



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 5.1.2e
5.1.2e@rivm.nl

Datum
10 februari 2021

Ons kenmerk
DPV_229

Auteurs
5.1.2a 5.1.2g 5.1.2e
5.1.2b 5.1.2e
5.1.2a 5.1.2e

memo

DPV_229 Memo datakwaliteit maatregelen

1. Inleiding

1.1 Aanleiding

Vanaf de start van het vaccinatieprogramma in januari 2021 heeft CIMS de eerste dataleveringen ontvangen en worden in het kader van beheer van datakwaliteit analyses uitgevoerd op de ontvangen data. Juist nu, bij het begin van de dataleveringen, is het van belang deze analyses nauwgezet uit te voeren, want deze analyses leveren veel informatie op over aard en omvang van de fouten in de aanleveringen. Zo kunnen mogelijke kinderziektes in applicaties zichtbaar en opgelost worden en wordt duidelijk of processen en procedures aangescherpt moeten worden.

1.2 Scopebepaling beheer datakwaliteit

In het bestek van deze memo wordt onder beheer datakwaliteit verstaan:

- controleren of de data die vanuit de bronregistraties wordt aangeleverd overeenkomstig de hiervoor geldende specificaties zijn en verwerkbaar in CIMS;
- treffen van maatregelen indien hiervan wordt afgeweken.

Onderscheid kan worden gemaakt tussen *technische* en *inhoudelijke* fouten.

Technische fouten hebben betrekking op het formaat, de lengte of het ontbreken van verplichte elementen. In document DPV_214 Technische foutafhandeling CIMS bestandsverwerking van 1 februari 2021 zijn de maatregelen ter zake beschreven. In de voorliggende memo worden de technische fouten niet verder uitgewerkt en wordt verwezen naar dit document.

Inhoudelijke fouten hebben betrekking op de kwaliteit van de data en leiden in de regel tot uitval en verhinderen de registratie van de vaccinatie in CIMS. Over deze inhoudelijke fouten gaat de memo.

Voor de goede orde: de controle op de kwaliteit beperkt zich per definitie tot hetgeen ook daadwerkelijk gedetecteerd kan worden. Indien vaccinatiegegevens in het bronregister zijn gekoppeld aan de verkeerde persoon dan valt dit in CIMS niet vast te stellen.

Datum
10 februari 2021

Ons kenmerk
DPV_229

1.3 Doel en inhoud memo

Deze memo heeft ten doel duidelijk te maken wie verantwoordelijk is voor de kwaliteit van data die aangeleverd wordt aan CIMS en afspraken te maken over de maatregelen die bij afwijkingen in deze data worden getroffen.

Nadrukkelijk hoort hier ook de vraag bij welke begrenzing wordt aangebracht bij het interpreteren van aangeleverde data in een poging de data toch in CIMS te verwerken.

De memo gaat achtereenvolgens in op verantwoordelijkheden voor datakwaliteit en de maatregelen die genomen kunnen indien de aangeleverde data afwijkt en/of niet verwerkbaar is in CIMS. Het derde deel van de memo bevat de eerste bevindingen m.b.t. de kwaliteit van de aangeleverde data en de maatregelen die ingezet (kunnen) worden om de kwaliteit te verbeteren en zo nodig fouten bij de bron te laten corrigeren.

2. Verantwoordelijkheden m.b.t. datakwaliteit

Allereerst: de bron (vaccinatietoediener) is verantwoordelijk voor:

- juistheid en compleetheid van de registratie aan de bron;
- verwerking van informed consent van de geregistreerde;
- correcte, rechtmatige en tijdige aanlevering van de data aan RIVM;
- gepersonaliseerd data (t.b.v. CIMS, bij opt in) of geanonimiseerde data (t.b.v. RIVM, bij geen opt in).

Datavastlegging aan de bron en verstrekking van data aan RIVM is gespecificeerd in de RIVM documenten DPV_161, DPV_220 en DPV_210.

RIVM is verantwoordelijk voor het beheer van de data in CIMS. Beheer van datakwaliteit maakt hier onderdeel van uit.

