



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

5.1.2e 5.1.2e
5.1.2e 5.1.2e

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

memo

Inventarisatie en beoordeling van beschikbare dataverzamelingen voor het onderzoeken van de gezondheidsimpact van de coronapandemie

KvK Utrecht 30276683

T 030 274 91 11

5.1.5@rivm.nl

5.1.2e en 5.1.2e

Datum

31 januari 2022

Ons kenmerk

Uw kenmerk

Behandeld door

5.1.2e

Levensloop en Gezondheid

T 5.1.2e

5.1.2e

@rivm.nl

Kopie aan

Bijlage(n)

Bijlage 1 Overzicht dataverzamelingen_def

Inleiding

Het coronavirus en de getroffen overheidsmaatregelen om het virus te bedwingen hebben een breed spectrum aan directe en indirecte effecten op gezondheid en leefstijl. Om de impact van de coronapandemie op landelijke schaal in kaart te brengen, zijn betrouwbare landelijke cijfers/feiten nodig. Naar aanleiding van de coronapandemie zijn zowel landelijk als lokaal diverse dataverzamelingen opgezet om deze effecten te onderzoeken. Ook bestaan er lopende dataverzamelingen (monitors, cohorten) waarbij al data verzameld werden vóór de pandemie uitbrak en waarbinnen vervolgens extra data m.b.t. de pandemie zijn verzameld. Om zicht te krijgen op welke dataverzamelingen er allemaal zijn, is deze inventarisatie gedaan.

Het doel van deze inventarisatie is het in kaart brengen van dataverzamelingen waarmee voor heel Nederland de impact van de coronapandemie op gezondheid (zowel lichamelijk, mentaal als sociaal welzijn) en leefstijl (zoals roken, alcohol, bewegen) onderzocht kan worden. Hiermee kunnen we zicht krijgen op de effecten van de huidige coronapandemie op de leefstijl en gezondheid van de mensen in Nederland, waarmee we kennis en inzichten opdoen die ook relevant zijn voor toekomstige pandemieën.

De dataverzamelingen die wij geschikt achten zouden de volgende vraagstelling moeten kunnen beantwoorden: 'Wat is de directe en indirecte gezondheidsimpact'¹

¹ Onder 'directe impact' verstaan we de effecten van de COVID-19 infectie zelf. Onder 'indirecte impact' verstaan we de effecten van de getroffen overheidsmaatregelen of van de coronapandemie in het algemeen. Gezondheid is onderverdeeld naar algemene gezondheid, fysieke, mentale en sociale gezondheid. Leefstijl nemen we ook mee als belangrijke determinant voor gezondheid.

van de coronapandemie in Nederland, en welke subgroepen zijn het meest getroffen?'.

Datum

31 januari 2022

Methode

Ons kenmerk

Selectiecriteria dataverzamelingen

We hebben gezocht naar registratie- en monitordata die aan de volgende *selectiecriteria* voldoen:

- **Relevantie:** de data bieden inhoudelijk de mogelijkheid om de gezondheidsimpact van corona in kaart te brengen.
- **Periode:** hebben minimaal betrekking op het jaar 2020 en/of 2021 en bij voorkeur ook meerdere jaren daarvoor ter vergelijking.
- **Dekking:** de data zijn landelijk verzameld in Nederland, niet lokaal of regionaal.
- **Type:** kwantitatieve data², waaronder: 1) cross-sectionele data (er wordt iedere keer een nieuwe steekproef getrokken, personen worden *niet* gevolgd over tijd), 2) longitudinale data (dezelfde personen worden gevolgd over tijd) mits die in het kader van Covid-19 zijn verzameld³, en: 3) registratiedata.
- **Omvang:** de dataverzameling betreft minimaal 100 personen.
- **Herkomst:** de data zijn afkomstig van landelijke kennisinstututen en registratiehouders.

Eigenschappen dataverzamelingen

De gevonden dataverzamelingen en bijbehorende eigenschappen, zoals het aantal deelnemers, de doelgroep en de uitvraagfrequentie, staan in een Exceloverzicht (bijlage 1) beschreven. Er is daarbij in kaart gebracht of data verzameld werden op de volgende *gezondheidsonderwerpen*:

- Algemene gezondheid (bijv. ervaren gezondheid, kwaliteit van leven)
- Fysieke gezondheid (bijv. COVID klachten, ziekenhuis/IC opnames, beperkingen)
- Mentale gezondheid (bijv. angst/depressie, stress)
- Sociale gezondheid (bijv. eenzaamheid, sociaal netwerk)

² Langlopende cohorten, zoals de Doetinchem cohort studie, LASA of Lifelines, laten we buiten beschouwing, omdat cohorten in een ander onderdeel van deze opdracht worden geïnventariseerd.

