
Van gegevens verzekerd

Kennis over de volksgezondheid in Nederland nu en in de toekomst

RGD

Aan de minister van Volksgezondheid, Welzijn en Sport
De heer dr. A. Klink
Postbus 20350
2500 EJ Den Haag

Onderwerp: aanbieding advies *Van gegevens verzekerd*
Uw kenmerk: PG/FBI 2.767.075
Ons kenmerk: -773/NdN/ts/834-E
Bijlagen: 1
Datum: 29 oktober 2008

Geachte minister,

In mei 2007 verzocht u de Raad voor Gezondheidsonderzoek een advies voor te bereiden over de maatregelen die nodig zijn om de beschikbaarheid van empirische gegevens over de Nederlandse volksgezondheid, ten behoeve van volksgezondheidsbeleid en wetenschap, te verzekeren.

Het gevraagde advies bied ik u hierbij aan, gehoord de beraadsgroepen Gezondheidsethiek en -recht en Maatschappelijke gezondheidszorg. In het advies, dat is voorbereid door een door mij ingestelde commissie van deskundigen onder voorzitterschap van prof. dr. D.E. Grobbee, formuleert de raad een aantal aanbevelingen. Voor drie daarvan vraag ik uw bijzondere aandacht.

Ten eerste beveelt de raad aan om de mogelijkheden tot koppeling van gegevens die in verschillende bestanden over dezelfde personen zijn verzameld, te optimaliseren. De raad bepleit in dat verband om gehoor te geven aan het verzoek van de Federatie van Medisch Wetenschappelijke Verenigingen en de Koninklijke Nederlandse Akademie van Wetenschappen om gebruik van het burgerservicenummer voor wetenschappelijk onderzoek mogelijk te maken.

Een tweede aanbeveling geldt de structurele financiering van longitudinaal epidemiologisch onderzoek. Het advies maakt duidelijk dat dit soort onderzoek essentieel is om waargenomen trends in de volksgezondheid te kunnen verklaren en aangrijpingspunten te vinden voor het beleid. Intermediaire organisaties als ZonMw en NWO bieden tot nu toe voor de noodzakelijke structurele financiering van longitudinaal epidemiologisch onderzoek nauwelijks mogelijkheden.

De derde aanbeveling die ik onder uw aandacht wil brengen betreft de opzet van een zogenaamd herhaald dwarsdoorsnede-onderzoek. Een dergelijk onderzoek, van voldoende omvang, uitgevoerd in de algemene bevolking, is een onmisbare bron van gegevens over risicofactoren voor ziekten en over functionele beperkingen. De raad constateert op dit punt in de huidige Nederlandse situatie een duidelijke leemte, waarin in principe goed te voorzien is.

Postadres

Postbus 16052
2500 BB Den Haag
E-mail: bureau@rgo.nl
website: www.rgo.nl



Bezoekadres

Parnassusplein 5
2511 VX Den Haag
Telefoon: 070-3407521
Fax: 070-3407524

Realisatie van de aanbevelingen zal een krachtige inzet vereisen van alle betrokkenen – beleidsma-
kers, onderzoekers en onderzoeksfinanciers – maar zal ertoe leiden dat volksgezondheidsbeleid,
gezondheidszorg en wetenschap in de toekomst verzekerd zullen zijn van de gegevens die ze nodig
hebben.

Ik heb dit advies vandaag tevens aangeboden aan de minister van Onderwijs, Cultuur en Wetenschap
en aan de minister van Economische Zaken.

Hoogachtend,

A handwritten signature in black ink, consisting of several loops and a long horizontal stroke at the end.

prof. dr. P.J. van der Maas
voorzitter RGO

Van gegevens verzekerd

Kennis over de volksgezondheid in Nederland nu en in de toekomst

aan:

de minister van Welzijn, Volksgezondheid en Sport

de minister van Onderwijs, Cultuur en Wetenschap

de minister van Economische Zaken

Nr. 58, Den Haag, 29 oktober 2008

De Raad voor Gezondheidsonderzoek (RGO) maakt deel uit van de Gezondheidsraad en heeft tot taak de ministers van Volksgezondheid, Welzijn en Sport (VWS), van Onderwijs, Cultuur en Wetenschap (OCW) en van Economische Zaken (EZ) te adviseren over prioriteiten in het gezondheidsonderzoek, in het zorgonderzoek en de technologieontwikkeling in deze sector, evenals over de daarbij behorende infrastructuur. Het maatschappelijk perspectief is daarbij voor de RGO steeds het uitgangspunt.



De Gezondheidsraad is lid van het European Science Advisory Network for Health (EuSANH), een Europees netwerk van wetenschappelijke adviesorganen.



INAHTA

De Gezondheidsraad is lid van het International Network of Agencies for Health Technology Assessment (INAHTA), een internationaal samenwerkingsverband van organisaties die zich bezig houden met *health technology assessment*.

U kunt het advies downloaden van www.gr.nl.

Deze publicatie kan als volgt worden aangehaald:
Raad voor Gezondheidsonderzoek. Van gegevens verzekerd. Kennis over de volksgezondheid in Nederland nu en in de toekomst. Den Haag: Gezondheidsraad, 2008; RGO nr. 58.

Preferred citation:
Advisory Council on Health Research. Securing the data supply. The availability of population health information in the Netherlands, now and in the future. The Hague: Health Council of the Netherlands, 2008; RGO no. 58.

auteursrecht voorbehouden

all rights reserved

ISBN: 978-90-5549-730-0

Inhoud

Samenvatting 9

Executive summary 17

1 Inleiding 25

1.1 Informatiebehoefte 25

1.2 Benodigde infrastructuur 26

1.3 Over dit advies 27

2 Bronnen van gegevens 29

2.1 Herhaald dwarsdoorsnede-onderzoek 29

2.2 Longitudinaal epidemiologisch onderzoek 31

2.3 Registraties 33

3 Gegevensbehoeften en beschikbaarheid 35

3.1 Behoeften van het beleid 35

3.2 Beschikbaarheid van gegevens voor het beleid 36

3.3 Verhouding tussen beleidsbehoefte en beschikbaarheid 39

3.4 Gegevensbehoefte en beschikbaarheid vanuit wetenschappelijk perspectief 42

3.5 Wetenschappelijke gegevensbehoeften in de komende tien jaar 43

3.6 Belanghebbenden 44

4	Gegevens beter benutten	45
4.1	Methoden voor betere benutting	45
4.2	Factoren die het delen van gegevens belemmeren	47
4.3	Technische voorwaarden	48
4.4	Juridische voorwaarden	50
4.5	Financiële voorwaarden	53

5	Aanbevelingen	55
5.1	Bevorder een efficiënt gebruik van gegevensverzamelingen	56
5.2	Hef tekorten aan gegevens op en voorkóm nieuwe tekorten	60
5.3	Creëer meer ruimte voor verzoeken om collectieve financiering van nieuwe gegevensverzamelingen en beoordeel deze systematisch	61

	Literatuur	63
--	------------	----

	Bijlagen	67
A	De adviesaanvraag	69
B	Raad voor Gezondheidsonderzoek	71
C	De commissie	75
D	Inventarisatie Nederlandse populatiecohorten	77
E	‘Eigendom’ en delen (in de zin van <i>sharing</i>) van onderzoeksgegevens	85
F	Afkortingen	99

Samenvatting

Effectief volksgezondheidsbeleid en succesvol wetenschappelijk volkgezondheidsonderzoek zijn beide afhankelijk van gegevens over de Nederlandse volksgezondheid. Op dit moment laat de beschikbaarheid van die gegevens te wensen over. De gegevens die er zijn worden bovendien niet altijd goed benut.

Dat is, in een notendop, de achtergrond van dit advies van de Raad voor Gezondheidsonderzoek (RGO). De raad analyseert de behoefte aan empirische gegevens over de gezondheid van de Nederlandse bevolking. Hij gaat na in hoeverre bestaande gegevensverzamelingen hierin kunnen voorzien, op welke punten ze tekortschieten en wat er moet gebeuren om beschikbare en nieuw op te zetten gegevensverzamelingen zo efficiënt mogelijk te gebruiken. Het advies mondt uit in aanbevelingen die ertoe moeten leiden dat volksgezondheidsbeleid, gezondheidszorg en wetenschap in de toekomst verzekerd zullen blijven van de gegevens die ze nodig hebben.

Gegevens over volksgezondheid zijn essentieel voor beleid en wetenschappelijk onderzoek

Om een adequaat volksgezondheidsbeleid te kunnen voeren moet de overheid informatie hebben over trends in het vóórkomen van ziekten, risicofactoren, functionele beperkingen, zorggebruik en sterfte. Gegevens over risicofactoren voor veelvoorkomende ziekten zijn bijvoorbeeld nodig om ziektelast en gezonde levensverwachting in de toekomst te kunnen schatten.

Gegevens over de gezondheid van de algemene bevolking zijn daarnaast essentieel voor het wetenschappelijk onderzoek naar de oorzaken van ziekte en veroudering, en voor het verklaren van waargenomen trends in de volksgezondheid. Verklaringen van trends zijn van belang om aangrijpingspunten voor het beleid te vinden en om genomen maatregelen te kunnen evalueren.

De methoden om de noodzakelijke gegevens te genereren zijn beschikbaar

Voor het documenteren en onderzoeken van veranderingen in de volksgezondheid zijn vooral drie manieren om gegevens te verzamelen van belang: (herhaald) dwarsdoorsnede-onderzoek, longitudinaal epidemiologisch onderzoek en registraties. Herhaald dwarsdoorsnede-onderzoek in de algemene bevolking, van voldoende omvang, is een goede manier om trends in risicofactoren en functionele beperkingen vast te stellen. Prevalenties van ziekten en beperkingen zijn te schatten op basis van gegevens uit herhaald dwarsdoorsnede-onderzoek en registraties. Registraties kunnen voorts informatie verschaffen over sterfte, doodsoorzaken en zorggebruik. Het longitudinaal epidemiologisch onderzoek – waarin in de loop van de tijd gegevens over dezelfde mensen verzameld worden – is vooral van belang voor het wetenschappelijk onderzoek, met inbegrip van het voor het beleid onmisbare onderzoek ter verklaring van waargenomen trends.

Lopend onderzoek kan niet goed voorzien in de beleidsbehoefte aan gegevens

Er is op dit moment geen bron van actuele en representatieve gegevens die zicht geven op trends in belangrijke risicofactoren in Nederland. Er zijn bijvoorbeeld geen goede en actuele gegevens over hoge bloeddruk en serumcholesterol. De schattingen van toekomstige ziektelast en gezonde levensverwachting zijn daardoor in toenemende mate onbetrouwbaar. De *Health Examination Survey 2008* is een aanzet om in dit tekort te voorzien, maar is beperkt van opzet en vooralsnog ontbreekt de (financiële) continuïteit. Ook voor trends in prevalenties van functionele beperkingen bestaat er in ons land geen robuuste gegevensbron. Trends in prevalenties van veel ziekten en aandoeningen lijken in Nederland wel te schatten op basis van zorgregistraties, mits de continuïteit ervan gewaarborgd is. De onzekerheid rond het voortbestaan van de Landelijke Medische Registratie (LMR) laat zien dat die continuïteit niet vanzelfsprekend is.

Longitudinale gegevensverzamelingen vergen onderhoud

Longitudinale epidemiologische gegevensverzamelingen zijn essentieel voor het wetenschappelijk onderzoek ter verklaring van ziekte, ziektebeloop, veroudering en dragen bij aan het verklaren van trends in de volksgezondheid. Maar ze zijn kostbaar, door de grootschaligheid, de langdurigheid en de kwaliteitseisen waaraan de gegevens moeten voldoen. Onderzoeksgroepen ondervinden al jaren problemen bij de instandhouding van de infrastructuur die voor dit soort gegevensverzameling noodzakelijk is. Wil Nederland zijn goede onderzoekers en sterke wetenschappelijk positie behouden en actief kunnen bijdragen aan zorginnovatie, dan moet de infrastructuur van succesvolle longitudinale epidemiologische gegevensverzamelingen gewaarborgd zijn.

Beschikbare gegevens kunnen beter worden gebruikt

De verzameling van gegevens over de Nederlandse volksgezondheid moet dus worden onderhouden en uitgebreid, maar het is ook belangrijk om de gegevens die er zijn, zo goed mogelijk te gebruiken. Op dat punt is winst te boeken.

De belangrijkste manieren om gegevens beter te benutten zijn:

- Secundaire analyse: het opnieuw analyseren van gegevens met een andere vraagstelling dan waarvoor ze oorspronkelijk zijn verzameld.
- Samenvoeging (*pooling*) van gegevens die op verschillende plaatsen op dezelfde manier bij verschillende mensen verzameld zijn. Het doel is het aantal onderzoeksdeelnemers en daarmee de betrouwbaarheid van het onderzoek te vergroten. Internationaal wordt hard gewerkt aan het combineren van biobanken en aan datapooling van longitudinale epidemiologische dataverzamelingen. Geharmoniseerde onderzoeksmethoden zijn vereist.
- Koppeling van gegevens over dezelfde mensen die in verschillende bestanden verzameld zijn. Een voorbeeld is het koppelen van gegevens uit het doodsoorzakenregister aan gegevens in een populatiecohort. Er is een ‘koppelingssleutel’ nodig om de gegevens over dezelfde persoon aan elkaar te kunnen koppelen.

Delen van gegevens is nodig, maar zorgvuldigheid is geboden

Om gegevens goed te kunnen benutten moet de instantie of onderzoeksgroep die ze verzameld heeft, ze vaak delen met anderen. Dat kan natuurlijk alleen als de privacy van de onderzochte personen gewaarborgd is, in overeenstemming met de daartoe strekkende wet- en regelgeving. Maar ook als aan die voorwaarde is

voldaan, is het delen van gegevens niet altijd vanzelfsprekend. De wetenschappelijke competitie werkt hier soms belemmerend.

De raad meent dat met collectieve middelen* gegenereerde gegevens toegankelijk moeten zijn voor onderzoek dat van nut is voor het algemeen belang, ook als dat wordt uitgevoerd door anderen dan degenen die de gegevens verzameld hebben. Maar er moet tevens zorgvuldig recht worden gedaan aan de belangen van de onderzoekers zonder wier kennis, vaardigheden en inspanningen het gegevensbestand er niet geweest zou zijn.

Aanbevelingen

Op grond van het voorafgaande komt de RGO tot drie aanbevelingen. Realisatie van deze aanbevelingen zal een gezamenlijke inzet vergen van alle betrokkenen: beleidsmakers, onderzoekers en onderzoeksfinanciers. In het bijzonder vragen ze op verschillende punten om een gezamenlijk optrekken, zowel beleidsmatig als financieel, van overheid en onderzoeksfinanciers.

1 Bevorder een efficiënt gebruik van gegevensverzamelingen

a *Stel een register van gegevensverzamelingen in*

De raad beveelt aan een register in te stellen van bestaande en nieuwe gegevensverzamelingen op het gebied van de volksgezondheid, zodat iedereen kan nagaan wat er al is aan gegevens, waar ze zijn en wie ze beheert.

b *Zorg voor een optimale toegankelijkheid van gegevens*

De RGO pleit voor het opstellen van een *gedragscode* voor het delen van gegevens. De raad ziet hier een belangrijke taak weggelegd voor de Stichting Federatie van Medisch Wetenschappelijke Verenigingen (FMWV) en/of de Koninklijke Nederlandse Academie van Wetenschappen (KNAW). *Financiers* van onderzoek kunnen hun bijdrage leveren door aan financiering van bestaande en nieuwe gegevensverzamelingen een viertal voorwaarden te verbinden:

- aanmelding bij het register
- onderschrijving van de gedragscode voor het delen van gegevens

* De raad verstaat hieronder niet alleen de publieke middelen (subsidies van overheid en gezondheidsfondsen) die worden aangewend voor onderzoek, maar ook de overheidsmiddelen en collectieve premies die geïnvesteerd worden in de registraties van bijvoorbeeld CBS, ziektekostenverzekeraars en zorgaanbieders.

- beschrijving, in de subsidieaanvraag, van een procedure voor het verkrijgen van toegang tot de gegevens
 - gebruik van gevalideerde standaardmeetmethoden, of goede argumenten voor het afwijken daarvan.
- c *Optimaliseer de mogelijkheden tot gegevenskoppeling*
De RGO beveelt aan gehoor te geven aan het verzoek van FMWV en KNAW aan de regering om gebruik van het *burgerservicenummer* (BSN) voor wetenschappelijk onderzoek mogelijk te maken. Voor de koppeling van gegevens uit verschillende bestanden is een *trusted third party* nodig, een instantie die onafhankelijk is van de bestandsbeheerders en van de gebruikers van de gekoppelde gegevens.
- d *Vergemakkelijk het delen van gegevens*
De raad vindt dat onderzoekers zoveel mogelijk gefaciliteerd moeten worden bij het delen van gegevens en beveelt daarom aan te voorzien in een onafhankelijke *gegevensmakelaar* die gegevenszoekers en –beheerders met elkaar in contact kan brengen en het delen van gegevens kan begeleiden, en bovendien in *praktische hulp en ondersteuning*. De raad beveelt financiers van onderzoek aan om voor deze vormen van facilitering middelen vrij te maken.

2 Hef tekorten aan gegevens op en voorkóm nieuwe tekorten

- a *Voorzie in herhaald dwarsdoorsnede-onderzoek van voldoende omvang*
De RGO stelt vast dat er in Nederland veel bruikbare gegevens zijn en worden verzameld, maar dat er, vanuit beleids oogpunt, sprake is van één duidelijk tekort: het ontbreekt aan een herhaald dwarsdoorsnedeonderzoek van voldoende omvang, dat kan dienen als bron voor periodieke actualisering van prevalenties van risicofactoren en functionele beperkingen. De raad beveelt aan om in Nederland, naar het model van de Amerikaanse National Health and Nutrition Examination Survey (NHANES studie), een dergelijk onderzoek op te zetten. Om in het beleid rekening te kunnen houden met de gezondheidstoestand van specifieke bevolkingsgroepen, zoals die van allochtone herkomst, zullen deze in dit onderzoek voldoende vertegenwoordigd moeten zijn.
- b *Garandeer de continuïteit van de registratie van zorggebruik in ziekenhuizen*
De registratie van zorggebruik in ziekenhuizen in de LMR is een belangrijke bron van informatie, niet alleen over zorggebruik, maar ook, in aanvulling op
-

andere gegevensbronnen, over de prevalentie van ziekten. De RGO onderschrijft het belang van het beleid van de minister van Volksgezondheid, Welzijn en Sport dat erop gericht is de continuïteit van de LMR, die door veranderingen in de organisatie van zorg in gevaar is gekomen, te garanderen.

c *Zorg voor adequate financiering van longitudinale epidemiologische gegevensverzamelingen*

De RGO meent dat longitudinale epidemiologische gegevensverzamelingen – essentieel als ze zijn voor het wetenschappelijk onderzoek en voor het verklaren van trends in de volksgezondheid – een publiek belang dienen. De raad beveelt daarom aan mogelijkheden te creëren om dit soort gegevensverzamelingen, wanneer ze actief benut worden en wetenschappelijk productief zijn, structureel te financieren. De implementatie van deze aanbeveling vereist een gezamenlijk optrekken van onderzoeksfinanciers als NWO, ZonMw en gezondheidsfondsen.

De raad meent dat de structurele financiering van longitudinale gegevensverzamelingen afhankelijk moet zijn van succesvol gebruik van de verzamelde gegevens. Nieuwe gebruikers van de gegevens zouden het hunne aan de instandhouding van de gegevensverzameling kunnen bijdragen door bij de verwerving van subsidie voor hun onderzoek een aparte post van serieuze omvang op te voeren als bijdrage aan de financiering van de infrastructuur van die gegevensverzameling. Dit vereist uiteraard bereidheid van onderzoeksfinanciers om langs deze weg aan de instandhouding van productieve gegevensverzamelingen bij te dragen.

3 Creëer meer ruimte voor verzoeken om collectieve financiering van nieuwe gegevensverzamelingen en beoordeel deze systematisch

Ook wanneer beschikbare gegevens optimaal worden gebruikt (aanbeveling 1) en (dreigende) tekorten worden opgeheven (aanbeveling 2) zullen er altijd nieuwe vragen uit wetenschap en beleid opkomen die nieuwe gegevensverzameling noodzakelijk maken. De RGO meent dat er ruimte moet zijn voor nieuwe initiatieven om te voorzien in nu nog grotendeels onbekende gegevensbehoeften. Daarbij moeten uiteraard de schaarse collectieve middelen doelmatig worden ingezet, al was het maar om zoveel mogelijk innovatieve ideeën te kunnen honoreren. De raad beveelt daarom aan alle verzoeken om collectieve financiering van nieuwe gegevensverzamelingen te beoordelen aan de hand van een aantal eisen van relevantie, noodzakelijkheid, kwaliteit en doelmatigheid (zie kader).

Ook verzoeken om subsidie voor nieuwe rondes van dataverzameling in lopend longitudinaal onderzoek kunnen aan de hand van deze eisen beoordeeld worden.

De RGO verwacht dat deze drie aanbevelingen ertoe kunnen bijdragen dat Nederland in de toekomst verzekerd zal zijn van de gegevens die voor effectief volksgezondheidsbeleid en hoogstaand wetenschappelijk volksgezondheids-onderzoek onontbeerlijk zijn.

Beoordelingskader voor verzoeken om collectieve financiering van nieuwe gegevensverzamelingen op het gebied van de volksgezondheid

Zijn de te verzamelen gegevens relevant?

De beoogde gegevensverzameling

- 1a is wetenschappelijk relevant, gezien de stand van de wetenschap en/of
- 1b voorziet in een gegevensbehoefte van het beleid.

Is de nieuwe gegevensverzameling noodzakelijk?

In de beoogde gegevensverzameling kan niet afdoende worden voorzien door

- 2 een bestaande of lopende Nederlandse gegevensverzameling.
- 3 een bestaande of lopende buitenlandse gegevensverzameling.
- 4 samenvoegen of koppeling van beschikbare Nederlandse gegevens.
- 5 samenvoegen of koppeling van beschikbare buitenlandse gegevens.
- 6 door aanvullende gegevensverzameling binnen een bestaande gegevensverzamelingsstructuur.

Voldoet de nieuwe gegevensverzameling aan eisen van kwaliteit en doelmatigheid?

- 7 De beoogde nieuwe gegevensverzameling voldoet aan de noodzakelijke kwaliteitseisen.*
- 8 De voorgestelde opzet is de meest kosten-effectieve manier om de beoogde gegevens te verzamelen.
- 9 De waarde van de nieuwe gegevensverzameling weegt op tegen de kosten.**
- 10 De voorgestelde opzet voldoet aan de voorwaarden voor financiering die door de RGO zijn voorgesteld ter bevordering van het delen van gegevens.***

* Voorbeelden van kwaliteitseisen zijn: met de voorgestelde onderzoeksopzet is het mogelijk de beoogde wetenschappelijke vraag te beantwoorden; de voorgestelde opzet is conform de huidige (inter)nationale stand van de wetenschap; controleerbaarheid, bijvoorbeeld toegankelijkheid voor audits. ** De waarde van een nieuwe gegevensverzameling kan behalve in de direct beoogde relevantie voor wetenschap of beleid gelegen zijn in, bijvoorbeeld, de toegevoegde waarde die ontstaat als de nieuw te verzamelen gegevens worden gecombineerd met, of gekoppeld aan, reeds beschikbare gegevens, of in de verwachting dat op basis van de nieuw te verzamelen gegevens toekomstige wetenschappelijke vragen te beantwoorden zullen zijn. *** Zie aanbeveling 1b in deze samenvatting.

Executive summary

Advisory Council on Health Research. Securing the data supply. The availability of population health information in the Netherlands, now and in the future. The Hague: Health Council of the Netherlands, 2008; RGO no. 58.

Effective public health policy and productive scientific research both depend on the availability of data on population health in the Netherlands. The supply of such data is currently not as good as it might be, however. Furthermore, the data that are available are not always utilised to best effect.

That is, in a nutshell, the background to this report by the Advisory Council on Health Research (Raad voor Gezondheidsonderzoek, RGO). In this report, the Council analyses the requirements for empirical data on the health of the Dutch population. The Council also seeks to clarify the extent to which existing data collections can satisfy these requirements, the shortcomings of such data collections and the steps that should be taken to ensure that existing and future data collections are utilised as efficiently as possible. The report concludes with a number of recommendations regarding ways of ensuring that, in the future, public health policy-makers, the health care sector and the scientific community have access to the data they require.

Data on population health are essential for policy management and scientific research

In order to pursue effective public health policies, the government needs to have information about trends in the prevalence of disease, risk factors, disabilities, care consumption and mortality. Data on the risk factors associated with common

conditions are necessary for various purposes, including the estimation of disease burden and healthy life expectancy in the future.

Furthermore, information about changes in the health of the general population is a prerequisite for scientific research into the causes of disease and aging, and for arriving at an understanding of the reasons for observed trends in population health. Such an understanding is necessary for the formulation of appropriate policies and for their subsequent evaluation.

Appropriate methods for generating the necessary data are already available

For the documentation and study of changes in population health, three methods of data collection are particularly important: (repeated) cross-sectional research, longitudinal epidemiological research and data registration. The repeated study of sufficiently large cross-sections of the general population is a good way of identifying trends in the prevalence of risk factors and disabilities. The prevalences of particular diseases and disabilities can be estimated using data from repeated cross-sectional research and registries. Registries can also yield information about mortality, causes of death and care consumption. Longitudinal epidemiological research – i.e. research in the context of which information is gathered on a given group of people at different points in time – is important mainly to support scientific research, including research designed to explain trends in population health.

Current research activities are unable to provide the data needed for policy management

There is presently no source of up-to-date representative data of a kind that can shed light on trends in important risk factors in the Netherlands, such as high blood pressure and serum cholesterol levels. As a result, the estimates of future disease burden and healthy life expectancy available to policy-makers are increasingly unreliable. The Health Examination Survey 2008 promises to go some way to providing appropriate data. However, it is limited in its scale and its (financial) continuity has yet to be assured. The Netherlands also lacks a robust source of data on trends in the prevalence of disabilities. By contrast, national trends in the prevalence of many diseases and medical conditions can be estimated from health care sector data registries, provided that their continuity is assured. The uncertainty that surrounds the future of the National Registry of

Hospital Admissions (Landelijke Medische Registratie, LMR) illustrates that continuity cannot be taken for granted.

