



Rijksinstituut voor Volksgezondheid
en Milieu
Ministerie van Volksgezondheid,
Welzijn en Sport

A. van Leeuwenhoeklaan 9
3721 MA Bilthoven
Postbus 1
3720 BA Bilthoven
www.rivm.nl

KvK Utrecht 30276683

T 030 274 91 11
info@rivm.nl

Datum
22 februari 2021

Ons kenmerk
DPV_210

memo

Aanlevering geanonimiseerde gegevens Covid-19
vaccinaties aan het RIVM

1. Inleiding

1.1 Doel memo

In de memo *DPV_161 Koppelingen CIMS t.b.v. covidvaccinatiegegevens* worden de manieren beschreven voor aanlevering van vaccinatiegegevens t.b.v. het landelijke register voor covidvaccinaties (CIMS). Deze specificatie voldoet voor gevaccineerden die middels een 'OPT-IN' toestemming gegeven hebben dat hun persoonlijke gegevens aangeleverd mogen worden aan het RIVM.

Om de landelijke vaccinatiegraad te kunnen bepalen wenst het RIVM, aanvullend op de gepersonaliseerde data, een beperkte dataset over de vaccinaties van personen die geen toestemming hebben gegeven te ontvangen. Vanzelfsprekend zonder persoonsgegevens.

M.b.t. deze geanonimiseerde data zijn de volgende uitgangspunten gehanteerd:

- De aanlevering aan CIMS blijft ongewijzigd. Hier worden alleen gegevens aangeleverd van gevaccineerden die middels 'OPT-IN' aangegeven hebben dat de gegevens aan het RIVM verstrekt mogen worden.
- De geanonimiseerde vaccinatiegegevens worden op een vergelijkbare manier aangeleverd (b.v. via SFTP) in afzonderlijke bestanden. Deze gegevens komen niet in CIMS maar worden via een alternatieve route ontsloten.
- Indien de leverancier niet beschikt over een BSN of geverifieerd BSN dan mag geen levering aan CIMS plaatsvinden en kan er per definitie geen sprake zijn van 'OPT-IN'. De vaccinatiegegevens worden dan meegenomen in de geanonimiseerde data.
- De keuze voor wel/geen 'OPT-IN' geldt in principe voor de vaccinatiecampagne. Indien teruggekomen wordt op 'OPT-IN', dan kan bij RIVM een verzoek om verwijdering uit CIMS worden gedaan.

1.2 **Contact bij RIVM**

Contact bij RIVM: 5.1.2e@rivm.nl.

Datum

22 februari 2021

1.3 **1.4 Begripsbepaling**

Ons kenmerk

DPV_210

In dit document wordt gesproken over de aanlevering van gegevens met betrekking tot een vaccinatie. Voor eenduidigheid over wat er bedoeld wordt volgt hieronder een lijst met definities.

Vaccinatie	Het enkelvoudig aan een persoon toedienen van een vaccin.
Vaccin	Vloeistof die gebruikt wordt om iemand in te enten.
Inenten	(Bij iemand) met een naald een vaccin in het lichaam spuiten.
Vaccineren	Het proces om tot een bepaalde stand van immuniteit te komen.
Injecteren	Het inbrengen van een vloeistof in een lichaam middels een holle naald.
Toediening	Injecteren met de bedoeling om tot een vaccinatie te komen.

2. Aanleveren van gegevens

Datum
22 februari 2021

2.1 Algemeen

Ons kenmerk
DPV_210

Per koppelende instantie zijn de volgende zaken bekend:

