Inventory Patient Registries in the Netherlands

DANS (M. Grootveld, M. Polhout, H. Tjalsma)
Mondriaan (S. Visser, W. De Bruijn)
RIVM (R. Nugteren, M. Weda)
NIVEL (R. Verheij, L. Van Dijk)

ISBN:978-94-6122-158-2
Utrecht, 2012

Content

Preface

This “Inventory Patient Registries in the Netherlands” focuses on the availability of data sources relevant to rational pharmacotherapy in the Netherlands. The study was written as preparation for the ZonMw Programme Rational Pharmacotherapy which started February 29, 2012. The study provides an overview of available registries, access to existing registries, gaps in the availability (and access to) registries and the organisation of (new) registries and was performed by the National Institute for Public Health and the Environment (RIVM), Mondriaan, Data Archiving and Network Services (DANS), and NIVEL, Netherlands institute for health services research. We would like to express our gratitude to the representatives of databases who participated in this study as well as to the ZonMw reading committee and the experts who granted us an interview or consultation. In addition, we thank the experts and representatives of meta initiatives we consulted for their input for chapters 5 to 7. Finally, we thank Margreet Bloemers en Benien Vingerhoed-van Aken, both from ZonMw, for their constructive support during the process of this study.

The authors
April 2012

1. Introduction

This report describes a preparatory study for the newly developed programme Rational Pharmacotherapy from the Netherlands Organisation for Health Research and Development (ZonMw). This study called "Inventory Patient Registries in the Netherlands" focuses on the availability of data sources relevant to Rational Pharmacotherapy in the Netherlands, on shortcomings in the availability of data as well as on conditions for effective use of registries.

1.1 The ZonMw programme Rational Pharmacotherapy

February 29 2012, the Netherlands Organisation for Health Research and Development (ZonMw) started a new programme called Rational Pharmacotherapy (RP).[1] Rational use of medicines requires that "patients receive medications appropriate to their clinical needs, in doses that meet their own individual requirements, for an adequate period of time, and at the lowest cost to them and their community" (WHO definition; http://www.who.int/medicines/areas/rational_use/en/). The aims of the ZonMw RP programme are:[2]

- Optimising effective, safe and efficient use of medications; this is accomplished by the systematic generation, dissemination and implementation of knowledge;
- Building an infrastructure that allows for adequate and independent ways to answer questions pertaining to rational pharmacotherapy;
- Increasing knowledge and expertise on rational pharmacotherapy in physicians, pharmacists and the general public;
- Enhancing knowledge and expertise concerning specific groups of patients (such as children, the elderly, pregnant women, patients with rare diseases, people of non-Dutch descent), as well as the dissemination and application of this knowledge;
- Strengthening the scientific foundation (evidence) of guidelines and policy decisions;
- Increasing practical knowledge which contributes to health care efficiency;
- Increasing practical knowledge on health care safety, reduction of hospital admissions (thus alleviating the burden on personnel), and increasing quality of care.

Structure of the RP programme

To realise the aims of the RP programme, ZonMW introduces three groups of activities (Figure 1). The first group refers to the research infrastructure. Infrastructure is an essential prerequisite for research: high quality data are needed. The RP programme has to strengthen and extend the existing RP research infrastructure in the Netherlands as well as to create conditions for optimal use of both existing and newly created registries. A second focus of the programme is research. Research within the RP

[1] http://www.zonmw.nl/en/programmes/pharmacotherapy/general-information/
Derived from: ZonMw. Proposal ZonMw Programme Rational Pharmacotherapy. Optimising effective, safe and efficient use of medicines in health care, The Hague, 2011.

programme needs to: a) fill existing gaps in the evidence needed for guidelines and healthcare standards; b) support policy makers in their decisions concerning medicines and pharmacotherapy and c) provide answers for questions and problems that arise in daily clinical practice. The third aspect is implementation which is the ultimate goal of all activities within the programme. Within this part of the programme, special attention is paid to: a) the systematic application of available knowledge on pharmacotherapy, disseminating news and information on best practices and keeping an eye out for any gaps in the evidence and b) the clustering of relevant knowledge (both existing and newly developed within the programme) and making it accessible.

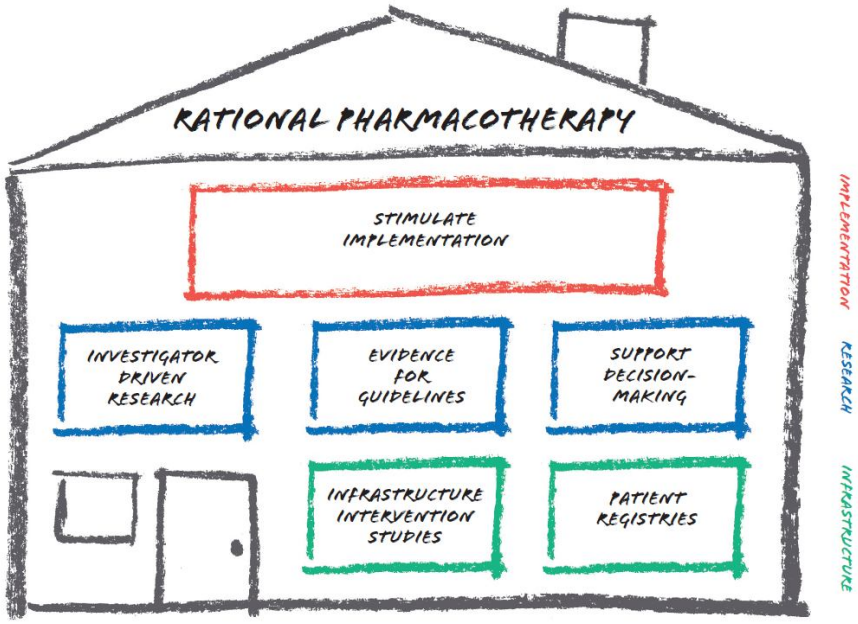

Figure 1.1: Summary of the RP programme (ZonMw 2011)

1.2 Why a preparatory study "Registries Rational Pharmacotherapy"?

The RP programme wants to enhance adequate use of existing infrastructures and to promote good practices in setting up infrastructures, both for intervention studies and observational research. The project described in this report focuses on observational research. The RP programme text clearly describes the relevance of this subject for the programme: "Observational research can be really close to daily practice. It is an excellent way to study trends and problems in an unbiased population. Automated data analysis allows for the recognition of patterns that would otherwise escape notice. This type of research requires a good infrastructure, clear protocols on data collection, storage and analysis and, of course, good hardware and a professional organisation".[3] These data that come together in patient registries that, for example, include data on utilisation of medicines, intended effects of medicines, unintended effects (adverse events, complications) but also on quality of life and costs. Different types of patient registries exist, for example automated routine registrations from health care

[3] ZonMw. Proposal ZonMw Programme Rational Pharmacotherapy. Optimising effective, safe and efficient use of medicines in health care, The Hague, 2011.

professionals, biobanks (a type of biorepository which stores human biological samples), but also survey data collected in patient cohorts. In fact, patient registries used for observational research include all possible data collected on a continuous/longitudinal basis in patient groups in uncontrolled settings.

Within the RP programme new patient registries can be developed, especially registries that focus on optimising pharmacotherapeutical treatments. As such, observational studies can be performed which are complementary to controlled clinical studies performed for obtaining registration of medicines. Outcomes of observational studies can be used in daily clinical practice, in clinical guidelines, in monitoring and in policy making processes. Therefore, the RP programme will enhance bundling knowledge and experience on patient registries as well as promoting that data from patient registries will become available for research. This requires good and trusted management and care of research data. In a report by the Dutch Advisory Council on Health Research[4] "*Van gegevens verzekerd (Securing the data supply)"* it was concluded that: "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." Instruments are needed to ensure that conditions for data sharing are optimal in terms of accessibility, privacy and incentives for parties who collect data.

The Netherlands already has quite a large infrastructure when it comes to patient registries and meta information about a number of patient registries is available (www.zorggegevens.nl). However, registries are heterogeneous and not all of them are accessible for (external) researchers. Moreover, it is not known, what the need is for new, additional data, both in terms of completely new registries and in terms of additional data collection within existing patient registries. Also, there is a need to know how governance of patient registries can be organised. Therefore, this preparatory study on patients registries was performed. It focuses on the availability of data sources relevant to rational pharmacotherapy in the Netherlands, on shortcomings in the availability of data as well as on conditions for effective use of registries. The study consists of four main topics :

1. Overview of available registries (chapter 4);
2. Access to existing registries (chapter 5);
3. Gaps in the availability (and access to) registries and good practices (chapter 6);
4. Organizing (new) registries (chapter 7).

In addition, it is important to know what type of research questions have to be answered within the RP programme in order to know what kind of patient data are needed.

1.3 Research questions

This study starts with an overview of the questions and topics the RP programme wants to address in relation to the type of data are needed to answer these questions in chapter 3. In the next four chapters the following research questions will be addressed.

[4] 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.

Part 1: Availability of patient registries

In the Netherlands there is a wide variety of patient registries that can potentially be used for research on rational pharmacotherapy. However, a systematic overview of such registries is lacking. Therefore, the first question of this study is:

- Which patient registries do exist in the Netherlands, that can be used for research on rational pharmacotherapy (within the ZonMw RP programme)?

Part 2: Access to existing registries

Although availability of patient registries is a condition sine qua non, availability alone is not sufficient: data also need to be accessible. Accessibility of patient registries for research can be hindered in several ways (for example because of costs or because registries do not allow data to be used by external researchers). No overview exists of the access to patient registries in the Netherlands. Therefore, this pilot survey will provide an inventory on this issue and identify bottlenecks in access to patient registries as well as potential solutions. The research questions of this part of the study are:

- How is the access to Dutch patient registries organised?
- What influence do developments with regard to so-called meta initiatives have on access to patient registries?
- How can access to existing patient registries be improved?
- Which lessons can be learned from the inventory on access to registries?

Part 3: New patient registries and extension of existing patient registries

This part of the study addresses the question what new data need to be collected in order to address to answer the questions relevant to the RP programme (see above). Moreover, we will look at good practices in current registries that can be used in setting up new registries. The research questions of part 3 are:

- To what extent can existing registries (and their access) be used to answer research questions relevant to the RP programme as well as be used to improve quality of care and, if not, what type of new data are needed?
- Which characteristics of existing registries can be considered as good practices that can be used as examples for new infrastructures?

Part 4: Governance in patient registries

The last part of the study deals with governance of patient registries. Up to now, registries seem to reinvent the wheel with respect to how to governance issues. A concept governance model may serve as a basis for new registries financed by the RP programme. The research questions for this part are:

- Which models regarding governance issues are used in existing patient registries?
- How do these models safeguard the interests of different stakeholders?
- To what extent do these models help to realise requirements regarding the accessibility of data?
- Can best practices be identified regarding governance?
- How can ZonMw promote such best practices (incentives)?

2 Methods

This chapter describes the methods we used in this study. More detailed information can be found chapters 4 to 7.

2.1 Questionnaire

This study consists of four parts:

1. Overview of available registries;
2. Access to existing registries;
3. Gaps in the availability (and access to) registries (including overview of the research questions of the RP programme);
4. Organising (new) registries.

This questionnaire is used in chapters 5 to 7. The aim of the questionnaire was to capture characteristics of registries in the Netherlands, to get insight in their accessibility and in the extent to which patient registries fulfil criteria for good practice. The questionnaire that was sent out to a total of 63 Dutch data patient registries, as well as another five questionnaires to Dutch umbrella organisations (below referred to as "meta initiatives"). We strived for a varied sample of registries. While some registries include a general patient population, other registries are disease-specific and some registries focus on specific age groups (children, the elderly). In addition, the way data within registries are collected differs: some registries use automated routine registrations from health care professionals, while others mainly collect data through questionnaires from patients. Also the type of data collected varies from health consumption data in general practice (including medication) to blood and urine samples to extended patient questionnaires. The questionnaire was sent out December 19, 2011. Registries were reminded one time either by mail or by telephone. Questionnaires could be sent in until the end of January 2012. Overall, 34 registries (54%) responded as well as three registries that were not approached themselves[5]. These registries reflect the variation that was found in the original sample of 63 registries with regard to type of database and type of patient population (see for more information Table 6.1). Furthermore, all meta initiatives responded.

The questionnaire covered the following topics:

- General information: name of the registry, aim of the registry, name of contact etc;
- Data collection and characteristics of the data: patient population (type and size), type of care included (primary care/ secondary care/ other), procedures for data collection, incentives for participants, and frequency of data updates, data format, type of data available (clinical data, medication use; demographics, etc);
- Privacy aspects: measures to ensure privacy, type and scope of consent;
- Linkage to other data registries: (how) can the data be linked to other registries, identifying/ linking variables;

5. These were two registries headed by the coordinator of a registry that we did approach and one that includes one of the registries that we approached (LINH) but that consists of more registries (NIVEL Zorgregistraties).

- Procedures to obtain data for external parties: conditions to use data (f.e. review committee, costs to obtain/ use data, parties who can obtain data, way of data exchange;
- Governance: ownership of the data, legal responsibility, and financing parties;
- Future perspectives for data registries.

The questionnaire for the meta initiatives consists of a subset of these topics. The full questionnaire is included in Annex 1 (in Dutch).

2.2 Other methods used in the study

Next to the questionnaire other methods are used. Table 2.1 provides an overview per chapter. In the respective chapters more details on the methods can be found.

Table 2.1 Overview of methods used in the following chapters

Chapter	Method(s) used
Chapter 3	To describes the main research areas of the Rational Pharmacotherapy programme the following method was used: *- Source verification based upon earlier work of ZonMw on Rational Pharmacotherapy*
Chapter 4	To construct an overview of patients registries the following methods were used: *- Search in www.zorggegevens.nl, a website collecting metadata about all the available registries for public health and health care in the Netherlands:* *- Search in PubMed*
Chapter 5	To provide an overview of the accessibility of patient registries in the Netherlands we used: *- Questionnaire for registries (Annex 1)* *- Questionnaire for metainitiatives*
Chapter 6	To provide an overview of good practices and to identify gaps in availability of databases the following methods were used: *- Questionnaire for registries (Annex 1)* *- Expert interviews*
Chapter 7	To describe the organization and governance of patient registries the following methods were used: *- Questionnaire for registries (Annex 1)* *- Expert consultation (interview, email)* *- Scan of websites of patient registries*

3 Research themes relevant for observational research

This chapter describes the main research areas of the ZonMw Rational Pharmacotherapy programme of ZonMw. These areas have been described in several ZonMw-reports. We focus on those research themes that can be answered using patients registries, i.e. observational studies.

3.1 ZonMw preparatory studies in 2009 and 2010

The introduction chapter described the overall aims of the ZonMW Rational Pharmacotherapy programme. ZonMw performed two studies in preparation of the programme. First, the Signal Rational Use of Medicines (Signal report) was published in 2009.[6] The second report provided an intensification of the Signal report in that it specified the problems signalled in the first report into research themes/questions ("Intensification").[7] Both reports provided the basis for the RP programme as it stands now. Therefore we will shortly describe the main results of these two studies. Next, we will describe the main research themes emerging from these reports which can be dealt with using observational data from patient registries.

Signal report

The 2009 Signal report aimed to describe the main gaps in evidence for rational pharmacotherapy. These gaps were identified after an invitational conference with representatives of relevant field parties and the academic field. In addition, (individual) interviews with twenty other representatives representing those two groups of expertise were held. The gaps emerging from this consultation were categorised into the following categories:

A. Is the medicine prescribed where it is needed?
B. Is the right medicine prescribed in the right dosage?
C. Are there other useful indications where the drug could be effective?
D. Is the drug used correctly?

More generally the Signal report stated that stimulating practice-oriented research on rational pharmacotherapy could lead to great benefits in terms of both quality and efficiency of care (ZonMw 2009, 2011).

Intensification report

After the publication of the Signal report. the Ministry of Health requested an intensification of the results from the Signal study, which were published in 2010. This 2010 report underlined and broadened the case made in the Signal report. ZonMW summarised the main conclusions of this second report as follows:[8]

1. In order to increase the quality of day-to-day health care, there is a structural need for research and implementation of knowledge on rational pharmacotherapy;

6 ZonMw. Signalement Goed Gebruik van Geneesmiddelen. The Hague, 2009.
7 ZonMw. Verdieping Goed Gebruik Geneesmiddelen. The Hague, 2010.
8 Cited from ZonMw, 2011.

2. Although there is some ongoing research, there are still major gaps in the evidence, including areas which are highly relevant to safety, efficiency and efficacy of pharmacotherapy in the Netherlands;
3. Scientific research on rational pharmacotherapy is important in order to generate reliable knowledge for daily practice;
4. Scientific research is an international endeavour, but to implement its conclusions within the Dutch healthcare system, additional fact-finding may be necessary. Basic pharmacology can be studied anywhere, but pharmacotherapy deals with the application of this knowledge within current praxis;
5. If research is attuned to the needs of daily practice, the application of its results is easier, leading to a more effective use of all available resources (ZonMw, 2010).

This second report also mentioned main research themes categorised according to the four themes mentioned above (A to D). In all four areas observational research can make an important contribution since data from observational databases provide the opportunity to study trends and problems in an unbiased population. As ZonMw states: "Automated data analysis allows for the recognition of patterns that would otherwise escape notice"[9]. In addition, observational data usually reflect clinical daily practice better than data from Randomized Clinical Trials (RCTs).

3.2 Gaps in evidence for rational pharmacotherapy and research themes emerging from the preparatory studies

This section describes the main research themes emerging from the two preparatory reports published by ZonMw in 2009 and 2010 (Table 3.1).[10] We will focus solely on those issues within research themes that can be addressed with observational data according to the ZonMw reports combined with our own views. The whole list of issues can be found in the 2010 ZonMw report.

Table 3.1 Main research themes mentioned in ZonMw's preparatory studies of 2009 and 2010

A. Is the medicine prescribed where it is needed?	B Is the medicine prescribed in the right dosage?	C Are there other useful indications where the drug could be effective?	D Is the drug used correctly?
1. Under- and overtreatment	1. Rationality and effectiveness	1. Off-label use	1. Adherence
2. Other interventions versus medicines	2.Guided-dosed pharmacotherapy	2.New indications	2. Polypharmacy
	3. Special patient populations		3. Unlicensed use
			4. Effects of lifestyle

9 ZonMw. Proposal ZonMw Programme Rational Pharmacotherapy. Optimising effective, safe and efficient use of medicines in health care, The Hague, 2011.

10 ZonMw. Signalement Goed Gebruik van Geneesmiddelen. The Hague, 2009 and ZonMw. Verdieping Goed Gebruik Geneesmiddelen. The Hague, 2010.

3.2.1 Gap A: Is the medicine prescribed where it is needed?

Patients do not always receive the medication they need. On the other hand, medicines may be prescribed to patients who do not need medication, for example because they would benefit more from a non-pharmaceutical treatment. Albeit these issues are addressed as being problematic, research is still needed and two themes emerged from the two preparatory ZonMw studies.

A1. Under- and overtreatment

The first issue is under- and overtreatment. The following themes can be addressed using observational data. First, the study of causes and consequences[11] of under- and overtreatment can improve knowledge about which problems arise from over- and undertreatment, especially in vulnerable patient groups such as the elderly, children, and patients with multimorbidity. A second theme is the knowledge about determinants of adherence to medication, since undertreatment (as well as overtreatment) can be caused not only by prescription behaviour of professionals but also by the way patients use their medication.

A2. Other interventions versus medicines

Often, alternatives exist for pharmaceutical treatment of medical problems. For example, life style interventions can be advised to reduce high blood pressure levels instead of prescribing antihypertensives. Often, a comparison of the effectiveness of pharmaceutical treatment versus non-pharmaceutical treatment is lacking. Relevant research questions include such comparison. Observational data can be used to compare how often pharmaceutical treatment is preferred to non-pharmaceutical treatment and also to study whether outcomes differ between these treatments.

3.2.2 Gap B: Is the right medicine prescribed in the right dosage?[12]

Patients sometimes receive a suboptimal pharmacotherapy, for example a medicine for which a better alternative exists or a dosage that is too high or low. This is undesirable from the perspective of rational pharmacotherapy. Three research themes were identified in the 2010 ZonMw report, two of which have elements that can be studied using observational data. These are described below.

B1. Rationality and effectiveness

Research on rationality and effectiveness is often performed before a medicine receives market authorisation. This research is sponsored by the pharmaceutical industry. However, many of the questions regarding rationality and effectiveness only arise after a product entered the market, because it is only then that large, unbiased patient populations receive the product showing its effect in an uncontrolled setting. Such information is highly relevant for several purposes such as patient safety but also for decisions on whether or not to reimburse the drug or whether or not a medicine should be recommended in professional guidelines. Therefore, research on (cost-)effectiveness after market authorisation is needed. In addition, research is needed on rational prescribing: do patients in daily clinical practice receive the care that best fits them? And if not, what are consequences in terms of health outcomes?

---

11 Consequences not mentioned in the two ZonMw reports but added by the research team of this study.

12 Theme B2 is not included because questions in this field are not expected to use observational data

B3 Special patient populations

ZonMw distinguishes five special patient populations: children, the elderly, pregnant women, patients with a rare disease and patients from a non-Western background. All these patient populations require a specific approach because "mainstream" knowledge on medicines does not apply to them. Observational research can contribute to increasing knowledge for each of these specific patient populations. For each group we provide *one example*. Special registries for children can provide knowledge whether or not these children are treated according the special formulary the Netherlands has for children and whether health outcomes differ between children treated according to the guidelines and children who are not. For the elderly observational data can provide more insight into the treatment of patients with multi-morbidity (see also D2 Polypharmacy), while for pregnant women effects of medication use on congenital anomalies can be monitored providing information as to which medicines preferably should not be prescribed during pregnancy. For patients with rare diseases data from patient registries can provide insight in the disease's progression and the effects of medicines on it. Finally, for patients from non-Western backgrounds observational data can provide insight in under- and overtreatment. It is for example known that non-adherence to medication is lower in patients from a non-Western background.[13]

3.2.3 Gap C: Are there other useful indications where the drug could be effective?[14]

C1. Off-label use

New drugs need to be authorised before market entrance. To receive authorisation, a favourable balance between beneficial and harmful effects has to be demonstrated. Drugs are usually registered for a limited number of clinical diagnoses and for adult patients only. However, in daily practice, drugs are often used in situations that do not adhere to the authorisation requirements: 'off-label prescribing'.[15] Since no pre-registration evidence is gathered for these indications, no data are available that show effectiveness and safety upon the moment of registration. While several years ago some studies have been published on off-label prescribing the extent of off-label prescribing in clinical practice remains largely unknown as are its consequences. Studying consequences is important in order to build evidence pro or against off-label prescribing. Observational data can contribute to answering these questions in case information on indications and prescriptions is included in the database.

3.2.4 Gap D: Is the drug used correctly?

D1. Adherence

It is well-known that patients do not always take their medication the way that was agreed upon with their health care professional. While many studies are performed in this field there still is a lack of knowledge as to why patients are non-adherent and how adherence can be improved. Moreover, consequences of non-adherence are not systematically studied. Much research ZonMw proposes has to be studied in intervention studies or concerns the development of instruments.[16] However, part of the

13 Van Dijk et al 2007

14 Theme C2 is not included because questions in this field are not expected to use observational data

15 For example Gijsen et al. 2009

16 Data from information systems can also be extracted in intervention studies but we do not include them here because we do not consider one-time extractions for specific studies as patient registries.

research questions can be answered using observational data from patient registries. These questions refer to the level of non-adherence for different medicines[17] and to further development of methods to measure non-adherence.

D2. Polypharmacy

A growing number of (elderly) patients uses multiple medications. Polypharmacy is an important reason for preventable hospital admissions[18], for example because of a complex network of interactions between different (combinations of) medicines. Moreover, guidelines for individual diseases sometimes provide conflicting advises causing uncertainty how to handle in case of a patient has a combination of diseases. Therefore, research is needed to study undesirable interactions between different (combinations of) medicines as well as to study how conflicting guidelines are dealt with in clinical practice (including the effects on health outcomes). Observational data from patient registries can be used to study these types of questions. In addition, ZonMw argues that methodological research needs to be done to improve the use of observational research in this field. Detailed adherence data need to be used for this purpose, such as electronic monitoring (Farmer 1999).

D3. Unlicensed use of medicines

The fact that not all patients can swallow tablets is an example where so-called unlicensed use cause problems for medication intake. Especially special patient groups such as children and the elderly may suffer from this. Albeit developmental and intervention studies seem to be more suited to answer questions within this theme, observational data can be used. Registries of patients who have problems taking there medication can be set up in order to study their problems and health outcomes can be studied in existing patient registries for example registries including children and the elderly.

D4. Lifestyle effects

Life style is an important determinant for health outcomes and many life style interventions are conducted. However, not much is known about the effect of lifestyle on the effectiveness of medicines. Do medicines work better in patients with a healthy life style? This question can be answered using observational data (combined with survey/cohort data on life style).

3.3 Research themes in the proposal for the ZonMw RP programme

Based upon the two preparatory studies in 2009 and 2010 ZonMw developed a proposal for the programme Rational Pharmacotherapy (see chapter 1). In this proposal three main research areas are identified, which are different from the main research themes from the preparatory studies:

A. Generating evidence for guidelines and health care standards to improve quality of care;
B. Input for decision-making (research for government agencies);
C. Practice-based research: research inspired by patients' needs and health care practice.

While those themes seems different from those in the preparatory studies they in fact are not, since most themes from the preparatory studies fit into one or more of these three research areas. For example, research on under- or overtreatment generates knowledge valuable to implement in guidelines and to serve as input for decisions on reimbursement (for example by trying to avoid overtreatment, such as was the case for benzodiazepines). Moreover, research in the field of under- and

[17] A start has made at www.therapietrouwmonitor.nl (Vervloet et al 2011).
[18] Leendertse et al 2008

overtreatment can lead to changes in daily clinical practice. Therefore, we will use the themes as stipulated in the two preparatory studies to study to what extent the current data infrastructure in the Netherlands can be used for answering these questions and what still needs to be developed. Table 3.2 shows what kind of data are needed to answer these questions (based upon the analysis above).

Table 3.2 Type of observational databases (minimally) needed to study main research themes mentioned in ZonMw's preparatory studies of 2009 and 2010

Type of database	Subjects for which this type of database is needed
Prescription databases in primary and secondary care, preferably including data on indications as well as other relevant patient characteristics and clinical outcome measures	A1. Under- and overtreatment A2. Other interventions versus medicines B1. Rationality and effectiveness C1. Off-label use D1. Adherence D3. Polypharmacy
Databases in primary and secondary care, including data on indications, prescriptions lifestyle interventions	D4. Effects of lifestyle
Patient cohorts (preferably from these databases) to collect extra information for example using survey data or clinical measurements	A1. Under- and overtreatment A2. Other interventions versus medicines B1. Rationality and effectiveness D1. Adherence D3. Polypharmacy D4. Effects of lifestyle
Special patient populations cohorts (if possible linked to existing databases) to collect information needed to answer specific research questions	B3. Special patient populations
Databases including detailed medication intake data (registering each medication intake)	D1. Adherence

3.4 Conclusion

Previous work of ZonMw pointed out numerous gaps in knowledge with regard to Rational Pharmacotherapy. Many of the remaining questions in this field can be answered using observational data. Databases suited for these type of research include for example prescription databases in both primary and secondary care and patient cohorts. In the next three chapters it will be studied what patients registries are available in the Netherlands, how accessible they are and whether they fulfil criteria for good practice. Chapter 6 will end with an overview of gaps in the availability of data.