The longitudinal data collections require maintenance

Longitudinal epidemiological data collections are essential for scientific research into the causes and courses of disease, aging and trends in population health. However, their scale and duration, and the quality requirements that the data must meet mean that such collections are expensive. Research groups have for many years found it difficult to maintain the infrastructure needed. If the Netherlands wishes to retain its best researchers, its strong scientific position and its ability to contribute actively to innovation in the care sector, the infrastructure of successful longitudinal epidemiological data collections will have to be secured.

Utilisation of the available data can be improved

For the reasons explained above, the collection of data on population health in the Netherlands needs to be continued and extended. However, it is also important to ensure that the data that *are* available are utilised to best effect. Data utilisation is not presently all that it might be.

The main routes to better data utilisation are as follows:

- Secondary analysis: re-using data to answer questions other than those that originally motivated collection.
- Pooling of data collected in the same way, but by different people in different places. This would have the effect of increasing the number of research participants and therefore the reliability of the research. Internationally, efforts are being made to combine biobanks and to pool other longitudinal epidemiological data collections. Pooling does, however, necessitate harmonised research methods.
- Linkage of data on various events in subjects' lives, which are recorded in different data files. Linking data in the register of causes of death to data on a population cohort provides an example. Linkage requires a 'linkage variable', to ensure that the correct individuals' data are being linked.

Data sharing is necessary, but caution should be exercised

The efficient use of data often depends on the body or research group that has collected the data sharing them with others. That is possible only if the privacy of the data subjects is protected, as required under the applicable legislation and

regulations. However, even when adequate protection can be provided, opportunities for data sharing are not always utilised – partly because of obstacles associated with competition within the scientific community.

The Council takes the view that data generated using public resources* should be available for research that is of public value, even if conducted by researchers unconnected to the original data collectors. Nevertheless, the interests of the researchers whose knowledge, skill and effort made the creation of a data file possible must be respected.

Recommendations

In light of the considerations described above, the RGO makes three recommendations to policy-makers, researchers and research funding bodies.

1 Promote the efficient use of data collections

a *Establish a register of data collections*

The Council recommends the establishment of a register of existing and new data collections in the area of population health, so that everyone can see what data are already held, and by whom.

b *Optimise access to data*

The RGO favours the formulation of a *code of conduct on data sharing*. The Council believes that the Dutch Federation of Biomedical Scientific Societies (FMWV) and/or the Royal Netherlands Academy of Arts and Sciences (KNAW) could play an important role in this context. Research funding bodies can contribute by making their support dependent on satisfaction of the following four conditions:

- Registration of the data collection.
- Subscription to the code of conduct on data sharing.
- The definition, in the funding application, of a procedure for providing outside access to the data.
- The use of validated standard test methods, except where properly justified.

* By 'public resources', the Council means not only public research funds (as provided by the government and health organisations), but also the government resources and premiums/contributions paid into collective schemes, which are invested in the data registration systems operated by, for example, Statistics Netherlands, health insurers and care providers.

- c *Maximise the scope for data sharing*
The RGO supports the appeal made to the government by the FMWV and KNAW, to allow the *Citizen Service Number* ('burgerservicenummer', BSN) to be used for scientific research purposes. The linkage of data from various sources would require the involvement of a *trusted third party*: a body that is independent both from the parties that maintain the separate data files and from the parties that make use of the linked data.
- d *Facilitate data sharing*
The Council believes that it should be made as easy as possible for researchers to share data. To this end, the Council would like to see the creation of an independent *data broker*, whose role would be to put parties seeking data in touch with parties in possession of data, and to assist the sharing of data. The Council also wishes to see the provision of *practical help and support* with the technical aspects of sharing, pooling and linking data. The Council recommends that research funding bodies make resources available for such facilitative activities.

2 Repair data shortages and prevent the development of new ones

- a *Provide for repeated cross-sectional health surveys on an appropriate scale*
The RGO recognises that useful data are generated and collected in significant volumes in the Netherlands. However, from the policy viewpoint, there is one clear shortcoming: no repeated cross-sectional surveys are carried out on a scale sufficient to provide a periodically updated picture of the prevalences of risk factors and disabilities. The Council accordingly recommends the establishment in the Netherlands of a programme modelled on the US National Health and Nutrition Examination Survey (NHANES). If policy is to take account of the health status of particular population groups, such as ethnic minorities, these groups must be properly represented in the research undertaken.
 - b *Ensure the continuity of systems for recording care consumption in hospitals*
The registration of care consumption in hospitals in the LMR serves not only as an important source of information about care consumption, but also as a supplementary source of disease prevalence data. The RGO supports the policy of the Ministry of Health, Welfare and Sports to secure the continuity of the LMR, which is threatened by changes in the way care is organised.
-

- c *Provide adequate funding for longitudinal epidemiological data collections*
Longitudinal epidemiological data collections are vital for scientific research into health. Without them, it is impossible to explain trends in population health. The RGO regards such data collections as serving an important public need and accordingly wishes to see systematic funding for those that are actively utilised and scientifically successful. The implementation of this recommendation would require coordinated action by research funding bodies such as the Netherlands Organisation for Scientific Research (NWO), Netherlands Organisation for Health Research and Development (ZonMw) and the health funds.

The Council believes that the funding of longitudinal data collections should depend on successful utilisation of the data. New users of the data could contribute to the maintenance of the collection they use by including in their funding application, as a *separate cost item*, provision for a generous contribution to the cost of the infrastructure that supports the collection. Of course, this requires that research funding bodies are willing to contribute in this way to the continuation of successful data collections.

3 Give more room to applications for the public funding of new data collections and assess them systematically

Even if existing data are put to optimal use (recommendation 1) and identifiable (existing and potential) blind spots are addressed (recommendation 2), the RGO believes that there will remain a need for initiatives designed to provide as yet unavailable data required by researchers and public health policy-makers. However, finite public resources need to be used efficiently, even if only to maximise the number of innovative ideas that are supported. The Council accordingly recommends that all applications for the public funding of new data collections should be assessed on the basis of a number of requirements relating to relevance, necessity, quality and efficiency. A proposed assessment framework is set out in the box below. Applications for the funding of new rounds of data collection within ongoing studies could be assessed on the basis of the same requirements.

The RGO expects that, together, these three recommendations will contribute to ensuring that in the Netherlands, in the future, the data necessary for effective public health policy and high-quality scientific health research are available.

Framework for the assessment of applications for the public funding of new public health data collections

Are the data to be collected relevant?

The proposed data collection:

- 1a is scientifically relevant, in the light of the scientific status quo and/or
- 1b meets a policy need.

Is the new data collection necessary?

The purpose to be served by the proposed data collection cannot be met adequately by:

- 2 an existing or current data collection in the Netherlands;
- 3 an existing or current data collection outside the Netherlands;
- 4 the combination or linkage of existing data in the Netherlands;
- 5 the combination or linkage of existing data outside the Netherlands; or
- 6 supplementary data collection within the context of an existing data collection system.

Will the new data collection satisfy the quality and efficiency requirements?

- 7 The proposed data collection satisfies the applicable quality requirements.*
- 8 The proposed data collection model is the most cost-effective option.
- 9 The cost of the new data collection will be justified by the benefit.**
- 10 The proposed model meets the criteria for funding proposed by the RGO to promote data sharing.***

* Applicable quality requirements include the following: the proposed research model must be capable of providing answers to the scientific questions addressed; the proposed research model must be consistent with the (inter)national scientific status quo; the proposed research model must provide scope for control, e.g. in the forms of audits. **The benefit of a new data collection may derive not only from its direct scientific or policy relevance, but also from, for example, the added value attainable by combination or linkage of the collected data with existing data, or from the expectation that the collected data will enable future scientific questions to be addressed. *** See recommendation 1b in this summary.

Inleiding

Begin 2007 signaleerden vertegenwoordigers van het ministerie van Volksgezondheid, Welzijn en Sport (VWS), de Nederlandse organisatie voor gezondheidsonderzoek en zorginnovatie (ZonMw), het Rijksinstituut voor Volksgezondheid en Milieu (RIVM) en de Raad voor Gezondheidsonderzoek (RGO) problemen met betrekking tot de beschikbaarheid van gegevens over de Nederlandse volksgezondheid. Gegevenstekorten die bleken bij de voorbereiding van de Volksgezondheid Toekomst Verkenning 2010, gebrek aan afstemming, samenwerking en synergie tussen gegevensverzamelingen en het ontbreken van een kader om aanvragen voor financiering van nieuwe gegevensverzamelingen op noodzakelijkheid te beoordelen vormden de hoofdmoot van die problemen. Deze signalen vormden voor de minister van VWS de aanleiding om een adviesaanvraag te richten tot de RGO (bijlage A).

1.1 Informatiebehoefte

De gezondheidstoestand van de Nederlandse bevolking verandert voortdurend onder invloed van bijvoorbeeld veroudering en migratie. De overheid ziet het als haar taak om het volksgezondheidsbeleid zodanig in te richten dat ongunstige trends in de volksgezondheid uitblijven of in de goede richting worden omgebo- gen. Om een dergelijk gericht volksgezondheidsbeleid te kunnen voeren is steeds actuele en betrouwbare informatie nodig over de volksgezondheid en veranderin-

gen daarin. Bovendien is kennis over de oorzaken van die veranderingen onontbeerlijk om aangrijpingspunten te vinden voor gericht beleid.

Op dit moment is de informatie die voor het beleid beschikbaar is ontoereikend. Recente informatie over trends in belangrijke risicofactoren voor veelvoorkomende ziekten ontbreekt. Zo zijn er bijvoorbeeld de afgelopen tien jaar geen epidemiologische gegevens meer verzameld over hoge bloeddruk en verhoogd cholesterol. Dit leidt tot toenemende onbetrouwbaarheid in de schattingen van ziektelast en gezonde levensverwachting in de toekomst. Voor andere informatie geldt dat de kwaliteit twijfelachtig is. Dat geldt bijvoorbeeld voor de gegevens over het vóórkomen van overgewicht in Nederland. Beleidsmakers moeten het voor de meeste leeftijdsgroepen doen met gegevens die gebaseerd zijn op zelfrapportage, terwijl bekend is dat dikke mensen geneigd zijn hun lichaamsgewicht te onderschatten. De vraag of er in ons land sprake is van een obesitas-epidemie en of Nederland in dit opzicht de Verenigde Staten achterna gaat, is hierdoor op dit moment in feite niet goed te beantwoorden. De snel veranderende samenstelling van de Nederlandse bevolking maakt het bovendien gewenst zicht te krijgen op trends in de gezondheidstoestand van specifieke bevolkingsgroepen, zoals jongeren, ouderen en etnische minderheden. Daarover is nu weinig bekend.

1.2 Benodigde infrastructuur

De gegevens over veranderingen in de volksgezondheid die nodig zijn voor het beleid komen vooral voort uit epidemiologische gegevensverzamelingen en uit registers waarin gegevens over bijvoorbeeld zorggebruik, sterfte en doodsoorzaken worden vastgelegd. Longitudinaal epidemiologisch onderzoek, in combinatie met biobanken, is bovendien essentieel voor het wetenschappelijk onderzoek naar de oorzaken en het beloop van ziekte en veroudering – een onderzoeksterrein waarop volgens de Raad voor Medische Wetenschappen (RMW) van de Koninklijke Nederlandse Academie van Wetenschappen (KNAW) voor Nederlands onderzoek goede kansen liggen.¹

Maar langlopend longitudinaal epidemiologisch onderzoek van voldoende omvang in de algemene bevolking of bij patiënten vergt investeringen. Om te beginnen van de burgers die, meestal steekproefsgewijs, worden uitgenodigd om aan het onderzoek mee te werken. En ook financieel is het een kostbare zaak, door de omvang en de duur van het onderzoek en door de hoge eisen die aan de kwaliteit van de gegevens moeten worden gesteld. Pas als het onderzoek jarenlang zorgvuldig is bijgehouden kan het licht werpen op veranderingen in de tijd. Langlopende gegevensverzamelingen van hoge kwaliteit zijn dan ook zowel voor volksgezondheidsbeleid en gezondheidszorg als voor het wetenschappelijk

onderzoek goud waard. De onderzoeksgroepen die in staat en bereid zijn dit soort onderzoek uit te voeren blijken in de praktijk echter al jaren moeilijkheden ondervinden bij de instandhouding van de infrastructuur van hun onderzoek.*

Uitgangspunt van de raad is dat een goede infrastructuur voor het verzamelen van gegevens over veranderingen in de Nederlandse volksgezondheid over de tijd voor de hele samenleving van groot belang is. Die infrastructuur zal de gegevensverzameling moeten ondersteunen, maar ook een doelmatig gebruik van de verzamelde gegevens. Daarbij gaat het niet alleen om gegevens uit onderzoek, maar ook om gegevens die in allerhande registers worden verzameld. De raad meent dat met collectieve middelen** gegenereerde gegevens die belangrijke inzichten kunnen opleveren, toegankelijk moeten zijn voor onderzoek, ook wanneer dat wordt uitgevoerd door anderen dan degenen die de gegevens hebben verzameld. Deze opvatting is in overeenstemming met de in 2007 door de Organisation for Economic Co-operation and Development (OECD) uitgebrachte principes en richtlijnen voor de toegang tot collectief gefinancierde onderzoeksgegevens.²

1.3 Over dit advies

Dit advies wordt uitgebracht door de RGO, die sinds 1 februari 2008 deel uitmaakt van de Gezondheidsraad (bijlage B). Het is voorbereid door een daartoe door de voorzitter van de RGO ingestelde commissie (bijlage C).

In het volgende hoofdstuk bespreekt de raad de typen gegevensverzamelingen waarop het advies in hoofdzaak betrekking heeft. In hoofdstuk 3 analyseert hij de behoefte aan empirische gegevens en aan kennis over veranderingen in de gezondheid van de Nederlandse bevolking vanuit het perspectief van volksgezondheidsbeleid en gezondheidszorg, en vanuit het perspectief van wetenschappelijk onderzoek. Hij gaat na in hoeverre bestaande gegevensverzamelingen hierin kunnen voorzien en op welke punten die tekortschieten. In hoofdstuk 4 staat de vraag centraal wat er moet gebeuren om beschikbare en nieuw te genereren gegevens beter te kunnen benutten. In het vijfde en laatste hoofdstuk doet de raad aanbevelingen die ertoe moeten leiden dat volksgezondheidsbeleid, gezond-

* Met infrastructuur wordt hier bedoeld: fysieke en organisatorische voorzieningen en menskracht die nodig zijn om gegevens en lichaamsmateriaal systematisch te verzamelen, te ordenen, te bewaren en weer terug te zoeken voor gebruik. Een dergelijke infrastructuur brengt altijd hoge kosten met zich mee.

** De raad verstaat hieronder niet alleen de publieke middelen (subsidies van overheid en gezondheidsfondsen) die worden aangewend voor onderzoek, maar ook de overheidsmiddelen en collectieve premies die geïnvesteerd worden in de registraties van bijvoorbeeld CBS, ziektekostenverzekeraars en zorgaanbieders.

heidszorg en wetenschap in de toekomst verzekerd zullen blijven van de gegevens die ze nodig hebben.

Bronnen van gegevens

Er zijn diverse bronnen die informatie kunnen verschaffen over de gezondheidstoestand van de Nederlandse bevolking en veranderingen daarin in de loop van de tijd. Niet alleen gegevensverzamelingen die direct betrekking hebben op risicofactoren, ziekte en beperkingen zijn van belang, ook onderzoek met een meer sociaal-wetenschappelijke invalshoek is een waardevolle bron van informatie. In dit hoofdstuk bespreekt de raad de bronnen van gegevens die met het oog op het documenteren en evalueren van veranderingen in de volksgezondheid het belangrijkste zijn: herhaald dwarsdoorsnede-onderzoek, longitudinaal epidemiologisch onderzoek en registraties.

2.1 Herhaald dwarsdoorsnede-onderzoek

In een dwarsdoorsnede-onderzoek vindt meting bij elke deelnemer slechts eenmaal plaats. Er is dus geen vervolgonderzoek bij dezelfde deelnemers. Wel kan men hetzelfde onderzoek van tijd tot tijd herhalen met telkens nieuwe deelnemers. Herhaald dwarsdoorsnede-onderzoek is een goede manier om trends in prevalenties van risicofactoren in de bevolking vast te stellen. Er moet dan wel aan een aantal voorwaarden zijn voldaan:

- de omvang van de steekproef moet voldoende groot zijn om uitspraken te kunnen doen over groepen in de bevolking die verschillen in kenmerken als leeftijd, geslacht, sociaal-economische status en etniciteit

- de respons van de mensen in de steekproef moet, in verband met de representativiteit van het onderzoek, voldoende zijn en in opeenvolgende herhalingen steeds van een vergelijkbaar niveau
- metingen van bijvoorbeeld lengte, gewicht en bloeddruk, en laboratoriumbepalingen, zoals serumcholesterol en bloedglucose, moeten volgens gestandaardiseerde methoden worden uitgevoerd.

Herhaald dwarsdoorsnedeonderzoek is informatief, maar niet geschikt om eventuele veranderingen, bijvoorbeeld in de prevalentie van risicofactoren voor veel-

Kader 2.1 Voorbeelden van herhaald dwarsdoorsnede-onderzoek

NHANES (the National Health and Nutrition Examination Survey in de VS) is een voorbeeld van herhaald dwarsdoorsnede-onderzoek. Het ging in 1959 van start. De eerste drie rondes vonden plaats in de periode 1959-62, 1963-65 en 1966-70. Dat waren zogeheten *Health Examination Surveys* met elk ongeveer 7500 deelnemers. Sindsdien zijn ook voedingspatronen onderzocht. Opeenvolgende surveys omvatten 16 000 tot 40 000 personen, soms met oververtegenwoordiging van bepaalde leeftijdsgroepen of bepaalde etnische groepen. Sinds 1999 is NHANES een doorlopend onderzoek. De verzamelde gegevens zijn afkomstig uit een interview met een lichamelijk onderzoek van de deelnemer en uit diens medische dossier. Het NHANES-onderzoek is vooral bekend door het aantonen van de trends in de prevalenties van overgewicht. NHANES heeft ook veel informatie opgeleverd over trends in prevalenties van gezondheidsgerelateerd gedrag, van diverse andere risicofactoren en van chronische aandoeningen. NHANES maakt tegenwoordig gebruik van een zogeheten *rotating sample design*. Dat combineert de voordelen van herhaald dwarsdoorsnede-onderzoek met individuele follow-up. Bij elke vervolgronde wordt een deel van de oorspronkelijke steekproef vervangen door een verse steekproef van dezelfde omvang. Na verloop van tijd is zo de hele oorspronkelijke groep vervangen.³

De Morgen studie ('Monitoring van Risicofactoren en Gezondheid in Nederland') is uitgevoerd in de periode 1993-1997 in Amsterdam, Doetinchem en Maastricht. Voorafgaand aan elk onderzoeksjaar werd een steekproef getrokken van mannen en vrouwen van 20-59 jaar uit de bevolkingsregisters. Personen die al eerder hadden meegedaan werden uitgesloten van deelname, behalve in Doetinchem, waar ook deelnemers aan een eerder project (Peilstations Hart- en Vaatziekten) werden uitgenodigd. In totaal zijn 23 000 personen onderzocht met vragenlijsten en lichamelijk onderzoek. De gegevensverzameling richtte zich op trends in prevalenties van biologische risicofactoren en leefstijl. Deze gegevens vormen essentiële input voor de wetenschappelijke modellering van de toekomstige ziektelast en ontwikkelingen in de gezonde levensverwachting.⁴

voorkomende ziekten, te *verklaren*. Daarvoor is onderzoek met herhaalde metingen op verschillende tijdstippen bij dezelfde personen noodzakelijk.

2.2 Longitudinaal epidemiologisch onderzoek

In longitudinaal epidemiologisch onderzoek ondergaan de deelnemers diverse metingen (tenminste twee) op verschillende tijdstippen. In de simpelste vorm betreft de eerste meting die van een factor waarvan men het verband met het ontstaan van een ziekte of andere gebeurtenis, bijvoorbeeld een botbreuk, wil onderzoeken (determinant) en de tweede, op een later tijdstip, het vaststellen of de onderzochte ziekte of gebeurtenis al of niet is opgetreden. Uitbreiding van het aantal te meten variabelen en het aantal meetmomenten, bijvoorbeeld om veranderingen in de determinant vast te stellen, maakt het onderzoek gecompliceerder, maar ook informatiever.

In een *klassiek cohort* onderzoek wordt de groep deelnemers in zijn geheel tot het eind van het onderzoek gevolgd. Anders dan bij een cohort, een gesloten populatie waaraan na het begin van het onderzoek niemand wordt toegevoegd, is bij een zogeheten *dynamische populatie* in- en uitstroom van onderzoeksdeelnemers mogelijk.

Een klassiek cohort wordt onvermijdelijk ouder en kleiner naarmate de tijd verstrijkt. Voor de bestudering van oorzaak-gevolg relaties is dat meestal niet van belang. Dat de groep deelnemers waarvoor vervolgmetingen beschikbaar zijn steeds minder representatief wordt voor de bevolking waaruit de steekproef is getrokken, is wel een bezwaar als men gegevens van vervolgmetingen uit dergelijk onderzoek wil gebruiken voor schattingen van prevalenties van risicofactoren voor ziekten en van functionele beperkingen. Dit bezwaar geldt niet voor dynamische populaties.

Longitudinaal epidemiologisch onderzoek is een van de beste methoden voor wetenschappelijk onderzoek naar interacties tussen genetische, gedrags- en omgevingsinvloeden op het ontstaan en het beloop van ziekten. Wanneer de onderzoekspopulatie bestaat uit patiënten met bepaalde ziektekenmerken spreekt men van een patiëntencohort. Er zijn in Nederland zeer veel patiëntencohorten. Voorbeelden zijn de Nederlandse Kankerregistratie⁵ en het recent gestarte Parel-snoer initiatief.⁶ Patiëntencohorten zijn vooral bruikbaar voor onderzoek naar het beloop van een bepaalde ziekte of naar de effectiviteit van interventies bij die ziekte. Ook voor het onderzoek naar neveneffecten van behandelingen en naar de kwaliteit van leven op de lange termijn worden ze steeds belangrijker. Omdat omgevings- en gedragsfactoren achteraf, als iemand de ziekte al heeft, niet meer betrouwbaar zijn vast te stellen, zijn patiëntencohorten minder geschikt voor de

Kader 2.2 Voorbeelden van longitudinaal onderzoek in de algemene bevolking

De Framingham Heart Study wordt beschouwd als de oervorm van het cohortonderzoek in de algemene bevolking. Het onderzoek stamt uit 1948 en had als doel de oorzaken van hart- en vaatziekten te achterhalen. Deelnemers waren 5209 mannen en vrouwen van 30 tot 62 jaar uit de Amerikaanse plaats Framingham. Na de eerste ronde (lichamelijk onderzoek en een leefstijlinterview) kregen zij elke twee jaar een oproep voor vervolgonderzoek. In 1971 werd een tweede-generatie-cohort gestart (5124 volwassen kinderen van de oorspronkelijke deelnemers en hun partners) en in 2005 het derde-generatie-cohort (4095 deelnemers). Met behulp van het Framingham-onderzoek zijn de belangrijke risicofactoren voor hart- en vaatziekten opgespoord. Meer dan 1200 wetenschappelijke artikelen zijn verschenen in toonaangevende tijdschriften. Onderzoekers uit de hele wereld werken samen met de Framinghamonderzoekers aan diverse vraagstukken⁷

Het Rotterdam onderzoek (ERGO) is gestart in 1990 onder 10 215 mannen en vrouwen van 55 jaar en ouder die in de Rotterdamse wijk Ommoord woonden, van wie er 7983 meededen, een respons van 78 procent. In 1999 is het cohort aangevuld met 3011 nieuwe deelnemers die inmiddels 55 geworden waren of 55-plussers die nieuw in Ommoord waren komen wonen. In 2006 is het cohort uitgebreid met ongeveer 4000 personen van 45-54 jaar. Het vervolgonderzoek vindt eens per 3-4 jaar plaats. Doel is de bestudering van oorzaken en beloop van veel voorkomende ouderdomsziekten: hart- en vaatziekten, diabetes, neurologische aandoeningen (ziekte van Parkinson, dementie), depressie en angststoornissen, osteoporose, en de oogziekten maculadegeneratie en glaucoom. Het onderzoek heeft belangrijke nieuwe inzichten in de oorzaken van deze ziekten opgeleverd, honderden wetenschappelijke publicaties en tientallen promoties.^{8,9}

Het Utrecht Gezondheidsproject (Leidsche Rijn) wordt uitgevoerd in een zogeheten dynamische populatie. Sinds 2000 wordt elke nieuwe inwoner die zich inschrijft bij een huisarts in de nieuwbouwwijk Leidsche Rijn bij Utrecht gevraagd deel te nemen. Van degenen die toestemmen wordt een individueel gezondheidsprofiel gemaakt op basis van lichamelijk onderzoek en gestandaardiseerde vragen. Ook wordt er lichaamsmateriaal opgeslagen. Verzameling van follow-up-gegevens vindt continu plaats via het elektronisch patiëntendossier van de huisarts en door koppeling aan zorgregisters. De respons is naar schatting 65 procent. Men verwacht jaarlijks 1500 tot 3000 nieuwe deelnemers. De follow-up duurt zolang als de deelnemers in Leidsche Rijn wonen.¹⁰⁻¹⁴

bestudering van factoren en hun interacties die van invloed zijn op het *ontstaan* van ziekten. Voor dat doel verdient het longitudinaal onderzoeken van groepen die niet zijn geselecteerd op met ziekte samenhangende kenmerken, zoals steekproeven uit de algemene bevolking, de voorkeur. De laatste jaren zijn voor de bestudering van factoren die van invloed zijn op het ontstaan van veelvoorko-

mende ziekten (hart- en vaatziekten, kanker) in binnen- en buitenland grote longitudinale onderzoeken in de algemene bevolking opgezet. In bijlage D staat een globaal overzicht van longitudinaal epidemiologisch onderzoek in de algemene bevolking in Nederland.

