

Hergebruik van routine zorgdata voor beleid en wetenschap: hoe het nog beter kan



NIVEL
Kennis voor betere zorg



Hergebruik van routine zorgdata voor beleid en wetenschap: hoe het nog beter kan

Robert Verheij (1)
Corine Witvliet (2)
Tessa Jansen (1)
Mariette Hooiveld (1)
Onno van Hilten (2)

1: Nederlands Instituut voor Onderzoek van de Gezondheidszorg, Nivel
2. Centraal Bureau voor de Statistiek, CBS

ISBN 978-94-6122-559-7

030 272 97 00
nivel@nivel.nl
www.nivel.nl

© 2019 Nivel, Postbus 1568, 3500 BN UTRECHT

Gegevens uit deze uitgave mogen worden overgenomen onder vermelding van Nivel en de naam van de publicatie. Ook het gebruik van cijfers en/of tekst als toelichting of ondersteuning in artikelen, boeken en scripties is toegestaan, mits de bron duidelijk wordt vermeld.

Voorwoord

Deze rapportage is tot stand gekomen op verzoek van de data-expert community van het ministerie van VWS/RIVM. De auteurs danken de opdrachtgever voor het in hen gestelde vertrouwen.

De auteurs, juli 2019

Inhoud

Voorwoord	3
Samenvatting	5
1 Inleiding	7
2 Aanpak	9
2.1 Afbakening	9
2.2 Voorbeelden: patiëntreizen	9
3 Bevindingen	12
3.1 Databronnen die van belang zijn voor beide cases	12
3.2 Kwetsbare ouderen	14
3.3 Kinderen van ouders met psychische problemen (KOPP) of van verslaafde ouders (KVO)	16
4 Conclusie	18
4.1 Belangrijkste bevindingen	18
4.2 Praktische beperkingen	19
4.3 Data koppeling	19
4.4 Juridische aspecten	20
4.5 Versnipperd datalandschap	21
4.6 Data kwaliteit	22
4.7 Aanbevelingen	23
Literatuur	26

Samenvatting

Steeds meer gegevens worden gedurende het zorgproces vastgelegd, in de vorm van elektronische patiëntendossiers, declaraties en andere administratieve processen. Die gegevens kunnen worden gebruikt om inzicht te krijgen in hoe de gezondheidszorg functioneert. Zorggebruik en behoefte aan zorg kunnen in beeld worden gebracht. Deze gegevens vormen daarmee in potentie een schat aan informatie voor beleid en statistiek.

In dit rapport wordt de vraag gesteld in hoeverre die gegevens ook daadwerkelijk worden gebruikt voor ondersteuning van beleid. Op basis van openbaar beschikbare informatie beantwoorden we die vraag als het gaat om beleid gericht op kwetsbare ouderen en kinderen van ouders met psychische problemen. Met andere woorden: in hoeverre zijn er routinematig in de zorg vastgelegde gegevens over deze groepen beschikbaar en in hoeverre worden die al in beleid en statistiek gebruikt?

Veel van de benodigde gegevens worden wel vastgelegd, maar nog niet altijd gebruikt voor beleid en statistiek. Het ontbreken van centrale landelijke of landelijk representatieve dataverzamelingen lijkt daar in belangrijke mate debet aan te zijn. Veel gegevens worden alleen nog maar gebruikt voor het doel waarvoor ze in eerste instantie werden verzameld: administratieve processen en het zorgproces zelf. In gevallen waar die gegevens centraal bij elkaar worden gebracht, worden ze veel meer gebruikt voor onderzoek en statistiek. Verder blijken er maar weinig databronnen voor onderzoek beschikbaar zonder praktische (waaronder financiële) voorwaarden voor hergebruik.

Onze bevindingen komen in belangrijke mate overeen met een internationaal vergelijkend onderzoek van de OECD uit 2015. Daarin wordt gesteld dat er in veel landen (waaronder Nederland) nog maar weinig gebruik wordt gemaakt van beschikbare zorggegevens ten behoeve van onderzoek en statistiek. De OECD roept in dat rapport lidstaten op tot een 'strong health data governance framework' om het gebruik van zorgdata beter te faciliteren en om daarmee beleidsdoelen te ondersteunen. Er zijn volgens het rapport te veel praktische en juridische obstakels om data te hergebruiken voor onderzoek.

Aanbevelingen

Uit deze conclusies komen de volgende aanbevelingen voort:

1. Data governance verdient meer aandacht. Omarm de aanbeveling van de OECD om te zorgen voor een 'strong health data governance framework'. Begin desnoods op kleine schaal, met een 'coalition of the willing' en laat het nut zien van hergebruik van gegevens. Kijk of best practices uit andere OECD-landen overgenomen kunnen worden.
2. Onderken het belang van centrale databases met landelijke en/of landelijk representatieve registraties, zoals ook de OECD aanbeveelt en wacht niet op de totstandkoming van personal health trains en gedistribueerde gegevensinfrastructuren.
3. Zoek oplossingen voor de kosten van het onderhouden van een centrale database, zodat (her-)gebruikers geen onoverkomelijke financiële drempels ondervinden.
4. Binnen de wettelijke kaders lijkt er meer mogelijk met zorgdata dan veel wordt gedacht. Geef daarom meer bekendheid aan de ruimte voor hergebruik die de WGBO nu al biedt en onderzoek de ruimte die de AVG biedt om voor specifieke hergebruiksdoelen de huidige WGBO-regels te verruimen.

5. De waarde en bruikbaarheid van zorgdata voor ondersteuning van beleid wordt sterk beïnvloed door dat zelfde beleid. Wees alert op de impact die beleidsmaatregelen (zoals schrapsessies om de registratiedruk en de regeldruk te verminderen, of veranderingen in het bekostigingssysteem) hebben op de kwaliteit en beschikbaarheid van gegevens. Ontwikkel een systematiek om bij belangrijke beleidswijzigingen de impact ervan op bruikbaarheid en beschikbaarheid van data inzichtelijk te kunnen maken.
6. Hergebruik van gegevens is ook mogelijk zonder dat er al sprake is van 100% eenheid van taal. Vaak is er een gemeenschappelijke basis waarbinnen die eenheid van taal er al wel is. Hergebruik van zorggegevens kan dus worden doorontwikkeld zonder dat de ontwikkeling van eenheid van taal is afgerond.
7. Hergebruik van routine zorgdata vereist mogelijk dat het proportionaliteitsbeginsel uit verschillende perspectieven wordt belicht. Het vastleggen en/of verzamelen van gegevens kan vanuit één bepaald perspectief niet nodig zijn, maar vanuit een ander perspectief wel. Toepassing van het proportionaliteitsbeginsel vanuit één perspectief kan leiden tot beperking van een gegevensverzameling, die vanuit een ander perspectief heel belangrijk kan zijn. Aanbeveling is om te onderzoeken in hoeverre dit een probleem vormt en wat er eventueel aan te doen valt.

1 Inleiding

Steeds meer zorggegevens worden routinematig elektronisch vastgelegd. In toenemende mate worden die gegevens niet alleen gebruikt voor hun oorspronkelijke doel, maar ook als bron van informatie voor onderzoek, spiegelinformatie, beleid, en verantwoording/transparantie naar derden. Het kan daarbij gaan om gegevens uit elektronische patiëntendossiers die routinematig worden bijgehouden door zorgprofessionals, declaratiegegevens, of gegevens die patiënten zelf bijhouden. Hergebruik van gegevens is één van de speerpunten van het Informatieberaad van VWS. Het gaat daarbij om outcomedoel 4: Eenmalig vastleggen en hergebruik, dat als volgt is geformuleerd:

Vanaf 1 januari 2021 vormt het primair zorgproces de basis voor gegevensvastlegging in de zorg en worden vastgelegde gegevens hergebruikt voor declaratie, onderzoek, kwaliteitstransparantie en governance.

Target: We spreken af dat vanaf 1 januari 2019 bindende afspraken zijn gemaakt voor "eenheid van taal" in de zorg die deze doelstelling mogelijk maken. Daarmee is de weg vrij om met eenduidig gedefinieerde data te werken aan meer kwaliteit, meer veiligheid en efficiëntere zorg.

Hergebruik van routine zorgdata is belangrijk om een 'lerend zorgsysteem' tot stand te brengen. Het American Institute of Medicine definieert een lerend zorgsysteem als een "integrated health system in which progress in science, informatics, and care culture align to generate new knowledge as an ongoing natural by-product of the care experience, and seamlessly refine and deliver best practices for continuous improvement in health and health care" (The Institute of Medicine, 2011; Friedman et al. 2015).