4. Rational Pharmacotherapy Registries in the Netherlands

This chapter is an inventory of available registries in the Netherlands for the programme Rational Pharmacotherapy (RP) from the Netherlands Organisation for Health Research and Development (ZonMw). The chapter focuses on the availability of data sources**,** registries relevant to Rational Pharmacotherapy in the Netherlands.

4.1 Definitions, methods, criteria and registries found

Within this project, the National Institute for Public Health and the Environment (RIVM) focused on the following aims of the RP program: *Building an infrastructure that allows for adequate and independent ways to answer questions pertaining to rational pharmacotherapy.* To realise the aims of the RP programme, ZonMW introduces three groups of activities (see Figure 1.1). This chapter focuses on the following: ***Infrastructure***: *Infrastructure is an essential prerequisite for research: high quality is needed. The RP programme has to strengthen and extend the existing RP research infrastructure in the Netherlands as well as to create conditions for optimal use of both existing and newly created databases.*

Methods

To investigate which data sources are available for building an infrastructure for research we used search terms to find relevant registries in the Netherlands. We searched through a website which collects metadata about all the available registries for public health and health care in the Netherlands, called www.zorggegevens.nl. We also searched through PubMed to find registries or other data sources that we would have missed if we had restricted ourselves to www.zorggegevens.nl. In this chapter we present the results of both. Similar selection criteria were used (see below) as were used to select registries for the Dutch website www.zorggegevens.nl. However, the focus of this RP project was on registries directly involved in, or (closely) related to collecting data or information about pharmacotherapy, pharmacy, pharmacists, medicine or drugs.

Selection criteria

Inclusion criteria:

- The data must be available to researchers or governmental organisations, if not freely available then either through special procedures, such as a medical-ethical committee, or in other ways (e.g. remote access via authorisation and authentication);
- The register is still operational, has a known contact person (able to fill in our questionnaires), and (preferably) contains 'enough' data to make it relevant for national public health research.

Other criteria relevant for selecting registries are whether or not the register is useful for national policy and whether or not there is a legal basis for the registries (mandatory registries).

We decided to include specific registries, e.g. the National Medical Register (LMR), which do not contain any data about medicine (use) themselves, but are useful because they have: 1. a national scope, 2.

contain unique information (e.g. about secondary care), and these registries 3. enable us to link to other health registries that do contain data about medicine use.
For practical reasons, for instance because registries enable us to link to other (patient) registries, some non-patient registries were also included.

Exclusion criteria:
Registries not collecting any data themselves, like classification systems, secondary databases or meta-databases, are excluded. Registries that have not been maintained for more then five years are also excluded.

To select the RP registries we used the aforementioned criteria and searched through the www.zorggegevens.nl metadata base with specific search terms using the search engine of the website. The complete search-results are reflected in Annex 2. Using the www.zorggegevens.nl website we found 37 data sources relevant for the RP programme.

The Dutch terms* we used were:

Farmacie (pharmacy)
Medicijnen (medicine)
Geneesmiddelen (drugs)
Bijwerkingen (side effects)
Apothekers (pharmacists)

We noted that using other (more) search terms did not add anything to the search results, there is a limited number of registries in the database (n=182).

Pubmed search for Dutch RP data sources

Inclusion criteria:
Pubmed search, period **included** 11-11-2011 en 29-11-2011 with the following search terms:

- register AND Netherlands AND medicine NOT cochrane NOT trial register NOT veterinary (period 1996-2011): 68
- prospective cohort AND Netherlands AND drugs NOT cochrane NOT trial register NOT veterinary (period 1996-2011): 176

Publications using registries or other data sources (monitors or cohort studies) are presented in Annex 3 and 4.

Exclusion criteria:
Excluded in this Pubmed search are:

- general practice registries
- case registries

This Pubmed search resulted in 23 data sources, mostly registries but also monitoring and cohort studies. The complete results are presented in Annex 5**.** Thirteen data sources were not found in www.zorggegevens.nl but were presented to the research team by ZonMW. Table 4.1 summarizes these registries.

Table 4.1 Type and number of registries

Type of registry*	**Number***
Type of data	
National databases (derived from systems of health professionals)	23
Regional databases (derived from systems of health professionals)	14
Cohort studies	10
Biobanks	3
Other / unknown	9
Type of patient population	
General (including general practice patient population)	25
Disease-specific registries (single or multiple diseases)	22
Elderly	4
Children	4
Pregnant women	1
Mental health	4

*Some registries fall into more then one category.

4.2 Conclusions en recommendations

In total 63 data sources were found in this inventory of data sources for the RP programme (see Annex 5), including registries contributed by ZonMw. There are many different registries that include data about pharmacy and medicine (or data related to these topics) and relevant for the RP programme. Therefore, there are potentially many useful sources but there also seems to be little coordination, many different registries that work more or less in isolation, divided between many different organisations. For instance, there are many primary care (General Practitioner (GP)) registries that collect data about medicine use in relation to information about the patient (diagnoses), but often there is no coordination between the different initiatives. It is not easy to compare data from these different sources and common rules for getting access to the data do not exist. There are also many different cohort and monitor studies that register data about the use of medicine by their participants, but it is not clear whether or not these data are comparable. Besides this, in the Netherlands there is a rather strict division between primary and secondary care registries. This also complicates matters because it makes the interpretation of the data from the different registries more complex, it is therefore hard to draw conclusions from either one the two types of registries. Given the above, this inventory of

available data sources has proven very useful. However, registries of RP data change quite often, and new data collections are initiated on a regular bases (very often independent of each other).

Recommendations

- We foresee a lot of progress in different types of RP research if there would be national coordination and standardisation efforts to make data from primary and secondary care registries (more) comparable. This would be true not only *within* the Netherlands, but also for data outside the Netherlands, i.e. within the EU. Currently, it is not clear whether the available data are comparable across countries, i.e. within the EU. This needs to be investigated further.
- Registries containing data about medicine use are often limited in scope. Technology is progressing very rapidly and linking different data sources is becoming easier (technically). Efforts to link different registries to each other are necessary to do research beyond the (limited) scope of the pharmaceutical databases. Issues related to health and health care: life style (nutrition, exercise); long-term care; being able to work and participating in society, are important to relate to, for instance, medicine use. Linking of different data sources is necessary and knowledge at a national level is needed to coordinate and implement safe linking of data.
- We recommend making an inventory of RP registries on a regular and continues basis, not just once. The inventory should also be done more systematically because there are so many different types of data sources involved in collecting data about pharmacy and medicine. Because of the very diverse nature of all the registries and studies, we probably missed relevant registries in this inventory. This is certainly true for the website www.zorggegevens.nl and it is advisable to update the website on a regular basis with information about data sources relevant for the RP programme and to dedicate a specific part of the website to give more in-depth information about RP data.

5 Accessibility of patient registries

This chapter investigates aspects related to the accessibility of registries and meta initiatives. Aspects such as permission of subjects to use data, enriching datasets by linking to other registries and request, permission and authorisation will be discussed.

5.1 Response

In total 63 registries were approached. Table 5.1 show their main characteristics. The so called "meta initiatives" were not asked to participate in the regular questionnaire; instead, they were approached separately and interviewed. Five of them participated.

Table 5.1 Type of patient registry participating in the survey

Type of register	Number in questionnaire
Type of data	
National databases (derived from systems of health professionals)	15
Regional databases (derived from systems of health professionals)	10
Cohort studies	4
Bio banks	1
Other / unknown	7
Type of patient population	
General (including general practice patient population)	18
Disease-specific registries (single or multiple diseases)	13
Elderly	1
Children	2
Pregnant women	1
Mental health	2

5.2 Definitions and criteria

In order to be able to perform scientific research it is necessary to have sufficient access to available research databases. Notably, the mere presence of a database is insufficient if accessibility is limited. There are many potential reasons causing insufficient data access: data might not be available to external researchers because of the policy of the registry, subjects might not give consent, data might be incomplete for a specific research question, or data might be too expensive to buy, to name a few reasons.

The aim of this chapter is to investigate accessibility of the databases included in the inventory of the previous chapter. In general and as introduced above, any registry's accessibility depends, amongst other things, on the following aspects:

- The registry does not exclude certain (research) groups a priori because of specific policies;
- Subjects have given consent for third parties to perform scientific research;
- Privacy / identifying data are dealt with sufficiently;
- Data be linked to data from another register;
- How is a formal request for data and subsequent authorisation dealt with.

Another aspect involves the influence of the so called “meta initiatives” on existing registries. These umbrella or network organisations can been seen as a portal to the underlying registries. Examples are Mondriaan and BBMRI. Participation in such initiatives obviously enhances accessibility. Below, we outline the findings in the questionnaire on these various criteria. Finally, some conclusions will be drawn.

5.3 Questionnaire outcomes

The following criteria for accessibility will be discussed in separate paragraphs:
- Permission of patients to use data;
- Enriching datasets by linking to other registries;
- Request, permission and authorisation;
- Meta initiatives;
- Future.

5.3.1 Permission of subjects to use data

In general there are two types of permission given by subjects to use data: informed consent and opt-out/no objection. Without any consent, only anonymous data may be used. In the case of informed consent, the subject explicitly states what is allowed to do with his/her data, and what is not. This type of consent is typically asked for in relatively small cohorts and if enough time and money is available to approach each individual. Examples include research cohorts for specific diseases, such as within randomized clinical trial settings. The opt-out or no objection approach is typically used in situations where large cohorts of subjects are included, for example in situations where GP data of many practices are collected. Privacy rules are very stringent with respect to the use of direct subject-identifying variables. Moreover, in these situations an opt-out should be available: subjects should be aware of the fact that their data are collected and used for research, including the possibility to object against the use of their data; they must have the possibility to ensure that their data are not included in the register. Of course there are other forms of consents, but those are usually based on the two forms described above and might involve a combination of both.

Table 5.2 presents the number (and percentage) of registries using the different forms. Note that the total sums up to over 100%. This is because two registries have multiple forms of consent: one has both informed consent and an opt out, and another register has an opt out combined with another mentioned, non-specified type of consent. Six registries have no consent at all from the subjects whose data are in the database.

Table 5.2 Type of permission granted by patients (n=37)

Type	N	%
Informed consent	18	49%
Opt out	14	38%
None	6	16%
Other	1	3%

Even when informed consent is given, it does not mean the data are allowed to be used for everything. It is possible that the consent only concerns the medical part of the data, but does not allow the use of personal identifying data for linking to other datasets. Table 5.3 gives an overview of the aspects covered by the different forms of permission. The percentages are relative to the numbers presented in table 5.2. Since all investigated registries (see the previous chapter) are somehow medically related in a broad sense, it is clear that almost always the medical data are allowed to be used (independently of the form of permission used). However, linking is relatively more allowed in registries using informed consent compared to registries using the opt out form. As, without informed consent, it is not allowed to collect directly identifying subject characteristics, linking is much harder when using the opt out form. This will be further investigated below. Moreover, within an informed consent permission on certain topics is explicitly requested. Since linking is important, it is usually part of the consent.

Table 5.3 Subjects for which subjects grant permission (n=31)

Permission	Medical		Linking		All research		Specific research	
Inf. Consent	12	67%	11	61%	10	56%	6	33%
Opt out	10	71%	5	36%	10	71%	1	7%
Other	1	100%	1	100%	1	100%	0	0%

Table 5.3 (last two columns) also shows the number (and percentage) of registries having acquired permission from subjects to use their data for all or specific research. Specific research may include analysis of biosamples (e.g. blood) and data obtained from specific questionnaires. Obviously, the relatively more generic registries generally employ the opt out design, whereas registries that are quite specific on various aspects and that may address detailed research questions often informed consent is present.

5.3.2 Enriching datasets by linking to other registries

Many scientific research questions require rich data sets. Often a dataset is lacking some of the necessary variables, when data was originally collected for other purposes. Therefore, next to the availability and accessibility of database, it is important to investigate whether data can also be used to enrich other datasets. So linking is an important issue. Table 5.4 summarizes the current state and future plans of the various databases investigated. The table shows answers to the question “whether the dataset can currently (now) be linked with other datasets on the subject level, based on subject variables available in the dataset, and if not whether there are plans to be able to do so in the future”.

Table 5.4 Possibilities of linking now and/or in the future

Linking now	N	Future	N	
No	11 (30%)	no	8	73%
		yes	3	27%
Yes	26 (70%)	no	5	19%
		yes	21	81%

The results show that, at this moment, data of about 70% of the registries can already be linked to other registries. It is interesting to note that from the registries that currently are not able to be linked, far the most also do not have the intention to do so in future. The question arises whether the databases that cannot be linked are in general containing a certain type of data, for example GP data. This is not the case (table 5.4). Databases unable to be linked are distributed over all types of databases. Notably, all types of databases are equally distributed in both the linkable group and non-linkable group, with the exception of databases on lifestyle, genetics and biosamples. These three types of data are more prevalent in the linkable group.

Currently, 11 registries are not able to link to other databases. From the three indicating they are taking into account future linking, two state it is because of international agreements; the data will also be used by for example European organisations. 26 registries are currently able to link. Of these databases, five will not take extra measures to enhance their possibility to link. However, far the most are playing an active role in working on an even better way to link. From the responses in the questionnaire it can be concluded that those registries that are already linking have accumulated good knowledge about the complex topic of linking. Based on the current and past experiences they want to improve wherever possible. This does not only include the technical side of linking, but also the legal side. Considering the legal side, (quite some) registries stated they might ask for informed consent to link. However there are situations where this might be practically impossible, for example, with databases collecting “generic” data of many subjects (e.g. SFK), unlike databases aiming at a very specific group of patients, usually a much smaller group, where asking each individual subject for informed consent is practically feasible. When asking for informed consent is unfeasible, collecting extra subject identifiers suitable for linking is the common approach. The use of the (hashed) social security number BSN has been mentioned. Another possibility is to make use of a trusted third party (TTP) as some of the registries which are currently linking do such as the Mondriaan meta initiative.

Table 5.5 gives an overview of the type of data available within all registries participating is this project. Note that one register can contain multiple types of data. In general, all types of data can be linked rather successfully (over 70% at least) with the exception of data originating from questionnaires, the latter not being very surprising given the specific way of data gathering.

Table 5.5 Type of data available in database divided by the possibility to link data

Type of data	Total	Link yes		Link no	
Drug	32	23	71.88%	9	28%
Clinical	30	21	70%	9	30%
other healthcare	19	14	74%	5	26%
Biomedical	18	13	72%	5	28%
Demographic	23	17	74%	6	26%
Lifestyle	19	15	79%	4	21%
Genetics	14	12	86%	2	14%
Biosample	15	13	87%	2	13%
Questionnaire	17	11	65%	6	35%
Other	6	5	83%	1	17%

Table 5.6 shows the variables used to perform the linkage of subjects. Date of birth and gender are used most often. In many types of linkages, these two are used as so-called “blocking” variables. This means that those variables should match, even when a probabilistic linking procedure is being used. From the NAW data the postal code is most frequently used. If possible the complete postal code is used, but often only the numerical part is available. Only few registries indicated that other NAW data like the first or last name of a subject, address or city name is used. In addition, some registries indicated that they use a trusted third party (TTP) for the linking process, or they want to use a TTP to link in future. Considering the complexity of linking (both technical and legal), it is advisable to use the expertise of organisations specialized in this field.

Table 5.6 Variables used for linking

Variable	N	%
BSN	10	38%
NAW	19	73%
Date of birth	21	81%
Gender	21	81%
Other subject variables	9	35%
Other non-subject variables	8	31%

5.3.3 Request, permission and authorisation

The fact that data has been collected for scientific research does not imply that everyone can use it. Policy of the register might dictate that data be denied for specific research questions or organisations, or the data might be available only for “local use”. This section shows how requests for data from a researcher are being dealt with.

When requesting data, about 90% of the registries require the researcher to write a protocol describing the research question (table 5.7). Typically this protocol will be judged by the board of the register, or by a special advisory board. All boards include at least a researcher to evaluate the protocol for its scientific part. Usually a special advisory board is consulted in case of sensitive issues, for example, relating to privacy or research for commercial parties. When a register requires a protocol describing the request

for data, the protocol or procedure to follow can be found online on the website of the register in almost 58% of the cases (table 5.7, right part).

Table 5.7 Protocol required before handing out data

Protocol required	N	%
No	4	11%
Yes	33	89%

Online	N	%
No	19	58%
Yes	14	42%

The participating registries were not asked whether they have different access policies with respect to raw data, 'cleaned or user friendly' data or frequent repeats of the same request. However from experience within Mondriaan generally the same procedure is applied for these types of requests. Occasionally registries are hesitant to give access to raw data due to chances on misinterpretations by investigators.

The participating registries were asked whether certain organisations or certain professions are refused a priori when requesting data (table 5.8). About 20% (7 registries) state they would not give permission a priori; especially "pharmaceutical industry" and "for profit organisations" are mentioned as parties not receiving permission. One register stated that organisations that might potentially be harmful to them are rejected.

Table 5.8 Permission rejected for some parties

No permission	N	%
No	26	70%
Yes	7	19%

* 4 registries did not answer this question.

When a protocol for requesting data has been approved by the board, the next issue is under what conditions the data will be available. Table 5.9 summarizes the different conditions used. The first three criteria all deal with (scientific) publication of the results of research using the requested data. The assumption here is that this research will result in a published article. Although it is unknown whether publishing is a requirement met in all cases, all interviewed registries have published articles in international journals or international reports. Any register included has published about 75 articles on average so far. Of course this number is to increase in future.

Since publishing is very important, most of the registries want to be involved in producing an article in different ways. Only 7 out of the 37 do not have as a minimal condition 1, 2 or 3 from table 5.9. Few require only prior review, but by far the most want to be involved as co-author.

Table 5.9 Conditions for using data

Condition	N	%
Peer-reviewed publication	20	54%
Prior review	21	57%
Co-authorship	24	65%
Along the code of conduct	10	27%
No marketing purposes	11	30%
Other	12	32%

About 30% of the registries require condition 4, 5 or 6 to be met when research data are requested. A code of conduct seems primary to be required in order to meet a good scientific research question, including the way of working with the data. An example of the latter may be that computer syntax files used for analyzing the data have to be stored so results can be checked or reproduced. Condition 6, "other", is rather similar to a code of conduct. Almost all registries mentioned under this condition that the research question should be sound, or that the research should show added value to the main purpose where the data was collected for in the first place. Finally, almost 30% of the registries a priori reject requests for data when the data will be used for marketing purposes.

When all conditions are met, the register can provide the data. The whole process of requesting data, the assessment by the board or committee, and finally the delivery of data takes on average about five weeks. Only four registries state that the process will take on average about three months or more.

5.3.4 Meta initiatives

Do Dutch meta initiatives have influence on the access to data?
Dutch meta initiatives (umbrella or network organisations) such as BBMRI, Mondriaan, RIVM and NFU work with independent organisations which are responsible for various types of datasets.
The themes on which the meta initiatives cooperate vary from the setup of research infrastructure (BBMRI, Mondriaan) to safeguarding the members' interests (NFU, Nefarma). Each on its own distinct manner, the meta initiatives have means to discuss and make agreements on the accessibility of datasets with the member organisations. This inventorial study tries to get an impression of to what extent these organisations influence the access to datasets in practice. Here BBMRI, Mondriaan, RIVM, NFU, Nefarma and the Ministry of health have been interviewed[19].

Impression of the inventory
As the nature of the meta initiatives varies, their activities to improve the accessibility of data vary. RIVM has setup and manages 'Zorggegevens.nl'. This website offers an overview of the larger, national datasets and hence improves findability of data. RIVM itself doubts whether it should be classified as a meta initiative or umbrella-organisation as the datasets within RIVM operate highly independently. Currently there is no central coordination. Nevertheless, a simple search on the internet shows that the datasets visited all have procedures for data requests. RIVM's intentions are to create more overview and coordination over the RIVM datasets. An important note is that RIVM and its datasets are cautious to cooperate with industrial parties.

[19] The following persons have been interviewed: NFU: mrs. C. Bouma; BBMRI: mrs. M. Brandsma; RIVM: mr. R. Nugteren; Mondriaan: mr. O. Klungel; Nefarma: mrs. N. Kraaijeveld

The Dutch Federation of University Medical Centres (NFU) has ongoing projects to enforce the position of the UMCs by bundling of research activities between UMC's on topics such as research infrastructure/ICT, clinical research, translational research and valorisation. The String of Pearls Initiative is the most prominent example. With respect to these projects the NFU pro-actively encourages its members to lower the access threshold for both profit and non-profit third parties to datasets. Nefarma has indicated that the topic of access to e.g. trial data to third parties has, on various occasions, been given thought. Unfortunately the stage of concrete plans has not been reached yet. Nefarma does not interfere with initiatives of individual members seeking cooperation with datasets of third parties such as university medical centres.

Mondriaan's primary aim is to improve the access to data in The Netherlands for scientific research. It therefore cooperates with various datasets as well as initiatives such as BBMRI. Mondriaan has a central access procedure through which data of the linked datasets can be requested. Data can be collected from each datasets through an ICT-infrastructure which is also able to link data on subject level. This infrastructure also contains an operational catalogue on subject level through which subject/data selections from the various datasets can be made.

BBMRI tries to improve the biobank research infrastructure in The Netherlands by facilitating several infrastructural projects which contribute to harmonisation, enrichment, linkage and cooperation of and between biobanks. BBMRI has taken a couple of measures regarding the access to data. Besides infrastructural efforts, BBMRI demands that under reasonable conditions data from BBMRI projects are accessible to all other BBMRI researchers. In addition, BBMRI has plans to setup a catalogue on subject level for biobanks from which researcher can select data or biomaterials.

Although not approached in this study, it is worthwhile to mention the FEDERA (Federation of Medical Scientific Associations). This organisation also affects the accessibility of datasets. The COREON committee of the FEDERA has a legal focus and tries to contribute to keep Dutch law suitable for research purposes. The committee issues codes of conducts which form the current standard for managing datasets for scientific research.

Conclusion & recommendations meta initiatives
Most meta initiatives have the topic 'access to data' clearly on their radar screen. They are aware that optimal access can advance science and innovation. For that matter no missionary work is needed. However the extent of the actual efforts applied considerably varies. Partly, of course, for sound reasons as the accessibility theme is not everyone's core business.
Currently RIVM (Zorggegevens.nl), Mondriaan, the NFU and BBMRI seem to be most direct and contributors to the improvement of the accessibility of datasets.

Progress can be made through further cooperation on uniform agreements on the access to datasets. Meta initiatives are pre-eminently suited to create broad acceptance of new standards.
Perhaps a start can be made by harmonizing the conditions that parties such as BBMRI, NWO, ZonMw, the Ministry of Health and perhaps in the future also Topsector Life Sciences & Health set to financing projects. ZonMw, through this study, has taken the first steps into that direction. Hopefully this will result in a guideline which, in cooperation with several or all of the abovementioned parties, will lead to a standard code of conduct.

5.3.5 Future

In the final part of the questionnaire we asked how the registries look at the future considering the accessibility of existing databases.

Most of the suggestions are not register specific but concern general issues. For example, there is obviously a need for having a good comprehensive overview of which registries exist in the first place. The lack of such an overview, or the unawareness of such overviews, can be solved by having a central access point. Another frequently heard suggestion is the need for better ways of linking to other datasets. Preferably, the use of the social security number (BSN), which is currently not allowed. Note the direct use of this number is not allowed unless explicitly stated in the consent, but other uses, like a hashed version, still belong to the possibilities as demonstrated in the Mondriaan project. Yet another point of interest involves the quality of the data. This is a topic where the implementation of standards (e.g. ICPC codes, ATC/DDD system, ICD10, see also chapter 7) can bring a quality boost. For example, in most GP systems traditionally a very large part of the information was stored as free text. Such data are very hard to be used for research, and moreover it is very error prone. Initiatives like "ADEPD registreren"[20] strive for uniform and structured registration of data, and will reduce the possibility for free text and increase the use of national or international coding systems.

5.4 Conclusions and recommendations

In the previous sections we investigated several important issues considering the accessibility of data. Regarding the request for data by a researcher, it is common practice that a protocol is needed, explaining the goal and methods to be used. This protocol is judged by a board usually including at least one person with a scientific background making sure the quality of the study will good. Codes of conduct play an important role. Most registries require that the study will result in a scientific article or report, and almost all want to be involved. By far the most want to be co author. In about 20% of the interviewed registries data are a priori denied when requested by specific types of organisations, usually "pharmaceutical industry" and "for profit organisations". Another reason when a data request is not approved is when it will be used for marketing purposes: about 30% of the registries do not allow that type of purpose.

The following recommendations can be made:

- For newly established registries with informed consent, it is very important that all possible uses of the data will be covered. In particular, it can be wise to explicitly ask for permission to link to other datasets. In case of opt out, sufficient variables should be collected to make linking possible. The linking itself can be done by specialized parties like TTP's.
- The use of standardized coding systems can increase the quality of data. Whenever possible such (inter)national standards should be used. Examples are ICPC codes, ATC codes and the use of the DDD system.
- Meta initiatives can play an important role in the awareness of all available registries as well as in the development of new standards on accessibility to datasets and/or the acceptance thereof. In addition, the expertise of these meta initiatives can be used for complex topics like privacy and technical issues like linking on subject level.

[20] ADEPD = Adequate Dossiervorming met het Elektronisch Patiëntendossier

6 Good practices and fit with ZonMw research themes

The focus of this chapter is threefold. First, we define what we consider good practice. Second, we describe good practices among existing patient registries so that newly developed databases can profit from these experiences. Third, we focus on which data are needed to be able to address the questions relevant to the RP programme. New data can both refer to complete new patient registries as well as to an extension of existing data registries. An analysis of how existing registries fit within the RP programme will be provided and gaps in availability and access will be identified.

6.1 Current patient registries

We captured the type of data collection and their patient population for 63 registries (see also chapter 4). Looking at what areas they cover the registries can be divided in different (partly overlapping) categories. The Netherlands has many national databases, collecting automated routine data from health professionals or health organisations. Examples include general practice databases (for example LINH/NIVEL zorgregistraties and IPCI), pharmacy databases (for example SFK and IADB) and claim databases (f.e. Vektis, GIP) as well as disease-specific databases (for example. HIV monitoring and the Netherlands Tuberculosis Registration). We also identified 14 regional databases, mostly regional general practice databases, but also some disease-specific databases. We identified ten cohort studies (such as the Doetinchem study and PIAMA), two bio banks initiatives (BBMRI and Parelsnoer) and nine other type of databases. Looking at the patient populations covered in these databases, two major groups can be identified: 25 registries include a general (practice) patient population and 22 are disease-specific registrations. These last registries include a variety of diseases such as diabetes, asthma, tuberculosis, HIV, rare diseases, and cancer. There are also some registries for specific patient populations such as children, the elderly and pregnant women. Twelve of the nationwide databases include a general patient population. These are mostly either primary care databases or databases that include claim or pharmacy data. Secondary care databases usually are disease-specific registries, focussing on one disease. The registries that participated in our survey (n=37) reflect the overall picture of databases. Information provided by these databases are used in the remaining of this chapter.