Gegeven de verantwoordelijkheid die belegd is bij de bronregistratie mag RIVM geen wijzigingen aan de brondata uitvoeren. Wel kunnen bepaalde datavelden geïnterpreteerd worden, zonder dat dit inbreuk maakt op de aangeleverde data. Een voorbeeld: het weglaten van spaties in een batchnummer. Voorwaarde is wel dat de oorspronkelijke waarde beschikbaar blijft.

Ook dit is nieuw voor het beheer van de datakwaliteit in CIMS, dus zal hier de komende tijd meer ervaring mee opgedaan moeten worden.

3. Welke maatregelen worden ingezet om te kunnen sturen op datakwaliteit en wie heeft hierin welke rol?

Datum
10 februari 2021

Ons kenmerk
DPV_229

M.b.t. inhoudelijke fouten wordt onderscheid gemaakt tussen preventieve en correctieve maatregelen. Preventieve maatregelen beogen in structurele zin herhaling te voorkomen. Correctieve maatregelen hebben in dit verband ten doel foutief aangeleverde data via correctie alsnog verwerkbaar te maken in CIMS.

Zoals hiervoor verwoord kunnen zowel preventieve als correctieve maatregelen worden ingezet, afhankelijk van de aard van de afwijkingen/fouten. Dus een goede foutanalyse ligt aan de basis van de in te zetten maatregelen, waarbij ook wordt gekeken naar de verschillende systemen. Zo vragen fouten gerelateerd aan de BRBA applicatie een andere aanpak dan de XIS-systemen.

Voorbeelden van preventieve maatregelen zijn validatie van invoergegevens aan de bron en gebruik van tabellen (zoals toegestane waarden), werkinstructies en voorlichting.

Voorbeelden van correctieve maatregelen zijn het verwerken van correcties in de bronsystemen op aangeven van RIVM door middel van zogenoemde signaleringslijsten. Na verwerking kan het aldus gecorrigeerde datarecord opnieuw worden aangeboden aan CIMS. Ook zijn handmatige herstelacties in CIMS, bv. na raadpleging van de overheidsbasisregisters, op zich mogelijk. Dit heeft evenwel twee nadrukkelijke aandachtspunten: daarmee wordt de fout in het (vaccinatie) bronregister niet gecorrigeerd en een correctie door de beheerder van CIMS kan al snel conflicteren met de verantwoordelijkheidsverdeling.

4. Eerste bevindingen en maatregelen

4.1 Aard en omvang van geconstateerde afwijkingen

Aard:

De volgende tekortkomingen worden in de gegevensleveringen geconstateerd:

Attribuut	Gesignaleerde fouten
Datum vaccinatie	Datum ligt voor startdatum campagne
Datum vaccinatie	Datum ligt in de toekomst.
Geboortedatum	Datum ligt in de toekomst
Geboortedatum	Datum ontbreekt
AGB-code	AGB-code ontbreekt
AGB-code	AGB-code van persoon gebruikt i.p.v. verantwoordelijke instelling
AGB-code	Onjuist gevuld (00000000)
Naam gevaccineerde	Naam hoort bij andere BSN-geboortedatum combinatie volgens BRP

BSN	BSN ontbreekt
BSN	BSN onjuist (voldoet bijvoorbeeld niet aan de 11-proef)
BSN	Gevaccineerde die hoort bij aangeleverde BSN is al geruime tijd geleden overleden volgens basisregistratie overheid (bijvoorbeeld BRP)
BSN	BSN in combinatie met geboortedatum is onjuist
Batchnummer	Invoerfout (tal van onjuiste schrijfwijzen)
Huisnummer (toevoeging)	Te lange waarden aangeleverd

Datum
10 februari 2021

Ons kenmerk
DPV_229

Omvang:

Er zijn tot zover 64.029 vaccinatieregistratierecords verwerkt die zijn aangeleverd (exclusief GGD-GHOR en BRBA). Hierin werden de volgende fouten aangetroffen:

Fouten in gegevens aanleveringen EPD-EVS en HIS:	Aantal	Perc	Perc totaal
Datum vaccinatie	151	0,76%	0,24%
BSN niet gevuld of onbekend	1013	5,12%	1,58%
AGB-code niet bekend of verkeerd gevuld	16236	82,10%	25,36%
Geboortedatum niet bekend	12	0,06%	0,02%
Batchnummer niet gevuld of verkeerd	2	0,01%	0,00%
Fout in huisnummer toevoeging	2362	11,94%	3,69%
totaal fouten	19776	30,89%	0,00%
totaal records	64029		

* Dit betreft niet alleen fouten, maar ook AGB-codes die nog niet in CIMS aanwezig waren doordat het een ander type zorginstantie betrof.