Hergebruik van dergelijke gegevens staat de laatste jaren sterk in de belangstelling. Recent stuurde de minister van VWS (2018) een brief naar de Tweede Kamer onder de titel 'Data laten werken voor gezondheid', met daarin een pleidooi voor beter gebruik van zorggegevens om de zorg beter te maken. Data FAIRness wordt in die brief gezien als de sleutel. Gegevens dienen Findable, Accessible, Interoperable en Reusable te zijn. Dit rapport sluit daar bij aan door aan de hand van enkele voorbeelden een antwoord te geven op de vraag:

In hoeverre zijn routine zorgdata daadwerkelijk herbruikbaar voor beleidsinformatie, en hoe kan het nog beter?

Het antwoord op die vraag is erg van belang. Om ontwikkelingen in de zorg te kunnen volgen en te kunnen sturen, zijn immers gegevens nodig. Het beleid is er steeds meer op gericht verschuivingen tot stand te brengen tussen de ene vorm van zorg naar de andere vorm van zorg. Notities als "Juiste zorg op de juiste plek" en "Langer thuis" illustreren dat, waarbij meteen duidelijk is dat er behoefte is aan informatie die zich uitstrekt over verschillende zorgvoorzieningen.

Deze rapportage focust op hergebruik van gegevens ten behoeve van beleid, inclusief de daarvoor benodigde statistiek en beleidsonderzoek. De resultaten zijn echter ook relevant voor andere vormen van hergebruik, zoals medisch wetenschappelijk onderzoek, waarbij ook steeds meer gebruik wordt gemaakt van *real world data*, het type data waarover dit rapport gaat.

Hergebruik van zorggegevens ten behoeve van declaratieprocessen, zorginkoop, keuze-informatie voor zorgverzekeraars en patiënten blijft in dit rapport buiten beschouwing.

Om onze hoofdvraag te kunnen beantwoorden is het belangrijk te weten aan welke beleidsinformatie behoefte is. Immers, pas vanuit het doelperspectief kan iets worden gezegd over de geschiktheid van gegevens (Verheij et al. 2018). Hieronder een aantal beleidsvragen die als voorbeeld kunnen dienen:

- Hoeveel mensen maken gebruik van een bepaalde zorgvoorziening?
- Wat zijn de kenmerken van die mensen?
- Welke zorg wordt geleverd?
- Welke andere vormen van zorg hebben die mensen verder nog ontvangen? Anders gezegd: welke zorgpaden/patiëntreizen zijn er afgelegd?
- Hoe is het verder gegaan met die mensen na afloop van het zorgtraject?
- In hoeverre past het zorgaanbod bij de behoefte aan zorg in een regio?
- Welke verschuivingen vinden er plaats van de ene vorm van zorgverlening naar de andere?

Dergelijke vragen kunnen nog maar zelden worden beantwoord op basis van één gegevensbron (Flinterman et al. 2018), ook al omdat bij de zorg voor een patiënt steeds vaker meer verschillende zorgverleners zijn betrokken. Bovendien vereisen de beleidsvragen vaak dat er enerzijds een landelijk beeld wordt gegeven en dat anderzijds ook ingezoomd kan worden op specifieke deelgroepen en regio's. Dit alles maakt dat hergebruik voor beleidsinformatie een 'veeleisende' vorm van hergebruik is.

Het is dus nodig om gegevens te kunnen koppelen, en wel op het niveau van al dan niet geanonimiseerde of gepseudonimiseerde gegevens van individuele patiënten en om die patiënten in de tijd te kunnen volgen¹. De mate waarin gegevens onderling koppelbaar zijn, is daarom een belangrijk aandachtspunt. Alleen dan is het immers mogelijk om zorgpaden van patiënten en burgers te volgen. Het doel is weliswaar om ten behoeve van beleid informatie te generen op bijvoorbeeld een landelijk of regionaal niveau, maar daar zijn op individueel niveau gegevens voor nodig die onderling koppelbaar zijn.

¹ De OECD verwoordt dit als volgt: "Essential to health care quality and performance assessment is the ability to track patients as they progress back and forth through the health care system from primary health care, to specialty care, hospital care, long-term care, home care and hospice care. This data should also provide information about underlying patient characteristics...." (OECD, 2015)

2 Aanpak

2.1 Afbakening

We beperken ons in dit rapport tot zorggegevens die in, door en ten behoeve van het zorgproces worden vastgelegd. Dat kan gaan over zorgdata zoals vastgelegd in elektronische patiëntendossiers en over declaratiegegevens.

Gegevens in de vorm van PROMS en PREMS (patiënt reported outcome/evaluation measures) worden in toenemende mate ook routinematig vastgelegd, bijvoorbeeld voor verantwoordingsinformatie richting zorgverzekeraars. Hoewel belangrijk om de effecten van zorg te kunnen meten, blijft dit soort gegevens in dit rapport buiten beschouwing, deels omdat routinematige verzameling nog niet overal 'standaard' is en deels omdat dit een veel grotere studie vereist dan waar hier sprake is. We verwachten dat een deel van de gegevens die worden vastgelegd in persoonlijke gezondheidsomgevingen (PGO's) in de nabije toekomst ook onder routine zorgdata komen te vallen, maar op dit moment is dat nog zelden of nooit het geval. Om die reden blijven ook PGO's buiten het kader van dit rapport.

Een andere inperking is dat deze rapportage gaat over gegevens en hun beschikbaarheid op microniveau. Dat wil zeggen op het niveau van individuen of zelfs behandelingen. Deze inperking komt voort uit de achtergrond van deze studie, namelijk de noodzaak om zorgpaden van individuen te kunnen verwerken tot informatie die voor beleid relevant is.

2.2 Voorbeelden: patiëntreizen

Binnen het bestek van dit onderzoek is het niet mogelijk om de onderzoeksvraag over de gehele breedte van de zorg en voor alle mogelijke vormen van beleidsinformatie te beantwoorden. We richten ons daarom op twee voorbeelden.

Het gaat om:

- a. Kwetsbare ouderen;
- b. Kinderen van ouders met psychische problemen (KOPP) of van verslaafde ouders (KVO)

Deze voorbeelden zijn gekozen omdat ze de complexiteit van hergebruik van zorgdata goed laten zien. Het zijn voorbeelden die beleidsmatig actueel zijn en waaronder omvangrijke en complexe zorgpaden schuilgaan. Gegevens over al die verschillende zorgpaden samen, vormen de basis van de informatie die nodig is voor het huidige beleid. Bij ieder zorgpad is sprake van een keten van gegevens, en als het gaat om hergebruik is die keten zo sterk als de zwakste schakel. Het ontbreken van zo'n schakel betekent dat zorgpaden niet meer volledig in beeld kunnen worden gebracht. Of dat erg is, is afhankelijk van de vraag die we met die gegevens willen beantwoorden.

Met deze in verschillende opzichten ver uit elkaar liggende voorbeelden dekken we een belangrijk deel van het zorgaanbod in Nederland. Daarbij is gekeken naar de zorgvoorzieningen die staan opgesomd in tabel 1. De volgorde is min of meer willekeurig. Sommige zorgvoorzieningen zijn relevant voor beide groepen, andere voor één daarvan.

Voor beide groepen is ook de sociaaleconomische situatie een belangrijk aandachtspunt en deze wordt dan ook meegenomen in de beschrijving van de resultaten.

