

Toets dataset geanonimiseerde data COVID-vaccinatie bestemd voor RIVM

5.1.2e | 1 februari 2021 | v 1.0

1 Inleiding

Personen die worden gevaccineerd, wordt toestemming gevraagd om de gepersonaliseerde vaccinatiegegevens te verstrekken aan het RIVM t.b.v. opname in CIMS. Teneinde RIVM in staat te stellen om onder meer de vaccinatiegraad te laten bepalen, wordt over de personen die wel gevaccineerd worden, maar geen toestemming geven voor dataverstrekking aan RIVM, een beperkte geanonimiseerde dataset verstrekt. Deze data komt niet in CIMS maar in een dataplatform van RIVM.

Over de in overleg met EPI vastgestelde dataset is twijfel of de herleidbaarheid naar personen wel is uitgesloten, zoals wordt vereist.

Met support van een deskundige op onderzoeksgebied is een analyse uitgevoerd op de dataset. Hierbij is gebruik gemaakt van de Handreiking anonimiseren¹.

2 Dataset geanonimiseerde data

De dataset betreft zogenoemde microdata, d.w.z. elk record bevat gegevens van een eenheid, zoals een persoon. De set bestaat uit de volgende gegevens:

Naam veld	Toelichting
Identificerend nummer voor de vaccinatierreeks gevaccineerde/priknummer	Uniek nummer zonder enig persoonsgegeven om 1 ^e en volgende prikken aan elkaar te kunnen relateren
Identificerend nummer toediening	Idem
Geboortjaar	
Geslacht	
Postcode eerste 2 cijfers gevaccineerde of Code voor Caribisch Nederland (AR, etc.)	
Selectie criterium	
Datum toediening	
HPK-code	
Batchnummer	
AGB code verantwoordelijke praktijk	

Ontleend aan Document DPV_210 v1.1

¹ Handreiking anonimiseren, Werkgroep anonimiseren Community van Data Experts Zorggegevens Informatieberaad Zorg, 27 mei 2019 <https://www.informatieberaadzorg.nl/binaries/informatieberaad-zorg/documenten/publicaties/2019/5/27/handreiking-anonimiseren-door-community-van-data-experts-zorggegevens-informatieberaad-zorg/Handreiking+anonimiseren+door+Community+van+Data+Experts+Zorggegevens+Informatieberaad+Zorg.pdf>

3 Handreiking anonimiseren

Hierboven genoemde Handreiking anonimiseren gaat in op het anonimiseren en aggregeren van data. Voor het kunnen publiceren van open data zijn anonimisering of aggregatie van data oplossingen, mits de herleidbaarheid naar personen is uitgesloten. De handreiking gaat niet in op pseudonimisering, want dit betreft nog steeds gepersonaliseerde data en valt daarmee onder de AVG.

De opgave zit in het vinden van de juiste balans tussen onthullingsrisico en informatieverlies.

4 Toets dataset met behulp van de Handreiking

Ook de Handreiking biedt vanzelfsprekend geen kant en klaar antwoord op de voorliggende vraag. Wel beschrijft de Handreiking stap 1 van anonimiseren, namelijk het verwijderen van direct te herleiden kenmerken, zoals naam, geboortedatum en adres. Kijkend naar de voorgestelde dataset vragen geboortjaar, geslacht en de eerste 2 cijfers van de postcode nadere analyse.

Naar mijn mening voldoet de nu gedefinieerde dataset niet aan het vereiste dat de herleidbaarheid naar personen is uitgesloten. In dun bevolkte gebieden zal uitsluiting van herleidbaarheid naar personen niet gegarandeerd kunnen worden, zeker niet omdat in deze microdata ook de AGB code meekomt. Dit kan een kleine instelling zijn in VV&T, VG of GGZ.

5 Voorstel aanpassing dataset

De combinatie van geboortjaar, geslacht en AGB code sluit in bepaalde gevallen de herleidbaarheid van de geanonimiseerde data niet uit en is daarmee niet toegestaan. De dataset vraagt om herziening, zoals bij handhaven van AGB-code het vervallen van geboortjaar en geslacht. De vraag of met de gegevens die uit de AGB code gehaald kunnen worden ook de twee eerste cijfers van de postcode toegevoegde waarde hebben.