Longitudinaal epidemiologisch onderzoek wordt vaak gecombineerd met het afnemen van lichaamsmateriaal en de opslag daarvan in zogeheten biobanken. Materiaal uit een biobank is later opnieuw te analyseren om nieuwe onderzoeksvragen te beantwoorden.

2.3 Registraties

Bepaalde gegevens die van belang zijn voor de bestudering van trends in de volksgezondheid worden standaard geregistreerd. Van belang in dit verband zijn de registraties van sterfte en doodsoorzaken (kader 2.3), en diverse zorgregistraties, zoals huisartsenregistraties (kader 2.3 en bijlage D), registraties van ziekenhuisopnames en apotheekregistraties. Veel longitudinaal epidemiologisch onderzoek maakt gebruik van koppeling aan bestaande registraties voor de follow-up van de onderzoeksdeelnemers en voor de bestudering van het zorggebruik. Ook is het mogelijk door koppeling van bestaande registraties patiëntengeschiedenissen te reconstrueren en zo op basis van registratiegegevens een cohort op te bouwen.^{14,15}

Kader 2.3 Voorbeelden van registraties

Doodsoorzakenstatistiek

Bij elk sterfgeval vult een arts een doodsoorzaakverklaring in. Alle doodsoorzaakverklaringen worden door het CBS (Centraal Bureau voor de Statistiek) bewerkt. Ze vormen de basis voor de doodsoorzakenstatistiek, die het CBS wettelijk verplicht is bij te houden (art. 12a van de Wet op de Lijkbezorging). Ook in World Health Organisation (WHO) verband heeft Nederland zich hiertoe verplicht. De doodsoorzakenstatistiek bevat gegevens over de primaire doodsoorzaak van alle in het Nederlands bevolkingsregister ingeschreven overledenen. Geaggregeerde gegevens zijn kosteloos te raadplegen op de website van het CBS.¹⁶ Onder voorwaarden is het voor wetenschappelijk onderzoekers mogelijk om de gegevens on site op het CBS of via remote access te analyseren. Het register van doodsoorzaken is van onschatbare waarde voor beleid en wetenschap op het gebied van de volksgezondheid.

Landelijk Informatie Netwerk Huisartsenzorg (LINH)

LINH is een netwerk van 85 huisartspraktijken met bijna 340 000 ingeschreven patiënten. De LINH-huisartsen verzamelen continu gegevens over aandoeningen, aantallen contacten, voorschriften van geneesmiddelen en verwijzingen van deze patiënten, zoals die worden vastgelegd in het Elektronisch Patiënt Dossier (EPD) van de huisarts.

Het doel van LINH is representatieve continue informatie over de door Nederlandse huisartsen geleverde zorg te genereren, voor de onderbouwing van het landelijk beleid van de huisartsen en voor wetenschappelijk onderzoek, zowel ten behoeve van het kwaliteitsbeleid van de beroepsgroep als voor algemene (beleids)vragen op het gebied van de volksgezondheid.¹⁷

Gegevensbehoeften en beschikbaarheid

De Nederlandse overheid heeft behoefte aan gegevens over de volksgezondheid, maar aan welke precies? En aan welke de wetenschap? Welke gegevensbronnen zijn beschikbaar en waar zijn of dreigen tekorten? Wie, afgezien van wetenschap en overheid, hebben er nog meer belang bij gegevens over de volksgezondheid? Over die vragen gaat dit hoofdstuk.

3.1 Behoeften van het beleid

Voor het voeren van een adequaat volksgezondheidsbeleid moet de overheid kunnen beschikken over gegevens die actueel, betrouwbaar en representatief zijn. Om het beleid richting te geven is het bovendien van belang op de hoogte te zijn van de gezondheid van specifieke groepen, zoals de inwoners van bepaalde regio's, mensen met diverse sociaal-economische status en groepen van verschillende etnische herkomst. Ook is informatie nodig over de kosten en de effecten op de volksgezondheid van preventieve en curatieve interventies. Goede voorbeelden van beleidsrelevante informatie bieden de Volksgezondheid Toekomst Verkenning¹⁸ en de Zorgbalans.¹⁹ De waarde van deze documenten is echter afhankelijk van regelmatige actualisering van de gegevens waarop ze zijn gebaseerd.

Om een goed beeld te hebben van de ontwikkelingen in de gezondheid van de Nederlandse bevolking is het nodig informatie te hebben over trends in:

- de prevalentie van bekende risicofactoren voor grote volksgezondheidsproblemen zoals hart- en vaatziekten, kanker, aandoeningen van het bewegingsapparaat en chronische luchtwegaandoeningen: roken, serumcholesterol, bloedglucose, hoge bloeddruk, longfunctie, overgewicht, bewegingsarmoede en ongezonde voeding
- de prevalentie en incidentie van lichamelijke en psychische ziekten en aandoeningen en de sterfte
- de prevalentie en incidentie van functionele beperkingen en het zorggebruik.

Om vervolgens beleidsmaatregelen te kunnen toespitsen en om genomen maatregelen te evalueren is het nodig om inzicht te hebben in de factoren (determinanten) die de waargenomen trends verklaren.

3.2 Beschikbaarheid van gegevens voor het beleid

Prevalenties van risicofactoren

Er is op dit moment in Nederland geen gegevensbron die de beschikbaarheid garandeert van actuele en representatieve gegevens over prevalenties van risicofactoren voor veelvoorkomende ziekten. Wel is het RIVM begin 2008 begonnen met de voorbereidingen voor een nieuwe *Health Examination Survey* (HES) bij 8 000 personen van 18-70 jaar met als doel de prevalentie van overgewicht te achterhalen. Deze HES heeft enkele beperkingen. Zo is de steekproefomvang tamelijk klein als het erom gaat de prevalenties per sociaal-economische klasse betrouwbaar te schatten en zijn personen van allochtone herkomst schaars vertegenwoordigd. De gedachte is om deze gegevensverzameling te laten aansluiten bij een ander initiatief, namelijk dat van de Locale en Nationale Monitor Gezondheid.²⁰ Dit project houdt in dat lokale Gemeentelijke Gezondheidsdiensten (GGD-en) hun – in principe met gestandaardiseerde meetmethodes verzamelde – gegevens invoeren in een centrale database. Zo kunnen in de toekomst gegevens van grote aantallen mensen worden verzameld.

Grote longitudinale epidemiologische gegevensverzamelingen in de algemene bevolking die uit wetenschappelijke motieven zijn opgezet, hebben meestal niet als expliciete doelstelling trends in de prevalentie van risicofactoren te onderzoeken. Bepaalde onderzoeken die nu lopen in ons land zouden op dit punt een bijdrage kunnen leveren, maar de deelname aan deze onderzoeken ligt doorgaans in de orde van 30 tot 60 procent, wat een probleem is als het erom gaat trends in prevalentiecijfers voor risicofactoren in de algemene bevolking te

schatten. Herhaald dwarsdoorsnede-onderzoek (2.1) zou soelaas kunnen bieden, al is het niet vanzelfsprekend dat de respons in dat soort onderzoek hoger is.

Prevalenties van ziekten en aandoeningen

Voor zover prevalenties van ziekten en aandoeningen met patiëntenregistraties niet betrouwbaar genoeg gemeten kunnen worden, is herhaald dwarsdoorsnede-onderzoek in de algemene populatie in veel gevallen een goede methode om prevalenties van ziekten en aandoeningen in de algemene bevolking te schatten. Daar zijn wel enkele voorwaarden en beperkingen aan verbonden. De gestandaardiseerde meetmethoden die nodig zijn om vast te stellen of een persoon een bepaalde ziekte wel of niet heeft maakt het onderzoek arbeidsintensief en dus kostbaar. Bovendien zijn de meeste ziekten in de algemene bevolking relatief zeldzaam, zodat voor betrouwbare schattingen een zeer grote steekproefomvang nodig is. In de praktijk is herhaald dwarsdoorsnede-onderzoek daardoor alleen goed bruikbaar als methode om prevalenties te meten bij veelvoorkomende ziekten, zoals diabetes en hart- en vaatziekten bij ouderen, en luchtwegaandoeningen bij kinderen. Voor de minder frequente ziekten en aandoeningen is het niet haalbaar om trends in prevalenties te verkrijgen door middel van onderzoek in de algemene populatie.

De belangrijkste bronnen van gegevens over trends in prevalenties van ziekten en aandoeningen zijn op dit moment de zorgregistraties, zoals huisartsenregistraties, de Landelijke Medische Registratie (LMR), de Perinatale Registratie Nederland^{14,21} en patiëntenregistraties zoals de Nederlandse Kankerregistratie.²² Voor infectieziekten vormen surveillancegegevens de bron. Prevalentiecijfers die afgeleid zijn uit huisartsenregistraties verschillen overigens van die van ziekenhuizen, doordat ze betrekking hebben op verschillende selecties van patiënten en daardoor ook minder ernstige ziekte-episodes in beeld brengen. Voor goed omschreven ziekten en aandoeningen waarvan verondersteld kan worden dat vrijwel alle patiënten via zorgcontacten in een zorgregistratie terecht komen, zoals bij kanker, zijn epidemiologische kengetallen goed af te leiden uit zorgregistraties. Voor ziekten en aandoeningen waarbij kans op onderbehandeling bestaat of waarvoor mensen niet of niet altijd contact met de zorg zoeken, zoals gewrichtsaandoeningen of depressie, is aanvullende informatie nodig. Die kan bijvoorbeeld komen uit onderzoek in de algemene bevolking, op basis van zelfrapportage. De interpretatie van zulke gecombineerde gegevens kan lastig zijn.

Prevalenties van functionele beperkingen

Hiervoor geldt hetzelfde als voor trends in prevalenties van risicofactoren. Idealerweise worden functionele beperkingen gestandaardiseerd gemeten in herhaald dwarsdoorsnede-onderzoek in de (oudere) algemene populatie, inclusief mensen die verblijven in een ziekenhuis, verzorgings- of verpleeghuis. Exclusie van mensen die in zorginstellingen verblijven leverde bijvoorbeeld problemen op bij schattingen van de frequentie van multimorbiditeit²³ en visus- en gehoorbeperkingen.^{23,24} Evenals voor prevalenties van risicofactoren is er in ons land geen robuuste gegevensbron voor prevalenties van functionele beperkingen. Zorgregistraties zijn maar beperkt bruikbaar, omdat gegevens over functionele beperkingen hierin niet worden vastgelegd. RIVM, de universitair medische centra van Vrije Universiteit en Erasmus Universiteit Rotterdam en het Sociaal en Cultureel Planbureau (SCP) zijn bezig alle beschikbare gegevens over trends in beperkingen bij ouderen te combineren in het zogeheten TRENDSDS project.²⁵ Zij verwachten de eerste resultaten eind 2008.

Determinanten van trends in prevalenties

Longitudinaal onderzoek in de algemene bevolking, dus onderzoek met individuele follow-up, is de ideale onderzoeksopzet om licht te werpen op verklaringen van gesignaleerde veranderingen in prevalenties van risicofactoren, ziekten en aandoeningen en functionele beperkingen (2.1). Er zijn in ons land geen gegevensverzamelingen die zich specifiek richten op het verklaren van trends alleen ten behoeve van de evaluatie van volksgezondheidsbeleid. Het ligt ook niet voor de hand om dergelijke verzamelingen los van specifieke wetenschappelijke vraagstellingen op te zetten. Het verklaren van trends is immers een wetenschappelijke aangelegenheid.

Natuurlijk wordt daarbij teruggegrepen op bestaande wetenschappelijke inzichten, die voor een belangrijk deel gebaseerd zijn op eerder longitudinaal epidemiologisch onderzoek. Zo is het bijvoorbeeld denkbaar dat trends in serumcholesterol – gegeven de algemeen aanvaarde effecten op het serumcholesterol van voedingsgewoonten en cholesterolverlagende middelen – verklaard kunnen worden door gegevens over het voorschrijven van cholesterolverlagende middelen en gegevens uit de Voedselconsumptiepeilingen²⁶ met elkaar te combineren, zonder te beschikken over op die vraagstelling gericht aanvullend, in Nederland uitgevoerd longitudinaal onderzoek.

In het algemeen zijn echter de mogelijkheden om trends in de Nederlandse volksgezondheid puur te verklaren op grond van verbanden die in buitenlands

longitudinaal onderzoek zijn gevonden, beperkt. Elders geïdentificeerde determinanten van ziekten die losstaan van lokale cultuur en milieublootstellingen kunnen van toepassing zijn op de Nederlandse situatie. Maar voorzichtigheid is op zijn plaats. In het buitenland gevonden genetische determinanten, bijvoorbeeld, hoeven niet zonder meer relevant te zijn in de Nederlandse situatie, aangezien de genetische samenstelling van de bevolking van Azië of de Verenigde Staten anders is dan die van Nederland. Buitenlands onderzoek kan bepaalde verbanden tussen determinanten en uitkomsten wel aannemelijk maken, bijvoorbeeld een verband tussen veranderingen in voedings- en bewegingspatronen en trends die zich voordoen in de prevalenties van hart- en vaatziekten. Maar in aanvulling daarop is onderzoek naar de verdeling van die determinanten in de Nederlandse bevolking noodzakelijk om inzicht te krijgen in de relatieve invloed van juist die risicofactoren op de Nederlandse volksgezondheid. Bovendien kunnen in de Nederlandse bevolking specifieke determinanten en interacties in het spel zijn – gelieerd aan bijvoorbeeld in Nederland veelvoorkomende leefgewoonten en genetische kenmerken – die uit onderzoek in andere landen nog niet naar voren zijn gekomen. Aan de ontrafeling van die verbanden en interacties kan in Nederland uitgevoerd longitudinaal epidemiologisch onderzoek dan een belangrijke bijdrage leveren. Hoe op de gevonden verbanden beleidsmatig het best kan worden ingespeeld en of eerder beleid succesvol is geweest, kan alleen in de Nederlandse populatie worden onderzocht. Hetzelfde geldt voor effecten van het Nederlandse zorgsysteem.

3.3 Verhouding tussen beleidsbehoefte en beschikbaarheid

Willen gegevens over trends in prevalenties van risicofactoren, ziekten en aandoeningen, en functionele beperkingen bruikbaar zijn voor het beleid, dan moeten ze betrouwbaar en representatief zijn en regelmatig geactualiseerd worden. Uit de lopende longitudinale epidemiologische gegevensverzamelingen komen die gegevens niet zonder meer voort. Gegevens uit dwarsdoorsnede-onderzoek zijn informatief voor volksgezondheidsbeleid, maar geven geen inzicht in factoren die gesignaleerde trends verklaren. Deze beperking kan deels worden overvaren door een variant van herhaald dwarsdoorsnede-onderzoek waarbij minstens twee metingen per persoon worden gedaan, met een tijdsperiode ertussen en met geleidelijke verversing van de steekproef (de methode van de *rotating samples*, zie kader 2.1), en door de deelnemers toestemming te vragen voor koppeling van hun gegevens aan registraties van bijvoorbeeld ziekenhuisopnames en doodsoorzaken. Op die manier worden in feite longitudinale gegevensverzamelingen gecreëerd.

Gegevens uit longitudinale onderzoeken in de algemene bevolking zijn, afgezien van de uitkomsten van de eerste meting, doorgaans beperkt bruikbaar voor prevalentieschattingen. De bruikbaarheid neemt toe wanneer men bij de opzet van grote (dynamische) populatie-onderzoeken expliciet rekening mee houdt dat de gegevens ook bestemd zijn voor prevalentieschattingen ten behoeve van het beleid. Dat kan bijvoorbeeld door zorg te dragen voor oververtegenwoordiging van deelpopulaties met wier gezondheidstoestand men in het beleid rekening wil kunnen houden. Op deze manier kunnen beleidsmatige en wetenschappelijke motieven elkaar versterken en tot grotere doelmatigheid van investeringen leiden.

De beschikbare gegevensbronnen overziend stelt de raad vast dat er op dit moment geen bron is van actuele en representatieve gegevens die goed zicht geven op trends in risicofactoren van veelvoorkomende ziekten in Nederland. De in 3.2 genoemde HES is een begin, maar is beperkt van opzet en vooralsnog ontbreekt de (financiële) continuïteit. Ook kan zoals gezegd het combineren van gegevens uit verschillende bronnen soms soelaas bieden (3.2). Dit stelt hoge eisen aan de aard en kwaliteit van de gegevens. Bovendien moeten de gegevens, die met een ander doel verzameld zijn, toegankelijk zijn voor derden.

Aan het verzamelen van gegevens met als doel trends in risicofactoren te signaleren ter ondersteuning van het gezondheidsbeleid moeten de volgende eisen worden gesteld:

- De steekproefomvang moet groot zijn, mogelijk enkele tienduizenden deelnemers, zodat het mogelijk is de gegevens te analyseren naar leeftijd, geslacht, sociaal-economische status en etnische achtergrond; vanwege de ongelijke geografische spreiding vraagt dat laatste bijzondere aandacht.
- De waarnemingen moeten bestaan uit een combinatie van de uitkomsten van vragenlijstonderzoek, metingen aan het lichaam (lengte, gewicht, bloeddruk *et cetera*) en laboratoriumbepalingen aan lichaamsmateriaal; het is aan te bevelen het lichaamsmateriaal op te slaan in een biobank voor eventueel later onderzoek.
- Metingen moeten worden verricht volgens gestandaardiseerde meetmethoden.
- De uitvoering moet in handen zijn van daarin gespecialiseerde professionals, die niet gericht zijn op het zelf verrichten van wetenschappelijk onderzoek, maar wel in intensief contact staan met en begeleid worden door wetenschappelijk onderzoekers; dit bevordert de kwaliteit, de continuïteit en de beschikbaarheid van de gegevens.

- Nieuwe gegevens moeten periodiek, in verband met de beleidscyclus bijvoorbeeld eens in de vier jaar, beschikbaar komen; dat kan door doorlopend elk jaar de gegevens van 25 procent van de deelnemers te verzamelen.

Voor het schatten van trends in prevalenties van ziekten en aandoeningen is, zoals eerder besproken, het gebruik van gegevens uit zorg- en patiëntenregisters als basis voor beleidsinformatie niet zonder problemen. Onder voorbehoud van de in 3.2 gemaakte opmerkingen concludeert de raad met enige voorzichtigheid dat het schatten van trends in prevalenties van veel ziekten en aandoeningen op basis van zorgregistraties in Nederland goed mogelijk is, mits de continuïteit ervan gewaarborgd is. Veranderingen in organisatie en financiering van de zorg kunnen die continuïteit, en daarmee de bruikbaarheid van de registraties als gegevensbron voor epidemiologisch onderzoek, in gevaar brengen. In Nederland zijn huisartsenregistraties belangrijke gegevensbronnen. Anders dan in andere landen zijn de mensen hier op naam ingeschreven, zodat de omvang van de populatie waarin de geregistreerde ziektegevallen zich voordoen (de noemer van de ‘epidemiologische breuk’) bekend is. Dat is een groot goed, maar nu, bijvoorbeeld, patiënten zonder verwijzing van hun huisarts naar de fysiotherapeut kunnen gaan, geven prevalentieschattingen van aandoeningen van het houdings- en bewegingsapparaat op basis van huisartsenregistraties een onderschatting. Een ander voorbeeld van een zorgregistratie waarvan de bruikbaarheid voor onderzoeksdoeleinden lijdt onder organisatorische veranderingen in de zorg is de Landelijke Medische Registratie (LMR). Sinds de invoering van de Diagnose-Behandel Combinaties (DBC’s) zijn kwaliteit en continuïteit van de LMR niet meer gegarandeerd. Of met behulp van de DBC-registratie een alternatief van vergelijkbare kwaliteit en continuïteit ontstaat is nog niet duidelijk. Ook weten we niet welke kosten gemoeid zouden zijn met het gebruik van de DBC-gegevens voor dit doel. De raad gaat ervan uit dat de vergoeding voor nader gebruik van met collectieve middelen vergaarde gegevens (zie de tweede voetnoot bij 1.2) voor wetenschappelijke of beleidsdoeleinden zich beperkt tot de dekking van de extra kosten die de gegevensbeheerder met dat doel moet maken voor de bewerking van de gegevens.

Als alternatieve of aanvullende bron van gegevens over trends in ziekten en aandoeningen kan men ook denken aan verzekeringsgegevens. Zorgverzekeraars beschikken over veel gegevens die in dit opzicht van belang kunnen zijn, waarvan tot nu toe weinig gebruik wordt gemaakt. Ook het elektronisch patiëntendossier zou van waarde kunnen zijn, maar de realisatie daarvan verloopt moeizaam. Voor trends in prevalenties van functionele beperkingen biedt de aanpak van het TRENDDES-onderzoek mogelijk perspectief.^{24,25}

3.4 Gegevensbehoefte en beschikbaarheid vanuit wetenschappelijk perspectief

Empirisch wetenschappelijk onderzoek is weliswaar altijd in zekere mate tijd- en plaatsgebonden, maar het hoofddoel is vrijwel steeds het genereren van algemeen geldige kennis. Epidemiologisch wetenschappelijk onderzoek naar verklaringen voor het ontstaan en het beloop van ziekten en functionele beperkingen is gericht op het ontrafelen van de effecten van genetische factoren, gedrag en omgeving – waaronder medische zorg – en hun interacties. Longitudinale epidemiologische gegevensverzamelingen met biobanken zijn daarvoor noodzakelijk. Behalve biobanken die zijn opgezet om inzicht te krijgen in genetische oorzaken van ziekten, zijn bijvoorbeeld ook longitudinale onderzoeken met een meer sociaal-epidemiologische insteek van belang.

Voor onderzoek naar de oorzaken van veelvoorkomende aandoeningen die ontstaan door combinaties van genetische aanleg, ‘spontane’ mutaties in een of meer genen, omgevingsfactoren en leefstijl, de zogeheten multifactoriële aandoeningen, is er behoefte aan algemene biobanken die lichaamsmateriaal en informatie (over fenotype, gedrag, omgeving) van een deel van de bevolking bevatten en die zich niet beperken tot één ziekte. De KNAW meent dat er goede kansen en mogelijkheden liggen voor het Nederlandse onderzoek op dit gebied.²⁷

In ons land lopen flink wat longitudinale epidemiologische gegevensverzamelingen in de algemene bevolking, waarvan vele met een biobank, en gesystematiseerde registraties in huisartspraktijken (bijlage D).²⁸ De gegevens uit deze onderzoeken zullen nog vele jaren van belang blijven om wetenschappelijke vragen te beantwoorden. Een gedetailleerde inhoudelijke analyse van witte vlekken en overlappen in de lopende gegevensverzamelingen valt buiten het bestek van dit advies. In kwalitatieve zin merkt de raad wel op dat er verschillende onderzoeken onder jeugdigen lopen. Replicatie kan wetenschappelijk gezien meerwaarde hebben, maar door betere samenwerking en afstemming zou de opbrengst bij nagenoeg gelijkblijvende investeringen wel eens veel groter kunnen zijn. Een positief punt is dat in de longitudinale onderzoeken onder de jeugd kinderen van diverse etnische achtergrond goed zijn vertegenwoordigd. Longitudinaal epidemiologisch onderzoek onder volwassenen van allochtone herkomst ontbreekt.

3.5 Wetenschappelijke gegevensbehoeften in de komende tien jaar

Wetenschap kent een eigen dynamiek, die het onmogelijk maakt om nu te bedenken welke wetenschappelijke vragen zich over tien jaar zullen voordoen en welke gegevens daarvoor nodig zullen zijn. Wil Nederland zijn positie als kennisland behouden en een serieuze speler blijven op het Europese onderzoekstoneel, dan is het niet voldoende om lopende longitudinale epidemiologische gegevensverzamelingen te onderhouden, onderzoeksdeelnemers te blijven volgen en onderzoekspopulaties aan te vullen met bijvoorbeeld jongere leeftijdsgroepen. Er zullen ook nieuwe longitudinale epidemiologische gegevensverzamelingen nodig zijn voor nieuwe vragen.

Verzamelingen van longitudinale gegevens over de algemene bevolking en patiënten zijn essentiële instrumenten van wetenschappelijk onderzoek. Ze vormen een belangrijk onderdeel van de onderzoeksinfrastructuur die Nederland nodig heeft om zijn goede onderzoekers en sterke wetenschappelijke positie vast te houden en actief te kunnen bijdragen aan zorginnovatie.^{*29} Ook participatie, via nationale initiatieven als Parelsnoer³⁰ en Lifelines³¹ in internationale consortia zoals *Biobanking and Biomolecular Resources Infrastructure (BBMRI)*³² en *the Emerging Risk Factors Collaboration*³³ is goed voor de positie van de Nederlandse wetenschap.

De waarde van longitudinale epidemiologische gegevensverzamelingen in de algemene bevolking en bij patiënten is vooral groot als er gegevens van zeer hoge kwaliteit zijn verzameld gedurende vele jaren. Dat kan alleen als de continuïteit van zo'n gegevensverzameling gegarandeerd is. Het gebeurt nog te vaak dat voor een nieuwe vraagstelling een nieuw cohort wordt gecreëerd dat vervolgens onvoldoende wordt benut en waarvan de lange termijn follow-up moeilijk te financieren blijkt te zijn. Het moet mogelijk zijn om tot een meer geconcentreerde aanpak te komen waarbij de beschikbare middelen efficiënter en effectiever worden ingezet en onderzoekers gegevensverzamelingen gezamenlijk beheren en benutten. Afgezien van de garanties voor het onderhoud van lopende gegevensverzamelingen, is het daarvoor noodzakelijk condities te creëren die toegankelijkheid van gegevens vergroten. Ten slotte dient er uiteraard ruimte te blijven voor nieuwe en innovatieve initiatieven die niet met bestaande gegevensverzamelingen kunnen worden uitgevoerd. Door bestaande gegevensverzamelingen efficiënter te benutten, ontstaat er meer ruimte voor innovatie. Nieuwe

* De term 'onderzoeksinfrastructuur' verwijst hier dus niet naar de infrastructuur van de afzonderlijke (longitudinale) onderzoeken, zoals in de eerste voetnoot bij 1.2.

initiatieven zullen nodig blijven om toekomstige, nog onbekende, wetenschappelijke vragen te beantwoorden.

