
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.