Het voert te ver om voor beide groepen alle mogelijke patiëntreizen te beschrijven, maar wellicht maken onderstaande voorbeelden het wat duidelijker:

*Bij **KOPP/KVO-kinderen** zijn gegevens nodig van de huisarts, waar eventuele problemen als eerste kunnen worden gesignaleerd en geregistreerd. Als de problematiek te ingewikkeld is, zal de huisarts of de praktijkondersteuner de patiënt doorverwijzen naar de gespecialiseerde GGZ. Ook via jeugdzorginstanties en de jeugdgezondheidszorg kunnen kinderen in beeld komen bij hulpverleners. Preventie van KOPP/KVO-problematiek kan ingezet worden bij kinderen van wie de ouders contact hebben met hulpverleners. Dus zijn daar ook gegevens over nodig.*

*Bij **ouderen** kan het bijvoorbeeld gaan om dementie. Ook daarvan worden de eerste symptomen bij de huisarts gemeld. Op een gegeven moment gaat het niet langer meer in de thuisomgeving en volgt opname in een verpleeghuis. De tijd die daartussen verstrijkt is beleidsmatig van groot belang. De mate waarin het thuis nog lukt kan afhangen van de sociaaleconomische omstandigheden waarin mensen zich bevinden, dus die informatie dient ook gegenereerd te worden.*

Tabel 1: Relevante databronnen/zorgvoorzieningen voor de beide cases

	Kinderen van ouders met psychische problemen (KOPP/KVO)	Kwetsbare ouderen
Declaraties bij zorgverzekeraars	X	X
Huisartsenzorg	X	X
Huisartsenposten	X	X
Ziekenhuiszorg	X	X
Sociaal economische kenmerken	X	X
Ambulancezorg		X
Integraal bekostigde eerstelijnszorg		X
Laboratoriumuitslagen		X
Indicatiestelling langdurige zorg		X
Pathologielaboratoria		
Langdurige zorg registratie		X
Maatschappelijke ondersteuning		X
Paramedische zorg (fysio- oefentherapie, dietetiek)		X
Wijkverpleging		X
Verslavingszorg	X	
GGD, preventie	X	
Gespecialiseerde GGZ, verslavingszorg (DBC's GGZ)	X	X
Jeugdetentie	X	
Jeugdzorg	X	
Jeugdbescherming	X	

Voor elk van de zorgvoorzieningen wordt (voor zover bekend) beschreven:

- a. Wat wordt er vastgelegd in het zorgproces?
- b. Is er sprake van hergebruik, en zo ja, welke voorwaarden worden daaraan gesteld?
- c. Is koppeling op persoonsniveau met andere bronnen mogelijk?
- d. In hoeverre is er sprake van eenheid van taal?
- e. In hoeverre zijn de data vindbaar en toegankelijk?
- f. Wat is de kwaliteit van de gegevens?

De antwoorden op deze vragen geven een indruk van de herbruikbaarheid van gegevens die routinematig worden vastgelegd bij elk van de genoemde zorgverleners. De beschrijvingen van de verschillende databronnen samen, geven antwoord op de centrale vraag: in hoeverre kunnen zorggegevens worden hergebruikt voor beleidsinformatie.

We beantwoorden deze vragen op basis van beschikbare kennis bij de auteurs van het Nivel en het CBS en op basis van informatie die men kan vinden via openbare bronnen zoals Zorggegevens.nl en op de websites van registratiehouders.

3 Bevindingen

Elk van de bekeken zorgvoorzieningen wordt in deze paragraaf kort beschreven en de belangrijkste knelpunten worden weergegeven. Hierbij is geen volledigheid nagestreefd. Dat zou een veel omvangrijkere en diepgravender studie vereisen.

We gaan eerst in op de gegevens die niet specifiek betrekking hebben op de zorg voor één van beide groepen. Daarna bespreken we de databronnen voor zorg voor kwetsbare ouderen en tenslotte die voor kinderen van ouders met psychische problemen of verslaafde ouders.

3.1 Databronnen die van belang zijn voor beide cases

In deze paragraaf bespreken we een aantal meer algemene gegevensbronnen van belang voor zowel KOPP/KVO kinderen als voor kwetsbare ouderen.

3.1.1 Declaratie bij zorgverzekeraars

Alle zorg die wordt gedeclareerd onder de zorgverzekeringswet wordt geregistreerd bij zorgverzekeraars. Vektis is de organisatie die deze gegevens bij elkaar brengt in een landelijke database en gegevens beschikbaar stelt aan derden. Dit is een belangrijke landelijk dekkende databron, maar de inhoud is erg afhankelijk van het bekostigingssysteem. Zodra een bepaalde vorm van zorg niet meer vergoed wordt vanuit de zorgverzekeringswet, verdwijnt de betreffende informatie uit de bestanden. Zorg die mensen zelf betalen is niet zichtbaar in gegevens van Vektis en dat kan per individu verschillen, afhankelijk van de verzekeringspolis. Voor het beschikbaar stellen van gegevens brengt Vektis kosten in rekening en er is toestemming vereist van de verzekeraars. Gegevenskoppelingen zijn in praktische zin goed mogelijk en ze worden ook regelmatig voor onderzoek gebruikt.

3.1.2 Huisartsenzorg

Bijna alle huisartsen gebruiken een huisartsinformatiesysteem. Die systemen voldoen allemaal aan de eisen van een goed beheerd zorgsysteem. Diagnoses, medicatie en diagnostische tests worden over het algemeen op een uniforme manier vastgelegd, gebruik makend van codeersystemen als ICPC, LOINC, ATC² die ook bij andere disciplines en in andere landen worden gebruikt.

In de jaren 2012 en 2013 is de kwaliteit van de registratie met succes gestimuleerd met een financiële prikkel voor huisartsenpraktijken (van der Bij, 2013). Niettemin zijn er nog steeds verschillen tussen praktijken in de manier waarop zij gegevens vastleggen. Dergelijke verschillen lijken er ook nog te zijn tussen softwarepakketten. Daarvan zijn er in Nederland een achttal, met elk een eigen structuur Van der Bij, 2013.

² ICPC: international Classification of Primary Care een klassificatiesysteem waarmee symptomen en diagnoses kunnen worden geregistreerd en dat speciaal ontwikkeld is voor de huisartsenzorg; LOINC: Logical Observation Identifiers Names and Codes is een standaard voor het vastleggen van laboratoriumuitslagen. De in Nederland gebruikte codes van het NHG zijn op LOINC af te beelden; ATC Anatomical Therapeutical Classification, voor het vastleggen van de werkzame stof in geneesmiddelen.

Het systeem van inschrijving op naam zorgt er voor dat de gegevens niet alleen betrekking hebben op personen die in een bepaalde periode zorg gebruikten, maar ook op personen die dat in het verleden hebben gedaan of die dat in de toekomst nog gaan doen. Dat zorgt er voor dat gegevens van huisartsen kunnen worden gebruikt voor zogenaamde populatiestudies, studies waarbij hele populaties worden vergeleken. Een ander belangrijk voordeel is dat, aangezien er bij de huisarts burgerservicenummers (BSN) geregistreerd worden, gegevens goed te koppelen zijn met andere gegevens. Het BSN kan dan door een Trusted Third Party (TTP) gebruikt worden als basis voor pseudonimisering.

Een centrale landelijk database met huisartsgegevens is er niet. In plaats daarvan is er een diversiteit aan instanties die deze gegevens verzamelen voor onderzoek. Vakgroepen huisartsgeneeskunde van universiteiten en het Nivel onderhouden eigen dataverzamelingen. Er is een lange traditie van hergebruik van deze gegevens voor beleid. Een belangrijke leverancier van beleidsinformatie is Nivel Zorgregistraties eerste lijn, waarvan de gegevens onder meer worden gebruikt voor de Staat van de Volksgezondheid en de Volksgezondheid toekomstverkenningen.

Er komen steeds meer leveranciers van huisartsinformatiesystemen die zelf gegevens centraal opslaan, en partijen (bijvoorbeeld Pharmo, Proigia) die deze gegevens verzamelen ten behoeve van dienstverlening aan huisartsen en onderzoek. Waar ze ook bij elkaar gebracht worden, de huisartsgegevens worden veelvuldig voor onderzoek gebruikt.

Als kinderen en hun ouders deel uit maken van hetzelfde gezin, staan beiden vaak bij dezelfde huisarts ingeschreven. Dat maakt het mogelijk om bij de KOPP/KVO kinderen de huisartsgegevens van zowel ouders als kinderen te betrekken. Bij ouderen maakt dit het mogelijk om ook zorggegevens van de eventuele mantelzorger te analyseren.

3.1.3 Huisartsenposten

Huisartsenposten voorzien in huisartsenzorg buiten kantooruren. Zij gebruiken informatiesystemen waarin onder andere ingangsklachten worden vastgelegd, de diagnose, de urgentie en de verleende zorg. Het Nivel verzamelt gegevens die geregistreerd worden op de huisartsenpost in Nivel zorgregistraties eerste lijn. Deze zijn koppelbaar via een Trusted Third Party aan andere databronnen en opvraagbaar bij het Nivel (www.nivel.nl/zorgregistraties). Ongeveer tweederde van alle huisartsenposten in Nederland neemt deel aan deze gegevensverzameling. De gegevens worden veelvuldig voor onderzoek gebruikt.