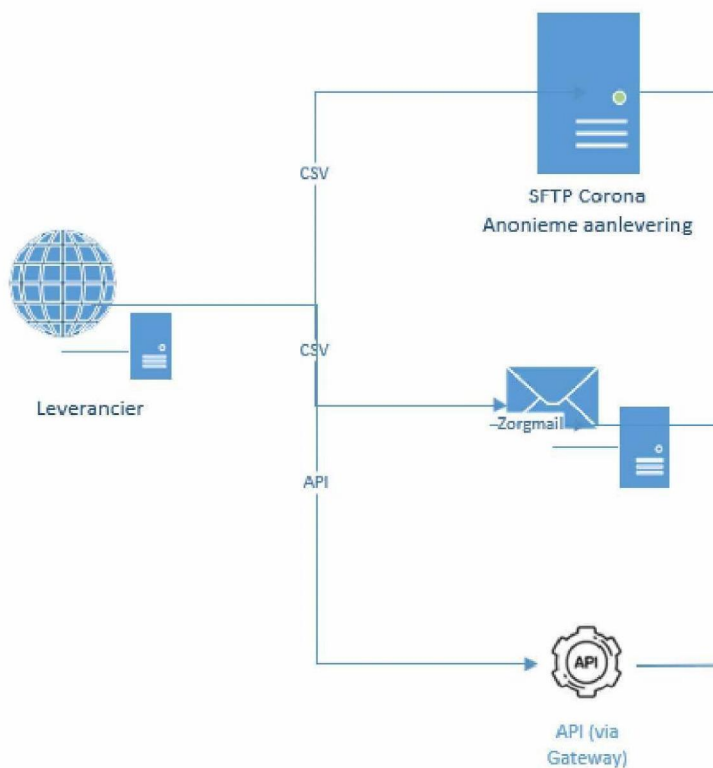
- Koppelende instantie
- Leverancierscode
- Adres
- Contactgegevens functioneel contactpersoon (telefoonnummer en e-mail adres)
- Contactgegevens technisch contactpersoon (telefoonnummer en e-mail adres)
- Koppelvlak (SFTP, Zorgmail, API)

Uitgangspunt: Bijbehorende tabel zal tijdig worden aangeleverd door de gegevensleveranciers.

In lijn met DPV_161 is verder gekeken hoe er een generieke aanlevering voor andere leveranciers gefaciliteerd kan worden. Hierbij wordt zo veel mogelijk aangesloten bij de methodieken zoals deze al voor de reguliere aanlevering gedefinieerd zijn in het document *DPV_161 Koppelingen CIMS t.b.v. covidvaccinatiegegevens*.

De volgende aanleverpaden worden onderscheiden (zie ook Figuur 1):

- Aanlevering als .CSV bestand via SFTP;
- Aanlevering als .CSV bestand via Zorgmail;
- [langere termijn]: aanlevering via een API.



Datum
22 februari 2021

Ons kenmerk
DPV_210

Figuur 1 generieke aanlevering anonieme vaccinatiegegevens

2.2 Aanlevering via SFTP als bestand

Het betreft hier automatische gegevensaanlevering. (De gegevensaanlevering vindt periodiek automatisch plaats zonder menselijk ingrijpen.)

2.2.1 Bericht

Het CSV bestand heeft de layout conform hoofdstuk 3.

2.2.2 Hoe

-Per bron (huisartsen, verpleegtehuizen, GGD's) worden hier aparte afspraken over gemaakt worden. Voorkeur heeft een gesloten netwerk.

2.2.3 Wanneer

- Aanlevering is op dagelijkse basis als er een toediening van een vaccin heeft plaats gevonden.

2.3 Aanlevering via Zorgmail met CSV als attachment.

Het vaccinatiebericht zal in deze variant via zorgmail als attachment in csv formaat worden verstuurd en vervolgens beschikbaar gemaakt worden.

Datum
22 februari 2021

Ons kenmerk
DPV_210

2.3.1 Bericht

Het CSV bestand heeft de layout conform hoofdstuk 3.

2.3.2 Hoe

- Via Zorgmail als attachment - Zorgmail Send Message (premium)

2.3.3 Wanneer

- De aanlevering is op dagelijkse basis als er een toediening van een vaccin heeft plaats gevonden.

2.4 Aanlevering via API (roadmap)

Het vaccinatiebericht zal in deze variant via een API worden aangeboden. Deze variant wordt in de toekomst ontwikkeld.