3.6 Belanghebbenden

Behalve de wetenschap en de overheid (rijk, provincie, gemeente) zijn er tal van andere instanties en personen die belang hebben bij een adequate verzameling van gegevens over de Nederlandse volksgezondheid: zorgaanbieders, gezondheidsfondsen zoals KWF Kankerbestrijding en de Nederlandse Hartstichting, zorgverzekeraars, bedrijven, patiënten en consumenten, de Nederlandse burger (vertegenwoordigd in het parlement) en steeds vaker ook Europese gremia zoals Eurostat. De belangstelling van patiëntenorganisaties om zelf longitudinaal epidemiologisch onderzoek met patiënten op te zetten, al of niet met een biobank, neemt toe, zoals blijkt uit de landelijke hemofiliestudie,³⁴ het International Pompe Register^{35,36} en registers van patiënten met cystische fibrose (taaislijmziekte).^{37,38}

De genoemde partijen hebben niet allemaal behoefte aan dezelfde gegevens, maar ze hebben wel allemaal belang bij goede voorwaarden voor gegevensverzameling en bij een efficiënt gebruik van gegevens. De commissie signaleert dat het gebruik door commerciële partijen van onderzoeksgegevens die met collectieve middelen verzameld zijn, specifieke vragen oproepen. Deze behoeven zeker aandacht, maar vallen buiten het bestek van dit advies.

Gegevens beter benutten

Er zijn verschillende manieren om gegevensverzamelingen beter te benutten. In de meeste gevallen zal het daarvoor nodig zijn dat degenen die de gegevens verzameld hebben ze delen met anderen. Om een betere benutting van gegevensverzamelingen te bevorderen zal in de technische, juridische en financiële voorwaarden voor het delen van gegevens moeten worden voorzien.

4.1 Methoden voor betere benutting

De belangrijkste manieren om beter gebruik te maken van gegevensverzamelingen zijn secundaire analyse, samenvoeging en koppeling.

Secundaire analyse

Dit is het opnieuw analyseren van gegevens met een andere vraagstelling dan waarvoor ze oorspronkelijk zijn verzameld. Dat is alleen mogelijk als de onderzoeker of de onderzoeksinstelling die de secundaire analyses wil doen, kan beschikken over de originele gegevens, ook als die oorspronkelijk door anderen zijn verzameld.

Samenvoeging

Een tweede manier om gegevensverzamelingen beter te benutten is het samenvoegen ('poolen') van gegevens die op verschillende plekken over hetzelfde onderwerp door verschillende onderzoekers of instellingen zijn verzameld. Het doel is het aantal onderzoeksdeelnemers en daarmee de betrouwbaarheid van het onderzoek te vergroten. Samenvoeging van bestanden is bijvoorbeeld aan de orde in het onderzoek naar gecombineerde visus- en gehoorbeperking in Nederland met gebruikmaking van gegevens uit acht onderzoeken onder ouderen,²⁴ onderzoek naar multimorbiditeit,^{23,39} het Parelsnoer Initiatief⁴⁰ en de Locale en Nationale Monitor Gezondheid.⁴¹ In de wereld van de biobanken is men zeer overtuigd van het nut van het combineren van gegevens. Er zijn diverse internationale projecten die harmonisering en samenvoegen van biobanken tot doel hebben, zoals het Public Population Project in Genomics (P3G)⁴² en de Biobanking and Biomolecular Resource Research Infrastructure (BBMRI).⁴³ Een ander voorbeeld is Gabriel, een onderzoek dat is opgezet om genetische en omgevingsdeterminanten van allergie en astma op te sporen. Geboortecohorten uit meer dan tien Europese landen dragen gegevens bij over omgevingsfactoren, diagnose en DNA die op identieke manier zijn verzameld. Het doel is genetische varianten te identificeren die in interactie met bepaalde omgevingsfactoren tot astma of allergie leiden. Het project bevat de gegevens van tienduizenden kinderen die deelnemen vanaf hun geboorte.⁴⁴

Koppeling

Koppeling van gegevens over dezelfde persoon die in diverse onderzoeken of registers zijn verzameld, is een derde manier om gegevensverzamelingen beter te benutten. Vaak worden gegevens uit een cohortonderzoek door koppeling aan gegevens uit zorgregistraties en doodsoorzakenregistratie uitgebreid met informatie over het optreden van ziekte of sterfte. Met dezelfde technieken voor gegevenskoppeling kunnen ook cohorten geconstrueerd worden op basis van gegevens uit bestaande registraties (zie bijvoorbeeld de Landelijke Perinatale Registratie Nederland).⁴⁵ Dat is ook het doel van het Mondriaan-project van Top Instituut Pharma.⁴⁶

Gegevenskoppeling is een manier om onderzoek te doen naar het optreden van late of relatief zeldzame bijwerkingen van veelvoorkomende ingrepen of blootstellingen aan bepaalde stoffen (kader 4.1). Het vermoeden van een attente behandelaar dat een bepaalde behandeling op de lange termijn wel eens schade-

Kader 4.1 Borstimplantaten. Voorbeeld van onderzoek dat alleen mogelijk was door koppeling van verschillende registers op individueel niveau

Het onderzoek naar mogelijke schadelijke gevolgen van siliconenimplantaten in Zweden begon met gegevens uit het ziekenhuis-ontslagregister. Eerst werden de gegevens geselecteerd van vrouwen van 15-69 jaar die in de periode 1965-93 een borstimplantaat hadden gehad. Door koppeling met Statistics Sweden (bevolkingsregister) via het unieke nationale registratienummer kon men die gegevens opschonen en door koppeling aan de kankerregistratie konden vrouwen die een implantaat kregen na een eerdere diagnose borstkanker worden uitgesloten.

Follow-up-gegevens van de 3521 vrouwen die een borstimplantaat kregen om een andere reden dan borstkanker, werden verkregen door koppeling met het doodsoorzakenregister. Uit de analyse bleek een verhoogd sterfterisico onder vrouwen die een borstimplantaat hadden gehad (er waren 85 sterfgevallen terwijl er 58,7 verwacht werden). Het verschil bleek toe te schrijven aan een verhoogd aantal gevallen van zelfmoord. Dit onderzoek was alleen mogelijk door koppeling op individueel niveau van vier registers, namelijk het ziekenhuisontslagregister, het bevolkingsregister, de kankerregistratie en het doodsoorzakenregister.⁴⁷

In Finland is een soortgelijk onderzoek uitgevoerd. Finse vrouwen die tussen 1970 en 2000 een borstimplantaat kregen om cosmetische redenen (n = 2166) werden geïdentificeerd uit operatiegegevens van alle ziekenhuizen die deze operaties verrichtten, inclusief die in Estland. Door koppeling aan het bevolkingsregister via het unieke persoonlijke identificatienummer werden de gegevens geverifieerd. Via koppeling aan de kankerregistratie werden vrouwen met een eerdere diagnose borstkanker uitgesloten. Sterfte naar doodsoorzaak werd verkregen door koppeling aan het doodsoorzakenregister. In het Finse onderzoek was het aantal sterfgevallen weliswaar overeenkomstig de verwachting, maar ook hier was het aantal zelfmoorden significant verhoogd.⁴⁸

In Nederland is een dergelijk onderzoek op dit moment misschien niet onmogelijk, maar wel buitengewoon moeilijk te realiseren.

lijke effecten zou kunnen hebben, kan men zo relatief snel, zonder langlopend cohortonderzoek, bevestigen of ontcrachten.

4.2 Factoren die het delen van gegevens belemmeren

Verzamelen van gegevens over de volksgezondheid vraagt inspanningen, vaardigheden en intellectuele investeringen. Beschikken over belangrijke gegevensbestanden is vervolgens van belang voor de productiviteit en werfkracht van onderzoeksgroepen en de carrière van onderzoekers. Gegeven de wetenschappe-

lijke competitie en de betrokkenheid bij de eigen gegevens, is het begrijpelijk dat voor degenen die ze zelf verzameld hebben, het delen van gegevens met anderen geen vanzelfsprekendheid is. Het wordt soms ervaren als weggeven zonder dat er iets tegenover staat. Aan de andere kant is het delen van gegevens met andere onderzoekers voor de gegevensbeheerder ook wetenschappelijk interessant, als de vraagstelling nieuw is en de analyses in een samenwerkingsverband kunnen worden gedaan en publicaties opleveren. Bovendien zullen onderzoekers die bereid zijn gegevens te delen, op hun beurt vaak ook toegang krijgen tot gegevens van anderen. Het is duidelijk dat onderzoeksgroepen en onderzoekers eerder bereid zullen zijn om gegevens te delen wanneer daarbij recht wordt gedaan aan hun intellectuele en praktische investeringen in de gegevensverzameling.

4.3 Technische voorwaarden

Om het delen van gegevens te bevorderen zal voorts in enkele technische voorwaarden voorzien moeten zijn. Allereerst moet degene die de gegevens mogelijk wil gebruiken, weten dat ze bestaan. De gegevens moeten vervolgens toegankelijk zijn voor onderzoek door anderen dan degene die de data beheert, eventueel in een samenwerkingsverband. Verder is voor het samenvoegen van gegevens harmonisatie van gegevensverzameling en –opslag noodzakelijk. Voor het koppelen van gegevens, ten slotte, is een koppelingsleutel nodig.

Hoe te weten welke gegevens er zijn: ‘metagegevens’

Een onderzoeker die op zoek is naar bestaande gegevens om een wetenschappelijke of beleidsvraag te beantwoorden moet toegang hebben tot actuele informatie over lopend onderzoek. Een toegankelijk, volledig en regelmatig geactualiseerd bestand met informatie over lopend onderzoek, bijvoorbeeld een verzamelwebsite met links naar websites van de afzonderlijke onderzoeken, kan een goede manier zijn om hierin te voorzien. In het bestand moet over elke gegevensverzameling een aantal ‘metagegevens’ zijn opgenomen, zoals onderzoeksopzet, populatiekenmerken, steekproefkenmerken, wat er op welke wijze is verzameld, of afstemming met andere studies plaats gevonden heeft en hoe, en contactinformatie. De websites die er nu zijn, bijvoorbeeld www.zorggegevens.nl en de onderzoeksdatabank van de KNAW, zijn geen van alle compleet en up-to-date. De website van het Finse Information Centre for Register Research biedt een voorbeeld van metagegevens over gegevensbestanden ten behoeve van gebruik door anderen in wetenschappelijk onderzoek, toegespitst op registraties. Via de site kan men ook praktische hulp krijgen bij het aanvragen en gebruiken

van de gegevens.⁴⁹ In Nederland kennen we DANS (Data Archiving and Networked Services), een instituut dat tot doel heeft opslag en blijvende toegankelijkheid van onderzoeksgegevens te bevorderen, in dit geval in de alfa- en gammawetenschappen.⁵⁰

Wanneer de metagegevens erop wijzen dat een gegevensverzameling voor de beantwoording van een bepaalde vraag bruikbaar is, heeft een onderzoeker vervolgens meer gedetailleerde documentatie nodig om de gegevens daadwerkelijk te kunnen gebruiken. Ontsluiting van de documentatie is een voorwaarde voor nader gebruik. Vaak is voor secundaire gegevensanalyse samenwerking met de primaire onderzoekers wenselijk of noodzakelijk.

Harmonisatie van gegevensverzameling en -opslag

Vergelijkbare en geharmoniseerde onderzoeksmethoden zijn een voorwaarde voor zinvol samenvoegen ('poolen') van gegevens. De wijze van rekruteren van onderzoeksdeelnemers die in de samen te voegen onderzoeken is toegepast, moet vergelijkbaar zijn. De gegevens moeten op vergelijkbare wijze zijn verzameld, met vergelijkbare meetinstrumenten en op vergelijkbare momenten in het leven van de onderzochte personen. Lichaamsmateriaal moet van vergelijkbare aard zijn en gestandaardiseerd zijn opgeslagen.

Harmonisatie kan soms ook achteraf gerealiseerd worden. Bij gegevens die al verzameld zijn, is men uiteraard afhankelijk van wat er verzameld is en hoe dat gedaan is, inclusief eventuele beperkingen voortvloeiend uit het *informed consent* van de deelnemers. Vergeleken met een nieuw op te zetten gegevensverzameling is het samenvoegen van gegevens voor bepaalde vraagstellingen technisch relatief eenvoudig, maar het onderling goed vergelijkbaar maken ervan blijkt in de praktijk vaak een hele klus, die bovendien niet altijd met succes bekroond wordt. Een voorbeeld waarin het is gelukt, is de constructie van een onderling vergelijkbare schaal voor het meten van algemene dagelijkse levensverrichtingen (ADL) op basis van gegevens uit verschillende cohorten.⁵¹

Bij nieuw op te zetten samenwerkingsverbanden kan men denken aan een model met een centrale opzet en regie, zoals in de Amerikaanse National Health and Nutrition Survey (NHANES; kader 2.1)⁵² en de US National Children's Study.⁵³ Een dergelijke opzet vereist bereidheid tot compromissen en veel voorbereiding. Een voorbeeld van een andere opzet is GA²LEN (Global Allergy and Asthma European Network), een verzameling eigenstandige cohortonderzoeken naar astma in diverse Europese landen met een beperkte gemeenschappelijke kern.⁵⁴

De technische voorwaarden voor het samenvoegen van gegevens uit verschillende bronnen zijn in Nederland niet optimaal. Er is weinig harmonisatie van methoden van gegevensverzameling en -opslag. Software-technisch kan het samenvoegen van gegevens uit verschillende bestanden soms moeilijk zijn, maar de software wordt wat dat betreft wel steeds beter. DANS heeft het DATA-keurmerk ingesteld om hoogwaardige en betrouwbare opslag en hergebruik van onderzoeksgegevens te garanderen.⁵⁵

Variabelen om te koppelen

Om gegevens van dezelfde persoon uit verschillende bestanden aan elkaar te koppelen is een koppelingsvariabele nodig. Dit kan bijvoorbeeld een persoonlijk identificatienummer zijn. Als er geen persoonlijk identificatienummer is (of het is er wel maar mag niet voor dit doel gebruikt worden), kan men deterministisch of probabilistisch koppelen met behulp van een combinatie van verschillende gegevens die ieder op zichzelf deels identificerend zijn en in combinatie een sterk onderscheidend vermogen hebben. Bij deterministisch koppelen is volledige overeenstemming van de koppelingsvariabelen vereist voor een match. Deterministisch koppelen heeft een hoge specificiteit (er worden weinig records ten onrechte als passend bij dezelfde persoon aan elkaar geplakt), maar de sensitiviteit is al gauw laag (omdat een terechte match niet gevonden wordt bij op zichzelf onschuldige verschrijvingen). Bij probabilistisch koppelen worden *records* al of niet aan elkaar gekoppeld op basis van een geschatte kans dat twee *records* van dezelfde persoon afkomstig zijn, waardoor de sensitiviteit toeneemt.⁵⁶ Het gebruik van de geboortedatum als één van de koppelingsvariabelen kan er toe leiden dat gegevens van bepaalde groepen migranten (met waarden als ‘onbekend’ of ‘1 januari 19xx’) moeilijker koppelbaar zijn, met selectieve beschikbaarheid van gegevens als gevolg.^{57,58}

4.4 Juridische voorwaarden

In verband met het delen van gegevens zijn drie juridische vragen van belang:

- Zijn er juridische gronden om het delen van gegevens te verhinderen dan wel af te dwingen?
 - Wordt voldoende rekening gehouden met privacy, geen-bezwaar-regeling en toestemmingsvereiste?
 - Mag het burgerservicenummer (BSN) gebruikt worden om gegevens van burgers te koppelen ten behoeve van wetenschappelijk onderzoek?
-

Juridisch kader van het delen van gegevens

De raad heeft hiernaar een achtergrondstudie laten uitvoeren (bijlage E). Samenvattend komt die analyse op het volgende neer. Een onderzoeksbestand is een samenhangende set van (ruwe) gegevens. Eigendom van gegevens bestaat als zodanig niet in juridische zin. Het auteursrecht is niet van toepassing op onderzoeksgegevens. Ook het databankenrecht is in beginsel niet van toepassing, ten eerste omdat onderzoeksgegevens in de regel geen databank vormen in de zin van het databankenrecht, en ten tweede omdat het databankenrecht bedoeld is om te voorkómen dat anderen gebruik maken van de gegevens uit de databank, terwijl het hier juist gaat om de vraag hoe gebruik van gegevens door derden bevorderd kan worden. De overdracht van de onderzoeksgegevens kan wel door externe financiers van onderzoek worden geëist in het kader van subsidievoorwaarden of de opdracht-overeenkomst.

Volgens deze analyse heeft een onderzoeksbestand dus geen ‘eigenaar’ in juridische zin. De onderzoeker (in feite de instelling waar deze werkt of werkte ten tijde van het onderzoek) heeft het bestand echter wel onder zich. Daarnaast is er geen verplichting om gegevens te delen of over te dragen, tenzij dat uitdrukkelijk in het kader van de opdracht tot het onderzoek is bedongen. Indien overdracht of beschikbaarstelling van gegevens niet als voorwaarde is gesteld bij de opdracht- of subsidieverlening, is er geen juridische basis om het delen van gegevens af te dwingen. Hoewel over deze uitleg van de toepasselijkheid van het eigendoms- en databankenrecht, zoals gepresenteerd in bijlage E, nog enige discussie mogelijk is, concludeert de raad dat een juridische benadering ons op dit punt niet verder helpt. Gegeven deze juridische patstelling, zou een gedragscode voor het delen van gegevens een goed middel kunnen zijn om een beter gebruik van gegevens te bevorderen. De toegankelijkheid van de gegevens voor derden, conform deze gedragscode, zou vervolgens door financiers in subsidievoorwaarden of in contractuele afspraken met onderzoekers kunnen worden vastgelegd. De raad komt hierop terug in het volgende hoofdstuk.

Privacy, geen-bezwaar-regelingen en toestemmingsvereiste

Uiteraard is het waarborgen van de privacy van de onderzochte personen in overeenstemming met de daartoe strekkende wet- en regelgeving een absolute voorwaarde voor het delen van onderzoeksgegevens. Leidend zijn de gedragscode van de VSNU (Vereniging van Universiteiten) voor gebruik van persoonsgegevens in wetenschappelijk onderzoek⁵⁹ en de gedragscodes ‘Goed Gebruik’ en ‘Goed Gedrag’ van de Stichting Federatie van Medisch Wetenschappelijke Ver-

enigingen (FMWV).^{60,61} De laatste code gaat over zorggegevens en is daarom hier relevant. Samengevat komt hij erop neer dat voor gebruik van direct identificerende gegevens in wetenschappelijk onderzoek in beginsel specifieke toestemming vooraf van de betrokkene vereist is. Indien het niet goed mogelijk is om toestemming te verkrijgen, bijvoorbeeld door de omvang van het bestand, mogen persoonsgegevens worden gebruikt indien ze voor de onderzoeker redelijkerwijs niet herleidbaar zijn en de betrokkenen tegen dat gebruik geen bezwaar hebben gemaakt. Aan gebruik van anonieme gegevens voor wetenschappelijk onderzoek zijn geen beperkingen of voorwaarden verbonden.

In het kader van samenwerking met het buitenland gelden de voorwaarden van de Europese Richtlijn 1995/46 EC. Binnen landen van de Europese Economische Ruimte (EER)* mogen persoonsgegevens worden gebruikt volgens de voorwaarden waaraan de verantwoordelijke, in casu de onderzoeker, in het land van herkomst is onderworpen. Voor gegevenstransport naar een land buiten de EER gelden aanzienlijk striktere regels. Persoonsgegevens mogen in beginsel niet naar een dergelijk land worden overgebracht, tenzij de Europese Unie heeft vastgesteld dat het beschermingsniveau in dat land gelijkwaardig is aan hetgeen binnen Europa met de genoemde Richtlijn is bereikt, of in het kader van een overeenkomst die aan de Europese normen voldoet.

Voor wetenschappelijk onderzoek op basis van gegevensbestanden die zijn samengevoegd om het aantal personen in de steekproef te vergroten, is het vaak niet nodig te beschikken over identificerende gegevens. Dat is anders bij gegevenskoppeling. Daarbij moet informatie over elk individu in het ene bestand gekoppeld worden aan gegevens over dezelfde persoon in een ander bestand. Het is voor het onderzoek niet nodig te weten *wie* die persoon is; het gaat er alleen om de gegevens die op dezelfde persoon betrekking hebben aan elkaar te koppelen. Daarvoor worden tot nu toe meestal combinaties van gegevens gebruikt, die samen identificerend zijn. Volledig anonimiseren is geen optie, want daarmee wordt het herleiden van verschillende *records* tot dezelfde persoon onmogelijk. Zogenaemde *privacy enhancing technologies* zijn hier wellicht van nut. Pseudonimisering is een vorm van versleuteling van identificerende gegevens waarmee identificatie van de persoon onmogelijk wordt, terwijl de gegevens nog wel tot één individu herleidbaar zijn. Het burgerservicenummer (BSN) is bijvoorbeeld in hoge mate identificerend, maar kan effectief versleuteld worden.

* Nederland, België, Luxemburg, Frankrijk, Duitsland, Italië, Groot-Brittannië, Ierland, Denemarken, Griekenland, Spanje, Portugal, Oostenrijk, Finland, Zweden, Noorwegen, IJsland, Liechtenstein, Cyprus, Estland, Hongarije, Letland, Litouwen, Malta, Polen, Slovenië, Slowakije, Tsjechië.⁶²

Gebruik van het burgerservicenummer voor wetenschappelijk onderzoek

Op 26 november 2007 is de Wet algemene bepalingen burgerservicenummer in werking getreden. Daarmee is het mogelijk geworden het BSN te gebruiken als unieke identificatie van de burger in diens contacten met de overheid. Inmiddels is het BSN ook in de gezondheidszorg ingevoerd. Op 19 februari 2008 stuurden de Koninklijke Nederlandse Academie van Wetenschappen (KNAW) en het bestuur van de FMWV een gezamenlijk verzoek naar de minister van Justitie om het gebruik van het BSN voor wetenschappelijk onderzoek mogelijk te maken door middel van een Algemene Maatregel van Bestuur.⁶³

In de brief wordt betoogd dat wetenschappelijk onderzoek met persoonsgegevens beter en efficiënter zou worden als het gebruik zou mogen maken van het BSN voor het koppelen van bestanden. Beter omdat met het BSN unieke koppelingen kunnen worden gemaakt en administratieve ‘meerlingen’ worden voorkómen. Efficiënter omdat het niet meer nodig zal zijn om gegevens te verzamelen die alleen nodig zijn voor de koppeling en niet inhoudelijk bijdragen aan het onderzoek. Door gebruik van het BSN wordt in Nederland onderzoek mogelijk dat nu niet uitgevoerd kan worden omdat het veel te omslachtig is, zoals het onderzoek in kader 4.1. Nederland zou zich dan in dat opzicht als onderzoeks- en kennisland weer kunnen meten met de Scandinavische landen. Daar is het uitgangspunt dat in onderzoek dat gebruik mag maken van persoonsgegevens, het nationale persoonsidentificatienummer gebruikt mag worden.

Voor toepassing in wetenschappelijk onderzoek moet het BSN versleuteld worden en voor de koppeling van bestanden is het nodig een *trusted third party* in te schakelen, een partij die onafhankelijk is van de instanties die de gegevensbestanden beheren en van de instantie die de gekoppelde gegevens gaat gebruiken.

Gebruik van het BSN voor wetenschappelijk onderzoek is nu niet toegestaan. De raad meent echter dat de mogelijkheid om het BSN te kunnen gebruiken voor wetenschappelijk onderzoek een noodzakelijke voorwaarde is om gegevensverzamelingen beter te benutten en om mee te kunnen blijven spelen op het Europese onderzoekstoneel.

4.5 Financiële voorwaarden

Het zal duidelijk zijn dat het delen van gegevens extra werk met zich meebrengt en daarmee ook geld kost. Daar staat tegenover dat het opnieuw verzamelen van gegevens bijna altijd duurder zal zijn. Extra werk gaat bijvoorbeeld zitten in het verzamelen en up-to-date houden van metagegevens over lopend en afgerond

onderzoek, het documenteren van onderzoeksbestanden, de activiteiten van de *trusted third party* bij gegevenskoppeling, enzovoort. Het is op dit moment moeilijk om financiering te verkrijgen voor documenteren en ontsluiten van gegevens voor derden.

Aanbevelingen

Gegevens over de volksgezondheid zijn essentieel voor zowel het beleid als het wetenschappelijk onderzoek op het gebied van de volksgezondheid. In de voorgaande hoofdstukken heeft de raad betoogd dat de lopende gegevensverzamelingen in Nederland niet in alle behoeften van beleid en wetenschap kunnen voorzien, dat het voortbestaan van sommige belangrijke gegevensbronnen niet vanzelfsprekend is en dat bovendien de gegevens die er zijn, beter kunnen worden benut. Er is dus werk aan de winkel, voor allen die in beleid, onderzoek en onderzoeksfinanciering bij het verzamelen van gegevens over de volksgezondheid en het gebruik ervan betrokken zijn. Dat moet er mede toe leiden dat een deel van de collectieve middelen die voor onderzoek naar de volksgezondheid beschikbaar zijn, kunnen worden ingezet om innovatieve plannen voor nieuwe gegevensverzamelingen te kunnen honoreren.

De raad komt op basis hiervan tot een drietal aanbevelingen die in dit hoofdstuk verder worden uitgewerkt:

- 1 Bevorder een efficiënt gebruik van gegevensverzamelingen.
- 2 Hef tekorten aan gegevens op en voorkóm nieuwe tekorten.
- 3 Creëer meer ruimte voor verzoeken om collectieve financiering van nieuwe gegevensverzamelingen en beoordeel deze systematisch.