³ Lifelines is naast een langlopend cohort een cohort dat een extra tak in het kader van COVID-19 heeft opgezet. Deze is in deze inventarisatie niet meegenomen, omdat de data niet landelijk dekkend zijn. Het is echter wel een waardevolle bron om de gezondheidsimpact van COVID-19 in te onderzoeken. Daarom wordt in andere onderdelen van deze opdracht een aantal vraagstellingen m.b.t. de gezondheidsimpact in Lifelines onderzocht.

- Leefstijl (bijv. alcohol, roken, bewegen)

Datum

De exacte definitie van fysieke, mentale, sociale en algemene gezondheid en leefstijl staan beschreven in Bijlage 1.

31 januari 2022

Daarnaast is beschreven wat voor informatie zij verzamelen ten aanzien van *corona*:

Ons kenmerk

- Infectie
- Overheidsmaatregelen
- Coronacrisis in het algemeen

Prioriteringscriteria

Vervolgens zijn de dataverzamelingen geprioriteerd aan de hand van de volgende *prioriteringscriteria*:

- Aantal indicatoren: zit er minimaal één indicator voor minimaal twee van onze onderwerpen (algemene, fysieke, mentale en sociale gezondheid en/of leefstijl) in de set? Dataverzamelingen met meer indicatoren krijgen een hogere prioriteit omdat we daarmee meer onderzoeksvragen tegelijk kunnen beantwoorden en dus tijdsefficiënter kennis kunnen opdoen.
- Doelgroep: betreft de doelgroep van de dataverzameling 0 of 18 jaar en ouder, én zijn qua achtergrondkenmerken tenminste leeftijd, geslacht en opleiding/inkomen (SES) beschikbaar? Zo krijgen we de algemene bevolking in beeld en kunnen we daarbinnen onderscheid maken naar diverse subgroepen.
- Steekproefgrootte: is de onderzoekspopulatie groter dan 2.400 personen? Voor deze afbakening is gebruik gemaakt van de formule voor het berekenen van de steekproefgrootte: als er over 17 miljoen mensen een uitspraak moet worden gedaan met een betrouwbaarheidsinterval van 95% en een foutmarge van 2% zijn ongeveer 2.400 personen nodig.
- Representativiteit: zijn de data landelijk verzameld en zijn alle groepen vertegenwoordigd qua leeftijd, geslacht en opleidingsniveau/SES met minimaal 100 personen?
- Koppelbaarheid: is de dataverzameling te koppelen aan andere data?

Wanneer een databron aan een criterium voldoet, is een 1 toegekend, indien deze niet voldoet aan het criterium een 0. Deze vijf scores zijn opgeteld tot een somscore van minimaal 0 tot maximaal 5 punten. De databronnen met een score van 4 of hoger worden in deze memo beschreven en vervolgens ten opzichte van elkaar op kwaliteit beoordeeld.

Resultaten

Van de in totaal 44 geïdentificeerde dataverzamelingen zijn er vijftien die een score van 4 of hoger hebben op de prioriteringscriteria. Deze dataverzamelingen zijn te onderscheiden in vragenlijstonderzoeken (N=9) en registratiedata (N=6).

Datum

31 januari 2022

Vragenlijstonderzoeken

Ons kenmerk

De eigenschappen van deze vragenlijstonderzoeken zijn samengevat in Tabel 1. Voor een uitgebreidere beschrijving, zie bijlage 1.

Tabel 1. Eigenschappen vragenlijstonderzoeken met een score van 4 of hoger.

Dataverzameling	Frequentie	N	Corona- onderwerpen	Gezondheids- onderwerpen	Score prioritering
Gezondheidsmonitor volwassenen en ouderen	Vierjaarlijks (extra in 2022 (onderdeel van de Integrale Gezondheidsmonitor COVID-19 van het Netwerk Gezondheidsonderzoek bij Rampen)	+/- 400.000	Infectie en coronacrisis	Algemeen, mentaal, fysiek, sociaal, leefstijl	5
CoronaGedragsUnit (CGU) vragenlijst-onderzoek	Zeswekelijks*	+/- 50.000	Infectie, coronacrisis en maatregelen	Algemeen, mentaal, fysiek, sociaal, leefstijl	4 (geen koppeling mogelijk)
CBS gezondheidsenquête/ Leefstijlmonitor	Jaarlijks**	+/-9.000	Infectie, maatregelen en coronacrisis	Algemeen, mentaal, fysiek, sociaal, leefstijl	5
Aanvullende module bewegen en ongevallen/ Leefstijlmonitor	Tweejaarlijks	+/-10.000	Infectie en maatregelen	Mentaal, fysiek, leefstijl	5
Aanvullende module middelen/ Leefstijlmonitor	Tweejaarlijks	+/-10.000		Algemeen, Mentaal, fysiek, leefstijl	5
Seksuele gezondheid in Nederland/ Leefstijlmonitor	Vierjaarlijks	+/-10.000		Algemeen, Mentaal, fysiek, leefstijl	5
Sociale Samenhang en Welzijn	Jaarlijks	+/-7.500		Algemeen, Mentaal, fysiek, sociaal	5
LISS panel	Maandelijks	+/-7.500	Coronacrisis	Algemeen, Sociaal, leefstijl	5
Nationaal Sportonderzoek (NSO)	Minimaal jaarlijks	+/- 2.600	maatregelen	Leefstijl	4 (te weinig gezondheidsonderwerpen)

