘131211 GBA\_Uniciteit\_pj-LMR 1995-2011V1’ is afgeleid van het bestand ‘131211 GBA\_Uniciteit\_compleet-LMR 1995-20130116V1’ en bevat een samenvatting van de uniciteit over hele kalenderjaren. Met dit bestand kan op eenvoudige wijze een groep personen geselecteerd worden die uniek koppelbaar was gedurende een vaste onderzoeksperiode (in hele kalenderjaren), de zogenoemde ‘altijd unieke populatie’. Er hoeft dan geen gebruik gemaakt te worden van de exacte datum waarop een persoon zijn uniciteit verliest. Om bijvoorbeeld de unieke populatie in de jaren 1999-2001 te selecteren, dienen de personen met status 1 in de variabelen UNI99, UNI00 en UNI01 geselecteerd te worden. Door alleen de uniek koppelbare personen in het onderzoek mee te nemen is vrijwel zeker dat alle in de LMR geregistreerde ziekenhuisopnamen van de onderzoeksgroep teruggevonden worden.

### **2.4 Representativiteit van de unieke populatie**

De uniek koppelbare populatie wijkt in enkele kenmerken af van de gehele bevolking. De kans dat een opname uniek gekoppeld kan worden aan een persoon in de GBA is

namelijk niet voor alle bevolkingsgroepen hetzelfde en kan veranderen in de tijd. Zo hebben personen die in dichtbevolkte postcodegebieden wonen en personen afkomstig uit landen waar de geboortedatum niet altijd exact geregistreerd is een kleinere kans om uniek gekoppeld te worden. Ook kan de koppelkans veranderen bij verhuizingen. Als men de representativiteit van een geselecteerde uniek koppelbare populatie wil bepalen ten opzichte van de gehele bevolking dan kan men deze groep personen koppelen met het ter beschikking gestelde GBA-referentiebestand. In dit bestand zijn van de gehele GBA-bevolking achtergrondkenmerken opgenomen. Met deze gegevens kan men een bepaalde onderzoeksgroep wegen naar de samenstelling van de gehele bevolking. Zo heeft het CBS voor de eigen statistieken van de LMR op persoonsniveau ophoogfactoren ontwikkeld ([\*Ophogen op persoonsniveau van gegevens van de Landelijke Medische Registratie gekoppeld met de GBA\*](#)). Deze zijn echter alleen voor de desbetreffende statistieken ontwikkeld en zijn niet algemeen bruikbaar.

## **2.5 Missende en niet koppelende records**

Bij de toepassing van deze methode voor analyses op persoonsniveau waarbij alleen de uniek koppelbare personen geanalyseerd worden, wordt ervan uit gegaan dat de registratie van ziekenhuisopnamen volledig is en dat alle opnamerecords, tenminste voor zover deze GBA-ingezetenen betreffen, uniek of meervoudig koppelbaar zijn met de GBA. Dit is echter niet helemaal het geval. Als eerste zijn er de niet in de LMR geregistreerde ziekenhuisopnamen, dit is gemiddeld 1,2% over 1995-2005 stijgend tot 10,5% in 2006 en 18,0% in 2011. Daarbij komen de LMR-opnamen die niet gekoppeld konden worden aan de GBA (gemiddeld 3,0% over 1995-2011). Deze beide groepen ziekenhuisopnamen worden bij deze analyses genegeerd; terwijl een deel van deze opnamen wel tot de (unieke) GBA-populatie zullen behoren. Niet-koppeling van een LMR-opname aan de GBA kan optreden als gevolg van administratieve fouten/vertragingen in de koppelvariabelen (b.v. het niet of later doorgeven van een verhuizing bij de gemeente of het ziekenhuis) en bij in het ziekenhuis opgenomen personen die niet in de GBA zijn ingeschreven (b.v. ambassadepersoneel; NAVO-militairen, illegalen). Omdat voor de niet in de LMR geregistreerde opnamen (missende records) en de niet koppelende records op persoonsniveau moeilijk valt te corrigeren, betekent dit dat voor met name de statistieken die relateren aan de hele bevolking de berekende cijfers een onderschatting van de werkelijkheid zullen zijn. Deze onderschatting van het aantal personen met opnamen in de bevolking kan niet precies gekwantificeerd worden, maar bedraagt maximaal het totale percentage van missende en niet gekoppelde records (gemiddeld 4,5% over 1995-2005 oplopend tot 19,7% in 2011). Omdat het aantal missende en niet gekoppelde records enigszins varieert per jaar is in tabel 3 een overzicht gegeven van de percentages in de periode 1995-2011. Hieruit blijkt dat het percentage missende records de laatste jaren beduidend hoger is dan in vorige jaren.