Let op: Er kunnen meerdere fouten voorkomen binnen 1 aangeleverd record. Het aantal fouten kan hierdoor hoger zijn dan het aantal aangeleverde records. Dit geldt voor alle data leveranciers.

Fouten in gegevens die klaar staan bij GGD-GHOR	Aantal	Perc	Perc tot
Datum vaccinatie	?		
BSN niet gevuld of onbekend	734	0,31%	0,34%
AGB-code niet bekend of verkeerd gevuld*	15000	6,35%	6,99%
Geboortedatum niet bekend	?		
Batchnummer niet gevuld of verkeerd	5972	2,53%	2,78%
Huisnummer niet gevuld	214.574	90,81%	99,99%
totaal fouten	236.280	110,11%	
totaal records	214.585		

* Dit betreft niet alleen fouten, maar ook AGB-codes die nog niet in CIMS aanwezig waren doordat het een ander type zorginstantie betrof.

Ge signaleerde fouten zijn nog niet verwerkt in CIMS. Huidige situatie is al beter dan de datakwaliteit in de initieel gereed staande gegevens. Echter, wij hebben nog meer (nog niet gekwantificeerde) fouten gezien dan hierboven benoemd. Het gaat dan om zaken zoals:

- Vaccinnaam verkeerd gevuld (anders dan in DPV_161)
- Selectie criterium niet gevuld (is optioneel veld, maar toch...)

- c. Problemen met karakterset: weergave diakritische tekens (anders dan in DPV_161)
- d. Voorletters in Voorvoegsel Geslachtsnaam (technische fout).

Datum
10 februari 2021

Ons kenmerk
DPV_229

Fouten in gegevens BRBA aanleveringen:	Aantal	Perc	Perc tot
Datum vaccinatie	71	0,52%	0,28%
BSN niet gevuld of onbekend	331	2,44%	1,32%
AGB-code niet bekend of verkeerd gevuld*	12351	90,90%	49,18%
Geboortedatum niet bekend	4	0,03%	0,02%
Batchnummer niet gevuld of verkeerd	830	6,11%	3,30%
Fout in huisnummer toevoeging	0	0,00%	0,00%
totaal fouten	13587	54,10%	
totaal aangeleverde records	25116		

* Dit betreft niet alleen fouten, maar ook AGB-codes die nog niet in CIMS aanwezig waren doordat het een ander type zorginstantie betrof.

4.2 Inmiddels getroffen maatregelen

In samenwerking met een aantal ICT-leveranciers (bijvoorbeeld de BRBA applicatie) worden nu al signaleren van inhoudelijke tekortkomingen in datakwaliteit besproken en worden correctieve maatregelen toegepast in applicaties. Nu wordt door een aantal ICT-leveranciers een RIVM hulptabel gebruikt die zorgt dat wordt gesignaleerd of in een vaccinatie registratie de gegevens van in omloop zijnde vaccin (met diens HPK-code, batchnummer, startdatum en expiratedatum) zijn gevuld in gegevensleveranties die klaar staan voor verzending naar RIVM.

Ook worden wekelijks de gesignaleerde tekortkomingen in de datakwaliteit informeel besproken met de ontwikkelaars van BRBA, zodat zij preventieve maatregelen kunnen treffen in de bronapplicatie.

Deze maatregelen zijn nog niet formeel voorgeschreven. Hierdoor blijven de gesignaleerde tekortkomingen in datakwaliteit plaatsvinden.