Deze drie aanbevelingen beogen een doelmatige besteding van de middelen die voor de verzameling van gegevens over de volksgezondheid kunnen worden ingezet. Ze vergen een gezamenlijke inzet van alle betrokkenen: beleidsmakers,

onderzoekers en onderzoeksfinanciers. In het bijzonder vragen ze, op verschillende punten, om een gezamenlijk optrekken, zowel beleidsmatig als financieel, van overheid en onderzoeksfinanciers, waaronder NWO, ZonMw en de gezondheidsfondsen, zoals KWF Kankerbestrijding en Nederlandse Hartstichting.

5.1 Bevorder een efficiënt gebruik van gegevensverzamelingen

a Stel een register van gegevensverzamelingen in

De raad beveelt aan een register in te stellen van bestaande en nieuwe gegevensverzamelingen op het gebied van de volksgezondheid, zodat iedereen kan nagaan wat er al is aan gegevens, waar ze zijn en wie ze beheert. Het register moet compleet en up-to-date zijn voor lopende en afgeronde verzamelingen en voor elke gegevensverzameling een nader vast te stellen set ‘metagegevens’ bevatten. Onderzoeksfinanciers kunnen volledigheid en actualiteit van het register waarborgen door aanmelding bij het register als voorwaarde voor financiering te stellen. Toekenning van een identificerend nummer aan iedere gegevensverzameling, vergelijkbaar met het ISBN voor boeken, zou het makkelijker maken gegevensverzamelingen te vinden en het gebruik ervan in kaart te brengen.⁶⁴

b Zorg voor een optimale toegankelijkheid van gegevens

De RGO pleit voor het opstellen van een *gedragscode* voor het delen van gegevens. De raad ziet hier een belangrijke taak weggelegd voor de Stichting Federatie van Medisch Wetenschappelijke Verenigingen (FMWV) en/of de Koninklijke Nederlandse Academie van Wetenschappen (KNAW). Uitgangspunt van de gedragscode zal moeten zijn dat met collectieve middelen gegenereerde gegevens toegankelijk moeten zijn voor onderzoek dat van nut is voor het algemeen belang, ook als dat wordt uitgevoerd door anderen dan degenen die de gegevens verzameld hebben. Het document zal tevens zorgvuldig recht moeten doen aan de belangen van de onderzoekers zonder wier kennis, vaardigheden en inspanningen het gegevensbestand er niet geweest zou zijn. De gedragscode moet onderschreven worden door onderzoekers, de instellingen waarbij zij in dienst zijn en onderzoeksfinanciers.

Financiers van onderzoek kunnen de toegankelijkheid van gegevens bevorderen door aan financiering van bestaande en nieuwe gegevensverzamelingen een viertal *voorwaarden* te verbinden:

- Aanmelding bij het register.

- Onderschrijving van de gedragscode voor het delen van gegevens.
- Beschrijving, in de subsidieaanvraag, van een procedure voor het verkrijgen van toegang tot de gegevens. Buitenlandse *data sharing policies*, zoals van MRC en NIH, kunnen tot voorbeeld dienen (zie kader 5.1). De beschreven procedure moet zijn goedgekeurd voor de subsidie verleend wordt, maar moet, zo meent de raad, niet betrokken worden bij de beoordeling en prioritering van een subsidieaanvraag. De criteria daarvoor zijn en blijven relevantie en wetenschappelijke kwaliteit.
- Gebruik van gevalideerde standaardmeetmethoden, of goede argumenten voor het afwijken daarvan. De raad beveelt aan harmonisatie van meetmethoden te bevorderen door een ‘gemeenschappelijke kern’ (*common core*) te creëren van gevalideerde meetmethoden voor belangrijke variabelen.

c Optimaliseer de mogelijkheden tot gegevenskoppeling

De mogelijkheid om gegevens uit verschillende gegevensbestanden te koppelen bevordert de kwaliteit en doelmatigheid van onderzoek, en het is goed mogelijk om bij het koppelen de privacy van mensen te waarborgen. De RGO beveelt aan gehoor te geven aan het verzoek van FMWV en KNAW aan de regering om gebruik van het *burgerservicenummer* (BSN) voor wetenschappelijk onderzoek mogelijk te maken. Hij beveelt voorts aan de vormgeving te onderzoeken van een *trusted third party* voor gegevenskoppeling, een instantie die zorg draagt voor de koppeling van de gegevens uit verschillende bestanden en die onafhankelijk is van de bestandsbeheerders en van de gebruikers van de gekoppelde gegevens.

d Vergemakkelijk het delen van gegevens

De raad realiseert zich dat de uitvoering van de regelingen die voor het delen van gegevens noodzakelijk zijn en het feitelijk delen van de gegevens, werk met zich meebrengt, dat niet tot de gangbare werkzaamheden van onderzoekers behoort, deels specifieke expertise vereist en hoe dan ook tijd en geld kost. De raad vindt daarom dat onderzoekers bij het delen van gegevens zoveel mogelijk gefaciliteerd moeten worden om het delen van gegevens mogelijk te maken en beveelt aan te voorzien in een onafhankelijke gegevensmakelaar en in praktische hulp en ondersteuning.

De *gegevensmakelaar* brengt gegevenszoekers en –beheerders met elkaar in contact en begeleidt het delen van gegevens. De makelaar krijgt zelf geen gegevens, maar kan bijvoorbeeld een onafhankelijke medebeoordeling geven van

Kader 5.1 Data sharing policies in het buitenland

Verenigde Staten: National Institutes of Health

De Amerikaanse National Institutes of Health (NIH) hebben sinds 2003 een actief beleid voor het delen van onderzoeksgegevens voor onderzoeksdoeleinden. 'In NIH's view, all data should be considered for data sharing. Data should be made as widely and freely available as possible while safeguarding the privacy of participants ...' 'By avoiding the duplication of expensive data collection activities, the NIH is able to support more investigators than it could if similar data had to be collected de novo by each applicant.' Sinds 2003 moet in iedere NIH subsidieaanvraag boven een bepaald bedrag expliciet een plan voor *data sharing* zijn opgenomen, of een passage die uitlegt waarom *data sharing* in een specifiek geval niet mogelijk zou zijn. In het plan moeten in ieder geval de tijdigheid, de privacy van de onderzoekspersonen, de methoden voor het delen van de gegevens, de manier van documenteren en het benodigde budget voor *data sharing* geadresseerd worden: Door er vooraf over na te denken, kunnen in het onderzoeks *design* al maatregelen genomen worden ten behoeve van *data sharing* straks, maar NIH schrijft geen *common core* van data of gestandaardiseerde metingen voor.

De NIH probeert datasharing te faciliteren door kennis ter beschikking te stellen. De website geeft voorbeelden van plannen voor data sharing plans en van succesvolle data sharing. Een *data sharing workbook* (soort handleiding) is beschikbaar. Aangeraden wordt een *data-sharing agreement* af te sluiten, vooral om ervoor te zorgen dat gegevens alleen voor de vooraf bepaalde doelen gebruikt worden.⁶⁵

Verenigd Koninkrijk: Medical Research Council (MRC) Policy on Data Sharing and Preservation

'Publicly-funded research data are a public good, produced in the public interest, and they should be openly available to the maximum extent possible'. Gegevens moeten tijdig beschikbaar worden gesteld. Gegevensbeheerders die hun gegevens ter beschikking stellen moeten daarvoor erkenning krijgen van subsidiegevers, de academische wereld en de gebruikers van de gegevens. Het secundair gebruik van gegevens gaat vaak het best in directe samenwerking met de onderzoeksgroep die de gegevens verzameld heeft. Een korte periode waarin de primaire gegevensverzamelaars de exclusieve gebruiksrechten hebben is redelijk. Sinds 2006 moeten ook alle onderzoeksaanvragen bij de MRC voorzien zijn van een paragraaf over *data sharing*. Additioneel aan informatie over tijdsplanning en technische aspecten moet daarin expliciet beschreven worden wat de toegevoegde waarde is van de nieuw te verzamelen gegevens ten opzichte van bestaande onderzoeken en registers, en hoe het delen van de gegevens zal bijdragen aan coördinatie en samenwerking in het onderzoeksveld.⁶⁶

De *Wellcome Trust* heeft een soortgelijk beleid.⁶⁷

Uit ervaringen met de genoemde *policies* blijkt dat deze alleen kunnen werken in de context van een actief implementatie traject, dat onderzoekers enerzijds de toegevoegde waarde van

data sharing laat zien en hen daarbij zoveel mogelijk helpt, en anderzijds het beleid om gegevens te delen consequent afdwingt.

In een commentaar in de JAMA worden enkele kanttekeningen geplaatst bij het beleid om onderzoeksgegevens snel ter beschikking te stellen voor secundaire analyses door derden.⁶⁸ Volgens dit commentaar laat het vrijgeven van de Framingham gegevens de wetenschappelijke wereld meeprofiten van de kwaliteitscontrole op de gegevens toen ze verzameld werden, maar gaat dit ten koste van de kwaliteitscontrole op de gegevensanalyse en op de productie van wetenschappelijke artikelen. De primaire onderzoekers zouden door de termijn waarop de gegevens moeten worden vrijgegeven onder extreme druk staan om snel te publiceren, terwijl er na het verstrijken van die termijn onvoldoende controle is op wie wat met de gegevens doet.

aanvragen voor het delen van gegevens en het proces van het delen van gegevens bewaken (tijdstermijnen, gebruik van gegevens voor het afgesproken doel, waarborgen van de rechten van de gegevensverzamelaars). Het is goed denkbaar de functie van gegevensmakelaar organisatorisch te combineren met die van de beheerder van het register van gegevensverzamelingen (zie aanbeveling 1a). Beide functies zouden bijvoorbeeld, misschien eerst voor een proefperiode van drie jaar, kunnen worden ondergebracht bij een onderzoeksfinancier of grote onderzoeksorganisatie. De raad is er geen voorstander van om nieuwe gebruikers ‘zomaar’ – bijvoorbeeld via een website – toegang te verschaffen tot gegevens die door anderen verzameld zijn. Contact tussen nieuwe gebruikers en gegevensverzamelaars ziet de raad als een belangrijke voorwaarde voor oordeelkundig gebruik van de gegevens.

Effectief nader gebruik van gegevens vereist kennis en kunde op het gebied van methoden voor het ontsluiten van gegevens voor derden, compatibel maken van gegevensbestanden, *privacy enhancing technologies*, enzovoort. De raad beveelt aan deze kennis voor iedereen die gegevens verzamelt goed toegankelijk en beschikbaar te maken, bijvoorbeeld via een website. Daarnaast is *praktische hulp en ondersteuning* noodzakelijk, bijvoorbeeld via een helpdesk of door de werkzaamheden die delen van de gegevens met zich meebrengt te laten verrichten door hierin gespecialiseerde professionals.

In longitudinaal epidemiologisch onderzoek wordt vaak gebruik gemaakt van koppeling van gegevens van de deelnemers aan bestaande registraties, bijvoorbeeld de Gemeentelijke Basisadministratie, de doodsoorzakenregistratie, de Landelijke Medische Registratie of het GezondheidsStatistisch Bestand. Dit is op zichzelf een buitengewoon efficiënte manier om follow-up gegevens te verzamelen, maar de procedures die voor deze gegevenskoppeling doorlopen moeten worden blijken in de praktijk soms niet eenvoudig. De gegevensmakelaar en de

praktische hulp en ondersteuning zouden ook hierin een faciliterende rol kunnen spelen. Vereenvoudiging van de koppeling aan bestaande registraties, kan voor beheerders van longitudinale gegevensverzamelingen een extra stimulans zijn om aan het delen van de eigen gegevens hun medewerking te verlenen.

De raad beveelt financiers van onderzoek aan om voor deze vormen van facilitering middelen vrij te maken.

5.2 Hef tekorten aan gegevens op en voorkóm nieuwe tekorten

a Voorzie in herhaald dwarsdoorsnede-onderzoek van voldoende omvang

De RGO stelt vast dat er in Nederland veel bruikbare gegevens zijn en worden verzameld, maar dat er, vanuit beleidsoogpunt, sprake is van één duidelijk tekort: het ontbreekt aan een herhaald dwarsdoorsnedeonderzoek van voldoende omvang, dat kan dienen als bron voor periodieke actualisering van prevalenties van risicofactoren en functionele beperkingen. De raad beveelt aan om in Nederland, naar het model van de Amerikaanse National Health and Nutrition Examination Survey (NHANES studie), een dergelijk onderzoek op te zetten. Het onderzoek moet voldoen aan de in paragraaf 3.3 geformuleerde eisen en voorzien in koppeling aan de doodsoorzakenregistratie. De opzet moet bovendien zodanig zijn dat het ook aan actuele wetenschappelijke gegevensbehoefte zoveel mogelijk tegemoetkomt. Om in het beleid rekening te kunnen houden met de gezondheidstoestand van specifieke bevolkingsgroepen, zoals die van allochtone herkomst, zullen deze in het onderzoek voldoende vertegenwoordigd moeten zijn.

b Garandeer de continuïteit van de registratie van zorggebruik in ziekenhuizen

De LMR is een belangrijke bron van informatie, niet alleen over zorggebruik, maar ook, in aanvulling op andere gegevensbronnen, over de prevalentie van ziekten. Kwaliteit en continuïteit van de LMR zijn echter niet meer gegarandeerd sinds de registratie van Diagnose-Behandel Combinaties (DBC's) is ingevoerd. De RGO onderschrijft het belang van het beleid van de minister van VWS dat erop gericht is de continuïteit van de LMR te garanderen. Ook het initiatief van Nederlandse Federatie van UMC's en de NVZ vereniging van ziekenhuizen tot de oprichting van Dutch Hospital Data is in dit verband belangrijk.

c **Zorg voor adequate financiering van longitudinale epidemiologische gegevensverzamelingen**

De RGO meent dat longitudinale epidemiologische gegevensverzamelingen – essentieel als ze zijn voor het wetenschappelijk onderzoek en voor het verklaren van trends in de volksgezondheid – een publiek belang dienen. De raad beveelt daarom aan mogelijkheden te creëren om dit soort gegevensverzamelingen, voor zover ze actief benut worden en wetenschappelijk productief zijn, structureel te financieren. Op dit moment zijn die mogelijkheden indicidenteel aanwezig bij sommige gezondheidsfondsen, bij de programmatisch werkende intermediaire organisaties zijn ze er formeel nu niet (ZonMw) of zeer beperkt (NWO).

De raad meent dat de structurele financiering van longitudinale gegevensverzamelingen afhankelijk moet zijn van succesvol gebruik van de verzamelde gegevens. Nieuwe gebruikers van de verzamelde gegevens zouden het hunne aan de instandhouding van de gegevensverzameling kunnen bijdragen door bij de verwerving van subsidie voor hun onderzoek een aparte post van serieuze omvang op te voeren als bijdrage aan de financiering van de infrastructuur van die gegevensverzameling. Dit vereist uiteraard bereidheid van onderzoeksfinanciers om langs deze weg aan de instandhouding van productieve gegevensverzamelingen bij te dragen.

Ook meent de raad dat het in het algemeen aanbeveling verdient om te bevorderen dat nieuwe gegevensverzameling waar mogelijk plaatsvindt in het kader van bestaand onderzoek. Uitbreiding van lopend onderzoek is immers vaak doelmatiger dan uit het niets met een nieuwe studie starten. De uitvoerder van de lopende longitudinale gegevensverzameling moet de mogelijkheid krijgen om een bijdrage te verwerven voor zijn dataverzamelingsinfrastructuur wanneer hij met andere onderzoekers samenwerkt bij het verzamelen van aanvullende gegevens.

5.3 Creëer meer ruimte voor verzoeken om collectieve financiering van nieuwe gegevensverzamelingen en beoordeel deze systematisch

Ook wanneer beschikbare gegevens optimaal worden gebruikt (aanbeveling 1) en (dreigende) tekorten worden opgeheven (aanbeveling 2) zullen er altijd nieuwe vragen uit wetenschap en beleid opkomen die nieuwe gegevensverzameling noodzakelijk maken. De RGO meent dat er ruimte moet zijn voor nieuwe initiatieven om te voorzien in nu nog grotendeels onbekende gegevensbehoeften. Daarbij moeten de schaarse collectieve middelen uiteraard doelmatig worden ingezet, al was het maar om zoveel mogelijk innovatieve ideeën te kunnen honoreren. De raad beveelt daarom aan alle verzoeken om collectieve financiering van nieuwe

gegevensverzamelingen zorgvuldig te beoordelen aan de hand van een aantal eisen van relevantie, noodzakelijkheid, kwaliteit en doelmatigheid (zie kader 5.2). Ook verzoeken om subsidie voor nieuwe rondes van dataverzameling in lopend longitudinaal onderzoek kunnen aan de hand van deze eisen beoordeeld worden.

De RGO verwacht dat deze drie aanbevelingen ertoe kunnen bijdragen dat Nederland in de toekomst verzekerd zal zijn van de gegevens die voor effectief volksgezondheidsbeleid en hoogstaand wetenschappelijk volksgezondheidsonderzoek onontbeerlijk zijn.

Kader 5.2 Beoordelingskader voor verzoeken om collectieve financiering van nieuwe gegevensverzamelingen op het gebied van de volksgezondheid

Zijn de te verzamelen gegevens relevant?

De beoogde gegevensverzameling

1a is wetenschappelijk relevant, gezien de stand van de wetenschap en/of

1b voorziet in een gegevensbehoefte van het beleid.

Is de nieuwe gegevensverzameling noodzakelijk?

In de beoogde gegevensverzameling kan niet afdoende worden voorzien door

2 een bestaande of lopende Nederlandse gegevensverzameling.

3 een bestaande of lopende buitenlandse gegevensverzameling.

4 samenvoegen of koppeling van beschikbare Nederlandse gegevens.

5 samenvoegen of koppeling van beschikbare buitenlandse gegevens.

6 door aanvullende gegevensverzameling binnen een bestaande gegevensverzamelingsstructuur.

Voldoet de nieuwe gegevensverzameling aan eisen van kwaliteit en doelmatigheid?

7 De beoogde nieuwe gegevensverzameling voldoet aan de noodzakelijke kwaliteitseisen.

8 De voorgestelde opzet is de meest kosten-effectieve manier om de beoogde gegevens te verzamelen.

9 De waarde van de nieuwe gegevensverzameling weegt op tegen de kosten.**

10 De voorgestelde opzet voldoet aan de voorwaarden voor financiering die door de RGO zijn voorgesteld ter bevordering van het delen van gegevens.***

* Voorbeelden van kwaliteitseisen zijn: met de voorgestelde onderzoeksopzet is het mogelijk de beoogde wetenschappelijke vraag te beantwoorden; de voorgestelde opzet is conform de huidige (inter)nationale stand van de wetenschap; controleerbaarheid, bijvoorbeeld toegankelijkheid voor audits. ** De waarde van een nieuwe gegevensverzameling kan behalve in de direct beoogde relevantie voor wetenschap of beleid gelegen zijn in, bijvoorbeeld, de toegevoegde waarde die ontstaat als de nieuw te verzamelen gegevens worden gecombineerd met, of gekoppeld aan, reeds beschikbare gegevens, of in de verwachting dat op basis van de nieuw te verzamelen gegevens toekomstige wetenschappelijke vragen te beantwoorden zullen zijn. *** Zie aanbeveling 1b in deze samenvatting.

Literatuur

- 1 KNAW. Multifactoriele aandoeningen in het genomics tijdperk. Amsterdam: 2006.
 - 2 OECD Principles and guidelines for access to research data from public funding. Paris (France): OECD; 2007.
 - 3 NHANES. www.cdc.gov/nchs/nhanes.htm (geraadpleegd op 26 augustus 2008).
 - 4 Blokstra A, Smit H, Bueno de Mesquita HB, Seidell JC, Verschuren WMM. Monitoring van Risicofactoren en Gezondheid in Nederland (Morgen project), 1993-1997. RIVM rapport 263200008/2005. Bilthoven: RIVM; 2005.
 - 5 Nederlandse Kankerregistratie. www.ikcnet.nl (geraadpleegd op 26 augustus 2008).
 - 6 Parelsnoer. www.parelsnoer.org (geraadpleegd op 26 augustus 2008).
 - 7 Framingham. www.framinghamheartstudy.org (geraadpleegd op 26 augustus 2008).
 - 8 ERGO. www.epib.nl/ergo.htm (geraadpleegd op 26 augustus 2008).
 - 9 Hofman A, Breteler MM, van Duijn CM, Krestin GP, Pols HA, Stricker BH e.a. The Rotterdam Study: objectives and design update. *Eur J Epidemiol* 2007; 22(11): 819-829.
 - 10 Grobbee DE, Hoes AW, Verheij TJM, Schijvers AJP, Van Ameijden EJC, Numans ME. The Utrecht Health Project: optimization of routine healthcare data for research. *Eur J Epidemiology* 2005; 20: 285-287.
 - 11 Molenaar EA, van Ameijden EJ, Grobbee DE, Numans ME. Comparison of routine care self-reported and biometrical data on hypertension and diabetes: results of the Utrecht Health Project. *Eur J Public Health* 2007; 17(2): 199-205.
 - 12 Rovers MM, Numans ME, Langenbach E, Grobbee DE, Verheij TJ, Schilder AG. Is pacifier use a risk factor for acute otitis media? A dynamic cohort study. *Fam Pract* 2008; 25(4): 233-236.
 - 13 Leidsche Rijn. www.lrgp.nl (geraadpleegd op 26 augustus 2008).
-

- 14 Stichting Perinatale Registratie Nederland. Perinatale zorg in Nederland 2005. Utrecht: Stichting Perinatale Registratie Nederland; 2008.
- 15 Meray N, Reitsma JB, Ravelli AC, Bonsel GJ. Probabilistic record linkage is a valid and transparent tool to combine databases without a patient identification number. *J Clin Epidemiol* 2007; 60(9): 883-891.
- 16 CBS. www.statline.nl (geraadpleegd op 2 september 2008).
- 17 LINH. www.linh.nl (geraadpleegd op 2 september 2008).
- 18 De Hollander AEM, Hoeymans N, Melse JM, van Oers JAM, Polder JJ. *Zorg voor gezondheid - Volksgezondheid Toekomst Verkenning 2006*. Bilthoven: RIVM/cVTV; 2006.
- 19 Westert GP, Verkleij H. *Zorgbalans – de prestaties van de Nederlandse gezondheidszorg in 2004*. Bilthoven: RIVM; 2006.
- 20 Lokale en nationale monitor gezondheid. www.monitorgezondheid.nl (geraadpleegd op 2 september 2008).
- 21 Perinatale Registratie Nederland. www.perinatreg.nl (geraadpleegd op 2 september 2008).
- 22 Nederlandse Kankerregistratie. www.ikcnet.nl (geraadpleegd op 26 augustus 2008).
- 23 Schram MT, Frijters D, van de Lisdonk E.H., Ploemacher J, de Craen AJM, de Waal MWM e.a. Setting and registry characteristics affect the prevalence and nature of multimorbidity in the elderly. *J Clin Epidemiol* 2008; e-pub ahead of print.
- 24 Vaal J, de Klerk MMY, Frijters DHM, e.a. Gecombineerde visus- en gehoorbeperking: naar schatting bij 30.000 - 35.000 55-plussers in Nederland. *Ned Tijdschr Geneesk* 2007; 151(26): 1459-1463.
- 25 Van Gool CH, Picavet SHJ, Deeg DJH, de Klerk MMY, Nusselder WJ, Hoeymans N. Oudere van nu gezonder dan 15 jaar geleden? Trends in lichamelijke beperkingen. *Tijdschrift voor Gerontologie en Geriatrie* 2008; 39: in druk.
- 26 Voedselconsumptiepeilingen. www.rivm.nl/vcp (geraadpleegd op 2 september 2008).
- 27 KNAW. *Multifactoriele aandoeningen in het genomics tijdperk*. Amsterdam: 2006.
- 28 Bijdrage Universitaire Afdelingen Huisartsgeneeskunde aan de maatschappelijke opgaven volksgezondheid en gezondheidszorg. Utrecht: Nederlands Huisartsen Genootschap; 2008.
- 29 Raad voor Gezondheidsonderzoek. *Translationeel onderzoek in Nederland – van kennis naar kliniek*. Den Haag: 2007: 55.
- 30 Parelsnoer. www.parelsnoer.org (geraadpleegd op 26 augustus 2008).
- 31 Stolk RP, Rosmalen JGM, Postma DS, de Boer RA, Navis G, Slaets JJP e.a. Universal risk factors for multifactorial diseases. Lifelines: a three generation population-based study. *Eur J Epidemiol* 2008; 23: 67-74.
- 32 BBMRI. www.biobanks.eu (geraadpleegd 2 september 2008).
- 33 The Emerging Risk Factors Collaboration. The Emerging Risk Factors Collaboration: analysis of individual data on lipid, inflammatory and other markers in 1.1 million participants in 104 prospective studies of cardiovascular diseases. *Eur J Epidemiol* 2007; 22: 839-869.
-