3.1.4 Ziekenhuiszorg

In alle ziekenhuizen wordt gebruik gemaakt van elektronische patiëntendossiers die in principe hergebruikt kunnen worden. Deze gegevens worden bijeen gebracht in de Landelijke Basisregistratie ziekenhuiszorg (LBZ). De LBZ is de registratie van medische, administratieve en financiële gegevens van patiënten die een klinische opname, dagopname of langdurige observatie achter de rug hebben of die poliklinisch behandeld werden. Gegevens worden bij elkaar gebracht door Dutch Hospital Data (DHD). Ze zijn via een TTP op persoonsniveau te koppelen aan andere databronnen en kunnen beschikbaar worden gesteld via het CBS. DHD vraagt een vergoeding voor het gebruik van de data.

Een andere bron van gegevens over ziekenhuiszorg (en feitelijk alle zorg die onder de zorgverzekeringswet valt) is de DBC registratie van medisch specialistische zorg. DBC staat voor diagnose behandeling combinatie en beschrijft met welke klacht en hoe een patiënt het ziekenhuis binnenkomt, welke diagnose er gesteld is, en het beschrijft de voorgenomen behandeling.

De LBZ-bestanden zijn gecodeerd volgens de ICD-10 en betreffen alleen de ziekenhuiszorg (dus bijv. niet de ZBC's). De DBC-bestanden bevatten alle contacten met medisch specialistische zorg (dus bijv. ook ZBC's), maar maken gebruik van een diagnoseclassificatie per specialisme om de diagnose te registreren. Vanaf 2016 zijn de DBC's MSZ grotendeels ook voorzien van een ICD-10-code, maar het is onbekend welk deel daarvan afgeleid is uit de DBC-diagnose, in welk geval de ICD-code dus niet aan de bron geregistreerd en niet betrouwbaar is (want de DBC-diagnoseregistratie is daarvoor vaak niet specifiek genoeg). Zowel de DBC data als de LBZ data worden al veelvuldig voor onderzoek gebruikt.

3.1.5 Sociaal economische kenmerken

Sociaaleconomische kenmerken zijn beschikbaar via het sociaal statistisch bestand bij het CBS. Het gaat daarbij om verschillende indicatoren voor inkomen, opleiding, migratieachtergrond etcetera. Veel van deze gegevens zijn beschikbaar voor de gehele bevolking. De gegevens worden veelvuldig voor onderzoek gebruikt.

3.2 Kwetsbare ouderen

In deze paragraaf beschrijven we in globale termen de beschikbaarheid van gegevens die betrekking hebben op de zorg aan kwetsbare ouderen.

3.2.1 Ambulancezorg

In de ambulancezorg worden verschillende registratieformulieren gebruikt, waarvan niet bekend is in hoeverre ze vergelijkbaar zijn tussen verschillende regionale ambulancevoorzieningen. Wel is er door Ambulancezorg Nederland een Basisset Ambulancezorg vastgesteld, waarmee is bepaald welke gegevens minimaal geregistreerd dienen te worden. Hergebruik van de gegevens vindt alleen nog op geaggregeerd niveau plaats via de monitor acute zorg van het RIVM. Er is geen centrale database waar gegevens op microniveau bij elkaar worden gebracht. Over de kwaliteit van registreren is weinig bekend. Aangezien er in principe BSNs worden geregistreerd, zijn gegevens in technisch opzicht op individueel niveau te koppelen aan andere databronnen. In sommige situaties worden een BSN niet altijd geregistreerd, zoals wanneer het gaat om een acute situatie waarbij een patiënt niet of verminderd aanspreekbaar is of geen identiteitsbewijs bij zich heeft.

Een belangrijk bron van gegevens bij ambulancezorg is de registratie op de meldkamer ambulancezorg. De wijze waarop die gegevens beschikbaar gemaakt kunnen worden verschilt tussen regio's. Een landelijke database ontbreekt. Wel is er één uniforme manier waarop meldingen worden geregistreerd. Deze zijn echter alleen aan andere bronnen te koppelen wanneer de gegevens worden doorgezet naar een ambulancedienst. Op de meldkamer zelf is immers vaak geen tijd om zaken als burgerservicenummers en of geboortedatum te vragen. Dit geldt met name voor 112 meldingen. Bij meldingen via huisartsen, verloskundigen, SEH, of andere ketenpartners is die tijd er meestal wel.

3.2.2 Integraal bekostigde eerstelijnszorg (ketenzorg)

Ketenzorg is programmatische eerstelijnszorg die wordt verleend aan mensen met een chronische aandoening door huisartsenpraktijken en paramedici als diëtisten, logopedisten, oefentherapeuten en fysiotherapeuten. Deze zorg wordt op persoonsniveau gedeclareerd door ketenzorgorganisaties. In veel gevallen wordt de zorg geregistreerd in een ketenzorginformatiesysteem, waarvan verschillende leveranciers bestaan. In andere gevallen wordt de zorg gedocumenteerd in de informatiesystemen van de individuele zorgverleners. Vaak is de werkwijze afhankelijk van regionale afspraken.

Deze versnippering bemoeilijkt het hergebruik van gegevens. Er is geen landelijk dekkende database, maar in principe zijn gegevens wel koppelbaar.

3.2.3 Laboratoriumuitslagen

Hier gaat het om de uitslagen van diagnostische tests. Deze worden uiteraard vastgelegd door de laboratoria waar die tests worden uitgevoerd. Er bestaan huisartsenlabs, die de gevonden gegevens doorsturen naar huisartsenpraktijken. Voor een aantal tests is het vermoeden dat verschillende laboratoria verschillende afkapwaarden hanteren. Omdat het hier gaat om zorg aan individuele patiënten en omdat er sprake is van declareerbare zorg, zullen gegevens in principe koppelbaar zijn. Er is geen centrale database. Een deel van de laboratoriumuitslagen komen terecht in de dossiers van huisartsen en zijn langs die weg te ontsluiten.

3.2.4 Indicatiestelling langdurige zorg

De Wlz biedt een wettelijk verzekerde recht op zorg voor de meest kwetsbaren in onze samenleving. Het gaat om mensen die sterk in hun mogelijkheden en regie zijn beperkt en blijvend zijn aangewezen op permanent toezicht of 24 uur per dag zorg in de nabijheid. De blijvende behoefte aan Wlz-zorg wordt op basis van objectieve criteria door het Centrum Indicatiestelling Zorg (CIZ) vastgesteld. De registratie gebeurt op uniforme wijze door één centrale instelling, de gegevens zijn dus volledig vergelijkbaar. Omdat er één organisatie is die dit regelt, is er ook sprake van landelijke dekking. De gegevens worden via CBS beschikbaar gesteld voor analyses.

3.2.5 Pathologielaboratoria

Alle pathologielaboratoria in Nederland zijn elektronisch verbonden met PALGA. PALGA levert gegevens uit de landelijke databank voor wetenschappelijk onderzoek, voor medische kwaliteitscontrole en voor de evaluatie en monitoring van de bevolkingsonderzoeken. PALGA biedt de mogelijkheid tot koppeling van gegevens via een TTP. Gegevens uit PALGA kunnen beschikbaar gesteld worden voor wetenschappelijk onderzoek. Uitgangspunt hiervoor zijn de PALGA-codes en PALGA termen uit de PALGA Thesaurus (een van SNOMED afgeleide terminologie). Onder strikte privacyvoorwaarden is het ook mogelijk om een onderzoekscohort te koppelen aan de PALGA-databank.

3.2.6 Langdurige zorg registratie

Vrijwel alle specialisten ouderengeneeskunde werken met een elektronisch patiëntendossier: Ysis (leverancier Gerimedica) en Ons (Nedap), MijnCaress (PinkRocade Healthcare) worden het meest gebruikt. Er is geen centrale dataverzameling van deze gegevens en er is weinig bekend over de kwaliteit van deze gegevens. Bij enkele universitaire vakgroepen ouderengeneeskunde bestaan netwerken waar onderzoek wordt gedaan met deze gegevens, afkomstig van een selectie van verpleeghuizen.

3.2.7 Maatschappelijke ondersteuning

Gemeenten zijn verantwoordelijk voor de verstrekking van zorg en ondersteuning vanuit de wet maatschappelijke ondersteuning. Het gaat hierbij om hulp bij huishouden, woonvoorzieningen, vervoersvoorzieningen, beschermd wonen GGZ, maatschappelijke opvang, ambulante ondersteuning en dagbesteding. Gegevens hierover worden door gemeenten bijgehouden. Deze gegevens worden bij het CBS bij elkaar gebracht, en zijn via het CBS koppelbaar aan andere gegevensbronnen. Gemeenten leveren deze data op vrijwillige basis aan in het kader van de Gemeentelijke Monitor Sociaal Domein. Niet alle gemeenten doen dat, het CBS heeft gegevens van iets meer dan 300 gemeenten. Van wege de grote beleidsvrijheid van gemeenten is het vaak moeilijk de gegevens goed te duiden zonder gedetailleerde kennis van de lokale praktijk.