Intern genomen maatregelen:

(bij geautomatiseerde verwerking in CIMS of door administratief medewerkers of beheerders van RIVM)

In een aantal gevallen komt in een vaccinatie registratie record **geen BSN voor of is het BSN onjuist**, maar bevat het record wel naam- en adresgegevens. CIMS doet op basis van de aangeleverde geboortedatum, naam- en adresgegevens een koppelvorschstel over welke persoon het zou kunnen betreffen. Een administratief medewerker beoordeelt dit voorstel en registreert de vaccinatie bij de juiste persoon. Ook in gevallen waar een niet voor de hand liggende geboortedatum van de BRP is ontvangen of ontbreekt worden gegevens geverifieerd door de administratief medewerkers en gekoppeld aan het clientdossier van de juiste persoon in CIMS.

In een aantal gevallen gaat het koppelen van een vaccinatie registratie in CIMS nog steeds mis doordat **nog geen persoonsgegevens zijn ontvangen van RNI, COA etc.** ook hier worden gegevens geverifieerd door de administratief medewerkers handmatig gekoppeld aan het cliëntdossier van de juiste persoon in CIMS.

Datum
10 februari 2021

Ons kenmerk
DPV_229

In een aantal gevallen zijn vaccinatie registratie records aangeleverd met een **vaccinatie datum vóór 2021 of na de systeemdatum of met een verkeerd formaat.** Door de administratief medewerkers is uitgezocht wat de echte uitvoerdatum is geweest en dit is/wordt na correctie in het cliëntdossier in CIMS geregistreerd.

In een aantal gevallen komt in een vaccinatie registratie record een **combinatie BSN en geboortedatum** voor die niet bekend is in CIMS of het **BSN behoort toe aan reeds overleden persoon.** Op basis van de signaleringsoverzichten gaan Administratief medewerkers in de GBA-V na om welke persoon het gaat en halen alsnog de cliëntgegevens van deze persoon vanuit BRP binnen. Vervolgens wordt een cliënt dossier voor deze persoon in CIMS aangemaakt en kan de vaccinatie geregistreerd worden.

Er zijn/worden enorm veel **onjuiste batchnummers** aangeleverd. Vaak waren dat schrijffouten en soms worden geheel ander gegevens gestuurd. CIMS verwijdert automatisch eventuele spaties in het batchnummer, transformeert kleine letters naar hoofdletters en registreert waar mogelijk (en dat evident is) het correcte batchnummer. In alle andere gevallen is de vaccinatie automatisch geregistreerd met 'batchnummer onbekend' en is het aangeleverde batchnummer in het memoveld geregistreerd. Het aantal niet geautomatiseerd te matchen batchnummer bedraagt nu nog 1689 (zie onderstaande tabel).

charge_in_cims	Values	
	Sum of aantal	Sum of aantal2
300042698	6419	10,03%
300042721	1	0,00%
charge onbekend	1689	2,64%
EJ6134	16	0,02%
EJ6795	31268	48,83%
EJ6796	7023	10,97%
EM0477	17613	27,51%
Grand Total	64029	100,00%

In een aantal gevallen wordt een **onjuiste AGB-code** (geautomatiseerd) gesignaleerd. Soms is de AGB-code ten onrechte onbekend en wordt dan alsnog handmatig door de Beheerder CIMS van RIVM opgevoerd en gekoppeld aan de vaccinatie registratie. Ook kan het gaan om de **persoonlijke AGB-code** van een arts in plaats van de AGB-code van de verantwoordelijke instelling of praktijk.

De AGB-code wordt handmatig door de administratief medewerker van RIVM opgezocht in het AGB-register van Vektis en (waar mogelijk) gecorrigeerd. Het resultaat is echter dat een vaccinatie die is uitgevoerd door bijvoorbeeld een ziekenhuis, nu nog geregistreerd kan staan op naam van een ingehuurde arts, waardoor het bij terugroepacties alsnog mis kan gaan.

Datum
10 februari 2021

Ons kenmerk
DPV_229

In een aantal gevallen worden **meerdere vaccinatie registraties records aangeleverd bij 1 combinaties BSN en geboortedatum** waar het niet voor de hand ligt dat dit een tweede vaccinatie betreft. Soms hebben deze vaccinatie registraties een volledig identieke inhoud. Als de gegevens al in het record zijn opgenomen, dan worden de volgende berichten genegeerd. Soms zijn er wel degelijk verschillen in de aangeleverde vaccinatie registraties records rond 1 combinatie van BSN en Geboortedatum. Zo kan de vaccinatiedatum of het batchnummer verschillen. In deze gevallen worden vaccinaties wél geregistreerd, waardoor een cliënt meerdere vaccinaties in het dossier krijgt. De vaccinatie registratie records op zichzelf zijn niet fout, maar er kan wel getwijfeld worden over de juistheid. Dit wordt wel zichtbaar gemaakt op een signaallijst. Er moet nog worden afgesproken hoe deze lijst wordt afgehandeld.