3. Bericht

Datum
22 februari 2021

Ons kenmerk
DPV_210

Het bericht dient ter aanlevering van gegeven toedieningen. Het bericht kan één of meerdere toedieningen bevatten van een of meerdere personen. Het bericht kan ook een bestand zijn. De generieke opbouw is 1 voorlooprecord en één of meerdere gegevensrecords.

Het bestand moet worden aangeleverd met als karakterset UTF-8. Het scheidingsteken wordt vastgesteld op de puntkomma (;). Tekstvelden worden omsloten door dubbel-quotes (""). Het laatste record van het bestand moet afgesloten worden met een linefeed (ASCII decimaal 10: LF).

Leverancierscode is de unieke identificatie van de gegevensleveranciers van het CSV-bestand en is gelijk aan de code die gebruikt wordt bij de aanlevering met persoonsgegevens zoals gespecificeerd in DPV_161 Koppelingen CIMS.

Dit kunnen b.v. de volgende soorten organisaties zijn (niet limitatief):

- HIS-leverancier
- Huisartsenpraktijk
- GGD-GHOR
- Ziekenhuis
- Verpleeghuis

De Naamgeving van het bestand is uniek volgens onderstaand format:

- AD – scheidingsteken - Leveranciers code – scheidingsteken - DATUM tijd (jjjjmmdduuss) – scheidingsteken - volgnummer (nnn)

Scheidingsteken: underscore (_)

Voorbeelden bestandsnaam:

Leverancierscode:
AD_H0007_20210104094844_001.csv

3.1 Voorlooprecord

Bij het aanleveren van een CSV-bericht/bestand hoort een voorlooprecord. Dit voorlooprecord kent de volgende structuur

- "Header" – apart label om onderscheid te maken met de vaccinatieregels
- *Scheidingsteken*
- "CIMS210-1.0" -Verwijzing naar deze bestandsspecificatie, met versienummer (huidige versie: 1) gevolgd door een punt en het releasenummer (huidig releasenummer: 0)
- *Scheidingsteken*
- Leverancierscode (5 posities, alfanumeriek, ANNNN)
- *Scheidingsteken*

- Datums van creatie van het bericht/bestand (formaat jjjmmdduumi)
- *Scheidingstekens*
- Aantal gegevensrecords in bericht/bestand

Datum
22 februari 2021

Ons kenmerk
DPV_210

Voorbeeld voorlooprecord: "Header";"CIMS210-1.0";
"H0007";202012040948;1

Momenteel is het versie en releasenummer: 1.0. Alleen in het geval dat een nieuwe versie van het bericht in gebruik wordt genomen, dan verandert dit nummer.

3.1 Gegevensrecord

Het bestand is voor meerdere gegevensleveranciers van toepassing en daarom generiek in opzet.

De vaccinatiegegevens van de personen die geen toestemming (dus geen 'OPT-IN') hebben gegeven om hun gegevens te delen met het RIVM worden aangeleverd. Ook vaccinatiegegevens van personen die wel toestemming hebben gegeven, maar waar een (geverifieerd) BSN ontbreekt, worden op deze wijze aangeleverd aan het RIVM.

Het bestand bevat per levering een incrementale dataset met anonieme vaccinatiegegevens van één leverancier. Onderstaande veldvolgorde is verplicht.

Naam veld	Verplicht (Registratie)	Type	Formaat (max)
Identificerende code voor de vaccinatierreeks gevaccineerde *	Ja	ALFANUMERIEK	1024
Geboortejaar**	Ja	NUMERIEK	Lengte 4
Herkomst: NED (=Nederland) CAS (=Curaçao, Aruba en Sint Maarten) BES (=Bonaire, Sint Eustatius en Saba)	Ja	ALFANUMERIEK	Lengte 3: AAA
Datum toediening	Ja	DATUM	DD-MM-JJJJ
HPK-code	Ja	ALFANUMERIEK	Lengte: 8
Batchnummer***	Ja	ALFANUMERIEK	Lengte: 50
Selectiecriteria****	Wens	Numeriek	1. Medisch 2. Beroeps 3. Leeftijd