3.2.8 Paramedische zorg

Bijna alle diëtisten, fysiotherapeuten, oefentherapeuten, logopedisten, ergotherapeuten leggen gegevens vast over de geleverde zorg. Dat biedt mogelijkheden voor hergebruik. Dat gebeurt dan ook, zij het nog niet bij alle disciplines en zij het dat nog slechts een beperkt deel van de zorgverleners deelneemt. Er is een landelijke database fysiotherapie en een database Keurmerk Fysiotherapie waarbinnen gegevens worden verzameld. Ook de oefentherapeuten en de logopedisten, werken aan de vorming van een dergelijk register. Een aantal daarvan worden door het Nivel beheerd. In principe zijn deze gegevens koppelbaar met gegevens uit andere bronnen (weer via een TTP). Dit gebeurt bijvoorbeeld binnen Nivel zorgregistraties eerste lijn (koppeling met huisartsendata).

3.2.9 Wijkverpleging

In 2017 publiceerde het Nivel een onderzoek naar de bruikbaarheid van gegevens die wijkverpleegkundigen vastleggen voor hergebruik in Nivel Zorgregistraties. Dit onderzoek wees uit dat er veel wordt geregistreerd, dat dat nog op veel verschillende manieren gebeurt, maar dat de mogelijkheden voor hergebruik zeker aanwezig waren. Belangrijkste aanbeveling was: begin klein en bouw de gegevensverzameling langzaam op, te beginnen met een aantal voorlopers. Op dit moment vindt hergebruik van gegevens die in de wijkverpleging worden geregistreerd voor zover bekend niet op grote schaal plaats.

3.3 Kinderen van ouders met psychische problemen (KOPP) of van verslaafde ouders (KVO)

In deze paragraaf bespreken we databronnen die van belang zijn bij het analyseren van zorgpaden van kinderen van ouders met psychische problemen of van verslaafde ouders.

3.3.1 Verslavingszorg

Instellingen voor ambulante en klinische verslavingszorg en particuliere instellingen leveren gegevens omtrent hun cliënten aan het Landelijk Alcohol en Drugs Informatiesysteem (LADIS), dit is een continue aanlevering. De gegevens worden bij elkaar gebracht door IVZ, de stichting Informatievoorziening Zorg. Koppeling van LADIS gegevens aan andere databronnen is mogelijk onder voorwaarden. De gegevens zijn momenteel beschikbaar t/m 2015. Publicatie en verstrekking van actuelere gegevens is opgeschort i.v.m. het regelen van een wettelijke grondslag voor het verzamelen van de gegevens.

3.3.2 GGD'en en zorgorganisaties die jeugdgezondheidszorg leveren

Na ieder contact met een jeugdige en/of zijn ouders noteert de jeugdarts, jeugdverpleegkundige of assistente de relevante bevindingen in het dossier. Om deze vastlegging te standaardiseren is de Basisdataset (BDS) JGZ ontwikkeld, welke aangeeft wat op welke wijze genoteerd moet worden. Door gestandaardiseerd te registreren, kunnen dossiers overgedragen worden, zodat zorg gecontinueerd kan worden – en daarmee ook de registratie. Het is echter niet duidelijk welke gegevens deze basisdataset bevat en of deze beschikbaar is voor hergebruik.

3.3.3 Gespecialiseerde GGZ, verslavingszorg (DBC's GGZ)

Deze via CBS en NZA beschikbare, koppelbare bestanden bevatten informatie over gespecialiseerde GGZ-trajecten met gegevens over looptijd, zorgtype, primaire diagnose en declaratiegegevens. Het gaat hier om declaratiegegevens. Informatie over jeugd is echter na 2015 niet meer in de bestanden opgenomen: met ingang van 2015 maakt de jeugd GGZ onderdeel uit van de jeugdzorg via de gemeenten (zie verder bij 3.3.5).

3.3.4 Jeugddetentie

Bij deze vrijwillige, centrale registratie van dossierinformatie afkomstig uit de justitiële jeugdinrichtingen is het onbekend hoe goed de gegevens koppelbaar zijn en er is goedkeuring nodig voor eventueel hergebruik. De databron bevat gedetailleerde informatie uit de dossiers van justitiële jeugdinrichtingen, waaronder ook informatie over eerdere uithuisplaatsingen, abnormale gezinssituatie, hulpverleningshistorie, opvoeders, psychische stoornissen, verslaving en incidenten tijdens het verblijf in de inrichting.

3.3.5 Jeugdzorg

Gegevens over alle vormen van jeugdhulp worden door alle aanbieders van jeugdhulp aan het CBS geleverd, in het kader van de Beleidsinformatie Jeugd. Deze leveringen zijn wettelijk verplicht en vloeien voort uit de Jeugdwet. De data zijn koppelbaar en geschikt voor hergebruik maar bevatten vrij basale informatie over de jeugdzorgtrajecten. Informatie over de indicatie voor het jeugdzorgtraject, de gestelde diagnose (indien van toepassing), details over de aard van het jeugdzorgtraject zelf en de uitkomst daarvan ontbreken. Deze gegevens worden wel vastgelegd door de jeugdhulpverleners, maar worden niet centraal bij elkaar gebracht.

3.3.6 Jeugdbescherming

Alle jeugdbeschermingstrajecten (voogdij en ondertoezichtstelling) die onder verantwoordelijkheid van de gemeente zijn uitgevoerd door de Gecertificeerde Instellingen worden aan het CBS geleverd in het kader van de Beleidsinformatie Jeugd. Deze leveringen zijn wettelijk verplicht en vloeien voort uit de Jeugdwet. Net als bij het bestand met jeugdzorgtrajecten geldt dat de data koppelbaar zijn en geschikt voor hergebruik, maar dat de inhoudelijke informatie over jeugdbeschermingstrajecten beperkt is. Die gegevens worden vastgelegd door de aangesloten jeugdhulpverleners, maar worden niet centraal bijeengebracht. De beschikbare data (van de gemeenten) geven dus vooral een overzicht van de aanwezigheid en loopduur van eventuele jeugdbeschermingstrajecten.

4 Conclusie

4.1 Belangrijkste bevindingen

Uit ons overzicht blijkt dat een groot deel van de databronnen al wordt gebruikt voor beleidsinformatie. En koppeling met andere databronnen vindt in veel gevallen ook al plaats. Er ontbreken echter schakels of er is geen landelijke dekking, waardoor het volledig in beeld brengen van de variëteit aan zorgpaden niet altijd mogelijk is. Daarnaast zijn er praktische en financiële barrières.

Ons overzicht laat drie groepen van zorgvoorzieningen zien:

1. De gegevens zijn er wel, worden geregistreerd in elektronische dossiers, maar worden niet verzameld ten behoeve van gebruik voor andere doeleinden dan ondersteuning van de zorgverlening zelf.
Voorbeelden zijn:
 - wijkverpleging
 - verpleeghuiszorg
 - jeugdzorg
2. Bij een deel van de zorgaanbieders worden de gegevens wel verzameld ten behoeve van gebruik voor andere doeleinden, maar er is geen sprake van landelijke dekking en er is sprake van versnippering. Vaak is hier sprake van meerdere verzamelaars en systemen waarin de data zijn samengebracht, maar geen daarvan heeft een landelijke dekking.
Voorbeelden zijn:
 - huisartsenzorg
 - ketenzorg
 - huisartsenposten
 - paramedische beroepenDeze dataverzamelingen zijn niet landelijk dekkend, maar kunnen wel landelijk representatief zijn.
3. Er is een landelijke registratie, met (nagenoeg) landelijke dekking.
Voorbeelden zijn:
 - ziekenhuiszorg (LBZ)
 - indicatiestelling voor de WLZ
 - declaraties zorgverzekeraars

Een eerste belangrijke bevinding is dat hergebruik op enige schaal vooral lijkt plaats te vinden als data zijn samengebracht in een centrale landelijke of landelijk representatieve database (groep 2 en 3). In dit verband is de uitspraak van de minister in de recente nota “Data laten werken voor gezondheid” interessant: “Ik hecht aan verantwoord gebruik van data en het – waar mogelijk - voorkómen van landelijke dataverzamelingen”.