- 34 Plug I, van der Bom JG, Peters M, Mauser-Bunschoten EP, de Goede-Bolder A, Heijnen L e.a. Mortality and causes of death in patients with hemophilia, 1992-2001: a prospective cohort study. *J Thrombosis Haemostasis* 2006; 4: 510-516.
- 35 Hagemans ML, Hop WJ, van Doorn PA, Reuser AJ, van der Ploeg AT. Course of disability and respiratory function in untreated late-onset Pompe disease. *Neurology* 2006; 66(4): 581-583.
- 36 Hagemans ML. Pompe disease in children and adults: natural course, disease severity and impact on daily life [Proefschrift]. 2006.
- 37 Gunn E, Baker M, Larsen A. The UK CF Registry; a succesful transition to a web-based system. *J Cyst Fibros* 2008; suppl 2: S117.
- 38 McCormick J, Sims EJ, Green MW, Mehta G, Culross F, Mehta A. Comparative analysis of Cystic Fibrosis Registry data from the UK with USA, France and Australasia. *J Cyst Fibros* 2005; 4(2): 115-122.
- 39 Schram MT, de Waal MWW, de Craen AJM, Deeg DJH, Schellevis FG. Multimorbiditeit: de nieuwe epidemie. *TSG* 2008; 86: 28-30.
- 40 Parelsnoer. www.parelsnoer.org (geraadpleegd op 26 augustus 2008).
- 41 Lokale en nationale monitor gezondheid. www.monitorgezondheid.nl (geraadpleegd op 2 september 2008).
- 42 P3G. <http://www.p3gconsortium.org/> (geraadpleegd 2 september 2008).
- 43 BBMRI. www.biobanks.eu (geraadpleegd 2 september 2008).
- 44 Gabriel. <http://www.gabriel-fp6.org/> (geraadpleegd op 2 september 2008).
- 45 Meray N, Reitsma JB, Ravelli AC, Bonsel GJ. Probabilistic record linkage is a valid and transparent tool to combine databases without a patient identification number. *J Clin Epidemiol* 2007; 60(9): 883-891.
- 46 Mondriaan project. www.tipharma.com (geraadpleegd 2 september 2008).
- 47 Koot VCM, Peeters PHM, Granath F, Grobbee DE, Nyren O. Total and cause-specific mortality among Swedish women with cosmetic breast implants: prospective study. *BMJ* 2003; 326: 527-528.
- 48 Pukkala E, Kulmala I, Hovi SL, Hemminki E, Keskimaki I, Lipworth L e.a. Causes of death among Finnish women with cosmetic breast implants, 1971-2001. *Annals of Plastic Surgery* 2003; 51(4): 339-342.
- 49 Finnish Information Centre for Register Research. www.retki.stakes.fi/EN/index.htm (geraadpleegd op 2 september 2008).
- 50 DANS. www.dans.knaw.nl/nl/over_dans/ (geraadpleegd op 15 september 2008).
- 51 Pluijm SMF, Bardage C, Nikula S, Blumstein T, Jylha M, Minicuci N e.a. A harmonized measure of activities of daily living was a reliable and valid instrument for comparing disability in older people across countries. *J Clin Epid* 2005; 58: 1015-1023.
- 52 NHANES. www.cdc.gov/nchs/nhanes.htm (geraadpleegd op 26 augustus 2008).
- 53 National Children's Study. www.nationalchildrensstudy.gov/Pages/default.aspx (geraadpleegd op 2 september 2008).
-

- 54 Keil T, Kulig M, Simpson A, e.a. European birthcohort studies on asthma and atopic diseases: I. Comparison of study designs – a GA2LEN initiative. *Allergy* 2006; 61: 221-228.
- 55 DANS. www.dans.knaw.nl/nl/over_dans/ (geraadpleegd op 15 september 2008).
- 56 Meray N, Reitsma JB, Ravelli AC, Bonsel GJ. Probabilistic record linkage is a valid and transparent tool to combine databases without a patient identification number. *J Clin Epidemiol* 2007; 60(9): 883-891.
- 57 Nielen M, Verheij R, De Bakker D, Deville W. Vooronderzoek naar verbetering kwaliteit huisartsenzorg in achterstandsgebieden. Utrecht: NIVEL; 2007.
- 58 Volkers A, Bus A, Uiters E. Zorggebruik in de huisartspraktijk. In: Jennissen RPW, Oudhof J, editors. Ontwikkeling in de maatschappelijke participatie van allochtonen. Een theoretische verdieping en een thematische verbreding van de integratiekaart 2006. Den Haag: WODC/CBS; 2007:
- 59 VSNU. Gedragscode voor gebruik van persoonsgegevens in wetenschappelijk onderzoek. 2005: <http://www.vsnul.nl/web/show/id=69988/langid=43/> (geraadpleegd op 8 september 2008).
- 60 FMWV. Code Goed Gebruik. 2002.
- 61 FMWV. Gedragscode Gezondheidsonderzoek (Goed gedrag). 2003.
- 62 Europese Economische ruimte. www.ce-marking.nl/nl/helpdesk/faq2.html (geraadpleegd op 15 september 2008).
- 63 KNAW-advies Gebruik burgerservicenummer (BSN) bij wetenschappelijk onderzoek met gegevens. www.knaw.nl/cfdata/publicaties/detail.cfm?boeken_ordernr=20081016 (geraadpleegd op 2 september 2008).
- 64 Kauffmann F, Cambon-Thomson A. Tracing biological collections: between books and clinical trials. *JAMA* 2008; 299(19): 2316-2318.
- 65 NIH Data sharing policy. http://grants.nih.gov/policy/data_sharing/data_sharing_guidance.htm (geraadpleegd op 4 oktober 2007).
- 66 Medical Research Council's data sharing policy. <http://www.mrc.ac.uk/PolicyGuidance/EthicsAndGovernance/DataSharing/PolicyonDataSharingandPreservation/index.htm> (geraadpleegd op 18 maart 2008).
- 67 Wellcome Trust, data sharing policy. <http://www.mrc.ac.uk/PolicyGuidance/EthicsAndGovernance/DataSharing/PolicyonDataSharingandPreservation/index.htm> (geraadpleegd op 18 maart 2008)
- 68 Psaty B, Arnett D, Burke B. A new era of cardiovascular disease epidemiology. *JAMA* 2007; 298(17): 2060-2062.
-

-
- A De adviesaanvraag
-
- B Raad voor Gezondheidsonderzoek
-
- C De commissie
-
- D Inventarisatie Nederlandse populatiecohorten
-
- E 'Eigendom' en delen (in de zin van *sharing*) van onderzoeksgegevens
-
- F Afkortingen

Bijlagen

De adviesaanvraag

Bij brief d.d. 10 mei 2007 richtte de Directeur-Generaal van de Volksgezondheid zich tot de RGO met een adviesaanvraag omtrent ‘verbetermogelijkheden bij bestaande monitoring-, cohortstudies en biobanken’:

Financiers van de publieke gezondheidszorg hebben monitoring- en cohortstudies nodig voor besluitvorming. Ook innoverend onderzoek is toenemend afhankelijk van goede cohorten in combinatie met biobanken. Beide toepassingen zijn gebaat bij efficiëntie. De RGO zou kunnen inventariseren wat op dit terrein gebeurt aan studies, wie de studies financiert en welk belang de betrokkenen hebben bij de studies. De vraagstelling is dan hoe het beter kan: wat is er méér uit de bestaande monitoring en cohortstudies te halen? Waar zijn witte vlekken in het geheel van de gegevensverzamelingen, en wat is nodig voor de toekomst?

Langs diverse kanten wordt de financiering van cohortstudies aan de orde gesteld op het politieke niveau. Van recente datum is de discussie over een rijksbijdrage in de financiering van bijvoorbeeld de studie ‘Generation at R’ (GGD-Rotterdam, EUR en EMCR) en de Doetinchemstudie. Een helder beoordelingskader voor de legitimering van de rijksfinanciering van dergelijke studies is hierbij belangrijk.

Raad voor Gezondheidsonderzoek

-
- prof. dr. P.J. van der Maas, *voorzitter*
hoogleraar maatschappelijke gezondheidszorg, Erasmus Medisch Centrum, Rotterdam
 - prof. dr. J.M. Bensing, *vice-voorzitter*
hoogleraar gezondheidspsychologie, Universitair Medisch Centrum Utrecht, directeur Nederlands instituut voor onderzoek van de gezondheidszorg (NIVEL)
 - prof. dr. W.J.J. Assendelft
hoogleraar huisartsgeneeskunde, Leids Universitair Medisch Centrum
 - dr. A. Boer
lid Raad van Bestuur College voor zorgverzekeringen (CVZ)
 - prof. dr. J.M.W. Hazes
hoogleraar reumatologie, Erasmus MC, Rotterdam
 - prof. dr. J.W. Hofstraat
vice president Philips research
 - prof. dr. J. Kievit
hoogleraar medische beslistkunde, Leids Universitair Medisch Centrum
 - dr. R. van Olden
medisch directeur GlaxoSmithKline
 - prof. dr. S.A. Reijneveld
hoogleraar sociale geneeskunde, Universitair Medisch Centrum Groningen
-

- dr. ir. H.A. Smit
hoofd Centrum voor Preventie en Zorgonderzoek, Rijksinstituut voor Volksgezondheid en Milieu, Bilthoven
- dr. C. Smit
vertegenwoordiger patiënten en consumenten
- prof. dr. A.E.M. Speckens
hoogleraar psychiatrie, Universitair Medisch Centrum St. Radboud, Nijmegen
- dr. M.J. Trappenburg
politicoloog, Utrechtse School voor Bestuurs- en Organisationswetenschappen
- dr. L. van 't Veer
moleculair bioloog, Nederlands Kanker Instituut, Amsterdam
- prof. dr. R. Vos
hoogleraar gezondheidsethiek en wijsbegeerte, Universiteit Maastricht
- prof. dr C. van Weel
hoogleraar huisartsgeneeskunde, Universitair Medisch Centrum St. Radboud, Nijmegen
- prof. dr. E.G.E. de Vries, *waarnemer*
voorzitter Raad voor Medische Wetenschappen
- prof. dr. H.M. Pinedo, *waarnemer*
vice-voorzitter ZonMw
- ir. M.W. Horning, *adviseur*
directie Markt en Innovatie, ministerie van Economische Zaken
- drs. J. Lelij, *adviseur*
directie Onderzoek en Wetenschapsbeleid, ministerie van Onderwijs Cultuur en Wetenschap
- dr. ir. C.M. Vos, *adviseur*
directie Macro-Economische Vraagstukken en Arbeidsvoorwaarden, ministerie van Volksgezondheid Welzijn en Sport
- dr. J.N.D. de Neeling, *secretaris*
Gezondheidsraad/RGO, Den Haag

De Gezondheidsraad en belangen

Leden van Gezondheidsraadcommissies – waaronder sinds 1 februari 2008 ook de leden van de RGO – worden benoemd op persoonlijke titel, wegens hun bijzondere expertise inzake de te behandelen adviesvraag. Zij kunnen echter, dikwijls juist vanwege die expertise, ook belangen hebben. Dat behoeft op zich geen bezwaar te zijn voor het lidmaatschap van een Gezondheidsraadcommissie.

Openheid over mogelijke belangenconflicten is echter belangrijk, zowel naar de voorzitter en de overige leden van de commissie, als naar de voorzitter van de Gezondheidsraad. Bij de uitnodiging om tot de commissie toe te treden wordt daarom aan commissieleden gevraagd door middel van het invullen van een formulier inzicht te geven in de functies die zij bekleden, en andere materiële en niet-materiële belangen die relevant kunnen zijn voor het werk van de commissie. Het is aan de voorzitter van de raad te oordelen of gemelde belangen reden zijn iemand niet te benoemen. Soms zal een adviseurschap het dan mogelijk maken van de expertise van de betrokken deskundige gebruik te maken. Tijdens de installatievergadering vindt een bespreking plaats van de verklaringen die zijn verstrekt, opdat alle commissieleden van elkaars eventuele belangen op de hoogte zijn.

De commissie

-
- prof. dr. D.E. Grobbee, *voorzitter*
hoogleraar klinische epidemiologie, Julius Centrum, Universitair Medisch Centrum Utrecht
 - prof. dr. D.H. de Bakker
hoogleraar structuur en organisatie van de gezondheidszorg, in het bijzonder de eerstelijnszorg, Universiteit van Tilburg, en Nederlands instituut voor onderzoek van de gezondheidszorg (NIVEL), Utrecht
 - prof. dr. J.W.W. Coebergh
hoogleraar kankersurveillance, Erasmus MC, Rotterdam
 - prof. dr. D.J.H. Deeg
hoogleraar epidemiologie van veroudering, VU medisch centrum, Amsterdam
 - prof. dr. P.J. van der Maas
hoogleraar maatschappelijke gezondheidszorg, Erasmus MC, Rotterdam, voorzitter Raad voor Gezondheidsonderzoek, Den Haag
 - prof. dr. D.S. Postma
hoogleraar longziekten, Universitair Medisch Centrum Groningen
 - dr. D. Ruwaard, *adviseur*
directeur Publieke Gezondheid, ministerie van Volksgezondheid, Welzijn en Sport, Den Haag
-

- dr. C. Smit
vertegenwoordiger patiënten en consumenten in de Raad voor
Gezondheidsonderzoek, Den Haag
- dr. ir. H.A. Smit
hoofd Centrum voor Preventie en Zorgonderzoek, Rijksinstituut voor
Volksgezondheid en Milieu, Bilthoven
- prof. dr. C. van Weel
hoogleraar huisartsgeneeskunde, Universitair Medisch Centrum St. Radboud,
Nijmegen
- dr. M.L. Essink-Bot, *secretaris*
Gezondheidsraad/RGO, Den Haag

Inventarisatie Nederlandse populatiecohorten

Bij de tabellen:

B = Biomateriaal afgenomen en bewaard A = Antropometrie (gemeten waarden voor bijv. lengte, gewicht, longfunctie, serumcholesterol, glucose, of andere aan het lichaam gemeten kenmerken)

Tabel 1 Lopende populatiecohorten.

Naam	Leeftijdsgrenzen bij inclusie / geboortecohorten	Omvang	Inclusie-periode	Follow-up duur	Eerstkomende dataverzameling	B	A	(Contact)informatie
ABCD	Zwangeren en kinderen vanaf 14 weken zwangerschapsduur	8 266	2003-2004	20 jaar	2008	ja	ja	www.abcd-study.nl Dr. T. Vrijkotte, T.Vrijkotte@amc.uva.nl; Dr. M. van Eijnsden, mveijnsden@ggd.amsterdam.nl
Amsterdam Growth and Health	13 jaar / (geb. 1963, 64, 65)	600	1976	2007: 41 jaar (n=350)	Nog niet bekend	ja	ja	www.emgo.nl/research_infra/amsterdam.index.asp Prof. J. Twisk, JWR.Twisk@vumc.nl; Prof. W. van Mechelen, w.vanmechelen@vumc.nl
Doetinchem	20-59 jaar / (geb.1928-1967)	12 500	1987-1991	4 ^e ronde 2003-2007 N=4500	2008-2012	ja	ja	www.rivm.nl/vtv/object_document-/o3354n19085.html Verschuren et al. Int J Epidemiol 2008 Dr. S. Picavet, Susan.picavet@rivm.nl; Dr. W.M.M. Verschuren, wmm.Verschuren@rivm.nl; Dr. H.A. Smit, jet.smit@rivm.nl

Tabel 1 Vervolg.

Naam	Leeftijdsgrenzen bij inclusie / geboortecohorten	Omvang	Inclusie-periode	Follow-up duur	Eerstkomende dataverzameling	B	A	(Contact)informatie
DOM cohort (Diagnostisch Onderzoek Mammacarcinoom Utrecht; inmiddels: Doorlopend Onderzoek Morbiditeit en Mortaliteit)	Vrouwen 50-70 in 4 cohorten	50 000	1974-1986	1996: 10-22 jaar	Jaarlijkse koppeling NKR	ja	ja	Petra Peeters, P.H.M.peeters@umcutrecht.nl; Collette Lancet 1984, EurJCancer 1992, Miltenburg BrJCancer 1998; http://www.juliuscentrum.nl/julius/Research/cohorts/DOM/tabid/168/Default.aspx
Eindhoven Study (EPOS)	I: Vrouwen 46-51 jaar (cohorten 1941-47) II: Uitbreiding met cohort t/m 1951	I: 6 000 II: 4 000	I: 1994-5 II: 1998	doorlopend	2008	ja	ja	Johannes L.P., Thyroid 2005;15(11):1253-9; Nyklicek I. Psychiatry Res 2004;128(2):111-6; Hagemans ML. Hum Reprod 2004;19(9):2163-9; Prof. dr. V. Pop, v.j.m.pop@uvt.nl; C. Wijnands, c.wijnands@pozob.nl
Erasmus Rucphen Familiestudie (ERF)	18 jaar en ouder	ERF: plm. 3 000 nakomelingen van 20 paren in Rucphen in 1850	Vanaf 1999	doorlopend	Via registers	ja	ja	www.epib.nl/erf/erf_index.html ; c.vanduijn@erasmusmc.nl; Aulchenko et al. Eur J Hum Genet. 2004 Jul;12(7):527-34 Onderdeel van Genetic Research in Isolated Populations (GRIP)
ERGO	ERGO I: ≤ 1934 (1990) ERGO plus: 1938-1947 (1999) ERGOjong: < 1962 (2006)	ERGO I n=7983 ERGOplus n= 3011 ERGO jong n=4 000 (totaal n=15 000)	1990, 1999, 2006	onbeperkt	Elke 3-4 jaar metingen + continue door koppeling aan registraties	ja	ja	www.epib.nl/ergo.htm ; Hofman A. et al. Eur J Epidemiol 2007;22:819-829 secretariaat.epib@erasmusmc.nl; J. Heeringa, j.heeringa@erasmusmc.nl
ERSPC	Mannen 50-75 jaar	42 376	1983-1998	20 jaar	2008	ja	ja	www.erspc.org ; De Koning et al, Int J Cancer 2002; 98(2):268; Dr. M. Roobol, m.roobol@erasmusmc.nl; Prof. dr. H.J. de Koning, h.dekoning@erasmusmc.nl
Generation R	3 mnd. zwangerschapsduur	9 778	2002-2006	30 jaar	1* 4 jaar via JGZ Vanaf 5 jaar om de 3 jaar	ja	ja	www.generationr.nl ; Paediatr Perinat Epidemiol. 2004;18:61-72; Eur J Epidemiol. 2006;21:475-84; Eur J Epidemiol. 2007;22:917-23; Dr. V. Jaddoe, V.Jaddoe@erasmusmc.nl; info@generationr.nl

Tabel 1 Vervolg.

Naam	Leeftijdsgrenzen bij inclusie / geboortecohorten	Omvang	Inclusieperiode	Follow-up duur	Eerstkomende dataverzameling	B	A	(Contact)informatie
Globe	15-74	18 973	1991	2 rondes: 1997 en 2004-05	Nog niet duidelijk	nee	nee	Van Lenthe et al.. Eur J Public Health 2004;14:63-70, F.vanlenthe@erasmusmc.nl
Hoorn	I: 1914-1939 (50-75 jaar in 1989) Nieuw Hoornstudie: 1940-1965 (40-65 in 2005)	I: 2 484 Nieuwe Hoornstudie: 3 000	1989 en 2005	onbeperkt	doorlopend	ja	ja	www.emgo.nl/research_infra / hoorn/index.asp Prof. dr. G. Nijpels, G.Nijpels@vumc.nl; Prof. r. J. Dekker JM.Dekker@vumc.nl
Hongerwintercohort	Index = levendgeboren in Wilhelmina Gasthuis Amsterdam tussen 1 november 1943 en 28 februari 1947; + hun moeders + hun kinderen (transgeneratiecohort)	Index n= 2 414	1994	Tot uitsterven	Via registers. 2000-2004 medisch onderzoek	Ja (lft 58-60)	ja (lft 58-60)	www.hongerwinter.nl; De Rooij et al, Am J Clin Nutr 2007;86:1219-24; Painter RC et al. Eur J Epidemiol 2005;20:673-76; Raveli AJC et al. Lancet 1998;351:173-77; Dr. T. Roseboom, t.j.roseboom@amc.uva.nl
KOALA	14 weken zwangerschapsduur	2 800	2000-2003	onbeperkt	2008 (lftd 6-7 jaar)	ja	ja	Kummeling I, Thijs C et al. Ped Allergy and Immunology 2005; 16(8):679-684; http://www.birthcohorts.net/Cohort.Show.asp?cohortid=41 Dr. C. Thijs, c.thijs@epid.unimaas.nl
LASA	I: 55-83 (geb. 1908-37) II: 55-65 (geb. 1938-47)	1: n=4 494 (in 2005/06: 1 257) 2: n=1 002 (in 2005/06: 908)	1991 en 2002	In ieder geval tot 2012	2008-09	ja	ja	www.lasa-vu.nl Prof. D. Deeg, djh.deeg@vumc.nl; M. vd Horst, mhl.vanderhorst@vumc.nl
Leiden 85-plus	I: 85 jaar (geb. 1901 en eerder) II: 85 jaar (geb. 1912-1914)	I: 977 II: 599	I: 1986-89 II: 1997-1999	Tot uitsterven	Via koppeling GBA, CBS	ja	ja	www.lumc.nl/ouderengeneeskunde; Dr. T. de Craen, A.J.M.de_Craen@lumc.nl
Leiden Lang Leven Studie	Families bestaande uit minimaal 2 langlevenden (mannen \geq 89 jr, vrouwen \geq 91 jaar), hun broers en zussen, kinderen, en de partners van de kinderen	420 families: 956 langlevende broers & zussen, 1 750 kinderen, 758 partners	2002-2005	Onbeperkt	2006-2008	ja	ja	Schoenmaker et al. Eur J Hum Genet. 2006;14:79-84; www.langleven.net; Dr. ir. D. van Heemst, D.van_Heemst@lumc.nl; Dr. T. de Craen, A.J.M.de_Craen@lumc.nl

Tabel 1 Vervolg.

Naam	Leeftijdsgrenzen bij inclusie / geboortecohorten	Omvang	Inclusieperiode	Follow-up duur	Eerstkomende dataverzameling	B	A	(Contact)informatie
Leidsche Rijn Gezondheidsproject	Alle leeftijden	20-30 000	Sinds 2001	onbeperkt	Baseline gezondheidsprofiel, f.up via EMD van huisarts	ja	ja	www.lgrp.nl; Grobbee et al, Eur J Epidemiol 2005; Dr. M.E. Numans, m.e.numans@umcutrecht.nl
LifeLines	3 generaties < 25, 25-50 (geb. 1956-1981), 50-75 (geb. 1931-1955)	165 000	Pilot 2007; 2008-2012	30 jaar 25-50, 50+	Fenotypering elke 5 jaar	ja	ja	www.lifelines.nl; lifelines@umcg.nl; Prof. dr. B.H.R. Wolffenbuttel, bwo@int.umcg.nl
Maastricht Aging Study (MAAS)	24-81 (1912-1969)	1 823	1993	12 jaar	2008	ja	ja	www-np.unimaas.nl/maas; Dr. M. van Boxtel, Maas@np4.unimaas.nl
Maastricht Cohort Study on Fatigue at work	Werknemers van 45 bedrijven	7 000	Vanaf 1999	?	2-3 vragenlijsten per jaar	nee	nee	www.fatigueatwork.nl; Dr. IJ. Kant, IJ.kant@epid.unimaas.nl
MORGEN (deel van EPIC NL)	20-59 jaar	22 715	1993-1997	Onbeperkt, via registraties	Nieuwe dataverzameling wordt voorzien.	ja	ja	Boker et al, Eur J Epidemiol 2001;17:1047; www.epicnl.eu; www.iarc.fr/epic; www.epicnl.eu/Default.aspx?PageContentID=55; dr. H. B. Bueno-de-Mesquita, hb.bueno.de.mesquita@rivm.nl; dr. J. van derLaan, jan.van.der.laan@rivm.nl); In 2007 is merging met PROSPECT tot stand gebracht: http://www.epicnl.eu/Default.aspx?PageContentID=56
Nemesis-2	18-65 jaar (geb. 1942-1989)	7 000	2007-2009	6 jaar (2 followup metingen)	2010	ja	nee	www.trimbos.nl; Ron de Graaf, rgraaf@trimbos.nl
NELSON	50-75	15 822	2003-2004	15 jaar	2008	ja	ja	www.nelsonproject.nl; Van Iersel et al, Int J Cancer 2007;120(4): 868. Prof. dr. H.J. de Koning, h.dekoning@erasmusmc.nl Dr. R. van Klaveren, r.j.vanklaveren@erasmusmc.nl
Nijmeegs (proef) bevolkingsonderzoek borstkanker	1975-1992: vrouwen 35-70. 1992-heden:vrouwen 50-75	2008: n= 53 370	Dynamische populatie sinds 1975	onbeperkt	Continue	nee	ja	Otten JD. J Epidemiol Community Health 1996Jun;50(3):353-8; Broeders MJ. J Med Screening 2002;9(4):163-7; A.L.M. Verbeek, a.verbeek@epib.umcn.nl

Tabel 1 Vervolg.

Naam	Leeftijdsgrenzen bij inclusie / geboortecohorten	Omvang	Inclusieperiode	Follow-up duur	Eerstkomende dataverzameling	B	A	(Contact)informatie
Nijmegen Bio-medical study	Vanaf 18 jaar (geb. 1983 en eerder)	9 500	2001-2003	Loopt tot heden	2008	ja	ja	Wetzels JFM et al. <i>Kidney Int</i> 2007;72(5):632-7; Gudmundsson J et al. <i>Nat Genet</i> 2007; 39(8): 977-983; Stacey SN et al. <i>Nat Genet</i> 2007; 39(7):865-9; Prof. dr. L.A.L.M. Kiemeny, b.kiemeny@epib.umcn.nl
Netherlands Cohort Study	55-69 jaar (geb. 1915-1931)	58 279 mannen en 62 573 vrouwen	1986	onbeperkt	Via registers	ja	nee	Van den Brandt et al, <i>J. Clin Epidemiol</i> 1990; 43: 285-95; Piet van den Brandt (contactpersoon), PA.vandenBrandt@epid.unimaas.nl; Sandra Bausch-Goldbohm, sandra.bausch@tno.nl; Dr. Leo Schouten, lj.schouten@epid.unimaas.nl
Peilstationsproject Hart- en Vaatziekten (PPHV)	20-59 jaar	36 588	1987-1991	onbeperkt	Via registers	ja	ja	Neth J Cardiol 1993;4:205-210; dr. JMA Boer, jma.boer@rivm.nl; dr. W.M.M. Verschuren, wmm.verschuren@rivm.nl; dr. J. van der Laan, jan.van.der.laan@rivm.nl
PIAMA	Zwangerschapsduur 7 maanden	4 146	1996-1997	8 jaar	2007	ja	ja	Wijga AH, Brussee JE, Smit HA. <i>RIVM rapport 260401002/2004</i> ; http://piama.iras.uu.nl ; Dr. H.A.. Smit, Jet.Smit@rivm.nl; Alet.Wijga@rivm.nl
Prevask	Moeders (prenatale inclusie van kinderen geb. 1997-2002)	843	1997-2002	6 jaar	2 jaarlijks	ja	ja	Kuiper S et al., <i>Pediatr Allergy Immunol</i> 2005;16: 321-31; Schonberger HJ et al. <i>Ann Allergy Asthma Immunol</i> 2003; 91(6):531-38. Prof. O. van Schayck, Onno.vanSchayck@hag.unimaas.nl
PREVEND	28-65 (geb. 1932-1969)	40 856 alg. pop., focuscohort: n=8592 met albuminurie	1997-1998	Baselinecohort via registers. Focuscohort 3-4 jaarlijks	Voor focuscohort: 2008	ja	ja	www.prevend.org Prof. dr. P.E. de Jong, p.e.de.jong@int.umcg.nl; r.t.gansevoort@umcg.nl

Tabel 1 Vervolg.