Table 6.1 Type of patient registries (approached and in survey)*

Type of register	Number identified in approached	Number in questionnaire
Type of data		
National databases (derived from systems of health professionals)	23	15
Regional databases (derived from systems of health professionals)	14	10
Cohort studies	10	4
Bio banks	3	1
Other / unknown	9	7
Type of patient population		
General (including general practice patient population)	25	18
Disease-specific registries (single or multiple diseases)	22	13
Elderly	4	1
Children	4?	2
Pregnant women	1	1
Mental health	4	2

* Table 6.1 is a combination of tables 4.1 and 5.1

6.2 Good practices

In order to be used for research as well as for the purpose to improve quality of care, patient registries need to meet quality requirements. Moreover, they need to be accessible. While chapter 5 focused on accessibility issues, this chapter looks at some of those issues as well but from a good practice point of view. As such, this section describes how the patient registries that participated in the study, scored on a number of (rough) indicators for best practices. We also include results from interviews with three experts. These experts gave more in-depth information on criteria for good practice (section 6.2.3).[21]

6.2.1 Indicators for good practice

The questionnaire was developed in such way that a number of questions could be used for the analysis of good practice. In several rounds the research team discussed – based on their experience and literature – what elements are crucial for good practice. The ZonMw review committee also got the opportunity to discuss the selection during a meeting with the project team. Finally, interviews performed by interviewing three experts on what they consider good practice on the issues below. In case these interview would result in a different view on good practice, we would have re-evaluated the indicators.

[21] Governance issues and privacy issues were not discussed in detail since those issues were addressed in interviews described in chapter 7.

For the following topics we defined indicators for good practice and posed questions referring to good practice:

1. Availability of protocols;
2. Incentives for delivering professionals and other delivering parties;
3. Opportunities to link the data to other patient registries;
4. Organisation of data requests;
5. Suitability for research.

Table 6.2 shows an overview of the information we used as well as the argumentation as to why these indicators were included in our study to represent good practice.

Table 6.2 Overview of indicators for good practice & requirements covered by the questionnaire

Indicator	Operationalisation	Motivation
Protocols available	There is a protocol for data collection	In order to enhance validity and reliability of the data collection protocols are needed to enable control on how data are collected
	There is a protocol for data delivery	In order to enhance quality, validity and reliability of the data delivered to external parties protocols are needed
Incentives for parties who deliver data	Feedback	It often takes delivering parties time to deliver data. Often these are health care professionals. Providing them with feedback makes them aware of their quality of care (and may provide information to improve quality of care) but also of the quality of their registration.
	Periodical meetings	Meetings can also be used for the above
Linking to other registries	Linking is possible	In order to answer complex research questions linkage of existing may be more efficient compared to collecting (all) new data
	Future opportunities for linkage are taken into account	Along with the motivation above it is generally more efficient to consider future (re)use in advance and to manage data accordingly
Handling of data requests	There is a review procedure to handle data requests	In order to enhance quality of research it is good to have a review procedure to see whether the question posed can be answered with the data
	Review is performed by different parties	To enhance an independent review process not only the "owners" of the database should review the requests
	Codebook available for other researchers	In order for other researchers to know which data are available and to use the database other researchers should be able to use the codebook
	Public publication	Public publication of research results based on particular data may stimulate reuse of these data and makes it possible for external parties to judge quality of data and research
	Involvement in performing research for example. preview in the publication or co-authorship	In order to enhance quality of research it is good to those who know the data well are involved in the research or at least review the publications before submission
	No or only handling costs for data (for publicly funded databases)	Publicly funded databases should be available for other researchers. High costs may limit accessibility
Suitability for research	Scientific publications about the data	This shows that scientific research is possible
	Multiple type of data available	Enhances the number of research questions that can be answered

We did not look at two categories of legal indicators we do not consider good practice but *requirements* for setting up a registry (see chapters 5 & 7):
- Measures taken with regard to privacy;
- Measures taken for informed consent.

6.2.2 Scores on indicators for good practice

Availability of protocols
Protocols are important in enhancing the quality of patient registries. Protocols describe, for example, how data collection and data control can be performed in a standardized way. Therefore, we asked whether the registry used protocols for data collection.[22] All but one registry answered to have such protocol. Three quarters of the registries (also) have a protocol describing how data have to be checked before they are delivered to other parties.

Incentives for delivering parties
The large majority of patient registries needs other parties to put effort in delivering data, for example physicians who have to register and extract (extra) data or patients who have to fill out questionnaires or participate in interviews or medical assessments (Table 6.3). In order to stimulate those parties to deliver data several incentives can be provided. We considered the following incentives to be good practice: providing feedback (either through the internet or on paper) and organising periodical meetings. Providing delivering parties (often health care professionals) with feedback or discuss relevant issues with them at a periodical meetings makes them not only aware of the – relative – quality of their care but also of the quality of their registration. In addition, they can use the feedback for other purposes such as negotiations with health insurance companies. Since financial incentives and presents do neither directly stimulate quality of care nor quality of registration we do not consider them as an indicator for good practice. Overall half of the patient registries provide feedback to parties who deliver data and 38% organise periodical meetings.

Opportunities to link the data to other patient registries
In order to answer new research questions using patient registries it may well be that existing registries each include part of the data needed. In that case it may be more effective to link databases than to start new data collections. Therefore, we consider it good practice if patient registries are able to link their data to other patient registries and/or when they enable future linking of their data to other registries. 70% of all registries included in this study can link to other registries now and two third says to be able to do that in the future. These two groups largely overlap. More information on this issue can be found in chapter 5.

Handling of data requests
Most of the issues belonging to handling of data requests are already discussed in chapter 5. We will discuss them here from a good practice point of view. The RP programme wants to create an infrastructure for research by stimulating the development of patient registries that not only can be

[22] We also considered asking for the content of the protocols. However, since these can differ very much between registries and in order to keep the length of the questionnaire limited, we decided not to include this.

used by those who start the registry but also by other researchers. Therefore, we asked the registries in our survey whether they deliver data to others and if so, under what conditions. Almost all registries participating in this study already deliver data to others than their own research group(s) and a large majority (89%) has a review procedure for data requests.
We also asked which parties are included in the review process of data requests. We argue that to enhance an independent review process not only those who (daily) use the data should review the requests but also other parties, for example relevant stakeholders. Less than half of all registries (43%) have other parties or a steering committee reviewing the data requests.
Finally, a large majority of all registries are prepared to share their codebook with other researchers.

Quality and independency of research can be enhanced by public publication since, in that case, results can be read and discussed by scientific peers. Therefore, requiring public publication of results from analyses using a patient registry is considered good practice. In total, 56% of the participating registries require such publication. In addition almost 80% of all registries require involvement in the publication process by demanding a preview of the publication and/or being a co-author of the publication. We consider this good practice for the following reasons. Parties managing a patient registry (often research institutes and/or medical centres) have much knowledge about and experience in analyzing their data. External parties who use these data can benefit from this knowledge in case the managing party is involved in the publication process. In addition, coordinating a registry requires effort, and being involved in scientific publications can be an important incentive for them.

Accessibility for others is an important asset of a patient registry. However, costs may be an important limiting factor for other parties to request. The Netherlands Organisation for Scientific Research (NWO) requires that all data collections funded by NWO are available for other researchers since the data collection is funded by public money. Since the RP programme of ZonMw will also (largely) be funded by public money, we consider it good practice that patient registries whose data collection is paid from public money provide data to others only asking financial compensation for data handling and advising. Looking at the answers of the participants in our study, it seems that this is already standing practice.

Suitability for research purposes
Patient registries in the RP programme will be used for scientific research. Publishing in scientific journals therefore is an indicator for good practice. Almost all patient registries (89%) participating in our study have 1 or more scientific publication: seven registries report that there are 1-10 scientific publications about their data, 14 registries report 11-100 publications and 12 registries were source for over 100 scientific publications. Another indicator may be the number of different types of data included in the database, since that enhances the number of research questions that can be answered. We asked for 10 different types of data such as data on utilisation of medicines, clinical data, questionnaire data, and bio samples (see also chapter 4). All databases included more than one type of data and on average they included five types.

Table 6.3 Percentage of patient registries scoring "yes" on indicators for good practice

	% patient registries 'yes' (n=37)
Protocols	
There is a protocol for data collection	95
There is a protocol to check data before delivering data	76
Incentives	
Feedback	51
Periodical meetings	38
Linking	
Linking is possible	70
Future opportunities for linkage are taken into account	65
Handling of data requests	
There is a review procedure to handle data requests	89
Review is performed by different parties	43
Codebook available for other researchers	89
Requiring public publication of the results	54
Involvement in performing research f.e. preview in the publication or co-authorship	78
No or only handling costs for data (for publicly funded databases)	100
Scientific research	
Scientific publications about the data	89
Multiple type of data available (range 1-10 types of data)	100 (mean: 5)

6.2.3 Distribution of scores across registries

We calculated how the individual registries scored overall on the indicators for good practice. Eleven indicators were included in this analyses out of the fourteen mentioned in table 6.4. We combined the indicators on linking because these two largely overlapped.[23] Moreover, we combined the scores on involvement as a co-author and/or having a preview of the publication. Finally, the indicator on financing was excluded because all registries seemed to fulfil our requirements, but the answers were not always fully clear. So in the end, we calculated a sum score that could range from 0-11. Of all 37 registries, there were 6 (16%) that met all criteria and another 7 (19%) that met 9-10 criteria. Five registries (14%) met less than half the criteria.

We looked whether there are differences in this score according to type of patient registries (national databases, regional databases, cohorts et cetera) and according to type of patient population in the registry. The highest average sum score on the eleven indicators was found for national databases (8.6; *95% CI: 7.5-9.7*), followed by "other databases" (8.0; *95% CI: 7.4-8.6*), and regional databases (7.5; *95% CI: 5.9-9.1*). Cohorts had relatively low scores (6.5; *95% CI: 4.7-8.3*). This may be caused by the fact that not all questions seem to be as relevant for cohorts, albeit we strived to include indicators that generally

[23] If a registry checked at least one of the two indicators on linking, this was counted as "yes".

reflect good practice.[24] Registries with a general patient population had a higher average score on good practice indicators (8.2; *95% CI: 7.2-9.1*) than registries including patients with specific diseases (7.5; *95% CI: 6.3-8.6*).[25]

Table 6.4 Number of times a registry scored "yes" on an indicator (range 0-11) (n=37)

	Number of patient registries with this sum score
Scored yes on **5 or less** items	5
Scored yes on **6** items	2
Scored yes on **7** items	9
Scored yes on **8** items	8
Scored yes on **9** items	5
Scored yes on **10** items	2
Scored yes on **11** items	6

We also looked at scores for the five different domains in our good practice index. For *Protocols* we included two items: the availability of a protocol for data collection and the availability of a protocol for data delivery. The majority of patient registries (73%) has both protocols. There were no differences between different groups of patient registries (according to type of data and according to patient population) in this. We also included two items for *incentives*: providing feedback and having periodical meetings. 40% of all registries do not provide feedback to those who deliver data nor organise periodical meeting, 30% does one of both and another 30% does both. Especially national and regional databases and registries with a general population provide these two types of incentives. Five items were included for *Handling data requests*. The average score on these five items across the registries was 3.5 (*95% CI: 3.1-3.9)* and 22% of all registries had the maximum score of five. Highest scores were found for national databases and "other" databases. No differences were found between registries with a general patient population and registries for specific diseases. Both for *Linking with other databases* and *Suitability for research* consisted of a single item. Almost 4 out of 5 registries have possibilities for linking (now or in the future); disease-specific registries have a slightly higher percentage that enables linking with other databases compared to registries with a general patient population. Nine out of ten registries have led to publications in international journals and there are no differences across different types of registries.

From the information above we can conclude that most patient registries fulfil the majority of (rough) criteria for good practice. Incentives for parties who deliver data are not always provided, while – in case the delivering party consists of health care professionals – this may make them aware of their own quality of care but also of the quality of their registration. Another criterion that is not always met is the requirement for open publication, while this enhances external review of data quality on the one hand and quality of analyses and interpretation on the other hand. Finally, in a minority of the patient registries, data requests are judged by parties other than the research institute itself, while this stimulates an independent review process. Scores for good practice are highest among national databases that include a general patient population.

[24] Our data included one bio bank, therefore this score is not included in the text.

[25] The other registries (children, pregnant women, elderly and mental health) had too few responses (1 or 2) to calculate a mean

6.2.4 Intensification of questionnaire results

To have more detailed input on good practice, we consulted three experts on in the field of (research with) databases: dr. M. Rookus, who was until recently the secretary of COREON, prof. dr. D. De Bakker, who was involved in setting up the LINH database, and has automated registration databases as one of his main research areas and prof. dr. Jan Raaijmakers, who is both a professor of pharmacy and vice president of European research collaboration with Glaxo Smith Kline Europe. Open interview questions were based upon the questionnaire for the registries, but posed in a more general way (see Annex 1). We asked the interviewees to reflect on what they think is important for ZonMW in order to finance a registry: what good practice criteria do these registries have to fulfil?

Arranging access for other researchers

As we stated earlier ZonMw wants researchers who receive funding for setting up a registry to provide others to data from this registry as well. We asked the interviewees what they consider important issues in arranging such access. Two experts stress the importance of a transparent procedure for data requests. For those researchers wanting to use the data it should be clear how data can be requested, who judges the data request, how long data handling will take and an indication of the costs.

A steering committee judging the data requests is needed according to the three interviewees. In line with what was considered good practice in this chapter, they argue that different parties should be represented in such committee. One of them argues that within such committee those members who potentially have a conflict of interest in a data request are not allowed to decide. Another expert states that external referees may be important to involve in order to avoid framed research questions. For reasons of neutrality one expert argues that a committee should be a mix of representatives of the data registry and external parties (including stakeholders). Those external parties should have the knowledge of what kind of research questions can be answered using the data in the registry. Such mixed committee is also important for reasons of transparency, image and acceptance. All three interviewees state that an important aspect for judging data requests by the patient registry is whether the data are suited to answer the research questions and have sufficient (scientific) quality. Therefore, all three interviewees find that those who request data should hand in a clear proposal. Moreover, they stress that doubling with other studies should be avoided and as such judgement procedures should include a check for this.

The interviewees find that registry holders are allowed to formulate requirements for data use, such as requiring public publications of the results and being involved as a co-author in the study in order to be able to overlook whether data are handled in the right way which was also included as an indicator in this chapter. In saying this, two interviewees do stress that in order to be a co-author an active role is needed; if not, the registry can be mentioned in the acknowledgements. Moreover, one of the interviewees argues that the patient registry should receive the dataset after the applicants have manipulated it (the dataset as it will be used for the analyses) and that preview in the manuscript may be helpful as well. According to one expert the bottom line is that patient registries always should be able to see that applicants use the data in the way that was agreed upon at the start. In addition, financial contributions can be asked such as paying (handling) costs or a financial contribution in order to keep the patient registry running and fees may be differentiated according to whom (and for what) data are deliverd (academia versus commercial parties). New registries that receive funding for setting up the organisation and data collection should immediately start using realistic fees (to be used once the funding stops) as pre-finance for the moment the funding ends.

According to the interviewees there may be valid reasons to refuse a data request. However, they differ when it comes to what these reasons are. Two of them mention that competing research is a reason to refuse a data request. One stresses that refusal to publish the results is a reason to refuse a data request and one interviewee finds that data requests of the pharmaceutical industry may be refused unless they have unique expertise. Some registries indeed stated not to deliver data to the pharmaceutical industry or pose restrictions. One interviewee states that in principal all requests should be honoured unless there are clear arguments not to do so and that the pharmaceutical industry is as well allowed to request and receive data under the same conditions as other parties (except for costs). Leading in this process is the research question and the data in the registry should be suited for that.

Visibility
According to the interviewees, patients registries should be registered at zorggegevens.nl and (all) other places where such registries are unlocked. Moreover, they should have an own website providing basic information about the registry and procedures for data requests. Moreover, a good catalogue with information is needed online for researchers as means of orientation. One interviewee states that the catalogue needs to be good but that more detailed information can be exchanged after a researcher has contacted the patient registry. Another interviewee states that making protocols requires time and regular updating since for various reasons databases can change over time making it necessary to change information needed in protocols. A good example are the so-called information leaflets made for the 2nd National Survey of General Practice where for each part of the database (for example prescription registration, morbidity registration, patient survey) a detailed description was provided about how data were collected, cleaned, what pitfalls the database had and what information was included. These were not published online but provided together with the requested dataset.

Technical and ethical aspects
With regard to privacy interviewees state that registries have to meet all legal requirements and one of them recommends registries to "take the safe side". According to all three interviewees a crucial requirement is that there has to be a wall between the researcher on the one hand and patient identifying information on the other hand. Moreover, no identifying factors should be delivered to external parties. One interviewee stipulates that registries educate their end users: these end users have to be aware of privacy aspects.

Linking of databases is becoming more common, a development strongly supported by the three interviewees. Two interviewees argue that the use of the BSN (through a TTP) for research purposes it would be of great benefit because this offers unique possibilities to link data. The other interviewee state that linking procedures should thoroughly be tested before becoming into use. All three interviewees agree that in order to improve recycling of data registries should use standard coding systems such as the ATC and ICPC or use coding systems used by large existing databases. Moreover they should provide information to researchers which protocols, variables and coding systems are used. One interviewee mentions that compatibility in coding between databases can be improved albeit registries sometimes cannot be blamed for this. This interviewee refers to the fact that while the new version of the ICPC-classification is available for quite some years, this is still not used in GPs' information systems which hampers the compatibility of the ICPC- and ICD-10 coding.

In archiving and protecting data protocols should be used as well, preferable a protocol that satisfies the NEN-requirements. Other important elements include: back up and auditing as well as using passwords wherever needed and a professional secrecy for employers working with privacy-sensitive data. Data should be transferred in a safe way, for example through a secured portal.

Incentives
Incentives for parties (health professionals, patients or others) who deliver data to a registry can be: co-authorship by rotation, access to own data, and feedback reports. Feedback reports becomes increasingly important for health professionals for several reasons: they work in increasingly more complex organisations and health insurers ask for scores on (quality and volume) indicators which can be provided in feedback reports derived from the registry. As such, participating in a patient registry can help health professionals to improve their registration but also provide them with information where they can improve their quality of care. Financial incentives are less valued by the interviewees. This is in line with the indicators used for good practice in this chapter. For parties who coordinate a registry incentives include co-authorship (2 interviewees explicitly state that in thatn case an active role is needed), acknowledgements, a reasonable financial contribution, and recognition of being a co-researcher and not just a "data-club".

From the interviews it became clear that the issues considered as good practice by the researchers largely overlap with the answers from the experts who provided, however, some more detailed information. Therefore, no changes were made in the initial set of indicators for good practice.

6.3 Do existing registries fit in the RP programme and can gaps be identified?

Chapter 3 described main research themes for the ZonMw RP programme. We identified those themes that were relevant for observational research (see tables 3.2 and 6.6). In chapters 4 to 6 we identified what types of databases are available in the Netherlands and whether they reflect good practice. From these chapters it appears that the Netherlands has quite a rich infrastructure of patient registries especially when it comes to automated routine registrations in primary care. Moreover, especially this last group of databases are highly accessible and, in case of public funding, costs seem limited to so-called "handling" costs. For secondary care, especially databases that include both data on health/indications and on drug utilisation/prescriptions seem to be scarce. There is, however, a wide variety of disease-specific patient registries which are often a combination of a cohort and a registration database: data of a cohort of patients are distracted from databases or from reports of professionals. However, these registries usually are not linked to other databases containing other information of the patient (for example utilisation of health care derived from automated primary care databases). The same holds for cohort studies that collect many data on patients through questionnaires and clinical measurements but lack connections to other databases. Moreover, accessible databases including detailed medication intake data are lacking. Table 6.5 summarises per type of database needed for the RP programme its current availability.

Table 6.5 Type of observational databases (minimally) needed to study main research themes mentioned in ZonMw's preparatory studies of 2009 and 2010

Type of database needed	Current availability
Prescription databases in primary and secondary care, preferably including data on indications as well as other relevant patient characteristics and clinical outcome measures	These databases are available and accessible for primary care but for secondary care availability is limited, especially with regard to databases including indications as well. Within these databases, extra data collections in patients are scarce and more clinical outcome measures could be included.
Databases in primary and secondary care, including data on indications, prescriptions and lifestyle interventions	See above, information on lifestyle seems not to be included in most databases.
Patient cohorts (preferably from these databases) to collect extra information for example using survey data or clinical measurements	Cohorts are available and accessible but they usually have limited linkages with databases in primary and secondary care.
Special patient cohorts (if possible linked to existing databases) to collect information needed to answer specific research questions	Cohorts for special patient populations are available (f.e. LASA) and accessible but they usually have limited linkages with databases in primary and secondary care.
Databases including detailed medication intake data (registering each medication intake)	These databases seem not to be available

6.4 Conclusion and recommendations[26]

Looking at the combination of research themes in the RP programme and the patient registries already available we conclude and recommend the following with regard to setting up new registries:

- Primary care databases are widely available and accessible and there is a good infrastructure to expand existing databases either by linking or by combining different type of data in one registry as will be done in the NIVEL Zorgregistraties where monodisciplinary databases (GPs, pharmacists, paramedics and psychologists) will be integrated in one network enabling to follow patients through the health care process. Detailed data on adherence (for example through electronic monitoring, could be added to such databases as well.
- There is a need for cohorts preferably nested within these existing automated routine databases in order to be able to fully answer part of the questions relevant for the RP programme; and the other way around: existing cohorts studies are recommended to use possibilities to link with data collected through automated routine registrations.
- There is a need for secondary care databases that combine information on indications as well as prescriptions (and other clinical information). Compared to primary care, the infrastructure for secondary care is still in its infancy. Therefore, we suggest not to aim to strive to come to a general infrastructure for secondary care including many secondary care settings and a general patient population (as in fact can be found in primary care) in the short run. We suggest to start with pilots in different settings to learn where problems lie such as in fact now has already with registries for specific diseases.

Based upon the good practices as well as on results from chapter 5 and chapter 7, we developed a checklist that ZonMW uses in the call for new registries. This checklist (in Dutch) includes criteria that are important in the referee rounds where the applications for new registries are judged and prioritized

[26] These recommendations are discussed in two of the expert interviews, who agreed upon this line of reasoning.

as well as a detailed instruction for issues that have to be addressed in the application for new registries on the following subjects: originality of the registry (not yet available), data sources used, ethics, governance, visibility, technique/data model (including linking). The list can be found here: http://www.zonmw.nl/fileadmin/documenten/Goed_Gebruik_Geneesmiddelen/Bijlage_1_Checklist_registers.pdf

7 Organisation of registries

The focus of this chapter is twofold. The first part is organisational and describes the situation of (medical) data governance and accreditation of registries. Following a presentation of the current situation a draft for a good governance model for the Rational Pharmacotherapy programma will be presented. The second part of the chapter focuses on the infrastructural side of the organisation of registries and takes stock of financial and technological aspects of access to patient data. Data discovery, the interoperability of registries, as well as ICT-supported access will be dealt with, ending again with recommendations. For this second part we studied the websites of a number of registries, in addition to the questionnaire.[27]

7.1 General introduction to governance and accreditation[28]

The Rational Pharmacotherapy (RP) programme is conceived as a public–private partnership between the Ministry of Health, pharmaceutical industry and health care insurers in the Netherlands. One of the main recommendations in the report "Verdieping Goed Gebruik Geneesmiddelen" (pages 8-9) was that funding for the RP programme should be "structural and independent". It was also recommended that research in this programme should be carried out in a integrated way as much as possible. Isolated execution of partial research projects in this field should be avoided. Related to this it was also advised that a solid infrastructure is necessary to enable a good RP programme.

Independency from both commercial interests (pharmaceutical industries) and policy management (health care insurers, government) is needed. At present the situation is unbalanced: The overwhelming majority of research in this field is funded by the pharmaceutical industry ("Verdieping Goed Gebruik Geneesmiddelen", p. 8). It was recommended in that report that all the different parties in this field or market should contribute, not only the pharmaceutical industries but also the health care insurers and the government, preferably in close cooperation. They are the main stakeholders and they should work more closely together.

In this context it is worthwhile to mention another conclusion from the report, which is that good research on the use of medicines (pharmacotherapy) is not only needed from clinical trials, but also by observational and translational methods. However, structural funding possibilities are mostly lacking for the latter types of research.

---

27 Expert reviews were given on the sections 7.1 to 7.6 by Mr. Evert-Ben van Veen (MedLawcConsult) and professor dr Mattijs E Numans, MD PhD (Primary Care Julius Center UMC Utrecht, Department General Practice VUmc Amsterdam and Surgery "Huisartsen Oog in Al" Utrecht). Their opinions have led to changes in the final version of the text. The registries involved in the website scan have been recommended to us by external experts or were already known as good practices; it is a modest and subjective selection. The findings of the website scan are documented in full in Annex 7

28 Results of a literature search on governance used for this chapter are included in Annex 6.

It is essential that the independency of the research outcome is guaranteed, in particular for the position of the patient. This includes good and trusted management and care of the research data. In a report by the Dutch Advisory Council on Health Research[29] "Van gegevens verzekerd (Securing the data supply)" it was concluded that "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."

To make these recommendations operational a clear model of data governance is needed. This is particularly relevant for the management of the data in a trustworthy way. A model is in itself not enough; its quality and trustworthiness should be assured as well. That can be done by formulating and applying *accreditation* standards. In the next six paragraphs we take a closer look at data governance and accreditation and formulate specific recommendations on these for the RP programme.

7.2 Data governance

For this programme we follow the view of the Dutch Advisory Council on Health Research 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 should be respected as well." In addition, the privacy of the data subjects has to be protected.

Certainly not all data are funded by public resources. This is reflected in the RP programme which is strongly conceived as a public-private partnership. The question here is under which circumstances databases can be regarded as publicly funded. Different views exist on the definition of public funding in the Netherlands and, consequently, on the nature and use of the data that are generated. Data that is specifically collected for research that is funded by the ZonMw or NWO is obviously publicly funded. However, with health data that were originally collected for other purposes (eg. billing, direct patient care, health service management) this is subject for debate. To what extent, for example, are data that are routinely collected by health care insurers or in pharmacies or GP practices or hospitals publicly funded? In the Netherlands there is a system of managed competition, where private insurance companies pay for health services, where health services are supposed to compete as well as to co-operate with each other and where patients pay insurance fees to insurance companies. In this situation it is not always clear to what extent data are publicly funded and to what extent they should be regarded as public. This also implies that part of this information is strategically as well as economically important for the entities involved. Sharing this information may lead to violations of the regulations of the Netherlands Competition Authority NMA. Consequently this information will, very probably, not be available for research purposes.

It follows from this that a "balanced arrangement" has to be made between the private and the public interest. As the recently published report "Data in public-private projects. Legal aspects" assigned by

[29] 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.