Tabel 3. Percentage missende en niet gekoppelde records per jaar in de LMR

<i>Jaar</i>	<i>Missende records *</i>	<i>Niet koppelende records*</i>	<i>Totaal missende en niet gekoppelde records*</i>
	%	%	%
1995	1,5	4,1	5,6
1996	1,4	3,8	5,2
1997	1,0	3,7	4,6
1998	0,9	3,6	4,6
1999	0,6	3,5	4,1
2000	0,9	3,4	4,3
2001	0,9	3,3	4,1
2002	0,5	3,1	3,5
2003	0,6	2,9	3,5
2004	1,1	2,8	3,9
2005	3,3	2,7	6,0
2006	10,5	2,4	12,9
2007	12,0	2,3	14,3
2008	12,0	2,3	14,3
2009	12,7	1,9	14,6
2010	14,0	1,9	15,9
2011	18,0	1,7	19,7

\* *Percentage van het totaal aantal LMR-opnamen dat jaar. In de cijfers zijn de opnamen met een ontbrekend registratiejaar en instellingsnummer, opnamen van buitenlandse patiënten en opnamen die niet in de LMR geregistreerd hadden moeten worden, niet meegerekend.*

Deze afnemende dekking van de LMR vanaf 2005 kan een probleem geven voor persoonsniveau-analyses, met name als de onderzoekspopulatie woonachtig is in de regio's met ziekenhuizen die geen microrecords hebben geregistreerd. Daarom zijn in de bijlage 'Randtotalen LMR 2011' randtotalen van alle ziekenhuisopnamen opgenomen, uitgesplitst naar o.a. soort ziekenhuis, provincie, zorgtype en gemeentecode. Dit biedt een onderzoeker de mogelijkheid om de bruikbaarheid van de LMR-bestanden voor zijn of haar onderzoek te beoordelen (bijvoorbeeld de dekking per regio), en om eventueel correcties voor de nonrespons toe te passen in de analyses.

De onderschatting van het aantal personen met opnamen in de bevolking kan nog enigszins variëren naar bijvoorbeeld leeftijd en diagnose. Zo komen niet gekoppelde LMR-records vaker voor bij jongvolwassenen en zeer ouderen en bij psychische stoornissen en ongevalletsels. Voor deze diagnoses en leeftijdsgroepen kan de onderschatting van het aantal personen met opnamen dus iets groter zijn. Bij de ontbrekende records in de LMR horen ook de opnamen van één kankerkliniek. Voor diagnoses in de groep nieuwvormingen kan er daardoor ook sprake zijn van een iets grotere onderschatting in de cijfers. Omdat van het totaal aantal opnamen voor nieuwvormingen echter maar een relatief klein gedeelte in deze kankerkliniek plaatsvindt

(ca. 4%) en een deel van de in deze gespecialiseerde kliniek opgenomen personen al eerder in een ander ziekenhuis zal zijn opgenomen, zal deze onderschatting toch beperkt zijn.

## **2.6 Kwaliteit**

Deze microdatabestanden zijn gebaseerd op de GBA. De kwaliteit van de GBA is afhankelijk van de manier waarop de gemeenten hun gegevens bijhouden. In de GBA worden mutaties dagelijks bijgehouden. Het hier besproken bestand is gebaseerd op een afslag van het GBA-bestand op een peilmoment.

De GBA\_Uniciteitsbestanden zijn geen publicatiebestanden. Dit betekent dat door het CBS geen controles en correcties uitgevoerd zijn.



### 3. Bestandsopbouw en toelichting

#### 3.1 Bestandsopbouw van het microdatabestanden

Onderstaand volgt een lijst van alle variabelen in de microdatabestanden.

##### *131211 GBA\_Uniciteit -LMR 1995-20130116V1*

<i>Nr</i>	<i>Variabele &amp; label</i>	<i>Format</i>	<i>Lengte</i>	<i>Decimalen</i>
1	RINPERSOON Identificatienummer persoon	Tekst	9	0
2	VOLGNR volgnummer	Getal	2	0
3	BEGGELY jaar begin geldigheid record	Tekst	4	0
4	BEGGELM maand begin geldigheid record	Tekst	2	0
6	EINDGELY jaar eind geldigheid record	Tekst	4	0
7	EINDGELM maand eind geldigheid record	Tekst	2	0
9	STATUS status uniciteit	Tekst	2	0

##### *131211 GBA\_Uniciteit\_compleet-LMR 1995-20130116V1*

<i>Nr</i>	<i>Variabele &amp; label</i>	<i>Format</i>	<i>Lengte</i>	<i>Decimalen</i>
1	RINPERSOON Identificatienummer persoon	Tekst	9	0
2	VOLGNR volgnummer	Getal	2	0
3	BEGGELY jaar begin geldigheid record	Tekst	4	0
4	BEGGELM maand begin geldigheid record	Tekst	2	0
5	BEGGELD dag begin geldigheid record	Tekst	2	0
6	EINDGELY jaar eind geldigheid record	Tekst	4	0
7	EINDGELM maand eind geldigheid record	Tekst	2	0
8	EINDGELD dag eind geldigheid record	Tekst	2	0
9	STATUS status uniciteit	Tekst	2	0