Naam	Leeftijdsgrenzen bij inclusie / geboortecohorten	Omvang	Inclusieperiode	Follow-up duur	Eerstkomende dataverzameling	B	A	(Contact)informatie
Prospect (deel van EPIC NL)	49-70 jaar	17 500 vrouwen 50-70 uit BO borstkanker	1993-1997	Vragenlijsten elke 3-5 jaar , registers	2010-2012	ja	ja	Boker I.K et al. Eur J Epidemiol. 2001;17(11):1047-53; www.juliuscentrum.nl/julius/Research/cohorts/ProspectEPICNL/tabid/169/Default.aspx http://www.epicnl.eu/English/Introduction/tabid/444/Default.aspx; www.iarc.fr/epic; Petra Peeters, P.H.M. peeters@umcutrecht.nl; In 2007 is merging met MORGAN tot stand gebracht: http://www.epicnl.eu/Default.aspx?PageContentID=56)
SHARE (Survey of health, ageing and retirement in Europe)	I: 50 jaar en ouder (geb. 1954 en eerder); in refreshment samples stromen nieuwe 50-jarigen in.	NL: n= 2 979 (totaal: 30 000 in 11 landen, later 16)	2004	Elke 2-3 jaar, zonder einddatum	2008-09	nee	ja	www.share-project.org; Prof. A. van Soest en M. de Groot (Netspar, U v Tilburg) a.h.o.vansoest@uvt.nl; I.M.deGroot@uvt.nl; Prof. J.P. Mackenbach, j.mackenbach@erasmusmc.nl
Studie naar Medische Informatie en Leefwijzen in Eindhoven: SMILE	12 jaar en ouder	40 000	Sinds 2002	onbeperkt	Twee keer per jaar EMD van de huisarts; Twee keer per jaar vragenlijsten	nee	Nog niet	http://www.smile-eindhoven.nl/; Van den Akker, M et al. BMC Public Health 2008;9:19; Dr. Marjan van den Akker, marjan.vandenakker@hag.unimaas.nl; Dr. Mark Spigt, m.spigt@hag.unimas.nl
Taaltrial	14 mnd (JGZ)	11 440	1995-1996			nee	nee	Van Agt et al. Pediatrics 2007;120(6):1317
TRAILS (Tracking Adolescents' Ind. Life Survey)	11 jaar / 1989, 90, 91	2 200	2001-2002	14 jaar (2015)	tweejaarlijks	ja	ja	www.trails.nl, Trails@med.umcg.nl
Tweede Nationale Studie / LINH (NSII/ LINH)	Alle leeftijden	390 000 patiënten ingeschreven bij 104 huisartspraktijken	2001-2002	onbeperkt	Via LINH (Landelijk Informatie Netwerk huisartsenzorg)	nee	nee	Schellevis et al, Huisarts & Wetenschap 2003; 46(1):7-12, F.schellevis@nive.nl
Vlagtwedde/ Vlaardingen Cohort Study	14-67 jaar	ca. 8 500 personen (50% mannen/50% vrouwen)	1965-1990	3-jaarlijkse follow-up onderzoeken (max. 8 er deelnemer) tot 1990	Via registers	ja	ja	Rijcken B et al. Am Rev Respir Dis 1993;147:1447-1453; Dr. H.M. Boezen, h.m.boezen@epi.umcg.nl
Zutphen	Mannen geb. 1900-1920	Plm. 1 000	1960 of 1985	50 jaar	2010	ja	ja	RIVM rapport 260854002 (Kalmijn et al); Prof. dr. D. Kromhout

Tabel 2 Afgesloten populatiecohorten

Naam	Leeftijdsgrenzen bij inclusie	Omvang	Inclusieperiode	Follow-up duur	B	A	(Contact) informatie
Arnhem Ederly study	65-85 jaar	N=1012 Bij 685 ook bloed en biometrie	1991	10 jaar (2000)	ja	ja	Giltay EJ, Geleijnse JM, Zitman FG, Hoekstra T, Schouten EG. Arch Gen Psychiatry 2004;61:1126-35, Marianne.geleijnse@wur.nl
DIMCA	>18	1 155	1992	Tot 2004	nee	ja	Afd. Huisartsgeneeskunde UMC Radboud, C.vanweel@hag.umcn.nl
EPOZ (Zoetermeer)	5-90 jaar (in 1975-1978) 5-25 jaar (cohorten 1950-1970)	10 000 (baseline) + 375 jongeren	1975-1978		ja	ja	
Gezondheidsmonitoring Vuurwerkkramp Enschede	Geen leeftijdsgrenzen	3 907 slachtoffers + referentiegroep (n=9 935) + hulpverleners	Retrospectief 010199-070502 Prospectief 130502-130505	6 jaar	nee	nee	Ijzermans et al, rapport NIVEL 2006, j.ijzermans@nivel.nl
GLAS	57-99 (in 1993) (1936 en daarvoor)	5 279	1993 (via HA netwerk)	?	nee	nee?	?
Krimpen Study	Mannen 50-75 jaar (geb. 1920-1945)	9 324	aug. 1995-jan. 1998	6,5 jaar (n=825)	ja	ja	Gericht op urologische aandoeningen (BPH, LUTS), Kok ET et al. J Urol 2005;174(3):1055, Blanker MH et al. Urol Int 2005;75(1):30
Monitoring gevolgen Nieuwjaarsbrand Volendam	Aanwezigheid in 't Hemeltje tijdens de brand + broers, zussen, ouders + voor kinderen en ouders gematchte controle-groep	300 jongeren, 303 broers en zussen, 1 966 controles; 499 ouders, 1 756 controles	Retrospectief 01012000-01012002 Prospectief 01012002-30062006	5,5 jaar	nee	nee	Diverse publicaties van Dorn & Ijzermans
Nemesis I	18-65 jaar (geb. 1931-1978)	7 076 (baseline)	1996	3 jaar (2 follow-up metingen)	nee	nee	Vollebergh W et al. Psychische stoornissen in Nederland – overzicht van de resultaten van Nemesis. Trimbosinstituut, 2003, Ron de Graaf, rgraaf@trimbos.nl
Nijmeegs Interventie Project (NIP)	Volwassenen uit 6 huisartspraktijken	5 945	1976-1978	1981, 1998	nee	ja	c.vanweel@hag.umcn.nl

Tabel 3 Huisartsennetwerken.

Naam	Aantal praktijken	Aantal patiënten	Dataverzameling	(Contact) informatie
Continue Morbiditeits Registratie, Peilstations Nederland (NIVEL)	44 (66 HA)	157 000	Wekelijks	www.nivel.nl/pdf/Continue-Morbiditeits-Registratie-2005.pdf , www.nivel.nl/griep sinds 1970 Mw. Dr. Gé Donker, g.donker@nivel.nl
Continue Morbiditeits Registratie Nijmegen	4	13 000	Dagelijks	Sinds 1971. Dynamische populatie. Dr. E. van de Lisdonk tel: 024 3615311 e-mail: e.vandelisdonk@hag.umcn.nl
GP-infonet	35 praktijken (Utrecht) + 118 huisartsen (Almere)	58 000 (Utrecht) + 140 000 (Almere)	regelmatig	www.gpinfonet.nl . Akkermans et al. Huisarts en Wetenschap 2008; 51(2): 90-95.
Huisartsen netwerk AMC (HAGNET)	34 HA	45 000	3-maandelijks	Henk Brouwer; http://www.amc.nl/index.cfm?pid=1282
Huisartsen Netwerk Utrecht	7	50 000	halfjaarlijks	Via huisartsinformatiesysteem (ELIAS). Sinds 1991. www.zorggevens.nl Julius Centrum, E.Hak
Integrated Primary Care Information database (IPCI)	200	>1 000 000	continue	www.ipci.nl . Via EMD, sinds 1992. Vlug AE, van der Lei J et al. Methods Inf Med. 1999 Dec;38(4-5):339-44; Van Wyk JT et al. Chest 2005;128:2521-2527; m.sturkenboom@erasmusmc.nl / j.vanderlei@erasmusmc.nl
LINH	85	340 000	Elk half jaar	www.linh.nl . Via EMD. Sinds 1997. Dr. Robert Verheij, r.verheij@nivel.nl
Nijmeegs Monitoring project	9	45 000 (basis populatie)	Jaarlijks	Sinds 1986. Proces en uitkomsten bij DM, hypertensie, COPD (w.degrauw@hag.umcn.nl)
Registratienetwerk Universitaire Huisartspraktijken Leiden en Omstreken (RNUH-LEO)		30 000	Jaarlijks	Sinds 1989. Via EMD. www.rivm.nl/vtv/object_document/03759n23465.html . LUMC afdeling Public health en Eerstelijngeneeskunde http://www.lumc.nl
Registratienet Huisartspraktijken (RNH) Maas-tricht		85 000	Elk kwartaal	Via MicroHIS. Sinds 1987. http://www.hag.unimaas.nl/rnh/database.html . Dr. Marjan van den Akker Registratienet Huisartspraktijken Capaciteitsgroep Huisartsgeneeskunde Care and Public Health Research Institute
Registratie Netwerk Groningen		50 000		Sinds 1989. www.zorggevens.nl . De praktijken zitten ook in LINH.
Rotterdam: HONEUR Huisarts Onderzoeks Netwerk EUR				Geen morbiditeitsregistratie, maar netwerk voor onderzoek en dataverzameling (bijv. Heintjes et al. Knee disorders in primary care. BMC Musculoskeletal Disorders 2005; Aug 23; 6:45), b.koes@erasmusmc.nl
Transitieproject				www.zorggevens.nl . www.transitieproject.nl C.van Boven, voorzitter transitieproject, wetenschappelijk medewerker van de afdeling huisartsgeneeskunde van de Radboud Universiteit Nijmegen, c.vanboven@hag.umcn.nl
VU MC Network of General Practitioners	29	70 000	halfjaarlijks	emgo@vumc.nl Petra Elders, p.elders@vumc.nl

‘Eigendom’ en delen (in de zin van *sharing*) van onderzoeksgegevens

Samenvatting

Onderzoeksbestanden bestaan uit de ruwe data die zijn verkregen in het kader van een onderzoek en waarop na interpretatie en analyse de gepubliceerde uitkomsten van het onderzoek worden gebaseerd. Onderzoeksbestanden kunnen door een veelheid van onderzoeksmethoden ontstaan; voorzover hier relevant, bijvoorbeeld door het bevragen van respondenten, door het verkrijgen van klinische gegevens en/of gegevens uit andere bestanden zoals doodsoorzaken, door analyse van lichaamsmateriaal en veelal door een combinatie van deze methoden. Deze onderzoeksbestanden kunnen los van de gepubliceerde uitkomsten ook betekenis hebben voor andere onderzoekers. Door de onderzoeksbestanden te combineren kunnen in beginsel¹ grotere relevante datasets ontstaan, meer zeldzame verbanden worden onderzocht, hoeft primaire dataverzameling niet te worden herhaald, kan een taakverdeling worden bereikt, etc. De wenselijkheid van het delen van delen van onderzoeksbestanden is daarmee een gegeven. De commissie Cohortstudies en registraties van de RGO was in dat kader ook geïnteresseerd in de juridische aspecten van dit delen. Dat betreft vragen als: is er een ‘eigenaar’ van de onderzoeksbestanden, kan delen worden afgedwongen, met welke aspecten moet daarbij rekening worden gehouden?

Deze juridische vragen worden in deze achtergrondnotitie behandeld. Daarbij wordt voortgebouwd op een eerder initiatief van de gecombineerde FMWV/VVE commissie regelgeving van onderzoek (Coreon).

Eerst komt een juridische analyse van het begrip onderzoeksbestand. Een onderzoeksbestand is een samenhangende set van gegevens. Eigendom van gegevens bestaat als zodanig in juridische zin niet, hetgeen overigens niet betekent dat gegevens soms niet door het recht worden beschermd. De onderzoeker heeft deze gegevens onder zich en het zou onrechtmatig zijn als een ander zich daartoe zonder toestemming toegang zou verschaffen. In een bepaalde vorm kunnen gegevens echter wel voorwerp zijn voor een intellectueel eigendomsrecht (IE). Hier is dan met name het databankenrecht van belang. Verdedigd wordt dat een onderzoeksbestand vrijwel nooit een databank is in de zin van het databankenrecht. Een bestand kan een dergelijke databank worden doordat middelen zijn aangewend om het onderzoeksbestand op een of andere manier openbaar te maken. De eigenaar van het databankenrecht op deze aldus ontsloten gegevens is degene die daarin heeft geïnvesteerd.

Het gaat hier echter om onderzoeksbestanden die (nog) niet openbaar zijn en dat ook niet hoeven te worden om te worden gedeeld met andere onderzoekers.

Er is dus geen ‘eigenaar’ van het onderzoeksbestand, maar de onderzoeker (of de instelling waaraan deze is verbonden) heeft bestand wel onder zich. Het zou onrechtmatig zijn om er zich toegang te verschaffen. Er is geen verplichting om te delen. In niet strikt juridische zin is deze onderzoeker of de instelling waaraan deze is verbonden, dus wel ‘eigenaar’. Een financier zou evenwel in de subsidievoorwaarden of de overeenkomst waarmee het onderzoek wordt gefinancierd kunnen stellen dat de onderzoeksbestanden aan de financier worden overgedragen. Dan heeft de financier voortaan dit bestand onder zich en de onderzoeker niet meer. Ook zou de financier het databankenrecht van het onderzoeksbestand kunnen claimen, zoals in de literatuur wel is geopperd.

Aldus blijkt in beginsel twee ‘eigenaren’ in niet strikt juridische zin in aanmerking komen:

- De onderzoeker die het onderzoeksbestand heeft samengesteld, c.q. de instelling waaraan de onderzoeker is verbonden;
- De externe financier van het onderzoeksbestand.

De onderzoeker en de instelling waaraan deze is verbonden vallen voor het doel van deze notitie samen. Indien het onderzoeksbestand in het kader van de werkzaamheden is samengesteld, is in beginsel de instelling de ‘eigenaar’. Blijven dus over de onderzoeker/instelling versus de externe financier als ‘eigenaren’.

¹ De gegevens moeten wel methodologisch vergelijkbaar zijn, ook kan de privacy of de toestemming van de betrokkene waaronder de gegevens zijn verkregen aan ‘nader gebruik’ van deze gegevens in de weg staan.

De claim van de financiers wordt zelden inhoudelijk onderbouwd. Als de claim is ‘wie betaalt, bepaalt’, is het de vraag of dit rechtvaardig is ten opzichte van de onderzoeker/instelling. De bijdrage van de financier is altijd partieel. Bovendien is er dan een probleem indien het specifieke onderzoek uit meerdere bronnen wordt gefinancierd. Als de achterliggende gedachte is dat ‘eigendom’ bij de financier veilig stelt dat onderzoeksbestanden worden gedeeld, kunnen daar – naast de genoemde argumenten - vraagtekens bij worden geplaatst. Is de externe financier wel de meest geschikte hoeder van het onderzoeksbestand? Heeft deze bijvoorbeeld wel voldoende kennis en mogelijkheden om het onderzoeksbestand voor ander onderzoek te ontsluiten?

Deze strikt juridische benadering biedt daarmee geen handvatten voor een evenwichtige oplossing. Belangrijker dan de strikt juridische benadering is een afweging van belangen. Inderdaad kan het onderzoeksbestand uit eerder onderzoek van groot belang zijn voor ander onderzoek, door deze of andere onderzoekers. Zeker indien het onderzoek publiek (of semi-publiek door de collectebusfondsen) is gefinancierd, is er een gerechtvaardigde claim dat dit bestand dan voor een volgend onderzoek in het algemeen belang kan worden gebruikt. Daartegenover staat echter dat het onderzoeksbestand niet uitsluitend dankzij deze financiering tot stand is gekomen. Er zit meer achter, zoals de infrastructuur van het onderzoeksinstituut, kennis en intellectuele arbeid van de onderzoekers en eventueel hun gegevens uit eerder onderzoek – wellicht van een geheel andere financier – waarop bij dit onderzoek is voortgebouwd.

De vraag kan daarom het beste worden gesteld als:

hoe kan worden bevorderd dat onderzoeksbestanden voor onderzoek kunnen worden gedeeld wanneer dat voor wetenschappelijk onderzoek in het algemeen (gezondheids)belang van nut is.

De ‘ruwe gegevens’ van een bepaald onderzoek zouden dan onder voorwaarden moeten kunnen worden opgenomen in de brede stroom van onderzoeksgegevens waar ook andere onderzoekers van kunnen profiteren.

Dat is waar het om draait. De vraag naar het exclusieve ‘eigendom’ van een onderzoeksbestand leidt slechts tot een vertroebeling van deze discussie.

Om het ‘delen onder voorwaarden’ te bereiken zouden gedragsregels of zou een Gedragscode kunnen worden opgesteld. In deze Gedragscode kan worden aangesloten bij reeds bestaande of zich ontwikkelende praktijken met betrekking tot het delen van bestanden met name in het buitenland. Zij zou breed moeten worden gedragen door financiers, onderzoekers en de instellingen waaraan zij zijn verbonden. De subsidievoorwaarden of de overeenkomst waarmee het onderzoek wordt gefinancierd, kunnen dan naar deze gedragsregels verwijzen. Daarmee worden zij bindend voor de onderzoeker/instelling. Enkele elementen voor die gedragsregels worden tot slot benoemd.

1. Inleiding

1.1 De aanleiding

Medisch-wetenschappelijk onderzoek of zorgonderzoek met gegevens zou kunnen worden gediend indien gegevens door onderzoekers worden gedeeld. Hoe groter het bestand, hoe beter genuanceerde verbanden kunnen worden gevonden, zeker als daarbij analyses op lichaamsmateriaal worden gebruikt. Of de primaire dataverzameling hoeft niet te worden herhaald, een taakverdeling kan worden bereikt, de enquêtedruk worden verminderd, etc. De te delen gegevens worden hier het onderzoeksbestand genoemd (zie ook sectie 1.3). Dit onderzoeksbestand moet worden onderscheiden van de gepubliceerde resultaten. Het gaat om de gegevens (antwoorden respondenten, analyses lichaamsmateriaal, etc) op basis waarvan die resultaten zijn verkregen.

Onderzoekers hebben daarom behoefte om bestanden van andere onderzoekers te kunnen hergebruiken of met andere te kunnen combineren voor nieuw onderzoek.

Hieraan zijn ook juridische aspecten verbonden. De RGO commissie Cohortstudies en registraties had daarom behoefte aan verkenning van de juridische vragen rond delen van onderzoeksbestanden.

1.2 De vraagstelling

De vraagstelling voor deze notitie heeft een strikt juridische en een meer rechtspolitieke kant.

Strikt juridisch is:

- Onder welke voorwaarden kunnen onderzoeksbestanden in juridische zin voorwerp van eigendom zijn;
- Bij wie berust dan het eigendom van onderzoeksbestanden?

Rechtspolitiek is:

- Welke claims op onderzoeksbestanden zijn gerechtvaardigd?
- Hoe kan een gedragscode voor het delen van onderzoeksbestanden worden bereikt die recht doet aan alle onderscheiden belangen?

Deze vragen worden in de onderhavige achtergrondnotitie behandeld.

1.3 Opzet en terminologische verduidelijking

Deze notitie bespreekt eerst de juridische aspecten. Die bespreking is niet uitputtend, maar behandelt de hoofdlijnen die van belang zijn om tot afweging over te kunnen gaan.

Vervolgens worden de aan de orde zijnde belangen besproken. Geconcludeerd wordt dat een evenwicht moet worden bereikt en dat een strikt juridische benadering hiertoe niet geëigend is. Tot slot worden de elementen besproken van de voorgestelde Gedragscode over delen van onderzoeksbestanden waarmee dit evenwicht naar mijn mening wel kan worden bereikt.

De term ‘onderzoeksgegevens’ is hier in een brede betekenis gebruikt. Het zijn alle gegevens waarover een onderzoeker ten behoeve van diens onderzoek kan beschikken en die in de loop van diens onderzoek verzameld worden. Met ‘het onderzoeksbestand’ wordt bedoeld dat deel van de onderzoeksgegevens die specifiek in het kader van een bepaald onderzoek zijn verzameld. Met een bepaald onderzoek wordt bedoeld op een onderzoek met een externe subsidie- of opdrachtgever die de beantwoording van deze bepaalde onderzoeksvraagstelling voor meer dan 50 % bekostigt.

Het is natuurlijk de vraag of het onderzoeksbestand wel los kan worden gezien van de onderzoeksgegevens. In technische zin meestal wel, maar voor de interpretatie van het bestand en voor de onderzoeksresultaten veelal niet. Het bestand dat een bepaald onderzoek oplevert, moet worden geïnterpreteerd. Los van de methodologie die daarbij wordt gebruikt (en waaraan weer mede gegevens ten grondslag liggen), wordt een onderzoeksbestand

veelal pas zinvol in vergelijking met andere onderzoeksgegevens.² Op die problematiek kom ik hierna terug.

2. Juridische aspecten van eigendom onderzoeksgegevens en onderzoeksbestanden

2.1 *Eigendom van gegevens in het algemeen*

Naar huidig Nederlands recht bestaat er niet de mogelijkheid van eigendom op iets onstoffelijks als data of gegevens.³

Dat betekent niet dat deze vogelvrij zijn. De stoffelijke gegevensdragers en de ruimten waarin deze zich bevinden, zijn wel degelijk voorwerp van eigendom. De bepalingen in het Wetboek van Strafrecht over verandering en wissen van de gegevens die in een 'geautomatiseerd werk' (art. 350a, 350 b WvSr.) zijn opgeslagen, veronderstellen daarboven toch een soort eigendomsrecht van degene die deze gegevens in het automatiseringssysteem heeft opgeslagen.

Ook het burgerlijk recht beschermt degene die bepaalde gegevens, bijvoorbeeld ten behoeve van bedrijfsprocessen, heeft verzameld. Een ander, die deze gegevens zonder toestemming verkrijgt en daarmee aan de haal gaat, bijvoorbeeld door deze aan een concurrent te verkopen, handelt in principe onrechtmatig.⁴ Een beding in de arbeidsovereenkomst dat de gegevens die een werknemer in het kader van diens dienstverband heeft verzameld, niet mogen worden meegenomen naar een nieuwe werkgever, is min of meer standaard. Zo dit niet expliciet is opgenomen, zal de rechter al snel oordelen dat zulk 'meenemen' onrechtmatig is.

Hoewel dus gegevens als zodanig geen voorwerp van eigendom zijn in de klassieke juridische betekenis van het woord (vandaar ook 'eigendom' in de titel van deze notitie), worden zij wel degelijk op verschillende manieren beschermd.

Dat is zonder meer het geval indien deze gegevens zodanig worden gepresenteerd dat zij voorwerp zijn van intellectueel eigendomsrechten (IE).

2.2 *IE op gegevens: het databankenrecht*

2.2.1 Inleiding

Twee IE zijn hier van belang: het auteursrecht en het databankenrecht. Over het auteursrecht kan ik kort zijn. Voor de gelding van het auteursrecht moet het werk oorspronkelijk zijn, in de zin van het persoonlijke stempel van de auteur dragen (en ook openbaar gemaakt). Een artikel geldt als een oorspronkelijk werk, een serie gegevens zoals een telefoonboek niet. Onderzoeksgegevens of het bestand vallen, hoe oorspronkelijk de methodologie ook was waarmee de gegevens zijn verzameld, niet onder het auteursrecht. Met dezelfde methodologie was men tot dezelfde uitkomsten gekomen. Daarom is het ook wetenschap. Het onderzoeksrapport valt weer wel onder het auteursrecht. Het bevat een oorspronkelijke presentatie van de feiten. Een andere onderzoeker zou het anders hebben opgeschreven.

Maar hier gaat het niet om het rapport, maar om de onderliggende gegevens, die niet alle in het rapport worden weergegeven.

Onder omstandigheden kan het databankenrecht daarop van toepassing zijn. Het databankenrecht is een apart recht waaraan een Europese Richtlijn ten grondslag ligt. Het is gecreëerd om de investeringen van de maker (in de Nederlandse vertaling 'de producent') van de

² Bovenberg noemt het concept van data streams waarin data eenheden nauwelijks kunnen worden onderscheiden. J. A. Bovenberg, Property Rights in Blood, Genes and Data, Naturally Yours?, Leiden/Boston, 2006, p. 173.

³ Bijvoorbeeld C. Prins, eigendom op informatie: economische realiteit maar juridische fictie, NJB 2005, 263.

⁴ Ch. Gielen, Bescherming van bedrijfsgeheimen, Zwolle, 1999, paragraaf 2.2.

databank te beschermen tegen commercieel hergebruik van een substantieel deel van die databank.

Bovenberg⁵ suggereert dat de ‘funding body’ (oftewel financier) de producent van de databank zou zijn. Hij noemt overigens ook enkele problemen bij dit standpunt. Er kunnen meerdere funding bodies zijn. De (eventuele) eigenaar van het databankenrecht zou dit recht dan moeten beheren, dus bepalen wie toegang heeft tot de databank en onder welke omstandigheden. Daartoe zijn funding bodies meestal niet toegerust. Belangrijker zijn echter de juridische bezwaren, die ik hieronder behandel.