Die opmerking is wellicht ingegeven door nieuwe technologische ontwikkelingen die in de brief worden genoemd en die het mogelijk moeten gaan maken om “data op hun plaats te laten blijven” en toch de gewenste analyses op die data te doen. Binnen het landelijke initiatief Health-RI (Research Infrastructure) worden de technische mogelijkheden van *personal health trains* ontwikkeld. De praktische toepasbaarheid voor (beleids)onderzoek wordt onderzocht (Zorginstituut Nederland, 2018). Het valt nog te bezien of dit concept de huidige centrale databases kan vervangen.

Temeer daar vanuit het oogpunt van onderzoek en statistiek het samenbrengen van microdata in registraties zeer grote voordelen heeft: het 'schonen' van de data gebeurt maar één keer en op één manier, resultaten van analyses zijn goed reproduceerbaar, en koppeling met niet-medische data (onder andere bij het CBS) wordt vergemakkelijkt.

Een andere belangrijke bevinding is dat veel gegevensbronnen in categorie 2 en 3 worden ontsloten via het CBS. Dat is echter niet met alle databronnen het geval.

Verder komt uit het overzicht naar voren dat van veel gegevensbronnen niet bekend is wat de kwaliteit ervan is. Binnen het bestek van dit rapport was het niet mogelijk dit helemaal in kaart te brengen, ook al omdat de kwaliteit van gegevens slechts kan worden beoordeeld in het licht van het doel waarvoor gegevens worden gebruikt. Het geeft wel aan dat gegevensbronnen veelal (nog) niet aan de FAIR-principes voldoen.

Het algemene beeld is dat er veel gegevens worden vastgelegd, maar dat hergebruik in veel gevallen nog maar beperkt plaatsvindt. Hieronder gaan we na welke factoren daaraan ten grondslag kunnen liggen.

4.2 Praktische beperkingen

Ook al zijn gegevens centraal (d.w.z. in één bestand/registratie) beschikbaar, in veel gevallen moet voor hergebruik voldaan worden aan een aantal eisen. Er worden kosten in rekening gebracht, of data mogen niet verplaatst worden, of mogen alleen gebruikt worden als de beheerder van een databron kan meekijken met analyses en rapportages, of er zijn commissies die toestemming moeten geven. Eigenlijk zijn er maar weinig databronnen waarvoor dergelijke eisen niet gelden. Ook: analyse is slechts op een beperkt aantal fysieke plekken mogelijk (CBS, universiteiten, sommige onderzoeksinstellingen). Derden (bijv. GGD'en) moeten dus contracten/samenwerkingsafspraken maken met deze instituten.

Deze zijn over het algemeen het gevolg van a) governance afspraken die door beheerders van de gegevens zijn gemaakt met de toeleveranciers van de gegevens en b) het werk dat er mee is gemoeid om gegevens geschikt en beschikbaar te krijgen voor hergebruik.

4.3 Data koppeling

Koppeling van databronnen is in veel gevallen technisch mogelijk. Bij nagenoeg alle zorgdisciplines worden burgerservicenummers geregistreerd. Dat maakt het in principe mogelijk om gegevens aan elkaar te koppelen, en dat gebeurt ook al veelvuldig.

Vaak wordt het CBS als koppelplatform gebruikt. De wet op het CBS maakt het mogelijk dat gegevens voorzien van burgerservicenummers direct naar het CBS worden gestuurd, om aldaar aan andere databronnen (ook buiten het zorgterrein) te worden gekoppeld. Dit heeft ook als voordeel dat er dan ook automatisch gebruik gemaakt wordt van de maatregelen die het CBS neemt om de privacy van burgers en bedrijven te beschermen. Vaak worden er pseudonimiseringstechnieken toegepast zoals die worden aangeboden door het bedrijf ZorgTTP.

Koppelen kan vaak in technisch opzicht, maar de praktijk wijst uit dat dit makkelijker is gezegd dan gedaan. De versnippering van gegevens is er de oorzaak van dat voor veel gegevenskoppelingen een groot aantal stakeholders moet worden overtuigd van het nut, van de juridische beperkingen (of de afwezigheid daarvan), en van de technische haalbaarheid. Het gaat hier vaak om een ervaren knelpunt en niet zozeer om een juridisch knelpunt. Adequate en transparante governance afspraken vormen de sleutel tot dit probleem, want daarin wordt vastgelegd wie wat onder welke voorwaarden met gegevens mag doen.

4.4 Juridische aspecten

Onbekendheid met de juridische mogelijkheden en beperkingen kan een belangrijke reden zijn waardoor hergebruik van gegevens niet van de grond komt. In dit verband is de CBS wet relevant, die zegt dat het CBS ten behoeve van statistische doeleinden gebruik mag maken van gegevens uit registraties die in verband met de uitvoering van een wettelijke taak worden bijgehouden bij overheidsinstanties en ZBO's. Hetzelfde geldt voor gegevens uit registraties die worden bijgehouden bij algemene maatregel van bestuur aangewezen rechtspersonen die een bij of krachtens de wet geregelde taak uitoefenen of geheel of gedeeltelijk, direct of indirect, worden bekostigd uit middelen van de Staat of uit de opbrengst van bij of krachtens de wet ingestelde heffingen. Het koppelen van registraties/databestanden van andere partijen en het aan daartoe gemachtigde onderzoeksinstellingen toegang verlenen tot de gekoppelde (met CBS-data verrijkte) gegevens, valt ook onder de wettelijke activiteiten van het CBS, met dezelfde waarborgen t.a.v. privacy en doelbinding (alleen voor statistiek en onderzoek) die gelden voor de data die het CBS voor haar eigen statistieken hanteert.

Daarnaast is de WGBO van belang. De hoofdregel is dat patiënten/burgers expliciet toestemming moeten geven. Toestemming van de patiënt is niet vereist, indien (artikel 7:458 lid 1 BW):

- 1. a. het vragen van toestemming in redelijkheid niet mogelijk is en met betrekking tot de uitvoering van het onderzoek is voorzien in zodanige waarborgen, dat de persoonlijke levenssfeer van de patiënt niet onevenredig wordt geschaad, of b. het vragen van toestemming, gelet op de aard en het doel van het onderzoek, in redelijkheid niet kan worden verlangd en de hulpverlener zorg heeft gedragen dat de gegevens in zodanige vorm worden verstrekt dat herleiding tot individuele natuurlijke personen redelijkerwijs wordt voorkomen.*
- 2. Verstrekking overeenkomstig lid 1 is slechts mogelijk indien: a. het onderzoek een algemeen belang dient, b. het onderzoek niet zonder de desbetreffende gegevens kan worden uitgevoerd, en c. voor zover de betrokken patiënt tegen een verstrekking niet uitdrukkelijk bezwaar heeft gemaakt.*

De AVG stelt in overweging 53 het volgende over het gebruik van bijzondere persoonsgegevens in het kader van gezondheidsdoeleinden:

“Bijzondere categorieën van persoonsgegevens waarvoor betere bescherming is vereist, mogen alleen voor gezondheidsdoeleinden worden verwerkt indien dat nodig is om die doeleinden te verwezenlijken in het belang van natuurlijke personen en de samenleving als geheel, met name bij het beheer van gezondheidszorgdiensten en -stelsels of sociale diensten en stelsels van sociale diensten, met inbegrip van de verwerking door de beheersautoriteiten en de centrale nationale gezondheidsinstanties van die gegevens met het oog op kwaliteitscontrole, beheersinformatie en het algemeen nationaal en lokaal toezicht op het gezondheidszorgstelsel

of het stelsel van sociale diensten, en bij het waarborgen van de continuïteit van de gezondheidszorg of de sociale diensten en grensoverschrijdende gezondheidszorg of voor doeleinden inzake gezondheidsbeveiliging, -bewaking en -waarschuwing of met het oog op archivering in het algemeen belang, wetenschappelijk of historisch onderzoek of statistische doeleind en op basis van Unierecht of lidstatelijk recht die aan een doelstelling van algemeen belang moet voldoen, alsook voor studies van algemeen belang op het gebied van de volksgezondheid”.