* deze code zorgt er voor dat de vaccinaties van één persoon aan elkaar te relateren zijn, maar mag nooit te herleiden zijn tot een natuurlijke persoon. Gebruik b.v een 'Salted' hash die gebaseerd wordt op een technische sleutel (géén persoonsgegeven) aangevuld met een 'salt' welke niet gedeeld wordt. Deze code is per leverancier uniek en wordt alleen gebruikt voor de communicatie met het RIVM. Deze code wordt niet opgenomen in de database van het RIVM, zie ook de uitleg in hoofdstuk 4.

** Om de anonimiteit van de aanlevering te garanderen dient voor personen ouder dan 90 als geboortjaar het huidige jaar -90 opgegeven worden. Voor BES en CAS dienen de eerste drie cijfers van het geboortjaar opgegeven te worden.

Datum
22 februari 2021

Ons kenmerk
DPV_210

Als synoniem voor batchnummer wordt ook wel partijnummer, lotnummer of charge gebruikt

Vanwege monitoring en rapportage wil het RIVM weten op basis waarvan toediening van het vaccin uitgevoerd is. Vanwege redenen van privacy is deze indeling aangehouden.

Voorbeeld gegevensrecord:

"753692ec36adb4c794c973945e"; 1963;"NED";20-01-2021;"01234567";"BN12345678";2

4. Waarborg bij anonimiteit gekoppelde vaccinaties

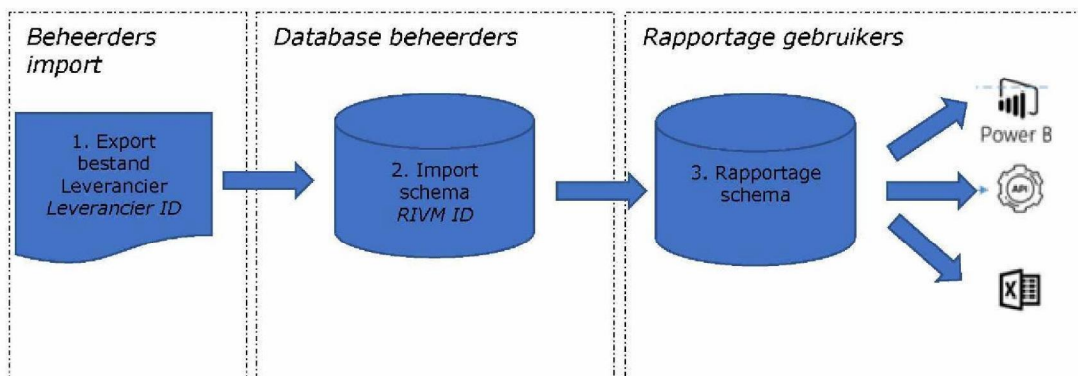
Datum
22 februari 2021

Ons kenmerk
DPV_210

Aangezien het voor de monitoring door het RIVM noodzakelijk is dat bekend is welke vaccinaties aan één persoon toegediend zijn is een technische sleutel toegevoegd aan de specificatie die niet aan een natuurlijk persoon te relateren is. Hierdoor is het mogelijk dat leveranciers de gegevens incrementeel aanleveren. Deze code is per leverancier uniek, is niet gebaseerd op een persoonsgegevens en wordt alleen gebruikt voor de communicatie met het RIVM.

Voor de verwerking van de bestanden is een methodiek bedacht waarbij het niet nodig is om deze technische sleutel van de leveranciers één op één op te slaan in de database van het RIVM. Hierdoor wordt het onmogelijk om aan de hand van de gegevens in deze database terug te gaan naar de persoonsgegevens in de systemen van de leverancier.

Een schematische weergave van de import en verdere verwerking van de data staat in Figuur 2.