2.2.2 Het toepassingsbereik van het databankenrecht

Onder een databank volgens het databankenrecht dient voorzover hier relevant (en daarmee sterk samenvattend) te worden verstaan:

- Een systematisch of methodisch geordende verzameling van gegevens;
- Die afzonderlijk toegankelijk is;
- Die in kwalitatief of kwantitatief opzicht getuigt van een aanzienlijke investering.

Een onderzoeksbestand voldoet in de regel in beginsel aan die omschrijving.

Om evenwel de werking van het databankenrecht te begrijpen dient men het volgende te weten.

Het databankenrecht is een bijzondere vorm van IE. Naast het genoemde auteursrecht is ook het octrooirecht een vorm van IE. De achtergrond van IE is dat de maker van een onstoffelijk werk wordt beschermd wanneer deze dat werk aan anderen beschikbaar stelt. Anders dan een ‘ding’ dat per definitie maar bij één persoon kan berusten, kan een onstoffelijk werk worden gekopieerd en dus bij meerdere personen berusten. Het IE bepaalt wanneer men een eigendomsrecht heeft op een onstoffelijk werk en onder welke voorwaarden een ander daar dan een gebruiksrecht op kan verkrijgen.

Het IE mist toepassing als men als het ware op het werk blijft zitten en het dus verder aan niemand ter beschikking stelt. Dat geldt ook voor het databankenrecht. Het databankenrecht begint pas te werken indien de onderzoeker/instelling zou overwegen het onderzoeksbestand aan een derde te beschikking te stellen.

2.2.3 De subsidiegever is geen producent in de zin van de Databankenwet

Omdat het databankenrecht is gebaseerd op Europese regelgeving⁶ kan het Europese Hof van Justitie (EHJ) uitspraak doen over de uitleg hiervan. Het EHJ heeft in een tweetal uitspraken de reikwijdte van het begrip investering in de databank (dus de beschermde databank) flink ingeperkt.⁷ Het komt er op neer dat het verzamelen van de gegevens, in het jargon van het EHJ ‘het creëren van de elementen’, niet onder de bescherming valt. Er moet meer aan de hand zijn, al is het omslagpunt moeilijk te bepalen.⁸ In de Nederlandse rechtspraak kreeg in ieder geval de NVM nul op het rekest toen deze stelde dat de Funda site een databank in de zin van de Nederlandse Databanken wet is.⁹ Deze wet heeft de Europese Richtlijn in het Nederlands recht geïmplementeerd. De beschrijving van de te koop staande huizen (creëren elementen) valt niet onder de bescherming van de Databanken wet. Het ontsluiten op een website zou vallen onder de normale bedrijfsactiviteiten en zou geen substantiële investering zijn.

Toegepast op een onderzoek betekent dit naar mijn mening dat het bestand in de regel niet als een beschermde databank kan worden aangemerkt. Het creëren van de elementen, bijvoorbeeld de bevraging van respondenten, wordt gesubsidieerd. Maar dat onderdeel van de databank wordt juist niet beschermd. Dan is er het eventuele meerdere, “de controle op”

⁵ o.c., nt. 1, p. 179.

⁶ Richtlijn 96/9/EG

⁷ Hof van Justitie EG 9-11-2004, C-203/02.

⁸ Namelijk “dat de verkrijging van deze elementen of de controle dan wel de presentatie daarvan [...] een in kwantitatief dan wel kwalitatief opzicht substantiële investering heeft vergd, los van de middelen die voor het creëren van deze elementen zijn aangewend”.

⁹ Hof Arnhem 04-07-2006, LJN AY0089.

etc. Het lijkt mij dat de subsidiegever daar juist niet investeert. De onderzoeker moet uitvoerige documentatie overleggen om aan te tonen dat het onderzoek aansluit bij eerder onderzoek en op basis van kennis en vaardigheden plaatsvindt die de onderzoeker daartoe bij uitstek geschikt maken. Met andere woorden, dat meerdere is er al en is de reden waarom, en niet waarin, de subsidiegever 'investeert' in de zin van een subsidie verstrekt.

In dit verband is ook het aan het slot van paragraaf 1 genoemde probleem relevant dat het bestand niet los kan worden gezien van onderzoeksgegevens. Zonder die inbedding zou het onderzoek geen relevante uitkomsten kunnen opleveren. Aan die onderzoeksgegevens heeft de opdracht- of subsidiegever voor een bepaald onderzoek echter niet of slechts gering bijgedragen.

Van een substantiële investering in meer dan het verkrijgen van de elementen zou uitsluitend sprake zijn indien de subsidie of opdracht er speciaal op is gericht om een publiek toegankelijk bestand te maken, en naast op het verwerven van de gegevens daartoe door de subsidie- of opdrachtgever ook speciaal wordt geïnvesteerd in de ontsluiting van het onderzoeksbestand, zoals de zichtbaarheid van bepaalde elementen van het bestand op een site, de zoekmethodiek op de site, etc.

Bij het overgrote deel van het gesubsidieerde onderzoek is dat echter niet het geval.

2.2.4 Bovendien levert subsidiegever = eigenaar databankenrecht op onderzoeksbestand een paradoxale situatie op

Daarnaast is er een tweede argument om de opdracht- of subsidiegever niet van rechtswege als de eigenaar van het bestand in de zin van het databankenrecht te beschouwen.

In de Nederlandse Databankenwet is het begrip 'maker van de databank' vervangen door 'producent'.¹⁰ Daarmee wordt de opdrachtgever bedoeld, indien die er is.¹¹ Deze wijze van omzetting van Richtlijn 96/9/EG is bekritiseerd¹² en nog niet door het EHJ getoetst.

Wat daarvan zij, de subsidiegever kan – in het algemeen – ook niet als opdrachtgever worden gekenmerkt, zoals bedoeld in definitie van de Nederlandse Databankenwet. Daarmee doel ik niet op het feit dat een 'opdracht' – dat is privaatrecht en veronderstelt een overeenkomst – wat anders is dan een subsidie, die in een beschikking uitmondt. Ik doel op een fundamentele kwestie.

De rationale achter de Richtlijn was het beschermen van de economische belangen van de producent. Een ander zou niet zonder passende compensatie van de investering in de databank mogen profiteren.

Subsidiegevers lopen echter geen enkel economisch risico. Het 'risico' voor hen is van geheel andere orde. Zij staan voor een publiek belang. In dat kader wordt het onderzoek gesubsidieerd. De resultaten in de zin van één of meer publicaties dienen openbaar te worden zodat wetenschap en beleid er op kunnen voortbouwen. Dat is standaard.

Met de claim op grond van het databankenrecht wensen subsidiegevers het tegenovergestelde van waar het begrip producent in de Databankenwet voor is bedoeld. Namelijk niet voorkómen dat een ander van die investering profiteert, maar voorkómen dat anderen – voorzover dat al niet door de publicaties gebeurt – er niet van kunnen profiteren. Dat publieke doel is op zich zelf legitiem, indien vertaald als veiligstellen dat het bestand onder voorwaarden kan worden toegevoegd aan de stroom van onderzoeksgegevens waardoor andere gekwalificeerde onderzoekers daarvan kunnen profiteren.

Echter, dat kan de subsidiegever als producent niet bereiken. De producent heeft het uitsluitend recht om toestemming te verlenen voor bepaald gebruik van de databank door een derde. Dat is immers ook de rationale. De producent kan de maker niet via het databankenrecht dwingen om de databank door anderen te laten gebruiken. Daarvoor is het databankenrecht als vorm van IE niet gemaakt.

¹⁰ Art. 1 1 onder b Databankenwet

¹¹ Kamerstukken 26108, nr. 3.

¹² Ch. Gielen, D.W.F. Verkade, *Intellectuele Eigendom*, Dordrecht, 2005, aantekening bij art. 1 Databankenwet, p. 157.

2.3 In het kader van de voorwaarden bij een opdracht of subsidie

De vorige paragraaf behandelde de claim dat het eigendom op grond van het databankenrecht zonder meer aan de opdracht- of subsidiegever zou toevallen. Een andere situatie is dat dit in het kader van een opdracht- of subsidieovereenkomst speciaal zou zijn bedongen. Indien men daar mee akkoord is gegaan, gaat het bovenstaande verhaal niet op.

2.3.1 De ARVODI en modelovereenkomst van het WODC

Zo'n beding komt bijvoorbeeld voor in de algemene voorwaarden van het rijk voor het verlenen van opdrachten tot dienstverlening (de ARVODI) en in de model opdracht overeenkomst van het WODC.

De ARVODI worden stelselmatig gehanteerd voor contractonderzoek waarbij het rijk opdrachtgever is. Het Wetenschappelijk Onderzoek en Documentatiecentrum (WODC) van het Ministerie van Justitie is een belangrijke opdrachtgever bij verslavings- en veiligheidsonderzoek. Niet helemaal onrealistisch is dat dit model ook zal worden gehanteerd voor subsidiegevers voor zorgonderzoek. Vandaar dat er hieronder verder op wordt ingegaan.

De WODC modelovereenkomst stelt in art. 4 dat “met het oog op eventueel toekomstig wetenschappelijk onderzoek (door het WODC) alle in het kader van het onderzoek verzamelde gegevens aan de staat worden overgedragen” Hier wordt niet gesproken van eigendom van de gegevens, omdat de opstellers zich hebben gerealiseerd dat eigendom van gegevens als zodanig juridisch niet bestaat. Met ‘overdragen’ van het bestand wordt evenwel hetzelfde bereikt. De staat (het WODC) heeft het, en de onderzoeker niet meer.

Indien het doel van het WODC uitsluitend zou zijn om dit bestand toe te voegen aan de eigen onderzoeksgegevens, zou dit ook op een minder drastische manier kunnen worden bereikt. Namelijk ‘beschikbaarstellen’ in plaats van ‘overdragen’. In dat geval zou de onderzoeker het bestand kunnen houden maar kan het WODC ook over het bestand beschikken. Gegevens kunnen immers worden gedupliceerd.

2.3.2 Subsidievoorwaarden

Ook in subsidievoorwaarden blijkt een dergelijk beding regelmatig voor te komen.¹³ De (algemene) Subsidieregeling VWS subsidies¹⁴ stelt dit overigens niet.

2.3.3 Analyse van deze voorwaarden

Indien een dergelijke voorwaarde wordt geaccepteerd, zit men er aan vast. Niet accepteren is in de praktijk lastig, met uitzondering van de overdracht van persoonsgegevens, zie hierna. In een aantal gevallen is het mogelijk gebleken om de scherpe kantjes er af te halen. Zo zou men kunnen onderhandelen en er op uit komen dat het onderzoeksbestand niet wordt overgedragen, maar beschikbaar wordt gesteld. Dan kan men het zelf nog kan gebruiken. Daarnaast heeft een kennisinstituut weten te bedingen dat het de gegevens als het ware voor de subsidiegever bewaart en mag opnemen in de onderzoeksgegevens van het instituut. De problematiek is daarmee verschoven tot het moment dat de langdurige subsidiering voor een reeks doorlopende projecten zou worden beëindigd.

De meer fundamentele vraag is of deze voorwaarde van de subsidie- of opdrachtgever wel gerechtvaardigd is. ‘Wie betaalt, bepaalt’ zal deze denken. Maar op zich mag dat uitgangspunt zeker in het publieke domein geen grondslag zijn om onredelijke eisen te stellen. Zo zal de publieke subsidiegever de wetenschappelijke integriteit van het onderzoek niet mogen beïnvloeden.¹⁵ Uit de analyse van de claim op grond van het databankenrecht is gebleken, dat de subsidiegever slechts bijdraagt aan het onderzoek waarmee het bestand kan worden gegeneerd. Dan komt de claim op het hele onderzoeksbestand overtrokken over. Tegelijk is die bijdrage voor dit bestand wel wezenlijk en is er ook een zekere rechtvaardigheid in de claim dat dit bestand voor verder onderzoek beschikbaar is. In paragraaf 3 zal daarop worden teruggekomen.

¹³ Uit mijn adviespraktijk hiervan een aantal voorbeelden.

¹⁴ Staatscourant 2006, nr. 116.

¹⁵ In ieder geval in het verleden is dat niet voor elke publieke subsidiegever een gemakkelijke opgave gebleken, zie F. Köbben, *De weerbarstige waarheid*, Prometheus, 1991.

2.4 Privacy en toestemming als randvoorwaarde

Bij het overdragen van de integrale databestanden speelt uiteraard ook een privacy problematiek. De randvoorwaarde voor delen is uiteraard de privacy van de betrokkenen. Indien onder onderzoeksbestand ook wordt begrepen de lijst met respondenten, voorzover van toepassing, zal daarmee nooit akkoord kunnen worden gegaan. Ook in andere gevallen zal de privacy van de betrokkenen aan de overdracht van het onderzoeksbestand in de weg kunnen staan. Voor de van toepassing zijnde regels zij verwezen naar de FMWV gedragscode ‘Goed Gedrag’.¹⁶

Indien men bestanden van (deels) dezelfde betrokkenen wil samenvoegen en daarbij administratieve meerlingen wil vermijden, is volledig anonimiseren geen optie. Dat geldt meestal ook indien men bestanden van nog lopende onderzoeken wil koppelen. Toepassing van PET (privacy enhancing technologies) kan er dan toe bijdragen dat zulke onderzoeksbestanden voor meerdere onderzoeken kunnen worden ingezet. Dat vereist echter een bredere aanpak dan de exclusieve eigendom van elk onderzoeksbestand door elke subsidiegever.

Een tweede probleem bij gebruik ten behoeve van vervolgonderzoek is dat soms de toestemming van de betrokkenen zich niet voor gebruik voor dit andere doel uitstrekt, ook indien hun privacy daarbij niet rechtstreeks in het geding is.

Uiteraard is de reikwijdte van de toestemming eveneens een voorwaarde voor hergebruik van het onderzoeksbestand. En pleit dit voor niet te beperkte toestemming (‘broad consent’) Deze kwestie wordt hier verder niet behandeld.

3. De relevante afweging en beleidsmatige aspecten

3.1 De aan de orde zijnde belangen

Van verschillende kanten is er de wens cq. noodzaak dat bestanden van het ene onderzoek ook voor ander onderzoek door eventueel andere onderzoekers kunnen worden gebruikt. Deze wens wordt met name door opdracht- en subsidiegevers ingegeven in verband met efficiënte besteding van onderzoeksgelden. De noodzaak wordt ook door onderzoekers gevoeld doordat onderzoek meer en gebruik moet maken van grote databestanden. Combineren van afzonderlijke onderzoeksbestanden en datasets kan daaraan bijdragen. Indien gegevens worden ontleend aan lichaamsmateriaal geldt bovendien dat dit lichaamsmateriaal schaars is (anders dan gegevens kan lichaamsmateriaal in de regel niet worden vermenigvuldigd¹⁷). Gegevens ontleend aan lichaamsmateriaal dienen dus zoveel mogelijk te worden hergebruikt.

Additioneel speelt ook mee dat de ‘enquêtedruk’ op de bevolking of patiënten grenzen heeft.

Daar staan echter andere wensen en belangen van onderzoekers en de instituten tegenover.

Onderzoekers willen erkenning voor hun inspanningen, ook als die in een brede stroom van onderzoeksgegevens worden opgenomen. Zoals Bovenberg ook opmerkt, wetenschappelijk onderzoek kan niet zonder voort te bouwen op de resultaten van anderen maar wetenschappelijke erkenning krijgt men echter als individu. Indien het voortbouwen gebeurt op basis van publicaties is er voor de ‘erkenning’ overigens niet direct een probleem. Dan blijkt het voortbouwen, tenminste voor de kenners, wel uit de referenties. Indien het voortbouwen er uit bestaat om bestanden op te nemen in onderzoeksgegevens, is het lastiger die erkenning te realiseren.

Voor onderzoekers en instituten vormen de ‘eigen’ onderzoeksgegevens ook een strategische waarde. Nieuw onderzoek kan daarop aansluiten, de exclusiviteit betekent dat men dergelijk nieuw onderzoek gemakkelijker zal kunnen verwerven.

Daarbij wordt wellicht wel eens vergeten dat niet-exclusiviteit twee kanten op werkt. De ‘concurrent’ krijgt een bepaalde toegang tot de ‘eigen’ bestanden, maar dat geldt andersom ook voor de bestanden van de ‘concurrent’.

¹⁶ Staatscourant 2004, nr. 82. Ook te vinden onder <http://www.federa.org/?s=1&m=60>

¹⁷ Cellijnen zijn uiteraard een uitzondering, maar daarom gaat het niet bij onderzoek met restmateriaal.

Overigens zitten aan hergebruik ook methodologische haken en ogen, zoals onvoldoende eenheid van taal, onduidelijkheid wat nu exact eerder is gemeten of gevonden, etc. Daarop wordt hier niet verder ingegaan.

3.1.1 Een brede beweging

Aandringen op het delen van gegevens is een brede beweging. De NIH in de VS stelt delen al langer als voorwaarde voor subsidiering.¹⁸ Het delen van bestanden was ook het uitgangspunt van een publicatie van de US National Research Council.¹⁹ Van belang is dat het hier gaat om voorwaarden voor het delen van bestanden. De NIH eigent zich de bestanden niet toe. De voorwaarden bevatten voorts talloze nuances en houden rekening met de privacy aspecten.

De OECD heeft eveneens aanbevelingen gedaan voor het delen van bestanden verkregen bij publiek gefinancierd onderzoek.²⁰

Het sociaal-wetenschappelijk onderzoek deelt al bestanden via het DANS instituut van de KNAW.²¹ Het lijkt hier overigens te gaan om bestanden van afgesloten onderzoek, die voor de betrokken onderzoekers geen actueel belang meer vertegenwoordigen.

Ook kan de beweging van de ‘creative commons’ worden genoemd die een alternatief wil bieden op exclusieve auteursrechtelijke claims van uitgevers. Voorafgaand aan delen is dat men weet wie welke bestanden heeft. In dit verband kan in Nederland de nieuwe door VWS opgezette site www.zorggegevens.nl worden genoemd.

3.2 De balans

Het lijkt mij een gerechtvaardigd publiek belang dat onderzoeksbestanden voor volgend onderzoek kunnen worden gebruikt, ook als dat vervolgonderzoek primair door andere onderzoekers plaatsvindt.

Het exclusieve eigendom van de subsidiegever van het onderzoek dat het ene bestand heeft gegeneerd, is daartoe echter een ongeschikt middel, om de volgende redenen:

- Een ander eveneens noodzakelijk bestand berust dan mogelijk bij een andere subsidiegever;
- Er zijn mogelijk meerdere subsidiegevers voor het ene bestand, of het is deels uit eigen middelen gefinancierd. Zie ook de bespreking van de rechtvaardigheid van een dergelijke claim.
- Het is de vraag of de subsidiegever wel de geschikte hoeder van het onderzoeksbestand is, met de mogelijkheden om het voor ander onderzoek te ontsluiten. Daartoe moet het onderzoeksbestand mogelijk worden bewerkt, moeten de methodologie, de gebruikte definities etc. exact bekend zijn. Anders worden appels en peren bij elkaar opgeteld. Men heeft dus bijna steeds de oorspronkelijke samensteller van het onderzoeksbestand nodig voor het vervolgonderzoek waarbij van dit bestand gebruik worden gemaakt.
- Het onderzoeksbestand zal slechts als anoniem (gestript van direct of indirect identificerende gegevens) of gecodeerd bestand kunnen worden overgedragen, terwijl voor zinvol koppelen soms meer nodig is;

Daarnaast is, zoals hiervoor besproken, de exclusieve claim ook onrechtvaardig tegenover de onderzoeker en diens instelling. Er zit veel meer achter dat onderzoeksbestand dan de financiële bijdrage van de subsidie- of opdrachtgever. Dat meerdere vindt geen erkenning bij exclusief eigendom van deze.

¹⁸ Zie: <http://www.med.nyu.edu/spa/policies/nih/DataSharing.html>

Voor een toepassing onder andere <http://datasharing.net/>

¹⁹ Sharing Publication-Related Data and Materials, Responsibilities of authorship in the life sciences, Washington DC, 2003.

²⁰ OECD, Acces to publicly financed research, 2000/2001.

²¹ www.dans.knaw.nl

Er zal daarom een evenwicht moeten worden bereikt. Delen van onderzoeksbestanden onder onderzoekers moet het uitgangspunt zijn, maar wel onder voorwaarden. Dat zou kunnen worden uitgewerkt in een ‘datasharing policy’ enigszins vergelijkbaar met die van de NIH.

De elementen voor een dergelijke policy en een mogelijke borging van deze worden in de volgende paragraaf uitgewerkt.

4. Hoe onderzoeksbestanden te delen?

4.1 De uitgangspunten

Hieronder worden, zonder volledig te willen zijn, de voornaamste uitgangspunten voorgesteld voor het delen van onderzoeksbestanden:

- Het uitgangspunt is dat onderzoeksbestanden tussen onderzoekers kunnen worden gedeeld;
- Dat wil zeggen dat de ene onderzoeker meewerkt om het bestand dat deze bij een onderzoek heeft gegeneerd te ontsluiten voor onderzoek waarbij dat bestand nodig is voor een andere onderzoeker;
- Onderzoekers zijn dus transparant over welke bestanden zij beschikken.
- De eerdere onderzoeker levert een bijdrage aan het volgende onderzoek, heeft immers deel van de dataverzameling gedaan, en wordt daarvoor gehonoreerd in de publicaties gebaseerd op het volgende onderzoek overeenkomstig de geldende wetenschappelijke gebruiken.
- De kosten voor het ontsluiten mogen niet meer zijn dan de reële kosten die de instelling van de eerdere onderzoeker daartoe moet maken;
- Indien de eerdere onderzoeker meer moet doen, zoals dat bijdragen aan methodologie, nieuwe bewerkingen op het bestand, moet dat meerdere ook worden gehonoreerd, zowel in de ‘eer’(opschuiven in de lijst auteurs of id.) als, waar nodig, financieel²²;
- Het eerste gedachtestreepje (en de uitwerking daarvan), geldt eerst na een bepaalde tijd na de totstandkoming van het onderzoeksbestand en de daarop gebaseerde met de financier overeengekomen publicatie(s). Tot die tijd mag de onderzoeker, het onderzoeksbestand exclusief gebruiken voor vervolgonderzoek (zich daarbij wel realiserend dat hij daar mogelijk bestanden van andere onderzoekers nodig heeft, voor wie dan hetzelfde geldt). Deze termijn mag enerzijds niet te kort, anderzijds niet te lang. Zij zal in de hierna te behandelen Gedragscode als richtlijn moeten worden opgenomen.
- Die termijn kan onder omstandigheden worden verlengd, indien verdergaande IE, namelijk een patentaanvraag, aan de orde is.
- Een en ander geldt uiteraard onder de voorwaarde van bescherming van de betrokkenen (respondenten, donoren van lichaamsmateriaal), zoals beschreven in sectie 2.4.

4.2 Uitwerking en borging

Uitwerking van het bovenstaande zal in een breed overleg van onderzoekers, de instituten waaraan zij zijn verbonden en subsidiegevers moeten plaatsvinden. Aldus kunnen gedragsregels of een Gedragscode tot stand komen.

Onderzoekers en hun instellingen zullen zich aan deze gedragsregels moeten verbinden. Subsidiegevers kunnen in hun subsidievoorwaarden opnemen dat de onderzoeker deze gedragsregels moet onderschrijven, en opdrachtgevers kunnen hetzelfde doen in de voorwaarden voor het verstrekken van de opdracht.

Onder omstandigheden zal er discussie kunnen ontstaan over toepassing in een concreet geval, bijvoorbeeld een stelling door de ene onderzoeker dat diens bestand voor een ander

²² Indien de vervolgonderzoeker meerdere bestanden wil combineren via PET en de inschakeling van een trusted third party worden die kosten uiteraard door het vervolgonderzoek gedragen.

onderzoek helemaal niet nodig is, of dat de methodologie niet deugt, of dat privacy voorwaarden delen in de weg staan, etc. Mogelijk kan hiertoe een arbitragecommissie in het leven worden geroepen.

24-1-2008

Evert-Ben van Veen

Afkortingen

<i>ADL</i>	Algemene Dagelijkse Levensverrichtingen
<i>BBMRI</i>	Biobanking and Biomolecular Resources Infrastructure
<i>BSN</i>	Burger Service nummer
<i>CBS</i>	Centraal Bureau voor de Statistiek
<i>DANS</i>	Data Archiving and Networked Services
<i>DBC</i>	Diagnose-behandel combinatie
<i>EPD</i>	Elektronisch Patient Dossier
<i>FMWV</i>	Stichting Federatie van Medisch Wetenschappelijke Verenigingen
<i>GA2LEN</i>	Global Allergy and Asthma European Network
<i>GGD</i>	Gemeenstelijke Gezondheidsdienst
<i>HES</i>	Health Examination Survey
<i>ISBN</i>	International Standard Book Number
<i>KNAW</i>	Koninklijke Nederlandse Academie van Wetenschappen
<i>LINH</i>	Landelijk Informatie Netwerk Huisartsenzorg
<i>LMR</i>	Landelijke Medische Registratie
<i>Morgen</i>	Monitoring van Risicofactoren en Gezondheid in Nederland
<i>MRC</i>	Medical Research Council
<i>NHANES</i>	National Health and Nutrition Examination Survey
<i>NIH</i>	National Institutes of Health
<i>NWO</i>	Nederlandse organisatie voor Wetenschappelijk Onderzoek
<i>OECD</i>	Organisation for Economic Co-operation and Development

<i>P3G</i>	Public Population Project in Genomics
<i>RGO</i>	Raad voor Gezondheidsonderzoek
<i>RIVM</i>	Rijks Instituut voor Volksgezondheid en Milieu
<i>SCP</i>	Sociaal en Cultureel Planbureau
<i>VSNU</i>	Vereniging van Universiteiten
<i>VWS</i>	Ministerie voor Volksgezondheid, Welzijn en Sport
<i>WHO</i>	World Health Organisation
<i>ZonMw</i>	Nederlandse organisatie voor gezondheidsonderzoek en zorginnovatie