SURFfoundation concludes: "In situations where a private party is involved as a co-funder or as a provider of some or all of the information required for the research, it is necessary for a balanced arrangement to be agreed on. This must take account, on the one hand, the interests of the private party concerned and, on the other, the Open Access policy pursued by [public funding councils] NWO and STW" [30].

To achieve this a range of recommendations is offered by SURFfoundation. Two points are of particular importance here. First, "If the private party acts as a co-funder, it may wish to make use of the research results" and "A private party involved in a public-private partnership may also be a *provider of research data*. When that external data [i.e. data not created in the research project] is "bought", a provision to protect that party's interests may be imposed to the effect that the external data it has provided must not be shared or reused any further". In health research there may be different views by the various stakeholders on what such a provision should entail as, again, the division between public and private funding is not seen by everybody in the same way.

The second issue is the prevention of intellectual property rights being an obstacle to the sharing and re-use of research data. Preventing this obstacle could be achieved by stipulating unambiguously that parties that receive funding must grant all the required consent for such sharing and re-use (in advance), or must waive the relevant intellectual property rights.

7.3 Accreditation: data access

As said earlier, accreditation is defined here as an instrument for quality assurance with respect to the management of and access to medical data for research, in particular patient registries. The question to be answered is: who is allowed to do what with medical data and how do persons and organisations fulfil accreditation standards? In the medical and health care world accreditation exists, as can be expected, in a number of fields and at different levels. Accreditation should not be confused with certification , the latter answering the question "who is qualified ("certified") to do what in health care?".

Assuming that the register data will be used for scientific research there are two general codes of conduct in this field. The first one is the *Nederlandse Gedragscode Wetenschapsbeoefening VSNU* (The Netherlands Code of Conduct for Scientific Practice). This code of conduct is published by the VSNU (Association of Universities in the Netherlands). It is aimed at scientific and academic researchers. This code is of a very general nature and has five basic principles: scrupulousness, reliability, verifiability, impartiality, and independence. It has no sanctions.

The other code of conduct is aimed specifically at researchers using datasets containing personal data. These researchers are required to comply with the *Gedragscode voor gebruik van persoonsgegevens in wetenschappelijk onderzoek* (Code of Practice for the use of personal data in scientific and scholarly research) published by the VSNU. This Code of Practice is obligatory for staff employed by the Dutch universities, NWO and KNAW institutes. It contains a complaints procedure. The DANS data archive has

[30] R.W. de Bruin, M. de Cock Buning and A. Ringnalda (CIER), *Data in Publiek-Private Projecten: Juridische aspecten* SURFfoundation 2011 p. 8
http://www.surffoundation.nl/nl/publicaties/Pages/DatainPubliek-PrivateProjectenJuridischeaspecten.aspx

made it compulsory for its registered users to comply with this latter code. In practice this means that only academic researchers can get access to datasets containing personal data. They are required to maintain the confidentiality of all personal data that they process. In publications persons should never be identifiable, in other words anonymisation is obligatory. This latter Code is based on the *Wet Bescherming Persoonsgegevens* (WBP: Data Protection Act), which states that personal data may only be used for "scientific, statistical or historical research". The restrictions do not apply if a dataset has been strictly anonymised. It should be noted that its period of validity has expired. The legal situation is not clear at the moment[31].

The law provides for various types of personal data, some of which have stricter regimes than others. According to the WPB, *special personal data* are data providing information, for instance, on sexual inclination, health, religion, race, etc. Stricter rules apply to the processing of such data than to the processing of data like name, address and residence. The law confines itself to living persons. Personal data of deceased persons therefore fall outside the sphere of influence of the WBP. However, this does not mean that all information on deceased persons can be made public without restrictions. Information on deceased persons must be handled carefully, the privacy of surviving relatives must be respected.

Medical data must be treated with the utmost reserve. Medical professional secrecy remains in force even after the subject has died. Utmost reserve applies also to datasets containing the BSN number (Dutch Social Service Number). For making datasets accessible that contain special personal data, such as data concerning religion, race, political opinion, criminal records or health, it is furthermore required that the licensor demonstrates that the parties concerned have given their consent for the archiving of their data and making these available by a data repository for example for statistical and scientific purposes. If the licensor cannot demonstrate this consent, the dataset shall only be accessible under very strict conditions, which, in accordance with Article 23 of the WBP, sets further requirements to the nature of the scientific research or the statistical processing, for the benefit of which use of the dataset is considered necessary. In practice it may not always be easy to determine who is a genuine "scientific researcher" who may be granted access and who is not.

Medical research is governed by the "Wet geneeskundige behandelingsovereenkomst" (WGBO -Dutch Medical Treatment Act) on top of the law and codes mentioned above. The "Federa Gedragscode Gezondheidsonderzoek" is based on the WBP and in particular on the WBGO. It is a code of conduct which regulates in more detail how researchers should handle medical data[32]. Its validity has also expired[33]. This law prescribes the care provider to protect and guarantee the privacy of the patient. All information should be treated confidentially. A care provider is not authorised to give any piece of directly identifiable information to other ("third") parties without the explicit permission of the patient or if the law requires the care provider to do so or for the purpose of scientific research and then only under strict conditions.

The medical professional secrecy is based on the patient's entitlement to confidentiality of his of her personal medical data. Death of the patient does not mean that his or her privacy is not protected any more. The protection remains in force. Access is only possible when the care provider knows for certain,

---

31 http://www.cbpweb.nl/Pages/ind_wetten_zelfr_gedr.aspx

32 http://www.federa.org/gedragscodes-codes-conduct-en
http://www.cbpweb.nl/Pages/ind_wetten_zelfr_gedr.aspx.

and can demonstrate, that the patient would not mind others having access to the medical data. The general term for preserving medical files is 15 years after which files must be destroyed. When data have been anonymised they can be preserved. When data have been anonymised it is impossible to identify individual persons. This exception is of great importance for scientific research.

It is furthermore important to note that it is possible to use medical data for scientific or statistical research without consent of the patient in two situations:

- It is not reasonably possible to ask for consent and the privacy of the patient is not unnecessarily jeopardised.
- Asking for consent is not feasible and the data arrive at the researcher in such a way that re-identification is sufficiently prevented.

In both cases three other conditions have to be met:

- The research serves a general interest;
- The research cannot be carried out without those data;
- The patient has not objected to such use of his or her data for research[34].

7.4 Accreditation: handling and preserving medical data

Seen in the light of the regulations for access to research data in general and to medical data in particular it will be clear that the management of how to preserve and give access to these data, particular patient registries, is an essential issue when setting up the RP programme. The question to be answered in this paragraph is what accreditation exists in this area, in other words: how are quality standards assured? There are two concepts here which need a closer look: trusted digital repositories and trusted third parties.

Trusted digital repositories

Trusted digital repositories (TDR) is a relatively new concept. TDRs come from the world of digital libraries and archives and should be seen in the context of digital preservation of research data. The word "trusted" here implies guaranteeing *long-term* preservation and access. In a TDR data are preserved in a safe place so that they can be accessed permanently in the same way now and in the future. Access to data itself should be reliable and based on existing and applicable laws, regulations and codes of conduct. A set of criteria is needed to be able to guarantee the quality and trustworthiness of a repository archiving digital data. This means that requirements have to be formulated to ensure that a TDR meets minimum standards of quality, traceability, accessibility, and usability. Ultimately this should lead to criteria and procedures for the certification of digital repositories.

Generally speaking the field of certification is very much in development at the moment; various organisations are in the process of specifying criteria for "trusted digital repositories". The guidelines and best practices being developed now range from very detailed audit checklists or methods focusing on the organisational and technical infrastructure to guidelines that are very general. One of the difficulties in this area is that, if guidelines contain technical specifications, they may be outdated soon

[34] WBP - Data Protection Act, article 23.2 and "Federa Gedragscode Gezondheidsonderzoek" p.16 - http://www.federa.org/gedragscodes-codes-conduct-en. Also: Evert-Ben van Veen, Patient data for health research. A discussion paper on anonymisation procedures for the use of patient data for health research p.29 (The Hague, October 2011) http://www.medlaw.nl/wp-content/uploads/patient-data-for-health-research.pdf

and have to be updated regularly because of the rapid technological developments. The most important, or promising, audit and certification developments at the moment are the Data Seal of Approval (DSA[35]) and the RAC "Audit and certification of trustworthy digital repositories ISO/DIS 16363"[36]. The DSA is a low level certification with a self-assessment system, concentrated on a data repository. It contains 16 short guidelines. The RAC on the other hand is a very extensive set of standards also focused on digital repositories. It is in the process of becoming a formal ISO standard.

An important consideration, and stimulus for the development of these certifications, is the possibility to store data in different places, not necessarily only in data archives or data libraries. This is also relevant for health research data. It means that data could be stored and made accessible at the place, the organisation, where they have been created. However, to guarantee (long-term) preservation, use and access to the data certification of such "local" digital repositories is necessary in accordance with all the legal obligations and codes of conduct. For this the certification models DSA or RAC could be used.

Trusted third parties

In health research some experience has been acquired already with the concept of "trusted third parties" (TTP). TDRs and TTPs are different organisations as they fulfil very different functions. The TTP is basically a service provider in handling register data, in particular patient data, *which have to remain anonymous*. They have do not have a task at all in storing or giving access to these data. The function of a TDR on the other hand is to play a vital role in saving the data as well as making them available. We return to this in the next section.

To follow a conclusion of Van Veen in his paper the TTP in health research "is an independent and reliable party which takes care of the controllable pseudonymisation of patient data into anonymous data". TTPs originated in the financial world to verify and guarantee the identity of a person in transactions between different parties. The main function of TTPs in the health care and health research world on the contrary is to make it impossible (or very difficult) to identify persons in a database, while preserving the possibility to link data from one source to data from another source on a patient by patient basis. In order to do that one individual who is present in both datasets is given the same unique *pseudonym* on the basis of – for example – this patient's identification number. Subsequently, researchers can work with the linked dataset without having access to the original patient identification number. This "privacy enhancing technology" makes it thus possible to link pseudonymised patient data, which can greatly increase the research possibilities. Patient data could for example be linked with data on socio-economic background. However, the use of TTP as an instrument for pseudonymisation should be based on a number of conditions[37].

An important problem here is that it is currently impossible to link data that are pseudonymised by different TTPs. It is impossible to link data that is pseudonymised by one TTP to data relating to that same individual that is pseudonymised by another TTP. It should be realised that this is exactly what pseudonymisation is supposed to do. This will however either increase costs (double pseudonymisation) or make data linkage impossible. A second issue here is to guarantee patients that it is impossible to

---

35 http://datasealofapproval.org/

36 http://www.crl.edu/Archiving%20%2526%20Preservation/Digital%20Archives/-Metrics%20for%20Assessing%20and%20Certifying-0

37 This section is for an important part based on Evert-Ben van Veen, Patient data for health research. A discussion paper on anonymisation procedures for the use of patient data for health research (The Hague, October 2011) http://www.medlaw.nl/wp-content/uploads/patient-data-for-health-research.pdf , in particular pp.11-15.

identify them from the data. This remains vital and is a concern because neither patients nor general practitioners would like to see patient data in the hands of others, including researchers, without proper anonymisation[38].

TTPs use digital certificates for securing the transmission of data[39]. Following Van Veen, however, there is not yet a complete audit system in place for checking the effectiveness of the data handling policies in health research. Who should carry out these audits, the data handling organisations themselves (hospitals, research institutes) or TTPs or regulatory bodies like the national Data Protection Authorities, like the CBP (College Bescherming Persoonsgegevens)? Auditing by these external bodies could become very expensive[40].

7.5 Data governance in the RP programme

In this section we will focus on governance at two levels: first, the level of patient data registries and second, the (national) level of the RP programme itself. Although there are obvious relations, a separation of concerns makes sense. Apart from the entities Trusted Third Party and Trusted Digital Repository introduced before, other entities enter the stage, such as the independent data broker.

7.5.1 Governance at registry level

In the questionnaire some questions were posed to make an inventory of various stakeholders. We start with providing some findings. Next, the importance of being a TDR will be stressed.

The answers given in the questionnaire show that most registry data seem to be managed by the bodies carrying out the research. The answers make clear that a variety of data "owners"[41] exists. In most cases this is one institute (like a university medical centre) or a group (a group of general practitioners participating in a research programme). In a small number of cases this is a combination of organisations. The same conclusion can be derived from the answers given to the questions on who is "legally responsible" for collecting, managing, storing and using the data or who has been funding the collection, management and storage of the data. The "owners" or the "legally responsible" organisations are sometimes the same, for example universities/university clinics or health insurers. Mostly mentioned as funders are the Ministry of Health (VWS), ZonMw and universities themselves. In particular VWS or ZonMw are as a rule not mentioned as "owner" or "legally responsible". Access to the data are mostly granted by the same organisations, i.e. the "owners" or the "legally responsible" organisations. In some cases the funding organisation is also the same as the owner, but this is not the case by research funded by VWS; for data resulting from VWS-funded research the research organisations are responsible. Furthermore, according to the answers given in the questionnaire, most organisations indicate that the data should be handled by a Trusted Third Party. This could be more or less expected.

---

38 This was explicitly remarked by medical data expert Numans.

39 See for example: https://www.zorgttp.nl/pages/certification_page

40 Van Veen, pp. 85-86.

41 In the questionnaire the term "owner" was used. Strictly legally speaking this should be:"rights holder". See R.W. de Bruin, M. de Cock Buning and A. Ringnalda (CIER), *Data in Publiek-Private Projecten: Juridische aspecten* SURFfoundation 2011 p.218 http://www.surffoundation.nl/nl/publicaties/Pages/DatainPubliek-PrivateProjectenJuridischeaspecten.aspx

Knowing these differences is interesting in itself, but because this preparatory study does not have the information to relate the different situations to for instance differences in quality or effectiveness, it is not possible to derive recommendations from the responses. Based on section 7.4, however, we recommend that registries adopt the concept of Trusted Digital Repositories. All participating research institutes or data creating institutes, generally speaking, could keep the data – either existing or to be created in the RP program – if they want to do so, on condition of some form of certification and then they can act as TDRs themselves. An important element in these conditions is access to the data of the RP program for at least other researchers. Data sharing and the waiving of property rights, at least within the RP program, are essential. Certification in a light (DSA) or heavy (RAC) way is also essential.

7.5.2 Governance at the national RP program level

It is useful to look at the recommendations of the Advisory Council on Health Research on the governance and the efficient use of research data and applying these to the RP programme. They distinguish three main issues:

Establish a register of data collections

What is needed according to the Council is "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". According to the Council the existing www.zorggegevens.nl is not sufficient: it does not cover all the available data (p. 45). The programme should establish registries of the existing or (at least) of the new data collections of the research data files generated in this programme. These registries, containing descriptions should be open for everyone.

Optimise access to data

The Council was strongly in favour of the formulation of a code of conduct on data sharing. 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:

- Public registration of the data collection, visible for everyone..
- Adherence 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.

Despite the fact that this code of conduct has never materialised until now these points are still valid and could to a large degree be enhanced by the programme.

It could be asked whether the existing laws and codes of conduct mentioned above – WBP, WGBO, the code of conduct VSNU for the use of personal data in academic research and the Federa code – are not sufficient. Possibly they are, but a code which would elaborate on the practice of handling and using medical data are still needed, according to the Health Council. This proposal should certainly be looked into when setting up the RP programme.

Stimulate data sharing

The Council believed 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. this service provider could act as an information broker in a number of ways. It could give information on the

availability, and accessibility of data in various registries. Furthermore it could assist in negotiating access to data between the various parties holding or accessing data. It could also play an important role in achieving standardisation of TTP protocols. The Council also wished to see the provision of practical help and support with the technical aspects of sharing, pooling and linking data. The Council recommended that research funding bodies make resources available for such facilitative activities.

Despite the important notion of independence, the data brokerage service might be provided by a TDR that is being perceived as "neutral", for instance CBS or DANS[42]. Neutrality can be seen as a subjective aspect of trustworthiness. Minimally it entails independence from funders and other stakeholders. If researchers would perceive a data broker who is related to a TDR as biased towards the data in that particular TDR, this broker could not play its role.
To draft a way how the RP programma might embed this data brokerage service, it is useful to briefly discuss the various stakeholders and their roles. There are several stakeholders involved in the RP programme. As funding partners can be distinguished pharmaceutical industries and policy management like health care insurers and governmental organisations. Furthermore, medical staff prescribing, pharmacists dispensing and patients using medicines are seen as vital "target groups" of the RP programme. We confine ourselves to the view that these stakeholders, as well as target groups, should be represented in a governing board or a steering committee of the program. This body should define and monitor the research programme as a whole, in principle for a number of years.

On the other hand, there should be a programme committee consisting of experienced medical researchers, preferably researchers from both the medical and the pharmaceutical faculties of universities[43]. They should not have close links with industry, the ministry or health insurers. This should guarantee academic independency. The programme committee is seen as the academic committee of the programme. It could make the final decisions on granting funds for research projects in this programme. More relevant and important in the context of the previous paragraphs is another function for this committee: It could ultimately be responsible for the management of the data in the RP programme, in particular on issues of access and preservation. The programme committee could delegate this latter task totally or partially to an independent data broker, as described before. It could and should enforce certification of the various – trusted and not yet trusted – data repositories participating in the RP programme. It could act as data access body overseeing the conditions for access to as well as preservation of the data. Sharing of the data and waiving of property rights, or at least access rights, should be taken care of by this body as well. It might also act as the organisation which actually decides on who, which researchers, will get access to which data. It would act as the data broker, including the tasks mentioned earlier. A code of conduct, to be developed for this purpose as suggested by the Health Council, could be of great help here as well as the standardisation of the various TTP protocols. The execution of all these tasks could be delegated to a central service provider in the RP programme, supervised by the programme committee.

42 DANS does not handle medical data at the moment. This would be a new task for DANS.

43 See for example the governance of the NWO programme "Geweld tegen psychiatrische patiënten", http://www.nwo.nl/nwohome.nsf/pages/NWOA_7NSKQB

7.6 Recommendations for data governance

To sum up we recommend the following:

1. The RP programme should establish a register of the existing or at least of the new data collections of the research data files generated in this programme. These registries, containing descriptions, should be open for everyone.
2. The RP programme should consider the formulation of a code which would elaborate on the practice of handling and using medical data as suggested by the Health Council to avoid uncertainties left by the existing codes of conduct and laws referred to in this chapter.
3. In a similar vein, the RP programme should stress the importance of renewing the codes of conduct that have expired.
4. The RP programme needs an "independent data broker". This should be a central service provider.
5. There should be special attention for the further development of certification for the TDRs and standardisation for the TTPs, either by a central service provider or a "programme committee".

In the next sections the perspective moves from the relatively high, organisational, level of governance to more practical levels of data access, i.e. financial and technological aspects.

7.7 Funding, incentives, and pricing

The survey carried out in this preparatory study distinguishes three situations where potentially money flows: from a registry funder to the registry (funding), from a registry to data providers (incentives), and from data users to a registry (pricing). Over time – and overly simplified – these flows look like this:

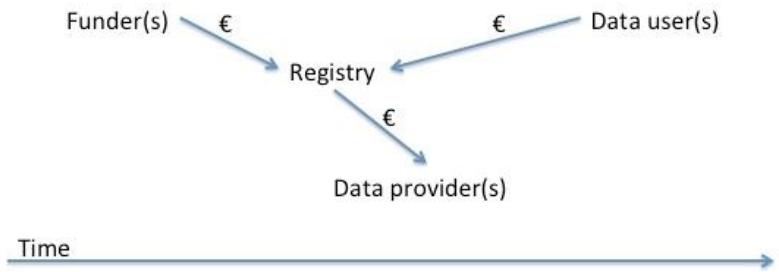

Figure 1 Money flows: funding, incentives, and pricing

At a more detailed level the current situation is as follows:
14 respondents mention public *funding* from VWS (Ministry of Health, Welfare and Sport), from other ministries, ZonMw, or CVZ (Health Care Insurance Board). 13 respondents mention funding by one or more hospitals or academic medical centres. Furthermore, 4 respondents mention funds or foundations among their funders (Q44, multiple answers possible). Merely 4 out of 37 respondents mention private organisations among their funders[44]:

- fee-charging: 1 registry funded by a health insurance company;

[44] As the statement cited from "Verdieping Goed Gebruik Geneesmiddelen" in section 7.1 refers to pharmaceutical industry as the main research funder, that does not say anything about who finance(s) registries.

- no charges: 1 registry funded by a public-private partnership; 1 registry being supported by three pharmaceutical companies; 1 registry funded by a medical centre a governmental organisation, and two pharmaceutical companies.

27 registries (73%) have *incentives* for data providers, while 10 do not. Incentives come in various forms: management-information (such as quality indicators) either on paper or online, financial compensation, and periodical meetings. Additionally, respondents mention that they support data providers with research or data collecting, as well as with newsletters and reports.

While at 15 registries the data are free of charge, 21 registries (58%) have a *pricing* system[45]. No single pricing system emerges from the answers, but overall the rate depends on various factors, often in combination:

- the amount and complexity of work that the registry has to carry out; a few respondents refer to this as "data management", although this may extend to preparation of the data for the specific researcher;
- whether access to the data are requested by data providers, funding body, internal researchers, students (typically: little or no costs), external non-commercial researchers (typically: cost price), or commercial use (in one case: twice the non-commercial fee);
- the extent to which the intended research is related to the registry's mission: If unrelated, the charges are higher;
- whether the researcher also contributes to maintenance of the infrastructure and/or the cohort data. In this case there seems to be a trade-off between money and effort for the requesting party.

Some registries charge a fixed fee, starting at €75-100 per hour. Other business models are based on €5.000-€10.000 per research publication or on an annual subscription fee ranging from €1.000 to €20.000 (commercial rate). Only 2 registries that charge fees indicate that information about rates is available at their website. The Perinatale Registratie presents a good practice: It informs about the hourly rate and mentions that any costs charged by the Trusted Third Party for data linking will be passed on to the researcher[46].

To conclude this section: nearly three quarters of the respondents offer incentives for providing data. These incentives predominantly consist in the exchange of information (identified as a good practice in chapter 6). The responding registries are predominantly funded from public means. The answers provided do not warrant any conclusion regarding a relation between the funding body or bodies on the one hand and whether registries charge a fee on the other hand.
Although the registries that do charge fees appear to be able to tailor to various kinds of requests, more harmonisation of pricing criteria and policies is advisable. For instance when a prospective researcher intends to request a subsidy but cannot tell the funding body beforehand what data access and reuse will cost. Alongside more harmonisation we recommend more online transparency about the pricing model.

---

45 One respondent did not provide information about pricing, so N=36 here.

46 http://www.perinatreg.nl/gegevensaanvragen?noCache=76;1327421553

7.8 Data discovery through metadata

In many cases researchers will have an inkling of existing registries that might be relevant for answering their research questions. Nevertheless, the process of discovering potentially relevant data and deciding whether they are really fit for use is strongly facilitated by rich metadata. Codebooks and other listings and explanations of variables used in a study can be considered as rich metadata as well.

From the answers to the questionnaire we find that 89% of the registries has a codebook available for researchers (in section 6.2 this is recommended as a good practice). We did not ask about the use of metadata per se. From the website scan carried out in this preparatory study we present three good practices regarding rich metadata:

- Zorggegevens.nl is an umbrella site that covers several registries and presents "data sources like health care registries, survey, monitors, and longitudinal (cohort) studies. Primarily it is aimed at professionals such as epidemiologists, researchers, policy makers and students"[47]. The site presents, for example, metadata regarding collection goals, data units, sample size, updating frequency, availability and contact.
- The Medical Research Council (MRC) in the United Kingdom provides a Research Data Gateway "for finding MRC-funded population and patient studies, datasets and variables; and for finding the procedures for requesting access to data[48]." The site presents, for example, metadata regarding study, time period, data collection event, variable, and contact.
- At the Centre for Health Equity Studies in Sweden (CHESS[49]) researchers from sociology, psychology and public health sciences work together on issues of health and inequality. Several data sources are available at CHESS, such as the Stockholm Birth Cohort Study with both survey and register information for 50 years on a cohort born in 1953[50]. The Stockholm Study site provides metadata and coding schemes, allowing for detailed interpretation of the data; all downloadable as PDF file. The other registries managed by CHESS also provide much information about the original studies.

What is not clear from most websites, is whether the metadata adhere to a standard. An exception is the UK Medical Research Council mentioned above, which advocates the use of DDI in its recent "Policy and Guidance on Sharing of Research Data from Population and Patient Studies"[51]. The Data Documentation Initiative (DDI[52]) is an international metadata standard to describe data from the social and behavioural sciences across the life cycle. The DDI metadata specification supports the research data life cycle, from data conceptualisation, collection, processing, distribution, discovery, analysis, re-use to archiving. Because the goals of the Medical Research Council and those of the RP programme coincide, the DDI standard is clearly worth looking into.

At the other websites within our scan such metadata are scarce or at least not presented in a similarly transparent manner as these good practices. However, most registries do provide information on the specifications or classifications that the data adhere to. This brings us to the next section.

---

47 http://www.zorggegevens.nl/kennisplatform/default.aspx

48 https://www.datagateway.mrc.ac.uk/node/14

49 http://www.chess.su.se/

50 http://www.stockholmbirthcohort.su.se

51 Medical Research Council, November 2011. Retrieved February 3, 2012, from http://www.mrc.ac.uk/Utilities/Documentrecord/index.htm?d=MRC008302

52 Data Documentation Initiative, http://www.ddialliance.org/

7.9 Interoperability of registries

Interoperability is the ability of systems – in our case patient data registries – to exchange information and to use the information hat has been exchanged, according to the definition of the Institute of Electrical and Electronics Engineers IEEE. Interoperability is a prerequisite for data linkage: In order to link a patient-based record in database A to a patient-based record in database B, the data must be expressed in interoperable formats.

On the one hand interoperability could be achieved through data formats or file formats. Therefore the questionnaire included the question *In what format are the data available*. 15 registries support a single format, the other registries support up to six formats. Statistical formats such as SPSS are popular (checked 27x), followed by spreadsheets such as Excel (18x). A few respondents answer that they can provide any format that a researcher or health care professional needs. From the point of view of the RP programme there is no reason to prefer one format to another. Note also that sustainability – will the format remain accessible over time? – was no issue in this preparatory study.
On the other hand, specification or classification of data is an instrument for comparing data from one registry to data from another; put differently, it makes registries semantically interoperable. Apart from patient data like social security number, sex, date of birth and address elements (not standardised), websites mention coding standards for diseases, diagnoses, treatments (e.g. ICD-10[53], ICPC[54], CTV3[55]), product specifications (e.g. ATC[56], GPK[57]), and identifiers for care professionals and insurance companies (e.d. UZI[58], UZOVI[59], AGB[60]).

A good practice here is presented by the Landelijke Basisregistratie Ziekenhuiszorg (National Hospital Care Basic Registration or LBZ), managed by the Dutch Hospital Data foundation[61]. From the website a data model can be downloaded which relates such classifications. LBZ itself claims that the data model has been developed for one-entry registration for multiple purpose. This is in line with the red tape reduction ambitions of the Dutch government ("Eenmalige registratie, meervoudig gebruik"[62]). Also NIVEL's Netherlands General Practice Research network[63] (LINH) provides ample information about the data that is collected and its representativeness.