Figuur 2 schematische weergave verwerking anonieme vaccinatiegegevens

De volgende stappen worden doorlopen tijdens de import:

1. De leverancier van de data levert het exportbestand aan via de voorgeschreven route. Dit bestand bevat per toegediende vaccinatie een technische sleutel (Leverancier ID) waarmee de vaccinaties van één persoon aan elkaar gerelateerd kunnen worden. Alleen de beheerders van de importroutine hebben toegang tot de map waar deze bestanden worden opgeslagen.
2. De data wordt verwerkt. Hierbij wordt de kwaliteit van de data gecontroleerd. De technische sleutel wordt vervangen door een intern RIVM ID en de anonieme vaccinatiegegevens worden opgeslagen in de database in een schema voor de anonieme

import. Omdat dit RIVM ID een andere code is dan aangeleverd is door de leverancier kan er vanuit de RIVM database nooit teruggegaan worden naar het systeem van de leverancier om met behulp van deze sleutel een persoon op te zoeken. Alleen de databasebeheerders hebben toegang tot dit schema.

Datum
22 februari 2021

Ons kenmerk
DPV_210

3. De gegevens worden verzameld in een rapportage schema waarbij de gegevens per gevaccineerde in één rij komen te staan: geboortjaar (>90 wordt als 90 weergegeven), selectie criterium, gegevens eerste toediening (datum, HPK-code, Batchnummer) en gegevens tweede toediening (datum, HPK-code, Batchnummer). Deze view bevat dus verder geen sleutels meer. De gebruikers van de rapportages hebben toegang tot dit schema.

De RIVM ID wordt bepaald door middel van een 'salted hash'. Een hash is een digitale vingerafdruk van een bestand of een stuk tekst en uniek voor het bestand of de tekst waarvan de hash berekend wordt: de kans dat er twee bestanden zijn die dezelfde hash code hebben is minder dan 1 op 4 miljard. Hierdoor kan je ervan uitgaan als je de hash code van twee bestanden of teksten vergelijkt en deze identiek zijn dit ook geldt voor de inhoud van het document.

Omdat de hash code niet de informatie bevat van het origineel, kan je vanuit een hash nooit het oorspronkelijke bestand of de oorspronkelijke tekst reconstrueren.

Je zou wél voor een reeks willekeurige teksten de hashcode kunnen berekenen en kunnen vergelijken met de waarden in de database: door een berekende hash te vergelijken met deze in de database zou je op die manier via een repeterend proces kunnen zoeken naar een match. Om dit te voorkomen wordt de oorspronkelijke technische sleutel van de leverancier uitgebreid met een 'salt' voordat de hash bepaald wordt. Dit is een stukje tekst dat alleen bekend is in de import. Zonder deze 'salt' is het onmogelijk een correcte RIVM ID te berekenen.

Nadat het RIVM ID gebruikt is om de vaccinaties van één persoon aan elkaar te relateren wordt deze niet meer gebruikt en is deze verder niet meer beschikbaar binnen de rapportageomgeving.

Voordelen van deze aanpak:

- De leveranciers ID is een code welke in het reguliere systeem van een leverancier niet te relateren is aan een persoon omdat deze alleen in de communicatie met het RIVM gebruikt wordt. Het wordt alleen gebruikt in het bestand dat aan het RIVM wordt aangeleverd.
- Het RIVM ID is de vingerafdruk van het leveranciers ID. Zowel in de overige systemen van het RIVM (b.v. CIMS) als de leveranciers is dit gegeven verder niet bekend en ook niet te relateren aan een natuurlijk persoon. We kunnen daarmee wel de vaccinaties van een persoon aan elkaar relateren, maar nooit terug naar het leveranciers ID.

- Het RIVM ID wordt alleen gebruikt bij de eerste stap van de import in de database. De tabel waarbij de gegevens van een persoon op één regel gecombineerd wordt bevat dit gegeven niet meer. Deze tabel wordt gebruikt voor de verdere verwerking. Hierdoor is zowel het Leveranciers ID als het RIVM ID niet beschikbaar op de plek waar de data daadwerkelijk wordt gebruikt.

Datum

22 februari 2021

Ons kenmerk

DPV_210