**Driewekelijks eerste 5 maanden, daarna zeswekelijks; ** Vanwege de coronapandemie zijn er in 2020 en 2021 kwartaalcijfers gepubliceerd naast de jaarcijfers.*

Hieronder beoordelen we de geschiktheid van de in tabel 1 genoemde dataverzamelingen aan de hand van enkele beoordelingscriteria:

Datum

31 januari 2022

Ons kenmerk*Aantal respondenten*

In dataverzamelingen met een groot aantal respondenten kunnen beter verschillen tussen doelgroepen binnen de populatie onderzocht worden dan bij dataverzamelingen met kleinere aantallen. Op deze manier kunnen betrouwbare uitspraken worden gedaan over welke doelgroepen het hardst getroffen worden door de coronapandemie. Een hoog aantal personen in een onderzoek is zeker van groot belang bij een pandemie, vanwege de relatief lage prevalenties van een infectie.

Meest geschikte dataverzamelingen:

- Gezondheidsmonitor
- CGU vragenlijstonderzoek

Meerdere indicatoren op gezondheid en leefstijl

De beschikbaarheid van meerdere indicatoren op verschillende onderwerpen in een dataverzameling is belangrijk omdat daarmee inzicht verkregen kan worden in een breed scala van leefstijlfactoren en gezondheidsuitkomsten binnen één populatie. Dit geeft een vollediger beeld van de gezondheidsimpact dan wanneer er een enkele indicator op één onderwerp beschikbaar is.

Meest geschikte dataverzamelingen:

- Gezondheidsmonitor
- CGU vragenlijstonderzoek
- Leefstijlmonitor: Gezondheidsenquête
- SCP/CBS Leven

Vragen over corona

Bij een crisis zoals de coronapandemie is het van belang dat er relevante informatie wordt verzameld die te relateren is aan de crisis.

Meest geschikte dataverzamelingen:

- CGU vragenlijstonderzoek
- Gezondheidsmonitor
- Leefstijlmonitor: Gezondheidsenquête
- SCP/CBS Leven
- LISS panel
- Leefstijlmonitor: Aanvullende module 'Bewegen en Ongevallen'
- NSO

Frequentie

Een crisis, zoals de coronapandemie, vraagt om frequente cijfers over de impact van een pandemie. Dit is voor beleidsmakers van belang om onderbouwde keuzes te kunnen maken bij het instellen van overheidsmaatregelen.

Datum

31 januari 2022

Met welke frequentie data worden verzameld is echter afhankelijk van de doelstelling en het verloop van een pandemie. In tijden dat infecties snel oplopen, of veranderingen plaatsvinden is een dataverzameling met een hoge frequentie (bijvoorbeeld om de 3 tot 6 weken) van belang. Wanneer de doelstelling is om de gezondheidsimpact in kaart te brengen over een langere periode waarin de pandemie heeft plaatsgevonden volstaat wellicht een dataverzameling die jaarlijks wordt uitgevoerd.

Ons kenmerk

Meest geschikte dataverzamelingen:

- CGU vragenlijstonderzoek
- Leefstijlmonitor: Gezondheidsenquête (kwartaalcijfers)

Koppelbaar

Het koppelen van de dataverzamelingen aan registraties die bijvoorbeeld infecties of ziekenhuisopnames registreren, is van belang om een zo'n objectief mogelijk beeld te krijgen van de besmettingen en de gezondheidsimpact daarvan. Als er in een dataverzameling bijvoorbeeld geen vragen over de pandemie of het doorlopen van een infectie zijn opgenomen, kan door het koppelen met dergelijke registratiebestanden die informatie worden aangevuld. Vervolgens kan gekeken worden hoe het doormaken van een infectie of het opgenomen zijn in het ziekenhuis van invloed is op de zelf-gerapporteerde (fysieke, mentale en sociale) gezondheid. Het koppelen van registraties aan dataverzamelingen heeft daarmee een grote meerwaarde bij het in kaart brengen van de gezondheidsimpact van de coronapandemie.