**120620 GBA\_Uniciteit\_pj-LMR 1995-2010V1**

<i>Nr</i>	<i>Variabele &amp; label</i>	<i>Format</i>	<i>Lengte</i>	<i>Decimalen</i>
1	RINPERSOON Identificatienummer persoon	Tekst	9	0
2	UNI95 status uniciteit 1995	Getal	1	0
3	UNI96 status uniciteit 1996	Getal	1	0
4	UNI97 status uniciteit 1997	Getal	1	0
5	UNI98 status uniciteit 1998	Getal	1	0
6	UNI99 status uniciteit 1999	Getal	1	0
7	UNI00 status uniciteit 2000	Getal	1	0
8	UNI01 status uniciteit 2001	Getal	1	0
9	UNI02 status uniciteit 2002	Getal	1	0
10	UNI03 status uniciteit 2003	Getal	1	0
11	UNI04 status uniciteit 2004	Getal	1	0
12	UNI05 status uniciteit 2005	Getal	1	0
13	UNI06 status uniciteit 2006	Getal	1	0
14	UNI07 status uniciteit 2007	Getal	1	0
15	UNI08 status uniciteit 2008	Getal	1	0
16	UNI09 status uniciteit 2009	Getal	1	0
17	UNI10 status uniciteit 2010	Getal	1	0
18	UNI11 status uniciteit 2011	Getal	1	0

### **3.2 Toelichting op de variabelen**

#### ***RINPERSOON***

De GBA bevat een uniek persoonsidentificatienummer, het A-nummer. In deze microdatabestanden is dit nummer vervangen door een intern persoonsidentificatienummer, het betekenisloze nummer 'RINPERSOON'. Met behulp van dit RINPERSOON-nummer kunnen de bestanden worden gekoppeld aan andere bestanden bij het CBS welke ook voorzien zijn van een RINPERSOON-nummer.

#### ***VOLGNR***

Bij wijzigingen in de GBA-uniciteitsgegevens van een persoon wordt een nieuw record aangemaakt en het volgnummer met 1 opgehoogd.

#### ***BEGGELY***

Het jaar van de datum waarop de geldigheid van het record inging.

#### ***BEGGELM***

De maand van de datum waarop de geldigheid van het record inging.

#### ***BEGGELD***

De dag van de datum waarop de geldigheid van het record inging.

#### ***EINDGELY***

Het jaar van de datum waarop de geldigheid van het record eindigde

#### ***EINDGELM***

De maand van de datum waarop de geldigheid van het record eindigde.

#### ***EINDGELD***

De dag van de datum waarop de geldigheid van het record eindigde.

#### ***STATUS***

De waarde '9' geeft aan dat een persoon niet aanwezig is omdat het de periode voor immigratie of na emigratie betreft.

#### ***UNIJJ***

Deze variabele geeft de status van de uniciteit in een jaar (JJ) weer. Onder status 1 vallen ook de personen die in een jaar geboren worden, na geboorte uniek koppelbaar zijn en dan overlijden. Onder status 2 vallen ook de personen die na geboorte en/of voor overlijden in het jaar niet uniek koppelbaar zijn.

### 3.3 Voorkomende waarden van de variabelen

In deze paragraaf worden per bestand voor de categoriale variabelen alle mogelijke scores en hun betekenis opgesomd.

#### *131211 GBA\_Uniciteit -LMR 1995-20130116V1*

1	RINPERSOON	Identificatienummer persoon
2	VOLGNR	Volgnummer
3	BEGGELY	Jaar begin geldigheid record
4	BEGGELM	Maand begin geldigheid record
5	EINDGELY	Jaar eind geldigheid record
6	EINDGELM	Maand eind geldigheid record
7	STATUS	Status uniciteit
1		persoon is uniek identificeerbaar
2		persoon is niet uniek identificeerbaar
9		niet aanwezig
-1		nog niet geboren
-2		al overleden

#### *131211 GBA\_Uniciteit\_compleet-LMR 1995-20130116V1*

1	RINPERSOON	Identificatienummer persoon
2	VOLGNR	Volgnummer
3	BEGGELY	Jaar begin geldigheid record
4	BEGGELM	Maand begin geldigheid record
5	BEGGELD	Dag begin geldigheid record
6	EINDGELY	Jaar eind geldigheid record

**7 EINDGELM**

Maand eind geldigheid record

---

**8 EINDGELD**

Dag eind geldigheid record

---

**9 STATUS**

Status uniciteit

---

- |    |  |
|----|--|
| 1  | persoon is uniek identificeerbaar      |
| 2  | persoon is niet uniek identificeerbaar |
| 9  | niet aanwezig                          |
| -1 | nog niet geboren                       |
| -2 | al overleden                           |

***131211 GBA\_Uniciteit\_pj-LMR 1995-2011V1***

**1 RINPERSOON**

Identificatienummer persoon

---

**2 UNI95 t/m 17 UNI11**

Status uniciteit 1995 t/m 2011

---

- |   |   |
|---|---|
| 0 | hele jaar niet aanwezig: al overleden of niet geboren   |
| 1 | hele jaar uniek óf na geboorte/voor overlijden uniek    |
| 2 | heel of deel jr. aanwezig, niet hele tijd uniek         |
| 9 | heel of deel jr. niet aanw. anders dan door geb./overl. |