As already mentioned in the previous chapter roughly two thirds of the respondents indicate that combining "their" data and data in other registries is possible, and that they take future opportunities for linking into account. That means that many registries are and will be interoperable. The most common hinge for interoperability of patient data is clearly patient information, in particular a social

53 International Classification of Diseases, http://www.who.int/classifications/icd/en/
54 International Classification of Primary Care, http://www.who.int/classifications/icd/adaptations/icpc2/en/index.html
55 Clinical Terms Version 3 from the Read Codes http://www.connectingforhealth.nhs.uk/systemsandservices/data/uktc/readcodes
56 Anatomical Therapeutic Chemical, http://www.whocc.no/atc/structure_and_principles/
57 Generieke Productcode, www.knmp.nl/producten-diensten/g-standaard/farmaceutische-gegevens/kenmerken-gpk-prk-en-hpk
58 Unieke ZorgverlenerIdentificatie, http://www.uziregister.nl/
59 Unieke ZorgverzekeraarsIdentificatie, http://uzovi.vektis.nl/
60 Algemeen GegevensBeheer Zorgverleners, http://www.agbcode.nl/MainPage/default.aspx
61 http://www.dutchhospitaldata.nl/
62 http://e-overheid.nl/onderwerpen/e-overheid
63 http://www.nivel.nl/linh/

service number and/or date of birth and address data (see also chapter 5). The availability of such information in a register makes it interoperable with other registries, no matter who – for instance a Trusted Third Party or a registry manager himself – pseudonimises the potentially identifying data before enabling the access.

It should be noted that the answers to the questions on available data formats and the possibility for combining data are unrelated. Therefore, the preparatory study neither supports nor undermines assumptions like "if a registry allows for combining data then it might facilitate this by offering a variety of data formats".

7.10 ICT-supported access

The process of requesting and permitting access to existing patient data is described in chapter 5 (the authorisation process) and 6 (good practices in handling data requests). The current section resumes the topic "access" from a more technological point of view.

In the application process itself the role of technology is limited to providing online information. A good practice here is the General Practice Research Database (GPRD) in the United Kingdom. "The GPRD is the world's largest computerised database of anonymised longitudinal medical records from primary care that is linked with other healthcare data." [64] The website describes the procedure for applying for data usage extensively and clearly. The information also includes the procedure for appealing for a process review, in case the request for data is turned down. An application form is provided as well.

In the questionnaire the respondents were asked how – possibly ICT-supported – access to the data has been arranged once permission to use them has been given. This yields a diverse picture. 17 registries provide on site access and 12 registries (also) deliver data on CD or DVD. Other options include internet access (7x), remote access via the CBS (Statistics Netherlands; 4x) and access via a trusted third party (1x).

Publication of research results, on the basis of existing data, is a relatively late step in the research process. Although registries make demands about (public) publications (see also chapter 6), in the website scan no formal requirements about referring to the data were found. The DANS research data archive illustrates how this might be organised. Persons to whom DANS makes available a dataset declare that they agree with the general conditions of use[65]. The first condition is clear acknowledgement of sources in any publication that is based on this particular dataset. Obligatory elements in the acknowledgement are the name(s) of the person(s) to whom the right of ownership to the dataset is registered, the title of the dataset, the data on which or the period in which the dataset was created, and a statement that the dataset originates from DANS. This way, a reference explicitly recognises the efforts of the data provider(s) and the registry – which could be an extra incentive for providing data. Furthermore, DANS demands that the acknowledgment contains a persistent identifier, which is a unique and permanent digital identification of the dataset involved. Persistent identification is an important instrument for sustainable access to digital objects. For instance, if datasets move to another registry in the course of time the persistent identifier will link to the actual location; should they

64 http://www.gprd.com/home/ In fact, the records are pseudonimised rather than anonymised.

65 http://www.dans.knaw.nl/en/content/data-archive/terms-and-conditions

be updated it will still refer to the original version as newer versions have their own persistent identifiers. This allows for reduplicating research. The concepts and technology behind persistent identification are largely domain-independent and recur in many research infrastructures[66].

7.11 Conclusions and recommendations

- In order to enhance data discovery and re-use of existing data, registries should provide rich metadata and make codebooks available online. Adherence to an existing metadata schema like DDI is certainly preferable.
- Among registry managers the awareness of the need for (re-)usable data and interoperability of registries is moderately high, given the responses: 60% of the registries support several data formats and a little more contain data that can be combined or extended, for that matter. However, this leaves room for improvement and therefore it is recommended to further analyse both possibilities (e.g. a stronger role for a Trusted Third Party?) and obstacles (e.g. are there limiting demands from funding parties?). Again, adherence to standards and coding systems is key to interoperability.
- In view of maximising Open Access to publicly funded research data it is recommended to promote that these data are managed in a sustainable way. In this context sustainable data formats and persistent identification should be paid attention to, preferably in conjunction with other research disciplines in order to benefit from generic infrastructural developments.

[66] http://ec.europa.eu/research/infrastructures/index_en.cfm

8 Summary and conclusions

In this chapter we summarise the main findings of the study and make recommendations. These are grouped according to the four research themes mentioned in chapter 1. We added a fifth theme, *Standards for data quality and access*, because this topic emerged from the analysis performed in the study.

8.1 Background

This year, 2012, the Netherlands Organisation for Health Research and Development (ZonMw) started a new programme called Rational Pharmacotherapy (RP). An important aim of this programme is to enhance adequate use of existing infrastructures and to promote good practices in setting up infrastructures for observational research. Currently there is a large data infrastructure in The Netherlands, holding many patient registries with reasonably well documented meta information (in general). However, these registries are heterogeneous and not all of them are accessible for (external) researchers. Moreover, it is not known, what is the need for new, additional data, both in terms of completely new registries and in terms of additional data collection within existing patient registries. Also, there is a need to know *how* governance of patient registries should be organised. Therefore, this study on patients' registries was performed by the National Institute for Public Health and the Environment (RIVM), Mondriaan, Data Archiving and Network Services (DANS), and NIVEL, Netherlands

8.2 Availability of patient registries

Many different patient registries include data on medicines or other data relevant to the RP programme. Therefore, there are potentially many useful data sources for the aims of the RP programme. However, there is little or no coordination and many registries work more or less in isolation. In addition, there is a rather strict division between primary and secondary care registries. Issues related to health and health care – life style (nutrition, exercise); long-term care; being able to work; participating in society – are important to relate to for instance medicine use (more about this in section 8.4) but such connections are currently not often made.

Much progress is expected in different types of RP research if there would be national coordination and standardisation (see section 8.6) as well as efforts to make data from primary and secondary care registries (more) comparable (or at least linkable). This is important, not only within the Netherlands but also for data outside the Netherlands, i.e. within the EU since for some purposes (for example in case of rare diseases and orphan drugs) combining registrations at the international level is important. As of now, it is not clear whether the available data are comparable within and across countries. This needs further investigation.

8.3 Access to existing registries

Although availability of patient registries is a condition sine qua non, availability alone is not sufficient: data also need to be accessible. First, data can only be used if patients give permission to use data or if data are anonimised. Second, many scientific research questions require rich data sets. As such, linking

is an important issue. At this moment, data of about 70% of the registries can already be linked to other registries. From the responses in the questionnaire, it can be concluded that those registries that are already linking data have accumulated substantial knowledge about the complex topic of linking data. Registries that cannot link their data to data from other registries, (usually) do not have the intention to do so in future. Third, the access policy regarding third parties to use the data is important for accessibility. Most registries have a request process. About 90% of the registries require the researcher to write a protocol describing the research question. Typically, this protocol is judged by the board of the register, or by a special advisory board. Usually a special advisory board is consulted in case of sensitive issues, for example, relating to privacy or research for commercial parties. When the board has approved a protocol for requesting data, the next issue is under what specific *conditions* the data will be available. Since publishing is very important, most of the registries want to be involved in producing an article in different ways: few require only peer review, they usually want to be involved as co-author. The whole process of requesting data, the assessment by the board or committee, and finally the delivery of data takes on average about five weeks.

The Netherlands has several good practices when it comes to accessibility of data. Procedures from these good practices should be adapted by ZonMw funded registries. ZonMw could refer applicants for grants for new registrations to these examples or make the procedures public in cooperation with the registries involved. Important elements are: a clear review procedure for data requests including independent parties (others than the institute that manages the database), applicants should provide a clear research protocol, costs for data requests should be only for handling the data and entirely transparant, requiring public publications, active involvement of the managing institute (also as an incentive for the original researchers to keep up data management). Examples of existing registries with such requirements are PRN and LINH/NIVEL zorgregistraties, as well as the British GPRD. These registries can easily be found on the internet, where they provide information on procedures, data that are available etc. Many registries are not really visible and transparant on the internet, while visibility is an important requirement for accessibility. A first step for patient registries is to register their information at Zorggegevens.nl and/or other websites where researchers can find information about what RP data are available in the Netherlands.

Another issue is data linking, which becomes increasingly important. We recommend as a first step towards safe linking to ask patients for permission to link their data to other datasets. In case of “opt-out”, sufficient variables should be collected to make linking possible. The linking itself can be done by specialised parties like a Trusted Third Party (section 8.6 contains more information on the technological side of access).

8.4 New patient registries and extension of existing patient registries for the RP programme

The Netherlands has a quite rich infrastructure of patient registries that fulfil important criteria for good practice, especially when it comes to automated routine registrations in representing general patient populations. Moreover, especially this last group of databases are highly accessible. However, to enhance quality of registration as well as quality of care some registries could provide more incentives to parties that deliver data. For secondary care, especially databases that include both data on health/indications and on drug utilisation/prescriptions seem to be scarce. There is, however, a wide variety of disease-specific patient registries, which are often a combination of a cohort and a registration

database: data of a cohort of patients are distracted from databases or from reports of professionals. However, these registries usually are not linked to other databases containing other information of the patient (for example utilisation of health care derived from automated primary care databases). The same holds for cohort studies that collect many data on patients through questionnaires and clinical measurements but lack connections to other databases. The same holds for databases measuring adherence on a detailed level. To answer research questions relevant to the RP programme improvements in the infrastructure seem necessary especially when it comes to the combination of automated routine registration with in-depth data from patients and to patient registries in secondary care combining complete information on both pharmaceutical treatment and indications.

We conclude that primary care in the Netherlands has a variety of automated patient registrations. The infrastructure in primary care can be improved by adding and integrating data. For example, there is a need for cohorts preferably nested within existing primary care databases in order to be able to fully answer specific questions relevant for the RP programme. Also the other way around seems to be true: existing cohort studies could better use possibilities to link with automated routine registrations. Parties who create new databases should consider linking their data to other relevant data, and should think about how they want to organize such data linking. Compared to primary care, the infrastructure for linking data to secondary care data is still in its infancy. As such, there is a need for secondary care automated routine databases that combine information about prescriptions and other clinical information (indications) for those diseases that are mainly treated in secondary care (such as rare diseases). Existing secondary care databases and relevant cohorts often do not contain information needed (for example no information on medicines) for linking data, and it is not always transparent how data are collected and can be used by other researchers. Therefore, we recommend running pilots in different settings to learn where potential problems lie and how essential elements can be organized in secondary care.

8.5 Governance in patient registries

Essentially, governance and accreditation of registries boil down to “who is permitted to do what with medical data?”. The responses to the questionnaire show quite some variation in the balance between ownership, legal responsibility, and funding of the registries’ data (i.e. between stakeholders). This means that in the Dutch situation various governance models exist (implicitly or explicitly). Furthermore, the influence of the so-called “meta initiatives” on existing registries has been studied. These meta initiatives can be seen as a portal to the underlying registries. Most meta initiatives have the topic ‘access to data’ clearly on their radar screen. They are aware that optimal access can advance science and innovation. Also among registry managers the awareness of the need for re-usable data and interoperability of registries is moderately high. However, there is room for improvement, therefore it is recommended to further analyse both possibilities (e.g. a stronger role for a Trusted Third Party?), and obstacles (e.g. are there limiting demands from funding parties?).

We also recommend that the RP programme should establish a register of the existing or (at least) of the new data collections of the research data files generated in this programme. In addition, the RP programme should consider the formulation of a code that would elaborate on the practice of handling and using medical data as suggested by the Health Council. Moreover, the RP programme needs an “independent data broker”: a central service that stimulates data sharing and could be connected to a trusted digital repository. Finally, in view of maximising Open Access to publicly funded research data it

is recommended to promote that these data are managed in a sustainable way. Therefore, attention should be paid to the further development of certification for trusted digital repositories and to standardisation of Trusted Third Parties.

The registries participating in this study are predominantly funded from public means. Nearly 75% offer incentives for providing data, which mainly consist in the exchange of (benchmarking) information. Nearly 60% of the registries have a pricing system, based on various factors such as costs for data management. More harmonisation of pricing policies among the registries is recommended, as well as more information at the registries' websites. Given the small share of private funding among the registries no conclusions can be drawn regarding a relation between the funding body or bodies on the one hand and whether registries charge a fee on the other hand. ZonMw will offer registries funding for setting up patient registries with the idea that once set up, these registries should be able to generate the money needed to continue the registry. As such, it is important for patient registries to develop a business model right from the start. These business models were beyond the scope of our project. More insight into current practice requires a new study.

8.6 Standards for data quality and access

Data quality can be increased by the use of standardised coding systems. Whenever possible (inter)national standards should be used. Examples are ICPC codes, ATC codes and the use of the DDD system. In addition, in order to enhance data discovery and thereby re-use of existing data, all registries should provide rich metadata and make codebooks online available; in the current study 90% of the registries already have a codebook available for researchers. Adherence to an existing metadata schema like DDI is certainly preferable, following the example from the British Medical Research Council. In this context, sustainable data formats and persistent identification should also be paid attention to. Preferably in conjunction with other research disciplines in order to benefit from generic infrastructural developments.

Not only for quality enhancement and for data discovery standards are important, but also for the interoperability of registries. Meta initiatives can play a role in the development of new standards on accessibility to datasets and/or the acceptance thereof. In addition, the expertise of these meta initiatives as well as the expertise of some complex patient registries can be used for complex topics like privacy and technical issues like linking data on patient level. A start can be made by harmonising the conditions that parties such as BBMRI, Mondriaan, NWO, ZonMw, the Ministry of Health and perhaps in the future also Topsector Life Sciences & Health set to financing projects, aligned with the above recommendations for governance. These parties could strive for a guideline in the field.

Literature & web links

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.

R.W. de Bruin, M. de Cock Buning and A. Ringnalda (CIER), Data in Publiek-Private Projecten: Juridische aspecten SURFfoundation 2011.

Gijsen R, Jochemsen H, Van Dijk L, Caspers P. Frequency of ill-founded off-label prescribing in Dutch general practice. Pharmacoepidemiology and Drug Safety, 2009, 18 (1), 84-91.

Leendertse AJ, Egberts AC, Stoker LJ, van den Bemt PM; HARM Study Group. Frequency of and risk factors for preventable medication-related hospital admissions in the Netherlands. Arch Intern Med. 2008 Sep 22;168(17):1890-6.

Van Dijk L, Heerdink ER, Somai D, Van Dulmen S, Sluijs E, De Ridder DT, Griens AFGM, Bensing JM. Patient risk profiles and practice variation in nonadherence to antidepressants, antihypertensives and oral hypoglycemics. BMC Health Services Research, 2007

ZonMw. Signalement Goed Gebruik van Geneesmiddelen. The Hague, 2009.

ZonMw. Signalement Goed Gebruik van Geneesmiddelen. The Hague, 2009 and ZonMw. Verdieping Goed Gebruik Geneesmiddelen. The Hague, 2010.

ZonMw. Proposal ZonMw Programme Rational Pharmacotherapy. Optimising effective, safe and efficient use of medicines in health care. The Hague, ZonMw, 2011.

Websites mentioned in the main body of the text:

http://datasealofapproval.org/
http://e-overheid.nl/onderwerpen/e-overheid
http://uzovi.vektis.nl/
www.agbcode.nl/MainPage/default.aspx
www.chess.su.se/
www.connectingforhealth.nhs.uk/systemsandservices/data/uktc/readcodes
www.crl.edu
www.datagateway.mrc.ac.uk/node/14
www.ddialliance.org
www.dutchhospitaldata.nl
www.federa.org/gedragscodes-codes-conduct-en
www.gprd.com/home/
www.knmp.nl/producten-diensten/g-standaard/farmaceutische-gegevens/kenmerken-gpk-prk-en-hpk
www.medlaw.nl/wp-content/uploads/patient-data-for-health-research.pdf
www.mrc.ac.uk/Utilities/Documentrecord/index.htm?d=MRC008302

www.perinatreg.nl/gegevensaanvragen?noCache=76;1327421553
www.stockholmbirthcohort.su.se
www.therapietrouwmonitor.nl
www.uziregister.nl
www.whocc.no/atc/structure_and_principles/
www.who.int/classifications/icd/en/
www.who.int/classifications/icd/adaptations/icpc2/en/index.html
www.who.int/medicines/areas/rational_use/en/
www.zonmw.nl/en/programmes/pharmacotherapy/general-information/
www.zorggegevens.nl

Abbreviations

AGB	:	General Data Administration Health Professionals
ATC	:	Anatomic Therapeutic Chemical
BBMRI	:	Biobanking and Biomolecular Resources Research Infrastructure
BSN	:	Social service number (in the Netherlands)
CHESS	:	Centre for Health Equity Studies in Sweden
COREON	:	Dutch federation of Biomedical Scientific Societies
CTV-3	:	Clinical Terms Version 3
CVZ	:	Health Care Insurance Board
DANS	:	Data Archiving and Network Services
DDD	:	Daily Defined Doses
DDI	:	Data Documentation Initiative
DSA	:	Data Seal of Approval
EU	:	European Union
FEDERA/FMWV	:	Federation of Medical Scientific Associations
GIP	:	Drug Information System of the Health Care Insurance Board
GP	:	General Practitioner
GPK	:	Generic Product Code
GPRD	:	General Practice Research Database
IADB	:	Inter Action Data Base
ICD-10	:	International Statistical Classification of Diseases (10^{th} version)
ICPC	:	International Classification of Primary Care
IEEE	:	Institute of Electrical and Electronic Engineers
IPCI	:	Interdisciplinary Processing of Clinical Information
KNAW	:	Royal Netherlands Academy of Arts and Sciences
LBZ	:	National Hospital Care Basic Administration
LINH	:	Netherlands Information Network of General Practices
MRC	:	Medical Research Council
NAW	:	Name, address, domicile (woonplaats)
Nefarma	:	Netherlands federation of Innovative Pharmacy industry
NFU	:	Netherlands Federation of University Medical Centres
NIVEL	:	Netherlands institute for health services research
NMA	:	Netherlands Competition Authority
NWO	:	Netherlands Organisation for Scientific Research
PIAMA	:	Prevention and Incidence of Asthma and Mite Allergy (project)
PRN	:	The Netherlands Perinatal Registry
RCT	:	Randomized Clinical Trial
RIVM	:	National institute for Public Health and the Environment
RP	:	Rational Pharmacotherapy
SFK	:	Foundation for Pharmaceutical Statistics
STW	:	Technology Foundation STW
SURFfoundation	:	Initiator of innovation in higher education and research
TDR	:	Trusted digital repositories
TTP	:	Trusted Third Party

UZI	:	Unique Health Care Professional Identification
UZOVI	:	Unique Health Insurance Identification
VSNU	:	Association of Universities in the Netherlands
VWS	:	Ministry of Health
WBP	:	Data Protection Act
WGBO	:	Medical Treatment Act
ZonMw	:	Netherlands Organisation for Health Research and Development

Annex 1: Questionnaire sent to registries

Inleiding op de vragenlijst

In 2012 start ZonMw een programma met de naam Goed Gebruik Geneesmiddelen (GGG). Dit programma heeft onder andere tot doel het optimaliseren van effectief, veilig en doelmatig gebruik van geneesmiddelen. Naast het systematisch verzamelen, verspreiden en implementeren van kennis is het opbouwen van een infrastructuur van belang, waarmee relevante vragen op het gebied van rationele farmacotherapie adequaat en onafhankelijk kunnen worden beantwoord.

Voordat dit ZonMw-programma van start gaat, wordt een aantal voorstudies uitgevoerd, waaronder een voorstudie over de benodigde infrastructuur voor observationeel onderzoek. In dit kader ontvangt u deze vragenlijst. De voorstudie "Registers GGG" richt zich op beschikbare data die geschikt zijn voor observationeel onderzoek, op lacunes in de beschikbaarheid en op de randvoorwaarden voor effectiever gebruik ervan.

Het RIVM, DANS, Mondriaan en het NIVEL voeren deze voorstudie uit met subsidie van ZonMw. Wij hopen op uw deelname. Met de gegevens ontstaat een overzicht van beschikbare gegevensbronnen (hierna: gegevensverzamelingen) in Nederland en van de mogelijkheden die deze gegevensverzamelingen bieden voor onderzoek al dan niet door koppeling aan andere gegevensverzamelingen. Ook kunnen hierdoor lacunes in beschikbare gegevensverzamelingen worden geïdentificeerd: op welke terreinen voldoen de bestaande gegevensverzamelingen (in hun huidige vorm) nog niet. Tot slot moet het onderzoek inzicht geven in de vraag hoe toegang tot de gegevens geregeld kan worden. Op basis van deze vragen krijgen zowel ZonMw als potentiële aanvragers van het programma GGG inzicht in:

1) waar nog nieuwe gegevensverzamelingen opgezet kunnen worden of waar men kan volstaan met het slim combineren van bestaande gegevens;
2) hoe nieuwe gegevensverzamelingen opgezet kunnen worden, aan welke voorwaarden moet worden voldaan en hoe gebruik van deze data goed geregeld kan worden.

ZonMw zal de resultaten van het onderzoek gebruiken voor de eerste oproep voor projectvoorstellen binnen het programma GGG die naar verwachting in maart 2012 zal uitgaan

Instructie voor het invullen van deze vragenlijst:
Vult u a.u.b. de vragenlijst zo compleet mogelijk in. Indien u een vraag niet kunt of wenst te beantwoorden,wilt u dit dan vermelden? De vragen hebben deels voorgestructureerde antwoordcategorieën. Wanneer u meer dan één antwoord kunt aankruisen, staat dit expliciet vermeld.

"Gegevensverzameling": alle mogelijke databronnen, inclusief biobanken
"Subject": patiënt of gezond persoon

Als u vragen heeft, kunt u contact opnemen met:
dr. ir. Liset van Dijk (L.vanDijk@nivel.nl)

Algemeen:

1. Naam gegevensverzameling:_____

2. Voor welk doel worden of zijn de gegevens of lichaamsmaterialen verzameld:_____

3. Contactpersoon:
   Naam: _____
   Telefoonnummer: _____
   E-mailadres: _____

Wijze van gegevensverzameling, gegevenskenmerken en beheer

4. Over wie worden gegevens verzameld (bijvoorbeeld patiënten of een bepaalde leeftijdsgroep). En gebeurt dit algemeen of bijvoorbeeld indicatie-specifiek? Indien van toepassing, alstublieft noteren welke indicaties zijn opgenomen:

   _____

5. Zijn er leeftijdsgrenzen? Zo ja, welke? _____ (in jaren)

6. Over welke periode zijn data aanwezig?
   Van _____ t/m _____ (in jaren)

7. Is er een gepland einde voor de dataverzamelingsactiviteiten?

   Zo ja, wanneer is dat?: _____ (jaartal)

8. Over welk type(n) organisatie(s) worden de data verzameld, en hoeveel in 2010 (of in laatste jaar van dataverzameling)? (meer antwoorden mogelijk)
   - ☐ Eerstelijnszorginstelling(en), nl.:
     - ☐ huisartspraktijken, aantal:
     - ☐ paramedische praktijken, aantal:
     - ☐ apotheken, aantal:
     - ☐ eerstelijnspsycholoog, aantal:
     - ☐ anders, namelijk_____
   - ☐ Tweedelijnszorginstelling(en), nl.:
     - ☐ academische ziekenhuizen, aantal:
     - ☐ overige ziekenhuizen, aantal:
     - ☐ anders, namelijk _____
     - ☐ thuiszorgorganisaties
     - ☐ verpleeghuizen
     - ☐ anders, nl.: _____

9. Is er een procedure c.q. protocol dat gebruikt wordt tijdens de dataverzameling?
   - ☐ nee
   - ☐ ja

10. Moet u zelf de gegevens invoeren in de database of gebeurt dit automatisch (vergelijk bijvoorbeeld het invoeren van vragenlijsten met het inlezen van gegevens uit zorgverlenersregistraties)?
    - ☐ een combinatie van zelf invoeren en automatische invoer
    - ☐ uitsluitend zelf invoeren
    - ☐ uitsluitend automatische invoer
    - ☐ anders, namelijk_____

11. Zijn er incentives voor de partijen die de data leveren?
    - ☐ nee → naar vraag 13
    - ☐ ja

12. Welke incentives zijn dat? (meer antwoorden mogelijk)
    - ☐ spiegelrapportage op papier
    - ☐ spiegelrapportage via internet
    - ☐ financiële vergoeding
    - ☐ periodieke bijeenkomsten
    - ☐ geschenken
    - ☐ anders, namelijk _____

13. Is er een procedure of protocol om de data te controleren voor de data uitgeleverd worden?
    - ☐ nee
    - ☐ ja

14. Hoe vaak komt een geüpdate versie van de database beschikbaar?
    - ☐ op continue basis
    - ☐ regulier, eens per ……
    - ☐ onregelmatig, maar gemiddeld met een tussenpoos van ….
    - ☐ anders, namelijk

15. Hoe heeft u de back-up van de data geregeld?

    _____

    _____

    _____

Privacy aspecten

16. Welke maatregelen heeft u genomen om de privacy van de patiënten te beschermen (bij uitlevering van data aan derden)? (meer antwoorden mogelijk)
    - ☐ (juridisch) een privacyreglement
    - ☐ (juridisch) aanmelding bij het CBP
    - ☐ (technisch) pseudonimisering van identificerende gegevens
    - ☐ (technisch) Identificerende informatie verwijderen
    - ☐ Anders, nl.: _____

17. Welke niveau van consent / toestemming moeten patiënten geven voordat hun gegevens gebruikt worden:
    - ☐ Informed Consent: _____
    - ☐ Geen bezwaar/opt-out regeling
    - ☐ Geen → naar vraag 19
    - ☐ Anders, namelijk → naar vraag 19

18. Welke vormen van gebruik zijn dan gedekt in deze toestemming? (meer antwoorden mogelijk)
    - ☐ gebruik van de medische gegevens
    - ☐ koppeling met (andere) medische registraties
    - ☐ alle onderzoeksonderwerpen
    - ☐ specifieke onderzoeksonderwerpen, namelijk

19. Hoeveel unieke patiënten bevat de gegevensverzameling in het laatste meetjaar?

    _____ (aantal patiënten)

20. Hoeveel unieke patiënten zitten er bij benadering in de gegevensverzameling over de gehele periode dat deze loopt/ liep?