Meest geschikte dataverzamelingen:

- Alle, behalve het CGU vragenlijstonderzoek.

Samenvattend

De vragenlijstonderzoeken die het vaakst worden beoordeeld als meest geschikt zijn het CGU vragenlijstonderzoek, de Gezondheidsenquête/Leefstijlmonitor en de Gezondheidsmonitor volwassenen en ouderen.

Registraties

De zes in kaart gebrachte registratiebestanden scoren op vier van de vijf prioriteringscriteria. De dataverzamelingen zijn koppelbaar, representatief, voldoende groot in omvang en hebben als doelgroep 0 of 18 jaar en ouder. Vanzelfsprekend gaan de dataverzamelingen maar over één onderwerp, waardoor

ze niet scoren op het criterium 'aantal indicatoren'. De registratiebestanden die relevant zijn voor onderzoek naar de gezondheidsimpact van corona zijn:

- NICE,
- LBZ,
- CoronIT,
- CBS doodsoorzaken,
- Nivel Zorgregistratie eerstelijns,
- Vaccinatieregistratie

Datum

31 januari 2022

Ons kenmerk

Dergelijke registraties kunnen natuurlijk niet alleen aan vragenlijstonderzoeken worden gekoppeld, maar ook aan elkaar. Door bijvoorbeeld vaccinatieregistraties en CoronIT aan de Nice data en/of CBS doodsoorzaken te koppelen kan bekeken worden hoeveel van de mensen die gevaccineerd zijn en/of corona kregen zijn opgenomen in het ziekenhuis en/of overleden aan corona. Deze dataverzamelingen zijn daarmee van grote waarde.

Beschouwing

Op basis van deze inventarisatie kan geconcludeerd worden dat van de vragenlijstonderzoeken de Gezondheidsmonitor, het CGU vragenlijstonderzoek en de Gezondheidsenquête het meest geschikt zijn om de gezondheidsimpact van de coronapandemie te onderzoeken. Deze drie dataverzamelingen bevatten een brede set aan vragen over algemene, fysieke, mentale en sociale gezondheid en leefstijlfactoren, en hebben voldoende hoge aantallen om uitspraken te doen over verschillende subgroepen in de Nederlandse bevolking.

Het CBS heeft al veel gepubliceerd over de gezondheidsimpact van de coronapandemie op basis van de Gezondheidsenquête⁴. Daarom zullen wij deze dataverzameling niet analyseren. Op basis van de CGU vragenlijstonderzoek en de Gezondheidsmonitor 2020 (eventueel gekoppeld aan CoronIT en NICE) zullen wij binnen ons project analyses doen die aanvullend zijn op wat er al bekend is en inzicht geven in de gezondheidsimpact van de coronapandemie op de algemene Nederlandse bevolking.

Voor de toekomst, waarin we mogelijk vaker te maken hebben met een pandemie, is het van belang dat er in bestaande vragenlijstonderzoeken makkelijk en snel vragen rondom een pandemie ingebouwd kunnen worden. Aanvullend is het van belang dat een dataverzameling een relatief hoge frequentie heeft om actuele cijfers te leveren aan beleidsmakers. Op dit moment lijkt het snel inbouwen van vragen over corona bij bestaande dataverzamelingen en het opleveren van cijfers met een hoge frequentie langzamer dan gewenst in een crisissituatie. Het in het kader van corona opgezette CGU vragenlijstonderzoek voldoet daar wel aan.

Daarnaast is het aan te raden dat data gekoppeld kunnen worden aan registraties waarin bijvoorbeeld besmettingen van de infectie staan beschreven. Het is daarbij dan wel belangrijk dat ook deze registratiedata snel beschikbaar zijn. De ervaring leert echter dat dit nog niet het geval is.

Datum

31 januari 2022

Ons kenmerk

Om de huidige data-infrastructuren beter te laten aansluiten bij de informatiebehoefte tijdens een crisis, is het aan te bevelen dat vragenlijstonderzoeken flexibele ruimte hebben om inhoudelijke aanpassingen te doen, een hoge frequentie hebben en/of de mogelijkheid hebben om frequent te publiceren over de data en dat data (zowel vragenlijstonderzoek als registratiedata) snel beschikbaar gemaakt kunnen worden voor onderzoek.