In de UAVG, de Nederlandse invulling van de Europese AVG is deze overweging verwerkt in artikel 9 van de AVG. Zie <http://www.privacy-regulation.eu/nl/index.htm>. Met name artikel 9 lid 2 sub i lijkt in onze situatie relevant. Daarin staat dat verwerking van persoonsgegevens over onder andere gezondheid niet is toegestaan tenzij:

*“de verwerking is noodzakelijk om redenen van **algemeen belang op het gebied van de volksgezondheid, zoals bescherming tegen ernstige grensoverschrijdende gevaren voor de gezondheid of het waarborgen van hoge normen inzake kwaliteit en veiligheid van de gezondheidszorg en van geneesmiddelen of medische hulpmiddelen**, op grond van Unierecht of lidstatelijk recht waarin passende en specifieke maatregelen zijn opgenomen ter bescherming van de rechten en vrijheden van de betrokkene, met name van het beroepsgeheim” [cursivering door de auteurs].*

Hiermee lijkt de AVG meer ruimte te bieden dan de WGBO. De WGBO is echter een zogenaamde “lex specialis” en krijgt daarom voorrang op de AVG die geldt als “lex generalis”. Artikel 9 lid 2 sub j, lijkt die ruimte te bieden:

*“de verwerking is noodzakelijk met het oog op archivering in het algemeen belang, **wetenschappelijk of historisch onderzoek of statistische doeleinden** overeenkomstig artikel 89, lid 1, op grond van Unierecht of lidstatelijk recht, waarbij de evenredigheid met het nagestreefde doel wordt gewaarborgd, de wezenlijke inhoud van het recht op bescherming van persoonsgegevens wordt geëerbiedigd en passende en specifieke maatregelen worden getroffen ter bescherming van de grondrechten en de belangen van de betrokkene” [cursivering door de auteurs]*

4.5 Versnipperd datalandschap

Uit ons overzicht komt naar voren dat het aantal potentiële gegevensbronnen groot is. Er zijn slechts enkele gegevensbronnen landelijke dekkend, en veel potentiële gegevensbronnen zijn niet of niet volledig ontsloten. Beide casussen laten zien dat er al veel gegevens zijn die hergebruikt (kunnen) worden, maar dat er schakels ontbreken waardoor het complete plaatje van zorggebruik niet goed te maken is. Initiatieven om data beter te gebruiken worden wel genomen, maar komen nog niet voldoende tot wasdom.

Bijna ieder zorgverlener in Nederland registreert in een elektronisch cliënten of patiëntendossier. Dat schakels ontbreken, heeft over het algemeen dus niet als reden dat gegevens niet geregistreerd worden. In veel gevallen is de reden dat deze gegevens niet bij elkaar gebracht worden om te kunnen dienen als informatie voor beleid.

Vaak blijft het beperkt tot (her)gebruik ten behoeve van de zorgverlening zelf. Zo wordt er nog onvoldoende gebruik gemaakt van gegevens die in de wijkverpleging, het sociaal domein, de jeugdzorg en in verpleeghuizen worden vastgelegd.

Dat beeld komt overeen met het onderzoek van de OECD uit 2015 naar “Health Data Governance” in 22 landen waaronder Nederland. Van de 22 landen staat NL op de 18e plaats als het gaat om het delen van en de toegang tot gezondheidsdata voor statistische en onderzoeksdoeleinden. Nederland hoort bij de landen met het meest versnipperde datalandschap. Daar staat tegenover dat Nederland (in 2015), samen met slechts drie andere landen, beschikt over een remote access structuur (bij het CBS) die onderzoekers beveiligde toegang geeft tot nationale datasets.

De OECD stelt vast dat er al een enorme groei is geweest in het verzamelen en dat tegelijkertijd de mogelijkheden om al die data te combineren en integreren toeneemt. Om die mogelijkheden te benutten is de OECD van mening dat overheden moeten zorgen voor “strong health data governance frameworks”. De OECD benoemt acht mechanismes die samen voor een dergelijk raamwerk zorgen. Eén van de aanbevelingen is dat hergebruik voor public health, onderzoek en statistiek toegestaan moet zijn, uiteraard met de benodigde gegevensbeschermingsmaatregelen.

4.6 Data kwaliteit

4.6.1 Eenheid van taal

Eenheid van taal is belangrijk om gegevens uit verschillende bronnen met elkaar te kunnen combineren. Dat geldt natuurlijk niet alleen voor hergebruik ten behoeve van onderzoek, maar vooral ook ten behoeve van de patiëntenzorg.

Gebrek aan eenheid van taal hoeft echter niet altijd te betekenen dat gegevens niet kunnen worden hergebruikt ten behoeve van onderzoek en beleid. Vaak is er wel een grootste gemene deler in het ‘taalgebruik’ waardoor het voor een aantal belangrijke kenmerken wel degelijk mogelijk is om gegevens te hergebruiken. Gebrek aan eenheid van taal hoeft hergebruik ten behoeve van onderzoek en beleid dus niet in de weg te staan.

Gebrek aan eenheid van taal in de uitwisseling van gegevens ten behoeve van patiëntenzorg is een veel groter probleem voor de uitwisseling van gegevens in de patiëntenzorg, dan in beleidsonderzoek dat met diezelfde gegevens plaatsvindt. Problemen met eenheid van taal kunnen vaak wel overbrugd worden door betreffende gegevens te hercoderen. Een voorbeeld is de huisartsenzorg, waarbij het, ondanks problemen met de ‘semantische integratie’ en ondanks het ontbreken van zorginformatiebouwstenen al sinds de jaren 1990 mogelijk is om gegevens voor beleid te gebruiken via wat nu Nivel Zorgregistraties heet. Het kost echter nog steeds veel inspanning om gegevens uit de verschillende softwaresystemen die in de huisartsenpraktijk worden gebruikt, semantisch te integreren.

4.6.2 Datakwaliteit en de kwetsbaarheid van het datalandschap

Binnen het bestek van deze studie was het niet mogelijk om van iedere databron vast te stellen wat de kwaliteit van de gegevens is. Bovendien is het lastig om daar uitspraken over te doen buiten de context van het doel waarvoor de gegevens gebruikt zouden moeten worden.

Wel is duidelijk dat de inhoud en de kwaliteit van gegevens die in het zorgproces worden vastgelegd afhankelijk is van de wijze waarop het zorgsysteem is ingericht (Verheij et al., 2018).

De kwaliteit (volledigheid, juistheid, tijdigheid) van gegevens verandert met de inrichting van het zorgstelsel. Dat geldt voor gegevens die in het primaire zorgproces worden vastgelegd, maar bij uitstek ook voor gegevens die gebruikt worden in het declaratieverkeer. De decentralisaties van de zorg in 2015 zijn hiervan een 'goed' voorbeeld. Doordat langdurige zorg thuis grotendeels overging van AWBZ naar Wmo zijn er nu geen indicatiedata meer voor die zorg. En doordat de jeugd-ggz is overgeheveld van de Zvw naar de jeugdwet, bevatten de declaratiedata van Vektis en de DIS-registratie van de NZa geen gegevens meer over deze belangrijke vorm van zorg.

Declaratiedata hebben als voordeel dat ze veelal landelijk dekkend zijn, en precies weergeven hoeveel geld welke zorg heeft gekost. Deze gegevens hebben echter ook de beperking dat ze vanzelfsprekend sterk afhankelijk zijn van de financieringsystematiek. Zo heeft bij sommigen lange tijd de indruk bestaan dat gegevens van DBC's zonder meer gebruikt kunnen worden om prevalenties te schatten. Er staat immers in voor welke ziekte een bepaalde behandeling is gedeclareerd. Veranderingen in de declaratiecodes hebben dan echter direct gevolgen voor de prevalentieschattingen die er op zijn gebaseerd. Dat betekent dat vragen als "wie krijgt welke zorg voor welke aandoening" beter kunnen worden beantwoord met inhoudelijke gegevens die voor en door de zorgverleners worden vastgelegd.

Gegevens die worden vastgelegd voor de directe patiëntenzorg zijn minder afhankelijk van het bekostigingssysteem dan gegevens die ten behoeve van het declaratieverkeer worden geregistreerd. In een recent OECD-rapport (OECD, 2017) wordt gesteld dat er in 16 van de onderzochte OECD-landen (Nederland deed niet mee) procedures zijn op nationaal niveau om data uit patiëntendossiers te gebruiken "for dataset creation and analysis". Verder wordt opgemerkt: "This study identified many challenges that countries are facing in their efforts to assure that electronic clinical records are complete, accurate, relevant, unique, interpretable and widely available. Solutions adopted in some jurisdictions, and which have been described in this study, could potentially address some of the data quality concerns that have been expressed. This would include setting terminology and interoperability standards, creating mandatory requirements or financial incentives to promote structure and standards, creating a standardised minimum dataset, and clinical data content quality auditing, among other strategies."