    _____ (aantal patiënten)

21. In welk format zijn de data beschikbaar? (meer antwoorden mogelijk)
    - ☐ Spreadsheets – bijvoorbeeld Excel
    - ☐ Statistisch bestanden – bijvoorbeeld SPSS
    - ☐ Databases – bijvoorbeeld Access
    - ☐ Mark-up-tekst
    - ☐ Tekstbestanden
    - ☐ Anders, nl: _____

22. Welk percentage van de datavelden is vrije tekst? _____ _____ _____

Koppelingen

23. Kunnen de gegevens op patiëntniveau gecombineerd worden met gegevens uit andere gegevensverzamelingen?
    - ☐ Nee: → naar vraag 25
    - ☐ Ja

24. Welke patiënt-identificerende variabelen gebruikt u/kunt u gebruiken om de gegevens te koppelen? (meer antwoorden mogelijk)
    - ☐ BSN
    - ☐ NAW-gegevens, te weten …. (NB: maak onderscheid tussen 4- en 6-cijferige PC)
    - ☐ Geboortedatum
    - ☐ Geslacht
    - ☐ Andere patiëntgegevens, te weten: _____
    - ☐ Andere niet-patiëntgegevens, te weten:

      _____

25. Welk type gegevens zijn aanwezig? (meer antwoorden mogelijk)
    - ☐ Geneesmiddelgebruik gegevens
    - ☐ Klinische gegevens
    - ☐ Overige zorgconsumptiegegevens (zoals verwijzingen en huisartscontacten)
    - ☐ Biometrie
    - ☐ Demografie
    - ☐ Life style
    - ☐ Genetische informatie
    - ☐ Biosamples, namelijk: _____
    - ☐ Vragenlijsten, namelijk:_____ _____

☐ Anders, nl.: _____

26. Hoeveel internationale wetenschappelijke publicaties of rapporten zijn bij benadering verschenen waarin deze database gebruikt is?

_____ (aantal)

27. Houdt u (of de bronorganisatie) bij het verzamelen van de data rekening met mogelijke toekomstige koppelingen aan andere databronnen bijvoorbeeld op technisch of juridisch gebied?
☐ nee
☐ ja -> naar vraag 29

28. Zo ja, hoe doet u dat?

_____

Beoordeling en afhandeling van een gegevensaanvraag

29. Is er een beoordelingsprocedure of –protocol voor onderzoekers of instanties die met de gegevens willen werken?
☐ nee -> naar vraag 34
☐ ja

30. Kan men deze procedure openbaar online vinden?
☐ nee
☐ ja

31. Zo ja, waar? Zo nee, hoe ziet deze procedure eruit?

_____

32. Welke organisaties, welk bestuurlijk orgaan of welke personen beoordelen externe aanvragen?

_____

33. Zijn er soorten organisaties of beroepsgroepen die op voorhand geen toestemming krijgen? Indien ja, welke (bijvoorbeeld for profit organisaties of zorgverzekeraars):
☐ nee
☐ ja, namelijk: _____

34. Welke voorwaarden stelt u aan het gebruik van de data (meer antwoorden mogelijk)
    - ☐ Geen marketingonderzoek
    - ☐ Eis van openbare publicatie
    - ☐ Voorinzage in publicatie
    - ☐ Betrokkenheid als auteur bij publicaties
    - ☐ Bepaalde *codes of conduct*, namelijk: _____
    - ☐ Anders, nl.: ____

35. Rekent u een prijs voor het gebruik van de data?
    - ☐ Nee: → naar vraag 40
    - ☐ Ja

36. Hoe berekent u de tarieven? (bijvoorbeeld op basis van de te verrichten handelingen)

    _____

37. Hanteert u verschillende tarieven voor verschillende typen organisaties?
    - ☐ Nee: → naar vraag 39
    - ☐ Ja, namelijk (toelichting) _____

38. Wat zijn de minimale en maximale tarieven?

    _____

39. Kan men informatie over tarieven online vinden?
    - ☐ Nee
    - ☐ Ja

40. Hoe is de toegang tot de gegevens geregeld? (meer antwoorden mogelijk)
    - ☐ Via CBS via remote access
    - ☐ Online via internet
    - ☐ Downloaden uit een online systeem
    - ☐ Toesturen op CD/DVD/andere gegevensdrager
    - ☐ On site (ter plekke bij de registratiehouder)
    - ☐ Anders, nl:_____

41. Hoe lang duurt het aanvraagproces gemiddeld tussen het moment van aanvragen en het aanleveren van de dataset?

    _____ (tijdsaanduiding)

Governance

42. Welke organisatie(s) is / zijn eigenaar van de gegevens?

_____

43. Welke organisatie(s) is/ zijn juridisch verantwoordelijk voor het verzamelen, beheer, opslag en gebruik van de gegevens?

_____

44. Welke organisatie(s) financiert / financieren de verzameling en beheer en opslag van de gegevens?

_____

45. Indien uw gegevens aangevraagd kunnen worden door externe organisaties, door welke organisatie worden de gegevens uitgegeven? (dit kan uw eigen organisatie zijn maar ook een andere organisatie)
    - ☐ Door één centrale organisatie, namelijk: _____
    - ☐ Door meerdere bronorganisaties, afhankelijk van de aangevraagde data, nl.:

---

46. Is er een codeboek / data dictionary beschikbaar met omschrijvingen van de aanwezige variabelen?
    - ☐ Ja.
        - Bent u bereid deze voor andere onderzoekers toegankelijk te maken?
            - ☐ Nee
            - ☐ Ja
        - Bent u bereid deze publiekelijk toegankelijk te maken?
            - ☐ Nee
            - ☐ Ja
    - ☐ Nee.
        - Bent u bereid of van plan dit te maken?
            - ☐ Nee -> naar vraag 47
            - ☐ Ja
        - Bent u bereid deze voor andere onderzoekers toegankelijk te maken?
            - ☐ Nee -> naar vraag 47
            - ☐ Ja
        - Bent u bereid deze publiekelijk toegankelijk te maken?
            - ☐ Nee
            - ☐ Ja

Toekomstperspectief

47. Hoe kunnen data van Nederlands gegevensverzamelingen vaker of beter worden hergebruikt? Wat is hiervoor nodig?

    _____

48. Hoe worden data van Nederlandse gegevensverzamelingen beter toegankelijk?

    _____

49. Heeft u nog aanvullingen?

    _____

Dank voor uw medewerking! U ontvangt van de onderzoekers een samenvatting van de resultaten van deze vragenlijst. Indien gewenst kunt u hieronder aangeven of u de volledige rapportage van het onderzoek wilt ontvangen (digitaal).

☐ Ik wil de volledige rapportage ontvangen.

Afgeleid van bovenstaande vragenlijst: Vragen gesprek best practice

Inleiding

Dit jaar start ZonMw een programma met de naam Goed Gebruik Geneesmiddelen (GGG). Dit programma heeft onder andere tot doel het optimaliseren van effectief, veilig en doelmatig gebruik van geneesmiddelen. Naast het systematisch verzamelen, verspreiden en implementeren van kennis is het opbouwen van een infrastructuur van belang, waarmee relevante vragen op het gebied van rationele farmacotherapie adequaat en onafhankelijk kunnen worden beantwoord.

Voordat dit ZonMw-programma van start gaat, is een aantal voorstudies uitgevoerd, waaronder een voorstudie over de benodigde infrastructuur voor observationeel onderzoek. De voorstudie “Registers GGG” richt zich op beschikbare data die geschikt zijn voor observationeel onderzoek, op lacunes in de beschikbaarheid en op de randvoorwaarden voor effectiever gebruik ervan. Het RIVM, DANS, Mondriaan en het NIVEL voeren deze voorstudie uit met subsidie van ZonMw. De studie is bijna afgerond. Wel willen we met een aantal verschillende stakeholders kijken naar de vraag : wat zijn “good practices” als het gaat om beschikbaar stellen van gegevens. Aan welke eisen moeten gegevensverzamelingen voldoen willen zij geld van ZonMW ontvangen? En hoe kunnen nieuwe gegevensverzamelingen opgezet worden, aan welke voorwaarden moet worden voldaan en hoe kan het gebruik van deze data goed geregeld worden? We willen een aantal punten met u doornemen.

Toegang tot het register voor derden

Als u een gebruiker was: hoe zou u dan graag zien dat de toegang tot een register geregeld is?
En wat vindt u reële eisen die een databeheerder kan stellen aan aanvragen?

Op welke gronden zouden registers gegevensaanvragen voor het register moeten beoordelen?
En wie zouden deze beoordeling moeten doen (inhoudelijk / juridisch
Welke voorwaarden mogen beheerders van gegevensverzamelingen volgens u stellen aan het gebruik van gegevens (denk aan co-auteurschap, voorinzage, gebruikslicentie, kosten)

Mogen beheerders bepaalde partijen weigeren gebruik te maken van de data als deze publiek gefinancierd zijn? Kunt u dit toelichten?

Zichtbaarheid en vindbaarheid

Welke maatregelen kunnen registers nemen om de vindbaarheid van het register te bevorderen? Denk aan opzet van een website, aanmelding bij Zorggegevens.nl en aansluiting bij de Mondriaancatalogus.
In hoeverre moeten dataverzamelingsprotocollen transparant zijn en toegankelijk gemaakt worden?
In hoeverre moeten datadictionary of codeboek aangemaakt worden met de beschrijving van alle beschikbare variabelen, hun betekenis en formaat (datum, getal, tekst, etc.) en toegankelijk gemaakt worden?

Techniek en ethiek

Aan welke eisen moet een register volgens u voldoen als het gaat om ethische aspecten? Denk hierbij aan privacy van deelnemers, voorwaarden waaronder gegevens gekoppeld mogen worden, toestemming die gevraagd wordt voor gebruik gegevens.

Welke maatregelen kunnen partijen nemen om de herbruikbaarheid van het register te optimaliseren? Denk aan (minimaliseren van het aantal vrije tekstvelden; gebruik van coderingstandaarden (ICD10, ATC, ICPC, Snomed, etc); Defined Daily Dose als getal registreren; labbepalingen).
Indien gekoppeld wordt aan welke juridische en technische eisen moet dan voldaan worden?
Hoe moet veiligheid gewaarborgd worden (denk aan eisen bij egbruik van online verbindingen, maar ook aan opslag van de data)?
Hoe moet archivering en opslag plaatsvinden?

Incentives

Welke incentives voor zorgverleners of anderen die data leveren aan een register zijn voglens u belangrijk (denk aan financiële vergoeding, spiegelrapportage etc)?
En welke incentives zijn belangrijk voor leveranciers van data (publicaties, geld om eigen onderzoek mee te financieren e.d.)

Annex 2: Search terms 'Zorggegevens.nl'

Results for 'bijwerkingen' (side effects)

Name register (8 results)	Abbreviation	Year	Administrator
Drug Information System of the Health Care Insurance Board (Genees- en hulpmiddelen Informatie Project)	GIP (Geneesmiddelen)	2008	CVZ
Drug Information System of the Health Care Insurance Board (Genees- en hulpmiddelen Informatie Project)	GIP (Hulpmiddelen)	2008	CVZ
Integrated Primary Care Information	IPCI	2009	Erasmus MC
The Netherlands Pharmacovigilance Centre (Landelijke Registratie en Evaluatie van (medicijn) Bijwerkingen)	LAREB	2008	Nederlands Bijwerkingen-centrum Lareb
Monitor Heamo/Tissue vigilance (Monitor Weefselvigilantie)	TRIP	2009	
Dutch Paediatric Surveillance Unit (Nederlands Signalerings Centrum Kindergeneeskunde)	NSCK	2009	TNO
HIV Monitoring Foundation (Stichting HIV Monitoring)	SHM	2011	Stichting HIV Monitoring
Transfusion Reactions in Patients (Transfusie Reacties in Patiënten)	TRIP	2009	Stichting TRIP

Results for 'geneesmiddelen' (drugs)	Abbreviation	Year	Administrator
Name register (15 results)			
Driving under the influence of drugs, alcohol and medicines	DRUID		Stichting Wetenschappelijk Onderzoek Verkeersveiligheid
European Registration Of Congenital Anomalies Twins	EUROCAT	2008	UMC Groningen
Pharmacy Information system (Farmacie Informatiesysteem)	FIS	2008	Vektis CV.
G-standard Pharmacy Products (G-Standaard apotheekproducten)		2007	Z-Index BV
Drug Information System of the Health Care Insurance Board (Genees- en hulpmiddelen Informatie Project)	GIP (Geneesmiddelen)	2008	CVZ
Drug Information System of the Health Care Insurance Board (Genees- en hulpmiddelen Informatie Project)	GIP (Hulpmiddelen)	2008	CVZ
PricesOfDrugs (Geneesmiddelenprijzen)	WGP	2006	Farmatec/aCIBG
Integrated Primary Care Information	IPCI	2009	Erasmus MC
Quality of Pharmacotherapy Consultation (Kwaliteit van Farmacotherapieoverleg)	FTO		DGV, IVM
Netherlands Pharmacovigilance Centre (Landelijke Registratie en Evaluatie van (medicijn) Bijwerkingen)	LAREB	2008	Nederlands Bijwerkingen-centrum Lareb
National Poisons Information Centre (Nationaal Vergiftigingen Informatie Centrum Monitor)	NVIC Monitor	2011	Nationaal Vergiftigingen Informatie Centrum, UMC Utrecht
Dutch Paediatric Surveillance Unit (Nederlands Signalerings Centrum Kindergeneeskunde)	NSCK	2009	TNO
PHARMO database (PHARMO-databank gebruik geneesmiddelen)	PHARMO	2009	PHARMO Institute
Pharmacy Register (Register personen artsenijbereidkunst (apothekers etc.))	Register-WOG	2007	IGZ
Foundation for Pharmaceutical Statistics (Stichting Farmaceutische Kengetallen)	SFK	2010	SFK

Results for 'apothekers' (pharmacists)

Name register (9 results)	Abbreviation	Year	Administrator
WHO European Health for All Database	WHO-HFA	2009	World Health Organisation
Foundation for Pharmaceutical Statistics (Stichting Farmaceutische Kengetallen)	SFK	2010	SFK
Pharmacy Register (Register personen artsenijbereidkunst (apothekers etc.))	Register-WOG	2007	IGZ
The Netherlands Pharmacovigilance Centre (Landelijke Registratie en Evaluatie van (medicijn) Bijwerkingen)	LAREB	2008	Nederlands Bijwerkingen-centrum Lareb
Quality of Pharmacotherapy Consultation (Kwaliteit van Farmacotherapieoverleg)	FTO		DGV, IVM
Generation R		2008	Erasmus MC afdeling Generation R
Drug Information System of the Health Care Insurance Board (Genees- en hulpmiddelen Informatie Project)	GIP (Genees-middelen)	2008	CVZ
Drug Information System of the Health Care Insurance Board (Genees- en hulpmiddelen Informatie Project)	GIP (Hulp-middelen)	2008	CVZ
G-standard Pharmacy Products (G-Standaard apotheekproducten)		2007	Z-Index BV
European Registration Of Congenital Anomalies Twins	EUROCAT	2008	UMC Groningen
Rotterdam Study (Erasmus Rotterdam GezondheidsOnderzoek (of Rotterdam Study))	ERGO (of RS)	2011	Erasmus MC
BIG register (Beroepen in de Individuele Gezondheidszorg)	BIG-register	2008	RIBIZ
Pharmacy Numbers CBS (Apotheken; aantal bedrijven, werknemers, exploitatie en investeringen)	Apotheken; kerncijfers	2008	CBS

Annex 3: Registries found through PubMed search

Nr	Register	Website with background information
1	Arthritis and Biologicals in Children Register = ABC-register	https://www.abc-register.nl/startpagina/home/home.php
2	Netherlands Twin Register biobank = Nederlands Tweelingen Register = NTR	http://www.tweelingenregister.org/
3	Netherlands Tuberculosis Register = Nederlands Tuberculose Register = NTR	http://www.kncvtbc.nl/
4	Psychiatric Case Registry Midden Nederland = PCR-MN	http://www.juliuscentrum.nl/julius/Research/Researchprojects/Cohorts/tabid/890/Default.aspx
5	Dutch Paediatric Surveillance Unit = Nederlands Signalerings Centrum Kindergeneeskunde	http://www.nvk.nl/Onderzoek/NSCK.aspx
6	Standard Diagnosis Register of Rheumatic Diseases	www.tno.nl
7	Hospital Discharge Register = Landelijke Medische Registratie	www.zorggegevens.nl
8	Nijmegen Continuous Morbidity Registration	www.zorggegevens.nl
9	InterActie DataBank = IADB	http://iadb.nl/
10	Groningen Initiative to ANalyse Type 2 diabetes Treatment = GIANTT	http://www.giantt.nl/
11	Registratie Netwerk Groningen = RNG	http://rng.med.rug.nl/websiteRNG/
12	Maastricht Mental Health case register = Psychiatrisch Casus Register = PCR	www.zorggegevens.nl
13	HNU (Huisartsen Netwerk Utrecht)	www.zorggegevens.nl
14	Leidsche Rijn Gezondheidsproject = LRGP	www.zorggegevens.nl
15	Amsterdam Cohort Studies	www.zorggegevens.nl
16	Prevention and Incidence of Asthma and Mite Allergy = PIAMA	http://piama.iras.uu.nl/en/
17	Doetinchem cohort	www.zorggegevens.nl
18	IPCI database	www.zorggegevens.nl
19	Rotterdam study	http://www.epib.nl/research/ergo.htm
20	Nijmegen Inception Cohort of Early Rheumatoid Arthritis	Part of DREAM
21	Dutch Rheumatoid Arthritis Monitoring = DREAM	http://www.dreamregistry.nl/
22	Longitudinal Aging Study Amsterdam = LASA	www.zorggegevens.nl
23	national registers of midwives and gynaecologists = PRN	www.zorggegevens.nl
24	National Neonatology Register – LNR = PRN	www.zorggegevens.nl
25	The Netherlands Perinatal Registry = PRN	www.zorggegevens.nl and http://www.perinatreg.nl/home_english

Annex 4 Literature search chapter 4

Exclusion:
- general practice registries
- case registers

A. Pubmed, 11-11-2011 en 29-11-2011
- register AND Netherlands AND medicine NOT cochrane NOT trial register NOT veterinary (period 1996-2011): 68
- cohort AND Netherlands AND medicine NOT cochrane NOT trial register NOT veterinary (period 1996-2011): 1308 (te veel hits)
- cohort AND Netherlands AND drugs NOT cochrane NOT trial register NOT veterinary (period 1996-2011): 403 (te veel hits)
- prospective cohort AND Netherlands AND drugs NOT cochrane NOT trial register NOT veterinary (period 1996-2011): 176

Literatuur:
a. Milder IE, Klungel OH, Mantel-Teeuwisse AK, Verschuren WM, Bemelmans WJ. Relation between body mass index, physical inactivity and use of prescription drugs: the **Doetinchem Cohort Study**. Int J Obes (Lond). 2010 Jun;34(6):1060-9. Epub 2010 Feb 2.

b. Willemen MJ, Mantel-Teeuwisse AK, Straus SM, Leufkens HG, Egberts AC, Sturkenboom MC. Cardiovascular and psychiatric risk profile and patterns of use in patients starting anti-obesity drugs. Pharmacoepidemiol Drug Saf. 2009 Jul;18(7):631-8.
A population-based cohort study was conducted in the **IPCI database** (1995-2007)

c. Haag MD, Hofman A, Koudstaal PJ, Stricker BH, Breteler MM. Statins are associated with a reduced risk of Alzheimer disease regardless of lipophilicity. The Rotterdam Study. J Neurol Neurosurg Psychiatry. 2009 Jan;80(1):13-7. Epub 2008 Oct 17.
6992 participants of the prospective, population-based **Rotterdam Study** (Ommoord) were followed, from baseline (1990-1993) until January 2005 for incident AD.

d. Kolesnyk I, Noordzij M, Dekker FW, Boeschoten EW, Krediet RT. A positive effect of AII inhibitors on peritoneal membrane function in long-term PD patients. Nephrol Dial Transplant. 2009 Jan;24(1):272-7. Epub 2008 Jul 30.
We analysed data from 217 long-term CAPD patients, participating in the **Netherlands Cooperative Study on Adequacy of Dialysis (NECOSAD)**.

e. Tromp AM, Smit JH, Deeg DJ, Bouter LM, Lips P. Predictors for falls and fractures in the **Longitudinal Aging Study Amsterdam**.

1. Otten MH, Prince FH, Armbrust W, Ten Cate R, Hoppenreijs EP, Twilt M, Koopman-Keemink Y, Gorter SL, Dolman KM, Swart JF, van den Berg JM, Wulffraat NM, van Rossum MA, van Suijlekom-Smit LW. Factors Associated With Treatment Response to Etanercept in Juvenile Idiopathic Arthritis. JAMA. 2011 Nov 6. [Epub ahead of print].
The **Arthritis and Biologicals in Children Register**, an ongoing prospective observational study since 1999, includes all Dutch Juvenile Idiopathic Arthritis patients who used biologic agents. All biologically naive patients who started etanercept before October 2009 were included, with follow-up data to January 2011.

2. Willemsen G, de Geus EJ, Bartels M, van Beijsterveldt CE, Brooks AI, Estourgie-van Burk GF, Fugman DA, Hoekstra C, Hottenga JJ, Kluft K, Meijer P, Montgomery GW, Rizzu P, Sondervan D, Smit AB, Spijker S, Suchiman HE, Tischfield JA, Lehner T, Slagboom PE, Boomsma DI. The **Netherlands Twin Register biobank**: a resource for genetic epidemiological studies. Twin Res Hum Genet. 2010 Jun;13(3):231-45.

In 2004 the Netherlands Twin Register (NTR) started a large scale biological sample collection in twin families to create a resource for genetic studies on health, lifestyle and personality. Between January 2004 and July 2008, adult participants from NTR research projects were invited into the study. Blood and urine samples were collected in 9,530 participants (63% female, average age 44.4 (SD 15.5) years) from 3,477 families. Lipid profile, glucose, insulin, HbA1c, haematology, CRP, fibrinogen, liver enzymes and creatinine have been assessed. Longitudinal survey data on health, personality and lifestyle are currently available for 90% of all participants. Genome-wide SNP data are available for 3,524 participants, with additional genotyping ongoing. The NTR biobank, combined with the extensive phenotypic information available within the NTR, provides a valuable resource for the study of genetic determinants of individual differences in mental and physical health. It offers opportunities for DNA-based and gene expression studies as well as for future metabolomic and proteomic projects.

3. Borgdorff MW, van den Hof S, Kremer K, Verhagen L, Kalisvaart N, Erkens C, van Soolingen D. Progress towards tuberculosis elimination: secular trend, immigration and transmission. Eur Respir J. 2010 Aug;36(2):339-47. Epub 2009 Dec 8.
This study aimed to determine to what extent tuberculosis trends in the Netherlands depend on secular trend, immigration and recent transmission. Data on patients in the **Netherlands Tuberculosis Register** in the period 1993-2007 were matched with restriction fragment length polymorphism (RFLP) patterns of Mycobacterium tuberculosis isolates.

4. Veerbeek MA, Pijl YJ, Driessen GA, de Vries SC, Pot AM. [Trends in the utilisation of Dutch mental health services by older adults between 1990-2004]. Tijdschr Gerontol Geriatr. 2009 Apr;40(2):45-53.
Information about the use of mental health services by older adults was retrieved from the **Dutch Psychiatric Case Registers.**

5. Mohangoo AD, Buitendijk SE, Hukkelhoven CW, Ravelli AC, Rijninks-van Driel GC, Tamminga P, Nijhuis JG. [Higher perinatal mortality in The Netherlands than in other European countries: the Peristat-II study]. Ned Tijdschr Geneeskd. 2008 Dec 13;152(50):2718-27.
Indicators ofperinatal mortality which were developed for Peristat-I were used again in Peristat-II. Data on perinatal mortality in 2004 were delivered by 26 European countries. The Dutch data originated from **national registers of midwives and gynaecologists** and the **National Neonatology Register**.

6. Schram MT, Frijters D, van de Lisdonk EH, Ploemacher J, de Craen AJ, de Waal MW, van Rooij FJ, Heeringa J, Hofman A, Deeg DJ, Schellevis FG. Setting and registry characteristics affect the prevalence and nature of multimorbidity in the elderly. J Clin Epidemiol. 2008 Nov;61(11):1104-12. Epub 2008 Jun 6.
We used data from three population-based studies, two general practitioner registries, one hospital discharge register, and one nursing home registry to estimate the prevalence of multimorbidity. Individuals aged 55 years and over were included.

7. van Laar JJ, Grishchenko M, van Wouwe JP, Stronks K. Ethnic differences in the timely diagnosis of children with Type 1 diabetes mellitus in the Netherlands: clinical presentation at onset. Diabet Med. 2007 Mar;24(3):296-302.
From a national register (**Dutch Paediatric Surveillance Unit** (DPSU)) we selected 3128 children aged < 15 years with newly diagnosed Type 1 DM. Ethnic differences in serum glucose, blood pH, bicarbonate, presence of ketonuria, level of consciousness, hydration status,and diabetic ketoacidosis were assessed by logistic regression.

8. Pijl YJ, Sytema S. The identification of trends in the utilisation of mental health services by elderly: a Dutch case register study. Int J Geriatr Psychiatry. 2003 May;18(5):373-80.
Details of elderly users and their use of community- and hospital-based services between 1990 and 1999 were retrieved from the Groningen case register.

9. Stronks K, Ravelli AC, Reijneveld SA. Immigrants in the Netherlands: equal access for equal needs? J Epidemiol Community Health. 2001 Oct;55(10):701-7.

Survey data were linked to an insurance register concerning people aged 16-64. Ethnic differences in the use of a broad range of health care services were examined in this group, with and without adjustment for health status and socioeconomic status, using logistic regression.

10. Hart HE, van der Wouden JC, Höppener P, van Schendel GJ, Knottnerus JA. General practice registration networks in the Netherlands: a brief report. J Am Med Inform Assoc. 1999 Mar-Apr;6(2):173-5.
In the Netherlands, several general practice registrations exist. Groups of general practitioners register elements of patient care according to agreed-upon criteria, and these data are collected in a central database. By means of a questionnaire the authors interviewed the managers of all **nine computerized registration networks** extensively about the possibilities and limitations of their registration. In addition, respondents answered some questions with data from the central database of their network. Various items are collected by nearly all the registration networks, while other items are collected by only one network. Answering questions with data from the central database turned out to be difficult. Organisation and manpower are the main obstacles.

11. Miedema HS, van der Linden SM, Rasker JJ, Valkenburg HA. National database of patients visiting rheumatologists in The Netherlands: the standard diagnosis register of rheumatic diseases. A report and preliminary analysis. Br J Rheumatol. 1998 May;37(5):555-61.
Standard Diagnosis Register of Rheumatic Diseases

12. Schlösser FJ, Vaartjes I, van der Heijden GJ, Moll FL, Verhagen HJ, Muhs BE, de Borst GJ, Tiel Groenestege AT, Kardaun JW, Reitsma JB, van der Graaf Y, Bots ML. Mortality after hospital admission for ruptured abdominal aortic aneurysm. Ann Vasc Surg. 2010 Nov;24(8):1125-32.
The mortality risks for 28-day, 1-year, and 5-year were derived from a retrospective nation-wide cohort study of patients who were first hospitalized for rAAA in 1997 or 2000, formed through linkage of the **Hospital Discharge Register** with the Dutch population register.