Extern genomen maatregelen:

Als gevolg van afstemming met de ontwikkelaars van de BRBA applicatie (waarin de tekortkomingen in datakwaliteit met hen is gedeeld) zijn de volgende maatregelen getroffen in de BRBA applicatie:

1. Gebruik RIVM hulptabel bij validatie van invoer van o.a. het toegestane batchnummer;
2. Gebruikers worden in staat gesteld (na signaleren binnen de applicatie) een ontbrekend BSN alsnog op te voeren, voordat de gegevens aan RIVM worden geleverd;
3. 11-proef controle is toegevoegd in de validatie van het invoerveld BSN;
4. AGB-code wordt geautomatiseerd ingevuld aan de hand van de accountgegevens van de gebruikersorganisatie (zodat hier geen vergissingen meer gemaakt kunnen worden);
5. Geautomatiseerde controle van de AGB-code (op bestaanbaarheid in AGB-register van Vektis) wordt m.i.v. de eerstvolgende release in gebruik genomen.

4.3 Nog te overwegen maatregelen

Correctieve maatregelen:

1. Verplichte toevoeging van validatie van invoer op de datumvelden: geboortedatum gevaccineerde en Vaccinatiedatum;
2. Verplicht gebruik van de RIVM hulptabel van in gebruik zijnde vaccins (met diens HPK-code, batchnummer, startdatum en expiratiedatum) in alle bronapplicaties. (nu doet alleen de BRBA applicatie dit);
3. Verplicht gebruik van met RIVM afgestemde AGB-codes van vaccintoedieners, waarbij het niet meer mogelijk is om persoonlijke AGB-codes te gebruiken, maar slechts de AGB-code van de instelling of praktijk onder wiens die verantwoordelijk toediening heeft plaatsgevonden;

4. Verplichte aanlevering van een geldige combinatie van vooraf gevalideerde BSN en geboortedatum, waarvan de BSN voldoet aan de zogenaamde 11-proef;
5. Toestaan van aanlevering en verwerking van te lange waarden in huisnummer (toevoeging) velden.

Datum
10 februari 2021

Ons kenmerk
DPV_229

4.4 Impact van datakwaliteit maatregelen bij BRBA

In onderstaande tabellen wordt zichtbaar hoe de datakwaliteit snel verbetert door het realiseren van relatief kleine aanpassingen in de bronapplicatie als een ICT-leverancier werkt aan de verbetering van de datakwaliteit van diens eigen applicatie / dienstverlening aan vaccin toedieners.

Verwerking tm 9 februari:

Fouten in gegevens BRBA aanleveringen:	Aantal	Perc	Perc tot
Datum vaccinatie	71	0,52%	0,28%
BSN niet gevuld of onbekend	331	2,44%	1,32%
AGB-code niet bekend of verkeerd gevuld	12351	90,90%	49,18%
Geboortedatum niet bekend	4	0,03%	0,02%
Batchnummer niet gevuld of verkeerd	830	6,11%	3,30%
Fout in huisnummer toevoeging	0	0,00%	0,00%
totaal fouten	13587	54,10%	
totaal aangeleverde records	25116		

Verwerking 10-2-2020:

Fouten in gegevens BRBA aanleveringen:	Aantal	Perc	Perc tot
Datum vaccinatie	5	0,72%	0,01%
BSN niet gevuld of onbekend	169	24,18%	0,40%
AGB-code niet bekend of verkeerd gevuld	202	28,90%	0,48%
Geboortedatum niet bekend	6	0,86%	0,01%
Batchnummer niet gevuld of verkeerd	2	0,29%	0,00%
Fout in huisnummer toevoeging	315	45,06%	0,74%
totaal fouten	699	1,65%	
totaal aangeleverde records	42345		