Niet alleen het bekostigingssysteem is veranderlijk onder invloed van beleid. Beleid kan ook gericht zijn op de registratie van gegevens zelf. Zo is er het beleid om de registratiedruk en administratieve lasten te verlagen. Daar kunnen goede redenen voor zijn, maar het heeft onvermijdelijk tot gevolg dat er minder gegevens worden geregistreerd. En in sommige gevallen waren die gegevens wellicht voor sommige doelen wel degelijk nuttig.

4.7 Aanbevelingen

Onze aanbevelingen komen hoofdzakelijk voort uit de hoofdconclusies van dit rapport:

1. Gegevens die in de zorg worden geregistreerd worden nog niet optimaal hergebruikt voor beleidsonderzoek en statistiek.
2. Het ontbreken van een centrale landelijke of landelijk representatieve dataverzameling lijkt daar in belangrijke mate debet aan te zijn.
3. Maar er zijn maar weinig centrale databases zonder praktische (waaronder financiële) voorwaarden voor hergebruik.

Onze aanbevelingen zijn:

1. Geef invulling aan de aanbeveling van de OECD om te zorgen voor een 'strong health data governance framework'. Governance is de sleutel tot vertrouwen van stakeholders in het belang van hergebruik. Dat betekent:
 - a. Begin met het Waaron, het omschrijven van doelen die geworteld zijn in het algemeen belang.
 - b. heldere omschrijving van verantwoordelijkheden bij beslissingen over het gebruik van gegevens;
 - c. transparantie over het gebruik van de gegevens;
 - d. adequate gegevensbeschermingsmaatregelen
 - e. het betrekken van alle relevante partijen (burgers/patiënten, zorgverleners, onderzoekers, software leveranciers).

Begin desnoods op kleine schaal, met een 'coalition of the willing' en laat het nut zien van hergebruik van gegevens. Kijk of best practices uit andere OECD-landen overgenomen kunnen worden.

2. Onderken het belang van gecentraliseerde landelijke en/of landelijk representatieve registraties, zoals ook de OECD aanbeveelt en wacht niet op de totstandkoming van gedistribueerde gegevensinfrastructuren en concepten als de Personal Health Train.
3. Zoek oplossingen voor de kosten van het onderhouden van een centrale database, zodat (her-)gebruikers geen onoverkomelijke financiële drempels ondervinden.
4. Geef meer bekendheid aan de ruimte voor hergebruik die de WGBO nu al biedt en onderzoek de ruimte die de AVG biedt om voor specifieke hergebruiks-doelen de huidige WGBO-regels te verruimen.
5. Wees alert op de impact die beleidswijzigingen hebben op de kwaliteit en beschikbaarheid van gegevens. Ontwikkel een systematiek om bij belangrijke beleidswijzigingen een B-DIA uit (Beleid-Data Impact Assessment), uit te kunnen voeren, om na te gaan wat het effect is van beleid op de kwaliteit en beschikbaarheid van data (bijvoorbeeld mogelijke effecten van schrapsessies om de registratiedruk en de regeldruk te verminderen, of veranderingen in het bekostigingssysteem).
6. Ons overzicht van gegevensbronnen maakt duidelijk dat hergebruik niet afhankelijk hoeft te zijn van eenheid van taal. Vaak is er een gemeenschappelijke basis waarbinnen die eenheid van taal er al wel is. Hergebruik van zorggegevens kan dus worden doorontwikkeld zonder dat eenheid van taal is afgerond.

Bij deze aanbevelingen maken we tot slot nog één kanttekening. In de media en in het beleid is veel aandacht voor registratiedruk en administratieve lasten in de zorg. In dit rapport ging het over gegevensbronnen die afhankelijk zijn van die administraties. Het is van groot belang om bij het bestrijden van administratieve lasten rekening te houden met gevolgen voor hergebruik. De impact van het schrappen van bepaalde registraties kan groot zijn en zich toch voor een belangrijk deel onttrekken aan het blikveld van de direct betrokkenen. Het vastleggen of verzamelen van bepaalde gegevens kan vanuit een bepaald perspectief niet nodig zijn, vanuit een ander perspectief kan dat wel degelijk het geval zijn. Het zogenaamde proportionaliteitsbeginsel in de AVG (geen gegevens verzamelen of bij elkaar brengen die niet nodig zijn voor een bepaald doel) kan een lerend zorgsysteem en hergebruik in de weg staan als die proportionaliteit alleen vanuit één perspectief of doel wordt beoordeeld. Hergebruik van gegevens houdt per definitie in dat er uiteenlopende doelen kunnen zijn en dat dat niet noodzakelijk doelen van één partij hoeven te zijn. Dit leidt tot een laatste aanbeveling:

7. Hergebruik van routine zorgdata vereist mogelijk dat het proportionaliteitsbeginsel uit verschillende perspectieven wordt belicht. Het vastleggen en/of verzamelen van gegevens kan vanuit één bepaald perspectief niet nodig zijn, maar vanuit een ander perspectief wel. Toepassing van het proportionaliteitsbeginsel vanuit één perspectief kan leiden tot beperking van een gegevensverzameling, die vanuit een ander perspectief heel belangrijk kan zijn. Aanbeveling is om te onderzoeken in hoeverre dit een probleem vormt en wat er eventueel aan te doen valt.

Hierboven werd al voorgesteld om een systematiek te ontwikkelen voor een zogenaamde B-DIA (beleid-data-impact-assessment). Een dergelijke B-DIA zou goede diensten kunnen bewijzen bij de vraag vanuit welk perspectief het proportionaliteitsbeginsel zou moeten worden toegepast.

Literatuur

Bij, S. van der, Verheij, R. Inzet variabiliseringsgelden 2013 leidt tot belangrijke verbetering EPD. SynthesHis: Tijdschrift voor Huisarts en Automatisering: 2013, 12(4), 16-17.

Bij, S. van der, Biermans, M., Levelink, H., Verheij, R. De kwaliteit van de elektronische verslaglegging door huisartsenpraktijken in de regio Nijmegen. Nivel, 2015.

Brief van het kabinet over elektronische gegevensuitwisseling in de zorg, 9 november 2018.
(<https://www.rijksoverheid.nl/documenten/brieven/2018/11/15/data-laten-werken-voor-gezondheid>)

Flinterman L, Groenewegen P, Verheij R. Zorglandschap en zorggebruik in een veranderende eerste lijn. Utrecht: Nivel, 2018.

Friedman C, Rubin J, Brown J, Buntin M, Corn M, Etheredge L, Gunter C, Musen M, Platt R, Stead W, Sullivan K, van Houweling D. Toward a science of learning systems: a research agenda for the high-functioning Learning Health System. J Am Med Inform Assoc. 2015 Jan;22(1):43–50.

<https://www.force11.org/group/fairgroup/fairprinciples>

Mark D Wilkinson, Susanna-Assunta Sansone, Erik Schultes, Peter Doorn, Luiz Olavo Bonino da Silva Santos, Michel Dumontier. A design framework and exemplar metrics for FAIRness. Scientific Data volume 5, Article number: 180118 (2018) doi: 10.1038/sdata.2018.118

Minister van VWS. Data laten werken voor gezondheid; Een kwestie van gewaarborgd vertrouwen, Den Haag, Ministerie van VWS, 2018. (<https://www.rijksoverheid.nl/documenten/brieven/2018/11/15/data-laten-werken-voor-gezondheid>).

OECD, Health Data Governance; privacy, monitoring and research. OECD Health Policy Studies, 2015.

OECD, Readiness of electronic health record systems to contribute to national health information and research, OECD Health Working Papers no.99, 2017.

Raad voor Gezondheidsonderzoek. Van gegevens verzekerd, 2008.

R. Nugteren e.a. DATA: DIGITALE DIAMANTEN Datamanagement voor onderzoeksprojecten gesubsidieerd door ZonMw RIVM, NIVEL, DANS, CentERdata en BBMRI-NL

The Institute of Medicine. Digital infrastructure for the Learning Health System: the foundation for continuous improvement in health and health care; 2011.

Verheij, R.A., Curcin, V., Delaney, B.C., McGilchrist, M.M., Possible sources of bias in primary care electronic health record data use and reuse. Journal of Medical Internet Research, vol. 20, 2018, nr. 5, e185 p

Zorginstituut Nederland. Eindrapportage praktijktoets FAIR Data.
(<https://www.zorginstituutnederland.nl/publicaties/rapport/2018/10/23/eindrapportage-praktijktoets-fair-data>).

Zorginstituut Nederland. Praktijktoets FAIR data & Personal Health Train; Eindrapportage Personal Health Train, versie 1.0, 24 juni 2018
<https://www.rijksoverheid.nl/documenten/rapporten/2018/11/15/eindrapportage-personal-health-train>).