B. Scopus, 28-11-2011

- register AND Netherlands AND medicine (period 1996-2011): 55
- cohort AND Netherlands AND medicine (period 1996-2011; Abstract): 65

Literatuur:
13. Van Weel, C. , Orbon, K. , Van Der Gulden, J. , Buijs, P. , Folgering, H., Thoonen, B. , Schermer, T. Occupational health and general practice: From opportunities lost to opportunities capitalised? Medicina del Lavoro, Volume 97, Issue 2, March 2006, Pages 288-294
Nijmegen Continuous Morbidity Registration;

f. Coutinho, R.A. The **Amsterdam Cohort Studies on HIV infection and AIDS**. Journal of Acquired Immune Deficiency Syndromes and Human Retrovirology. Volume 17, Issue SUPPL. 1, 1998, Pages S4-S8

C. Picarta, 28-11-2011

- register AND Netherlands AND medicine (period 1996-2011): 12

Literatuur:
-

D. Pharmaceutisch Weekblad

- register AND wetenschappelijk: 67
- cohort AND register: 12

14. Marieke H. Otten, MD, MSc; Femke H. M. Prince, MD, PhD, MSc; Wineke Armbrust, MD; Rebecca ten Cate, MD, PhD; Esther P. A. H. Hoppenreijs, MD; Marinka Twilt, MD, PhD, MSc; Yvonne Koopman-Keemink, MD; Simone L. Gorter, MD, PhD; Koert M. Dolman, MD, PhD; Joost F. Swart, MD; J. Merlijn van den Berg, MD, PhD; Nico M. Wulffraat, MD, PhD; Marion A. J. van Rossum, MD, PhD; Lisette W. A. van Suijlekom-Smit, MD, PhD, MSc. Factors Associated With Treatment Response to Etanercept in Juvenile Idiopathic Arthritis. JAMA. Published online November 6, 2011.

The **Arthritis and Biologicals in Children Register**, an ongoing prospective observational study since 1999, includes all Dutch JIA patients who used biologic agents.

15. J. Jentink, M.K. Bakker, F. Vroom, P.B. van den Berg, H.E.K. de Walle en L.T.W. de Jong-van den Berg. Voorschrijfpatronen voor, tijdens en na de zwangerschap voor chronische, incidentele en zwangerschapsgerelateerde medicatie in Nederland. PW Wetenschappelijk Platform. 2007;1(1):8-15.
Er is gebruikgemaakt van de **InterActie DataBank (IADB.nl).** Hierin staat informatie over afgeleverde receptgeneesmiddelen van 50 openbare apotheken in Noord- en Oost-Nederland met gegevens van een populatie van ongeveer 500.000 mensen.

16. E. Bettina Samre´n, MD,*† Cornelia M. van Duijn, PhD,† G. C. M. Lieve Christiaens, MD, PhD,‡ Albert Hofman, MD, PhD,† and Dick Lindhout, MD, PhD*§. Antiepileptic Drug Regimens and Major Congenital Abnormalities in the Offspring. Ann Neurol 1999;46:739–746
National Perinatal Data Base LVR

NTvG, 28 november 2011

- register OF cohort AND geneesmiddelen (period 1996-2011); onderzoek: 37

17. Charles Agyemang, Anton E. Kunst, Raj Bhopal, Paola Zaninotto, Nigel Unwin, James Nazroo, Mary Nicolaou, William K. Redekopen, Karien Stronks. Hypertensie in Nederlandse en Engelse etnische minderheidsgroepen. Ned Tijdschr Geneeskd. 2011;155:A3318.
Voor een gedetailleerde beschrijving van bloeddrukmetingen en onderlinge standaardisatie daarvan, antropometrie, sociaaleconomische status, fysieke activiteit en de definitie van etnische groepen verwijzen wij naar andere publicaties.17-19 De Britse gegevens kwamen van de Health Survey for England (HSE) en het Newcastle Heart Project (NHP). De gegevens over de Nederlandse etnische groepen waren afkomstig uit het **SUNSET-onderzoek**.

18. Jaco Voorham, Flora M. Haaijer-Ruskamp, Klaas van der Meer, Dick de Zeeuw, Bruce H.R. Wolffenbuttel, Klaas Hoogenberg en Petra Denig. Kwaliteit van de behandeling van type 2-diabetes. Ned Tijdschr Geneeskd. 2010;154:A775.

Voor deze observationele studie maakten we gebruik van gegevens uit 95 huisartspraktijken (124 huisartsen), die in 2007 deelnamen aan het **'Groningen Initiative to ANalyse Type 2 diabetes Treatment' (GIANTT)-project.**

19. W.J. van der Veen en G.Th. van der Werf. Gebruik van rofecoxib in de Noord-Nederlandse huisartspraktijk, 2000-2004: achtergronden en gevolgen. Ned Tijdschr Geneeskd. 2006;150:1016-21.
Met behulp van gegevens van huisartspraktijken (17 huisartsen), aangesloten bij het **Registratie Netwerk Groningen** met circa 30.000 patiënten in het noorden van het land, werden over de periode 2000-2004 incidenties en prevalenties van rofecoxibgebruik per kwartaal bepaald per 1000 patiënten. Overzichten betreffende duur, dosering en indicatiestelling werden samengesteld vanuit alle prescripties van rofecoxib.

Annex 5: List of registries approached for the survey in this study

Greyed-out rows indicate duplicates

National Databases	Name	website
SFK	Pharmacy Information System	http://www.sfk.nl/english
LAREB	The Netherlands Pharmacovigilance Centre	http://www.lareb.nl/?lang=en-GB
PHARMO	Pharmo	http://www.pharmo.nl/
Doetinchem cohort study	Doetinchem cohort study	http://www.rivm.nl/en/Topics/Topics/D/The_Doetinchem_Cohort_Study
PIAMA	The PIAMA research project	http://piama.iras.uu.nl/en/
LINH	The Netherlands Information Network of General Practice	http://www.linh.nl/
IPCI	Integrated Primary Care Information	http://www.ipci.nl/Framework/Frames.php
LMR (DHD)	National Medical register	http://www.dutchhospitaldata.nl/Registraties/LMR.php
Dutch CF-Registry	Dutch CF-Registry (cystic fibrosis)	http://www.ncfs.nl/
GIP	Drug Information System of the Health Care Insurance Board	http://www.gipdatabank.nl/
FIS	Vektis Pharmacy Information System	http://www.vektis.nl/
LCMR	Landelijke Centrale Middelen Registratie	http://www.sivz.nl/ivz-verslavingszorg/lcmr
PALGA	The nationwide network and registry of histo- and cytopathology in the Netherlands	http://www.palga.nl/palga/palgacms.nsf/viewdoc/eng-14
Dutch Growth Research Foundation	Dutch Growth Research Foundation	http://www.kindengroei.nl/site/index.php
ABC register	Arthritis and Biologicals in Children Register (ABC-register)	https://www.abc-register.nl/startpagina/home/home.php
NKR cancer registration	NKR cancer registration	http://www.iknl.nl/page.php?id=3244&nav_id=298
HIV/AIDS: HIV Monitoring Foundation	HIV/AIDS: HIV Monitoring Foundation	http://www.hiv-monitoring.nl/index.php/nederlands/
NTR	Netherlands Tuberculosis Register (NTR)	http://www.kncvtbc.nl/
ITP database	ITP database(Liset: welke?)	http://www2.hematologie-amc.nl/JHM-AIP-105_ITP
Psychiatric Case Register	Psychiatric Case Register	RGO centrum UMC Groningen
PRN	Perinatal Registration	http://www.perinatreg.nl
LADIS	National Alcohol and Drugs Information System	http://www.sivz.nl/ivz-verslavingszorg/ladis
NIVEL health care registries	NIVEL hcr	
IADB	Interaction Database	http://www.iadb.nl/
Achmea Health Database	Achmea Health Database	http://www.agisweb.nl/Voor_Zorgverleners/Zorgprogramma_s/Achmea_Health_Database

Biobanks		
BBMRI-NL	Dutch project for biobank collaboration	http://www.bbmri.nl/en-gb/home
LifeLines	LifeLines	http://lifelines.nl/
PSI	String of Pearls Institute	http://www.string-of-pearls.org/
Disease specific databases	-	
NKR cancer registration	NKR cancer registration	http://www.iknl.nl/page.php?id=3244&nav_id=298
Diabetes: bundled payments (DBC) for diabetes care	Experimenting with Bundled payments for **diabetes** care	http://www.rivm.nl
Nieuwe Hoorn studie		http://www.diabetes-zorg.nl/onderzoek_nieuwehoornstudie.html
SKION Child cancer AMC	SKION Child cancer AMC	http://www.skion.nl/
Antifungal register (SWAB, Pfizer, MSD en Pharmo)	Antifungal register (SWAB, Pfizer, MSD en Pharmo)	
Orthopedia register	Orthopedia register	
(Cardio)vasculair: SMART (UMCU)	Second Manisfestations of ARTerial Disease	http://www.umcutrecht.nl/subsite/Vasculair-Preventie-Programma/
Dementia: ERGO	Rotterdam Study	http://www.ergo-onderzoek.nl/wp/
DREAM	Dutch Rheumatoic Arthritis Monitoring (DREAM)	http://www.dreamregistry.nl/
NTR	Netherlands Tuberculosis Register (NTR)	http://www.kncvtbc.nl/
Kidney cancer (Pfizer, ZonMw)	Kidney cancer (Pfizer, ZonMw)	
Skin cancer (Galderma, ZonMw)	Skin cancer (Galderma, ZonMw)	
Maastricht breastcancer register (financiering Roche, ZonMw)	Maastricht breastcancer register (financiering Roche, ZonMw)	
Hemobase (haematology malignants, Leeuwarden hospital, Roche)	Hemobase (haematology malignants, Leeuwarden hospital, Roche)	
DSCA	Dutch Colorectal Audit	http://www.dccg.nl/colorectalaudit
ITP database	ITP database(Liset: welke?)	http://www2.hematologie-amc.nl/JHM-AIP-105_ITP
HIV/AIDS: HIV Monitoring Foundation	HIV/AIDS: HIV Monitoring Foundation	http://www.hiv-monitoring.nl/index.php/nederlands/
Neonatal databases	-	
PRN	Perinatal Registration	http://www.perinatreg.nl
InterActie DataBank IADB	IADB drug use research	http://iadb.nl/
The elderly	-	
LASA	Longitudinal Aging Study Amsterdam	http://www.lasa-vu.nl/index.htm
SHARE		
Rotterdam study, ERGO	-	http://www.ergo-onderzoek.nl/wp/
NPO	the-national-care-for-the-elderly-programme	http://www.nationaalprogrammaouderenzorg.nl/
Rotterdam Study (Ommoord)		http://www.erasmusmc.nl/47708/51019/ergo
Children		

ABC register	Arthritis and Biologicals in Children Register	https://www.abc-register.nl/startpagina/home/home.php
Dutch Growth Research Foundation	Dutch Growth Research Foundation	http://www.kindengroei.nl/site/index.php
SKION Child cancer AMC	SKION Child cancer AMC	http://www.skion.nl/
NSCK	Dutch Paediatric Surveillance Unit	http://www.nvk.nl/Onderzoek/NSCK.aspx
Mental health	-	
Psychiatric Case Register	Psychiatric Case Register	RGO centrum UMC Groningen
NEMESIS-2	Netherlands Mental Health Survey and Incidence Study-2	http://www.trimbos.nl/projecten-en-onderzoek/nemesis-2
LADIS	National Alcohol and Drugs Information System	http://www.sivz.nl/ivz-verslavingszorg/ladis
LCMR		http://www.sivz.nl/ivz-verslavingszorg/lcmr
Academic General Practitioner GP networks		
RNH	MU:, Registration network GP's Limburg (RNH)	http://rnh-maastricht.nl/
Maastricht study	Maastricht UMC: Maastricht study	http://www.demaastrichtstudie.nl/
Julius GP Network	UMC Utrecht; Julius GP Network	www.juliuscenter.nl
VUmc: Ac. Netw. Huisartsengeneeskunde	VUmc: Ac. Netw. Huisartsengeneeskunde	http://www.vumc.nl/afdelingen/huisartsgeneeskunde/kn/
Nijmegen Biomedical Study	UMC St. Radboud: Nijmegen Biomedical Study	http://www.nijmegenbiomedischestudie.nl/
TransHis	UMC St. Radboud: Transition project	http://www.transitieproject.nl
CMR Nijmegen	Continues Morbidity Registration Nijmegen	http://www.nivel.nl/oc2/page.asp?pageid=12135
GP Network AMC (HAG-net-AMC)	AMC: GP Network AMC (HAG-net-AMC)	www.amc.nl/web/Het-AMC/Afdelingen/Overzicht/Huisartsgeneeskunde
LUMC: RNUH Leo	Registration Network University Practices Leiden and Environment.	http://www.lumc.nl/con/4070/83231/901190557282221/90202045508221/
UMCG/RUG: Giantt	Groningen Initiative to Analyse Type 2 diabetes Treatment	http://www.giantt.nl/enoutline.htm
UMCG/RUG: RNG	GP Registration Network Groningen	http://rng.med.rug.nl/websiteRNG/
UMCG/RUG: IADB (see above)		http://iadb.nl/
Non-academic databases:		
Farma industrie: data uit clinical trials (MSD/Nefarma)		http://www.janssennederland.nl/?product=none
Ziekenhuisapotheken: Intramural geneesmiddelengebruik		
Rare diseases (off-label/orphan drugs):		
NSCK (see above)	Dutch Paediatric Surveillance Unit	http://www.nvk.nl/Onderzoek/NSCK.aspx
Pharos		http://www.pharos.nl/home/nieuws/3/379/
Dutch CF-Registry	Dutch CF-Registry (cystic fibrosis)	http://www.ncfs.nl
Orphanet	Orphanet (international, Tipharma project)	http://www.orpha.net/consor/cgi-bin/index.php?lng=EN

Metabolic Diseases	Metabolic Diseases	http://www.umcutrecht.nl/subsite/metabool/Kliniek/Bereikbaarheid.htm
Cohort studies		
LASA (see above)		http://www.lasa-vu.nl/index.htm
Doetinchem (see above)		http://www.rivm.nl/Onderwerpen/Onderwerpen/D/Doetinchem_Cohort_Studie
SMILE	Study of Medical Information and Lifestyle in Eindhoven	http://www.zorggegevens.nl/
EPIC-NL	European Prospective Investigation into Cancer and Nutrition	http://www.epicnl.eu/
Rotterdam Study (Ommoord)		http://www.epib.nl/research/ergo.htm
NTR	Netherlands Twin Register (NTR)	http://www.tweelingenregister.org/en/
PIAMA (see above)		http://piama.iras.uu.nl/
GenerationR		
Rotterdam study, ERGO		http://www.ergo-onderzoek.nl/wp/
LifeLines	LifeLines	http://lifelines.nl/
NEMESIS-2	Netherlands Mental Health Survey and Incidence Study-2	http://www.trimbos.nl/projecten-en-onderzoek/nemesis-2

Annex 6: Literature search chapter 7 on governance

1. Laurie G. Reflexive governance in biobanking: on the value of policy led approaches and the need to recognise the limits of law. Hum Genet. 2011 Sep;130(3):347-56. Epub 2011 Jul 16. http://www.springerlink.com/content/y8741t3236345pt7/
[UK/ Scotland]

Policies, commonly initiated by funders and other stakeholders, become an important means by which biobankers are held accountable. This article analyses this policy-driven phenomenon asking how effectively policy – often as an alternative to law – serves to police and to promote biobanking. It argues that a policy of reflexive governance – defined and developed herein – can best meet the challenges faced by many biobanks and without the need for recourse to law. An example of this is the UK Biobank which operates according to an Ethics and Governance Framework. According to Laurie, an outline **framework against which biobank policies and design can be built and measured**, should contain:

- designing-in interoperability with respect to scientific and governance approaches,
- designing-out approaches that are restrictive of sharing, cooperation, flexibility and mutuality,
- establishing policies and procedures to protect adequately the interests of participants,
- establishing policies and procedures to promote actively the use of the resource in keeping with its original purposes,
- ensuring the longevity of the biobank through carefully managed access policies and arrangements and stewardship of the resource, and
- ensuring that governance policies and mechanisms remain fit for purpose over time.

2. Williamson O.D., Cameron P.A., McNeil J.J. Medical registry governance and patient privacy. Med J Aust. 2004 Aug 2;181(3):125-6.
http://www.mja.com.au/public/issues/181_03_020804/wil10219_fm.html [Australia]

The authors advocate governing of medical registries that both safeguards individual privacy and allows the registries to continue to provide the foundations for quality-improvement programs and epidemiological research. Good **medical registry governance** involves developing a structure that

- includes stakeholders in management of institutions that analyze personal medical information;
- has a management independent of the institutions that provide healthcare;
- provides a research environment that maximizes scientific benefit to patients and the wider community; and
- receives adequate funding to ensure continuity of data collection and quality assurance.

PubMed query: "medical registry governance" > governance AND medical AND (registries OR registry OR register): 23

3. Fischer A.S., Mansmann U. A metadata-based patient register for cooperative clinical research: a case study in acute myeloid leukemia. Stud Health Technol Inform. 2011;169:857-61. [Germany]

In many medical indications clinical research is organised within study groups which provide and maintain the clinical infrastructure for their randomized clinical trials. **Sharing this data between**

study groups is not straightforward. We propose a metadata based patient register and describe a first prototype.

4. Lipscomb J., Gillespie T.W. State-level cancer quality assessment and research: building and sustaining the data infrastructure. Cancer J. 2011 Jul-Aug;17(4):246-56. [USA]

The United States still lack a population-based data infrastructure for accurately identifying cancer patients and tracking services and outcomes over time. The most effective pathway forward may be the development of **state-level** cancer **data systems**, in which **central registry data are linked to multiple public and private secondary sources**. These would include administrative/claims files from Medicare, Medicaid, and private insurers.

5. Bennett B., Deakin C.; Australian Gender Equity in Health Research Group. Registration of clinical trials: challenges for global regulation. J Law Med. 2009 Aug;17(1):82-94. [Australia]

In 2004 the International Committee of Medical Journal Editors (ICMJE) issued a statement indicating that from 1 July 2005 **registration in a publicly accessible trials registry would be a condition of publication** in an ICMJE member journal. The World Health Organisation is coordinating the International Clinical Trials Registry Platform (ICTRP) as a means of providing a standardised framework for registration.

6. Lyons R.A., Jones K.H., John G., Brooks C.J., Verplancke J.P., Ford D.V., Brown G., Leake K. The SAIL databank: linking multiple health and social care datasets. BMC Med Inform Decis Mak. 2009 Jan 16;9:3. [UK]

The SAIL (**Secure Anonymised Information Linkage**) databank has been established using disparate datasets, and over 500 million records from multiple health and social care service providers have been loaded to date, with further growth in progress. The aim of this work was to develop and implement an accurate matching process to enable the assignment of a unique Anonymous Linking Field (ALF) to person-based records to make the databank ready for record-linkage research studies.

7. Gilkes C.E., Casimiro M., McEvoy A.W., MacFarlane R., Kitchen N.D. Clinical databases and data protection: are they compatible? Br J Neurosurg. 2003 Oct;17(5):426-31. [UK]

In the current climate of clinical governance and audit an effective clinical database is an invaluable tool. The arrival of the **UK 1998 Data Protection Act has put many clinical databases and registries in jeopardy**, and introduced further bureaucracy to research. We discuss the Act and its interpretation by the General Medical Council, Medical Research Council, British Medical Association, Department of Health and our own trust with respect to databases and research.

PubMed query: registry[Title/Abstract] AND governance[Title/Abstract]: 20 (results overlap with the previous query)

8. Watterson D., Gabbe B.J., Cleland H., Edgar D., Cameron P.; Members of the Bi-NBR Steering Committee. Developing the first Bi-National clinical quality registry for burns-Lessons learned so far. Burns. 2011 Nov 10. [Epub ahead of print] [Australia and New Zealand]

The Australian and New Zealand Burn Association (ANZBA) commenced the **Bi-National** Burns **Registry** (Bi-NBR) in 2004. Institutional ethics approval has been obtained for 16 out of 17 sites and **a formalised governance process** has been developed. The minimum dataset was improved and includes clinical quality indicators.

9. Connelly L.B. The economic characteristics of registries and their policy implications. J Trauma. 2009 Feb;66(2):531-5. [Australia]

The typical production activities and cost structures of (trauma) registries are analyzed, along with the way registries generate social benefits. Assuming that the purpose of a trauma registry is to maximise the value or social good it creates, a number of **investment, governance, and pricing principles** are then proposed. Trauma registries are generally characterized by large and indivisible fixed, joint costs, and relatively low marginal costs. This implies that registries are subject to strong economies of scale and scope. The optimal price schedule for access to trauma registry data are likely to be zero, or close to zero, for some users.

PubMed query: "Registries/standards" AND governance: 6 (results overlap with the previous query)

10. Isasi R.M., Knoppers B.M. Governing stem cell banks and registries: emerging issues. Stem Cell Res. 2009 Sep-Nov;3(2-3):96-105. Epub 2009 May 25. [Canada]

Stem cell banks and registries are emerging as an essential resource for transnational access to quality-controlled and ethically sourced stem cell lines. We report the preliminary findings of a survey of stem cell banks participating in the International Stem Cell Forum's International Stem Cell Banking Initiative (ISCBI). The questionnaire circulated to all ISCBI members addressed both general issues surrounding research policies and issues relating to the **governance of stem cell banking projects**.

PubMed: related

11. McHale J.V. Accountability, governance and biobanks: the ethics and governance committee as guardian or as toothless tiger? Health Care Anal. 2011 Sep;19(3):231-46. doi: 10.1007/s10728-011-0195-7.

Biobanks raise many questions of the control of rights, of consent, of privacy and confidentiality and of property in human material. There has been a lively debate as to how biobanks should operate, the boundaries of participation and what **governance structure** they should adopt, a debate which has been engaged in across the academic community and by funders and researchers alike. Can ad hoc ethics and ethics and governance committees **long term** provide an **effective solution to the legal and regulatory challenges arising from biobanks**?

PubMed-query: "Confidentiality/legislation and jurisprudence"[MeSH Terms] and (register OR registry OR registries): 107 (results overlap with the previous query; mainly results about privacy)

12. Thygesen L.C., Daasnes C., Thaulow I., Brønnum-Hansen H. Introduction to Danish (nationwide) registers on health and social issues: structure, access, legislation, and archiving. Scand J Public Health. 2011 Jul;39(7 Suppl):12-6. [Denmark – strong resemblance to the Dutch situation]

Because all Danish citizens have a unique personal identification number, **linkage at the individual level between nationwide registers** and other data sources is possible and feasible. We introduce selected Danish registers and the data structure and requirements, not forgetting access to data at Statistics Denmark, which is the main provider of register data. We introduce the Danish Data Archive and briefly present the Act on Processing of Personal Data, which is the legal foundation for analyses of register-based data in Denmark.

Annex 7: Website scan of patient data registries (good practices)

In this annex you'll find some descriptions of a few Dutch and foreign patient data registries. The information has been retrieved from the registries' websites in January 2012 and where necessary translated for this study. The selection of registries is subjective: Based on earlier experience of the project team and on recommendations from external experts some registries were selected that demonstrate good practices. The study, being preparatory, does not aim at an exhaustive overview. In the same vein, the information taken from the websites is not complete and subjective, especially the sections headed "Good practice".

Below the following registries are presented (the order is arbitrary):

1. Landelijk Informatie Netwerk Huisartsenzorg (LINH)
2. Landelijke Basisregistratie Ziekenhuiszorg (LBZ)
3. Prevention and Incidence of Asthma and Mite Allergy (PIAMA)
4. Perinatal registration Netherlands (PRN)
5. SFK Data Warehouse
6. Nederlands Signalerings Centrum Kindergeneeskunde (NSCK)
7. Center for Health Equity Studies (CHESS, Sweden)
8. General Practice Research Database (GPRD, United Kingdom)
9. Medical Research Council (MRC, United Kingdom)

1. Landelijk Informatie Netwerk Huisartsenzorg (LINH), by NIVEL, IQ health care, NHG and LHV, to be part of NIVEL zorgregistraties

Website: http://www.linh.nl/ > English
Questionnaire received: yes

Introduction
LINH is the Netherlands Information Network of General Practice. The LINH database holds longitudinal data on morbidity, prescribing, and referrals of about 330.000 individuals. Data are collected in a representative network of about 84 general practices, spread throughout the Netherlands. LINH started in 1992 recording only referrals. It gradually developed into a system recording all patient contacts and all interventions, including diagnoses.

Aim
The aim of LINH is to develop and maintain a high quality longitudinal database on morbidity and GP care in the Netherlands and to use this database for health services research and quality of care research.

How does LINH work?
Data are extracted twice a year from the electronic medical records used in the practices to file patient information. Recording data for LINH hardly interferes with daily practice. Participants receive feedback reports comparing their own practice with the LINH average. They also receive a modest financial compensation and LINH organises an annual meeting.

What data are collected?
GPs record data on all patient contacts, including diagnoses, referrals and prescriptions. Diagnoses are coded using the ICPC (International Classification of Primary Care) and medicines according to the ATC (Anatomical Therapeutic Chemical). Patient, gender, age, type of health care insurance (before 2006) and place of residence are recorded as well.

How can I get access to the LINH database?
Request for (anonymised) data are judged by the LINH steering committee. Applications should provide a clear research question to the steering committee and guarantee that their findings be publicly available, regardless of the outcomes.

Financing
LINH's basic infrastructure is financed by the Dutch Ministry of Health Welfare and Sports

Good practice
LINH has a clear review procedure and requires public publication of studies using LINH-data. Data requests are reviewed both by researchers of different institutes and representatives of the medical profession, enhancing that both quality of research as well as relevance for policy and practice are taken into account. External parties pay no other costs than handling costs for the data. In addition, LINH requires to be a co-author in publication in order to enhance the quality of the data analyses and data interpretation. LINH can be linked to other databases, up to now based upon NAW-data, but in the near future through a TTP, also because LINH will be extended from a general practice database to a primary care database involving data from all types of primary care health providers.

Properties
1. *Permissions and conditions of use*

Authentication and authorisation of potential users of the data are a manual procedure mainly. The online form is the starting point for the offline procedure. In principal all requests are granted if they fulfil scientific quality criteria, and will be published. Sometimes additional requirements are posed (for example when the results may potentially harm the GP profession). Moreover, commercial parties do not receive data on the (anonymous) individual patient level, but tables on an aggregated level.

LINH request procedure:
The petitioner fills in the online form and sends it to the LINH request officer via the button 'Send'.
The petitioner has to add information, such the research proposal, and send it via e-mail to linh@nivel.nl.
NIVEL discusses the request in the project group LINH and the LINH-board. If necessary, the LINH request officer will contact the petitioner. This procedure will usually take about 5 weeks after the request has been submitted. If necessary, the request can be adapted.
If the request is granted, the LINH request officer sends a copy of the request form to the petitioner. He / she signs the form, and returns it to the LINH request officer.
When the form is received, the request will be processed.
The petitioner receives an user name from the LINH request officer by e-mail, and a password by SMS.
With respect to licencies, user conditions and codes of conduct, the petitioner needs to agree with the general conditions of use of LINH data, and with the procedure. Once he / she has ticked the box to agree, the form appears.
Apart from information about the petitioner, LINH wants to know more about the specific information the petitioner would like to use. The petitioner can narrow his request on the following subjects: Sex, age, national/regional, year(s), disease(s) identified by ICPD code.

The conditions for requesting access to LINH data are:
The data must be suitable to answer the research question.
The data can solely be used for the research question mentioned on the initial application form.
The intended research for which the data will be used, cannot interfere with (proposed) LINH activities.
The data should be used for a public publication.
The petitioner ought to send the concept publication to LINH, and offer LINH employees 2 weeks to read and comment it.
The petitioner should mention LINH as a source in the literature list of the publication. LINH doesn't use Persistent Identifiers or DOIs to (have) refer(red) to their data.

Available metadata
Descriptive metadata in LINH:
Sex
Age (exact, but not date of birth)
National sample of 84 practices across the country, information on ZIP-code
Concerning year / years … (1996-onwards)
Concerning illness(es). Please mention specific ICPC codes.
According to zorggegevens.nl (in Dutch) also: Contacts, illnesses, prescriptions, references, declarations, medical surgery, age, and sex.[67] BSN is not (yet) available, identification within the LINH-database based upon a patient identification number

Data formats
Not known.

[67] http://www.zorggegevens.nl/zorg/eerstelijnszorg/landelijk-informatie-netwerk-huisartsenzorg/

Interoperability
Interoperability of data are dependent on the data format used in both data files, and the presence of a communal metadata field, such as a Social Security Number (BSN in Dutch) or ICPC code. For linking NAW-data are used (age/data of birth, gender and Zip-code 4 or 6 positions; linking occurs usually through Statistics Netherlands (CBS).

---

2. Landelijke Basisregistratie Ziekenhuiszorg (LBZ), formerly known as Landelijke Medische Registratie (LMR), managed by Dutch Hospital Data

Website: http://www.dutchhospitaldata.nl/LBZ/LBZ_introductiepagina.php
Questionnaire received: yes

Introduction
Via the LBZ the NVZ Dutch Hospitals Association and the Nederlandse Federatie van Universitair Medische Centra aims to achieve transparency, based on good and comparable numbers. Next to that, LBZ intends to keep the registration load for hospitals to an absolute minimum.

Aim
To collect and provide access to information about patients in hospitals in order to monitor care production, management of hospitals, the health situation of the population, market information, complying with legislation (for instance CBS, RIVM, EU), and research and education.

Good practice
Once hospitals have uploaded their data these can be used for multiple goals, including LBZ, but also "societal users" like the organisations mentioned above ("single registration, multiple use"). For this purpose LBZ makes use of an explicit data model that supports both longitudinal data use and recent perspectives like DBCs (diagnosis-treatment combinations). This single registration approach supports a focal point the Duch government, viz. to reduce administrative burdens both for organisations and citizens.

What data are collected?
Hospital data, preferably from local "basisregistraties". LBZ provides a specification for data supply.

Properties
1. *Permissions and conditions of use*
   Regarding authentication and authorisation mechanisms, a petitioner needs to apply for the data by a written request. There's not much known about the process.
   Due to insecurity about supplying recognisable patient information to commercial parties service has been cancelled indefinitely.[68]
   The protocol is:
   1. Petitioner submits written request on company writing paper at DHD-window with the names of the petitioner, company, goal, requested data, proper care, health care provider, period concerned, a stated time-limit, consent with the General Conditions and the Protocol.
   2. The request is proposed to the "Orde van Medisch Specialisten", which decides.

[68] More information: http://www.dutchhospitaldata.nl/Gegevensloket/Protocol_DIS_LMR.php (in Dutch)

3. DHD-window informs and quotes the petitioner. The agreement is once only, and at a maximum of 12 months from the agreement date.
4. Petitioner pays according to quote
5. Petitioner informs DHD-window about outcome of research in which the data have been used, and mentions DHD in the source list of the research.
6. Hospitals can always use their own data without consent of DHD.
7. The owners of data, NFU, NVZ and the "Orde", can always use data without consent of DHD.
8. Communal parties within protocol keep an eye on compliance with privacy legislation.
9. When the proposal is not conform protocol, DHD and the "Orde" decide.
10. DHD evaluates protocol periodically, and adapts it, when necessary.

Petitioners need to refer to the used data in publications according to a fixed notation. There are not any Persistent Identifiers or DOIs used.

2. *Available metadata*
   Not known.

3. *Data formats*
   Not known.

4. *Interoperability*
   According to the website, DHD would like to link datasets to another.[69] DHD uses metadata to describe:
   Patients: DE1; Social Security Number (BSN)[70]; contact-ID (former survey number "opnamenummer"); patient-ID; COD046 (VEKTIS); postcodes; ISO-3166-1 (land); Origin and destination
   Health insurance companies: UZOVI-register VEKTIS
   Health care institution: WCC institutions list; institution serial number DIS, AGB code list VEKTIS
   Care traject: EI characterisation list DOT; ICD-10, version 2006[71]; care product chart DOT; price chart DOT; closing reason chart DOT; Claim code chart DOT; CVB file[72]; care activities chart; CBV; code list nature of care; CvZ80; CvV; diagnosis thesaurus; nature of contact; policlinics
   Practitioner: Caregiver specific "Zorgverleners-spec" COD016 VEKTIS; GPs (VEKTIS); LBZ specialisms

---

3: PIAMA by IRAS (UU), RIVM, Erasmus Medical Center and UMCGroningen

Website: http://piama.iras.uu.nl/en/
Questionnaire received: yes

Introduction
The PIAMA study is a multi centre birth cohort study conducted in the Netherlands since 1996. PIAMA stands for 'Prevention and Incidence of Asthma and Mite Allergy'.

[69] More information: http://www.dutchhospitaldata.nl/LBZ/Datamodel.php (in Dutch)
[70] BSN: http://www.dutchhospitaldata.nl/LBZ/BSN.php
[71] ICD-10: http://www.dutchhospitaldata.nl/LBZ/ICD10.php
[72] CVB-bestand: http://www.dutchhospitaldata.nl/LBZ/CBV_bestand.php

Aim
The study has two aims, one is to evaluate the effectiveness of mite impermeable mattress and pillow covers to reduce exposure to mite allergens, and to reduce incidence of allergic sensitisation and asthma. The other is to study the natural history of allergy and asthma in childhood, in relation to nutrition, familial factors, day care, pet ownership, air pollution, gas cooking, genetics etc.

How does PIAMA work?
PIAMA started with about 4000 pregnant women form the general population in 3 Dutch regions (North, Mid and South-West). Their children have been followed until their current age of 15. Different measurements are used (se below). Participating parents and children receive presents and regularly a newsletters and results from clinical tests. PIAMA has a cohort and an intervention study. The last study only includes children of allergic mothers, since those children are most likely to develop an allergy as well.

What data are collected?
Parents received an annual postal questionnaire each year. Questionnaires include information on life style, medication use, other health care utilisation, the environment etc. Besides, a part of the children were visited at home were exposure to mite allergens is measured by taking "monsters" of floor, mattresses and pillows. Most children (and their parents) took a medical examination including blood sampling to run allergy tests (for a more detailed overview of the data collection: see http://piama.iras.uu.nl/piama_project_overzicht.php).

How can I get access to the PIAMA database?
This procedure is not online. Everyone who is interested can contact PIAMA and discuss the opportunities for research. In performing the study there always has to be a cooperation between the external partner and the PIAMA research group.

Financing
Data are collected with funding from: RIVM, Asthma Foundation, ZonMw, NWO, Ministry of Health and Ministry of the Environment. The current data collection (2011/2012) is financed by RIVM. Data management and data "opslage" are subsidized by IRAS (Institute for Risk Assessment Sciences.

Good practice
PIAMA is a long running cohort study including children even from before their birth. Included children are up to 15 years old now. Parents receive clinical information about themselves and their child, which may stimulate participation. Use of these unique cohort data are open for everyone for low (handling) costs, but always in collaboration with the PIAMA team. Pharmacy data for over 800 participating children are available after informed consent was obtained by their parents. PIAMA has a privacy policy and pseudonimisation of data.

Properties
Permissions and conditions of use
Conditions or other pertaining information are not mentioned on the website. From the questionnaire PIAMA filled in it becomes clear that everyone who is interested can contact PIAMA and discuss the opportunities for research. In performing the study there always has to be a cooperation between the external partner and the PIAMA research group.

Available metadata
Descriptive metadata in PIAMA (from the questionnaire):
Sex
Age (exact)
Zip-code (6 positions) + street + number

Concerning year / years: 1995-onwards)
Concerning illness(es)": Asthma and allergies.
Other: medication use, health care utilisation, clinical information

Data formats
Not known.

Interoperability
Interoperability of data are dependent on the data format used in both data files, and the presence of a communal metadata field, such as a Social Security Number (BSN in Dutch) or ICPC code.
Up to now PIAMA has only been linked with pharmacy data. Linking is possible through date of birth, gender and zip-code (6 positions).

---

4. Perinatal registration Netherlands (PRN), by KNOV, LHV, NVOG, NvK

Website: http://www.perinatreg.nl/home_english
Questionnaire received: yes

Introduction
The Netherlands Perinatal Registry (PRN Foundation) is a joint effort of four professional organisations that provide perinatal care in the Netherlands:
KNOV (Royal Organisation of Midwives in the Netherlands)
LHV (National Organisation of General Practitioners)
NVOG (Dutch Association of Obstetrics & Gynaecology) and
NvK (Paediatric Association of the Netherlands).

Aim
The mission of the Netherlands Perinatal Registry is to improve the quality of health care by giving insight into the perinatal care process and outcomes.

How does PRN work?
As mentioned, four professional organisations participate in PRN: KNOV, LHV, NVOG and NvK. All have their own voluntary based medical registry:
LVR1-registry for midwives / KNOV
LVRh-registry for GPs / LHV
LVR2-registry for obstetricians / NVOG
LNR-registry for paediatricians and neonatologists / NvK
The LVR1, LVR2 and LNR registries are linked to one combined PRN-registry.
Participants receive feedback reports. PRN also organises an annual meeting.

What data are collected?
All participating professionals register extensive data on mother's and child's health (including problems), medication use, demographics, lifestyle etc. There are 3 ways in which to provide data: on paper, via e-mail, or via the web. More information can be found on www.perinatreg.nl.
Based upon this information PRN constructed one national databank using a uniform language throughout the whole chain for perinatal care and including an electronic patient dossier.

How can I get access to the PRN database?
A request for (anonymised) data are judged by the PRN-office, a privacy commission and the PRN board that is composed of an equal representation of the four professional organisations involved. Applications should provide a clear research question, which will be posted on the PRN website. PRN wants to be mentioned as a source when figures are published and applicants have to sign a form to agree upon this. Until 2011 applicants did not have to pay for the data. As of 2011, applicants have to pay for the hours PRN need to distract the data. Rates can be found on the website.

Financing
PRN is partly financed by the Ministry of Health, Welfare and Sport, partly by "premiegelden" and partly by "gebruikersgelden".

Good practice
PRN has a clear procedure for requesting data which is transparent. In addition, data requests are reviewed by three different bodies within the organisation. PRN can be linked on patient record level to other databases through a TTP. External parties pay nothing but handling costs for the data and, if applicable, costs to use the TTP. To enhance transparency, summaries of data requests are published on the PRN website. Codebooks are available on the website.

Properties
Permissions and conditions of use
Authentication and authorisation of potential users of the data are mainly a manual procedure. The online form is the starting point for the offline procedure. Data to fill out in this form include for example the aim of the projects, type of project (research, policy etcetera), data requested and research question for a summary for the PRN website, as mentioned earlier.
PRN request procedure:
The petitioner fills in the online form and sends it to the PRN through e-mail. The petitioner signs the form.
PRN discusses the request with the board and the privacy commission.
If the request is granted, the request will be processed.
Data are sent to the petitioner on a DVD or CD, through the internet or via CBS.

Available metadata
City, age and ethnicity of the mother in 7 categories
Gravidity en parity of the mother
Birth weight, birth place and sex of the baby
Single / multiple birth and birth sequence, if applicable
Pregnancy duration at birth
Condition of newly born baby (alive / deceased, Apgar score)
Details on pregnancy in codes
Delivery process (inducing labor or not, expulsion period, help with delivery) and the care given to the newly born baby
Information on referring from 1st to 2nd line, from midwife / GP / gynaecologist to paediatricians and neonatologists
Care provider

Data formats
Databases, exact format not known.

Interoperability
Interoperability of data are dependent on the data format used in both data files, and the presence of a communal metadata field, such as a Social Security Number (BSN in Dutch) or ICPC code.

The metadata which is used mostly to combine datasets, is BSN, name, address, city, date of birth, sex and other patient and non patient details.[73]

---

5. SFK Data Warehouse, managed by Foundation for Pharmaceutical Statistics (SFK): public pharmacies

Website: http://www.sfk.nl/english
Questionnaire received: no

Introduction
The Dutch Foundation for Pharmaceutical Statistics (SFK) has been collecting data about the use of pharmaceuticals in the Netherlands since 1990. The SFK directly gathers its data from a panel of pharmacies; more than 1,872 of the 1,980 community pharmacies in the Netherlands are represented on this panel. The pharmacies on the panel combined serve about 15.3 million people.

Aim
The main goals of the SFK are supporting the promotion of:
good pharmaceutical services;
scientific practice of pharmacy;
promotion of pharmacists' interests within the Netherlands.

Good practice
Communication towards the community and beyond is ample: the site contains an explicit list of hardware and software needed to consult the data in the Data Warehouse. There's also a good description of Data Warehouse online. Therefore the user can decide on the usability of the data before he or she has made a request. This is a great advantage, and very customer friendly. Moreover, SFK publishes a weekly update "Pharmacy in numbers" both in a pharmaceutical magazine and online.

Properties
Permissions and conditions of use
Logging in the system is only possible with an account of SFK. Unless you are a pharmacist, you cannot get an account.
According to the website, it is not allowed to distribute regional or national data from the Data Warehouse by publishing or other forms of distribution without a written concent of SFK.
Since the data are only available to pharmacists, the process of identification of a potential subscriber will probably take place by checking the apotheekregister ("pharmacy register")

Available metadata
Not known.

Data formats
The data format used in the SFK Warehouse cannot be found online, but there's a list of the hardware and software needed. The list is from 2004, so it's not very recent. It's a plus that the SFK

[73] See 2. Available metadata, and www.perinatreg.nl.

hardware communicates about the needed though. (at least, the 2004 version). A new user gets a CD-rom with the software needed.

Interoperability
For each dispensation, the SFK registers information about the drug supplied, the dispensing pharmacy, the health insurance company that does or does not reimburse the remedy, the prescribing doctor and the patient for whom the prescription was issued.
Linking to other registers through the ATC code seems possible.[74] With regards to the use of ontologies, at least ATC-codes and GPK numbering is used in the SFK Data Warehouse. "With regard to the prescribing doctor and the patient, the SFK only uses anonymously gathered data."

---

6. Nederlands Signalerings Centrum Kindergeneeskunde (NSCK), managed by NVK and TNO

Website:
http://www.nvk.nl/Onderzoek/NSCK/AlgemeneinfoNSCK/Hetmeldingssysteem/tabid/278/language/nl-NL/Default.aspx
Questionnaire received: yes

Introduction
In 1992 the Dutch Signalling Centre Paediatrics (NSCK) was initiated by the Dutch Association of Paediatrics (NVK). It is accommodated by TNO Prevention and Health (applied research institution). All paediatricians in general and academic hospitals participate in this Signalling system.

Aim
to gain more insight in prevalence of rare or new illnesses of children in the age of 0 to 18 years old.
to promote scientific research, aimed at backgrounds, diagnostics, treatments, prognoses and prevention of these illnesses.

Financing
Depending on the study NSCK is financed partly through the principal investigator, partly through TNO/NVK, and incidently by external sponsors.

Good practice
Good response from hospitals: general hospitals: 92%, academic hospitals: 83%. Thus it is widely supported by paediatricians nationwide. *However, this information dates back to 2001.*

Properties
Permissions and conditions of use
Not known. Paediatricians and other physicians who have direct or indirect contact with children and / or paediatrics mainly use the signalling system.[75] They contribute to the system by sending in a 'blue card' by e-mail or post monthly.[76]

[74] http://www.whocc.no/atc_ddd_index/ : look for 'paracetamol'
[75] More information on use of data from NSCK in new research:
http://www.nvk.nl/Onderzoek/NSCK/AlgemeneinfoNSCK/Onderzoek/tabid/276/language/nl-NL/Default.aspx
(in Dutch)

Available metadata
Not known.

Data formats
Not known.

Interoperability
Since neither data format, nor metadata information is mentioned online, it is not clear whether data in the NSCK Signalling System can be linked to other data or not.

---

7. Center for Health Equity Studies (CHESS, Sweden)

Website: http://www.chess.su.se/
Questionnaire received: not applicable: the questionnaire was only sent to Dutch registries

Introduction
CHESS hosts four data collections with their own websites:
The Stockholm Birth Cohort Study (SBC): The SBC was created in 2004/2005 by a probability matching of two comprehensive and longitudinal datasets. It consists of both registry data and survey data.
The Swedish Panel Study of Living Conditions of the Oldest Old (SWEOLD): SWEOLD analyses the living conditions of elderly people in Sweden. Since 1992 three surveys have been carried out, dealing with areas such as mobility and activities of daily living, as well as health, economy, family, political resources and leisure activities.
The Uppsala Birth Cohort Multigeneration Study (UBCoS): This study started in 2005 with combining data on a cohort of 14,192 males and females born in Uppsala from 1915-1929 (the Uppsala Birth Cohort: UBCoS) with information on descendants of the original cohort members obtained from routine registers.
The Swedish Work and Mortality Database (HSIA/WMB): This database is being used in the SBC. It has no separate website.

Aim
The aim of CHESS' research and education is to provide a broad knowledge on the links between society and health, including social policies and social stratification; psychobiological processes, stress and health behaviours; life course and developmental perspectives on health.

Financing
Although all separate websites acknowledge the funders of the original or ongoing studies, neither the CHESS website nor the other websites provide information about costs of data usage by third parties.

Good practice

---

[76] More information on the 'Blue card':
http://www.nvk.nl/Onderzoek/NSCK/AlgemeneinfoNSCK/Hetmeldingssysteem/tabid/278/language/nl-NL/Default.aspx

All data collections managed by CHESS provide ample imformation about the studies carried out. Especially SBC excels at the extensive Open Access provisioning of coding schemes et cetera, thereby allowing detailed interpretation of the surveys, the data, and the codebooks. It should be noted that the term "codebook" is used in a broader sense than usual: In the SBC context a codebook contains not just background information to the study, but also aggregated data, for instance in tabular form. This yields even richer information than usual.

Properties

Permissions and conditions of use

SBC: People interested in the data should contact the data manager.
SWEOLD: no information on the website
UBCoS: "Apply to the Multigen principle [sic] investigator by using the Research Proposal Form (available online). At least one member of the UBCoS Multigen team has to be listed as collaborator in the project. A scientific committee of the UBCoS Multigen study will review the proposal. If access to data are granted, the recipient must agree with terms of data handling and data security stated by the Multigen team. A pre-publication approval of the Multigen team is also needed before publication of data from this database." (Source: http://www2.chess.su.se/ubcosmg/info/access.htm)

Available metadata

SBC: This site excels in providing metadata under the heading "Resources", available in PDF format for download.
SWEOLD: For each of the three surveys the questionnaire and a codebook are available in PDF format for download.
UBCoS: Per data source – such as birth, conscript, or hospitalisation data – the website lists all variables. There seems to be no downloadable version.

Data formats

SBC, SWEOLD, and UBCoS: The format of the data cannot be derived from information on the respective websites.

Interoperability

SBC and UBCoS: The information in the databases is based on different ICD diagnostic codes for different periods.
SWEOLD: probably due to the fact that the data derive from interviews with citizens rather than professionals, the codes are not explicitly mapped to external code systems or vocabularies.

---

8. General Practice Research Database (GPRD, United Kingdom)

Website: http://www.gprd.com/
Questionnaire received: not applicable: the questionnaire was only sent to Dutch registries

Introduction

"The GPRD is the world's largest computerised database of anonymised longitudinal medical records from primary care that is linked with other healthcare data. Currently data are being collected on about 5 million active patients of research standard. These are from around 630 primary care

practices throughout the UK. The database is managed by the GPRD Group within the Medicines and Healthcare products Regulatory Agency[77], the UK's medicines and devices regulator."

Aim
The GPRD Division provides data, research and other services as well as tools to support medical and public health research in a variety of areas including clinical research planning, drug utilisation, studies of treatment patterns, clinical epidemiology, and drug safety.

Financing
The GPRD is operated on a self-financing not for-profit basis and the data are licenced for use exclusively for medical and health research purposes on a non-profit making basis. The cost of providing access to data varies with the complexity of the request (http://www.gprd.com/academia/servicescost.asp).

Good practice
The application procedure for data usage is thorough and explicit and includes a protocol review appeals process, forms and instructions. An independent advisory committee evaluates data usage requests. GPRD is also very explicit about what they expect from data contributors (general practitioners), which is laid down in GPRD Data Recording Guidelines (available for download). GPRD offers external record linkage to other National Health Services datasets. The lists of publications using the GPRD is update quarterly.

Properties
Permissions and conditions of use
Source: http://www.gprd.com/ISAC/default.asp
Anyone interested in conducting a study using GPRD data are required to gain approval from the MHRA's Independent Scientific Advisory Committee for MHRA database research (ISAC). Proposed studies should be submitted as a protocol to the ISAC secretariat (instruction and form are available). Where investigators wish to send a questionnaire to GPs as part of a study, this must also be submitted to ISAC.
ISAC members assess aspects such as protection of practice and patient confidentiality, whether there is a well-defined hypothesis and the GPRD is a suitable database for it, as well as the appropriateness and the clarity of the questionnaire. They provide their comments to the Chairman and through him to the secretariat.
ISAC may make proposals for modifications to study protocols/ questionnaires as a condition of approval.
ISAC requests that on publication, a copy of the published article is provided to the ISAC secretariat.
Available metadata
For each patient GPRD collects and makes available (source: http://www.gprd.com/academia/primarycare.asp):
Demographic information, including gender, year of birth, ethnicity and practice location (to Strategic Health Authority level).
All clinical information including diagnoses, symptoms, procedures, and medical history.
All prescriptions issued both acute and repeat, with dosage instructions (which can be translated to numeric form).
Referrals to secondary care including hospital specialty, urgency, and nature of the referral (e.g. day case).
Immunisation details, including status, stage, and type, route of administration, reason and batch number.

[77] http://www.mhra.gov.uk/

Tests results including qualitative and quantitative test result values and also normal ranges for the laboratory.
Lifestyle information including BMI, height, weight and details on smoking and alcohol consumption.
Patient registration details including historic registration details, which are used to generate start and end dates of longitudinal electronic recording for person-time calculations.
Appointment and staff details including duration of consultation and gender and role of the health professional concerned.
Adverse drug reaction details including certainty and severity assessments for ADRs and drug intolerance and allergy.
Data formats
GPRD's website does not mention in which file format(s) data are provided once permission to access the data has been granted. From our own experience we know, however, that GPRD supports formats such as Excel and SPSS.

Interoperability
Record linkage of GPRD data are possible to other NHS datasets: national data, hospitalisation records, disease registries, and socio-economic class data. "GPRD has gained ethics, scientific and confidentiality approval to enable record linkage of GPRD data with other healthcare datasets via the patient's NHS number, sex, date of birth and Post Code. The linkage is done by an external NHS group in a way that GPRD does not see the identifying details. The additional data are returned using the GPRD anonymised research level identifier." (Source: http://www.gprd.com/products/links.asp)
"GPRD has formed an alliance with IMS Health that enables both parties' clients to have easier access to anonymised, person-level data from both Europe and the USA. Data are currently available for the UK, France, Germany and the USA." (Source: http://www.gprd.com/products/global.asp)
"Clinical data recorded in primary care are coded using Clinical Terms version 3 terminology (Read Terms, CTV3[78]). GPRD provides a regularly updated dictionary based upon the Read Terms used within GPRD data. Prescribing data within GPRD is coded using the Multilex Product dictionary[79]. A regularly updated drug dictionary based upon the Multilex products used within GPRD data are available." (Source: http://www.gprd.com/services/online.asp)

---

9. Medical Research Council (MRC, United Kingdom)

Website: http://www.mrc.ac.uk/
Questionnaire received: not applicable: the questionnaire was only sent to Dutch registries

Introduction
The Medical Research Council is research organisation dedicated to improving human health. The research portfolio includes areas such as cancer, mental health, and neurology. The MRC also funds research through funding of research centres in partnership with universities and a range of grants including so-called career awards to scientists in universities and hospitals.

Aim
The MRC's mission is to
Encourage and support research to improve human health.

[78] http://www.connectingforhealth.nhs.uk/systemsandservices/data/uktc/readcodes
[79] http://www.firstdatabank.co.uk/8/multilex-drug-data-file

Produce skilled researchers.
Advance and disseminate knowledge and technology to improve the quality of life and economic competitiveness of the UK.
Promote dialogue with the public about medical research.
Financing
The MRC receives annual 'grant-in-aid' funding from Parliament. Although government-funded, the MRC is independent in its choice of which research to support. The costs of data re-use are unknown.

Good practice
The MRC provides a Research Data Gateway "for finding MRC-funded population and patient studies, datasets and variables; and for finding the procedures for requesting access to data[80]." The site presents, for example, metadata regarding study, time period, data collection event, variable, and contact. Note that the Gateway is a portal rather than a registry. The MRC also sets great store by knowledge transfer. This is the exploitation of research findings to benefit society – financially, or regarding health, the environment or public services. One of the instruments is a Public Panel consisting of lay people from all walks of life who have an interest in different aspects of medical research.

Properties
Permissions and conditions of use
The Gateway directory does not provide one with direct access to research data. Interested parties have to contact the relevant study directly if they wish to find out more about the datasets described in this directory, including the arrangements for requesting access to the data for research purposes.
Available metadata
For each study MRC presents the following (Source: https://www.datagateway.mrc.ac.uk/node/14):
Study: a programme of research whereby data from and about individuals representing a population group are collected and analysed.
Time Period: a wave, sweep, phase or time point within a longitudinal study
Data collection event: a survey, screening, interview series or clininc, being the event through which research data were collected; there can be various data collections events within a phase.
Variable: each data variable belongs to a particular study, phase and collection event
Contact: point of contact for a study
Resource: an information item for a study, e.g. a questionnaire form, report etc.
Data formats
MRC's website does not mention in which file format(s) data are provided once permission to access the data has been granted. This might differ across the studies disclosed via the Gateway.

Interoperability
The Gateway does not provide information about hinges for interoperability.

[80] https://www.datagateway.mrc.ac.uk/node/14