Record linkage to enhance data from perinatal registries
Tromp, M.

Citation for published version (APA):
Tromp, M. (2009). Record linkage to enhance data from perinatal registries

General rights
It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations
If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: http://uba.uva.nl/en/contact, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.
Introduction
Introduction

Perinatal health care concerns pregnancy, childbirth and the postpartum period; it is a good example of a multidisciplinary field in health care. Intensive collaboration between different disciplines like midwives, general practitioners, gynaecologists, paediatricians and neonatologists is regarded essential to achieve optimal care for the pregnant woman and her child(ren). This collaboration is commonly denoted as ‘chain’. To evaluate the performance of this perinatal care chain, data from all the separate links – that is the professional actors involved - have to be available in a connected form. In the Netherlands, the 4 different key disciplines involved in perinatal care all have their own registry in which they collect data on the perinatal care process and its outcomes. These medical registries are independent and no national register of pregnant women and their children was available. Generally, medical registries can address an array of clinically relevant questions (aetiology, prognosis, prevention, quality of care, treatment and costs). However, if care essentially rests on a professional chain – as here – separate profession based registries are unable to produce complete and correct figures on indicators of care, in this case perinatal care in the Netherlands. If data about a pregnancy, childbirth and the postpartum period are combined, double counting of pregnancies or incomplete perinatal outcome reports can be prevented, and the referral and admission process becomes visible. The proficiency of registry based epidemiology is apparent in the field of obstetrics and paediatrics as adverse events are relatively rare and outcomes cover a wide time horizon ranging from intrauterine (foetal death) to adult cardiovascular disease (Barker’s hypothesis).

Medical record linkage

In the absence of a unique identifier, record linkage techniques can be applied to combine datasets containing data of the same person or entity. In record linkage, a set of partially identifying variables such as name, address, date of birth and gender are combined to a powerful discriminating system to identify records that belong to the same person. Two general approaches to record linkage exist: deterministic record linkage (DRL) and probabilistic record linkage (PRL). In deterministic linkage the decision to quantify a record pair as link or non-link is based on a predefined number of variables (mostly all or all but one) that must agree among both records. Probabilistic record linkage is a more sophisticated method where a weight is assigned to each linking variable depending on the likelihood of agreement among matches and among non-matches. The total weight for a
record pair has to be above a certain threshold before a record pair is classified as a link. This threshold too is established by the linkage procedure.

Perinatal health care
Perinatal health care encompasses the entire period from first antenatal visit up to and including the first weeks of life of the newborn. The Dutch system of maternity care is different from other Western countries, based upon risk management with different caregivers according to estimated risk class. Primary obstetric care for the low risk group is provided by midwives and general practitioners with the option of homebirth or hospital birth. Women with an increased risk deliver under supervision of a gynaecologist. Referral as (candidate) high risk pregnancy is based on a guideline for consultation or referral agreed upon by both professions. In addition, neonatologists and paediatricians are involved if a child needs care after birth.

An important indicator of perinatal health care is perinatal mortality, defined as mortality from 22.0 weeks gestation onwards including the first week of life (0-6 days). Besides mortality, preterm birth (gestational age < 37.0 weeks), small for gestational age (< 10th percentile) and low APGAR score (APGAR score at 5 minutes < 7) are relevant outcome measures in perinatal health care. Due to the fact that women and children can be present in more than one registry, these outcome measures could not be determined straightforward on a national level in the Netherlands.

Registries in perinatal health care in the Netherlands
Information on (obstetric) history, pregnancy, delivery, postpartum period and provided care, is collected by caregivers according to the national professional protocols. The national registers on perinatal care include the national obstetric database for midwives (LVR-1), the national obstetric database for gynaecologists (LVR-2) and the national neonatal database for paediatricians and neonatologists (LNR) (general practitioners started registering in 2007). Since 2001 the profession based registries are assembled in the Netherlands Perinatal Registry (PRN). The purpose of the PRN is to improve the quality of care by collecting and analyzing data encompassing the whole perinatal care process. To enable these analyses on a national level, the individual perinatal registries (LVR-1, LVR-2, LNR) had to be linked. In 2003, the registers of the year 2001 had first been linked using deterministic and probabilistic strategies under the authority of the PRN. More information on the linkage procedure can be found in the appendix of this thesis.

Objective of this thesis
This thesis has two objectives:
1) To validate and refine existing record linkage techniques to optimally combine data from separate registries in perinatal care in the Netherlands.
2) To demonstrate the added value of the linked national dataset to answer relevant clinical or epidemiological questions on perinatal care.
Introduction

Research questions
The first part of this thesis focuses on methodological questions to improve the quality of a record linkage project:
- What is the difference in performance between deterministic and probabilistic linkage strategies under varying conditions?
- How to optimize the performance of probabilistic record linkage strategies in situations with different types of error in linking variables?
- How to handle dependencies among linking variables?
- How to validate a record linkage procedure?

The second part of this thesis describes a set of clinical relevant questions that can be answered using the linked perinatal datasets:
- How high is the perinatal mortality in the Netherlands, and is there a trend?
- Are there regional variations in perinatal mortality within the Netherlands and can these differences be explained?
- What is the influence of increasing maternal age on adverse outcomes?
- What is the added value of linking cohort data with national registry data?

Outline
After this introductory chapter, the first chapter of this thesis discusses the difference in performance between different linking techniques (probabilistic and deterministic) under varying conditions. Chapter 2 presents an extension of the standard probabilistic model to accommodate for frequently made (non random) errors in variables. The assumption of independence among linking variables in probabilistic linkage and a method to deal with violation of this assumption is presented in chapter 3. In the next chapter (chapter 4) an efficient and informative approach is described for validating the results of a probabilistic linkage where the selection of pairs for external validation is guided by the outcome of the probabilistic approach. The successful linkage of the Dutch perinatal registries produced a rich dataset available for research. A few examples of the type of research questions that can be examined with the linked datasets are presented in the next chapters. In chapter 5 an analysis of the trend over time in perinatal mortality in the Netherlands is presented. Chapter 6 presents the results of a study examining regional variation in perinatal mortality within the Netherlands. The consequences of the increasing age at which women become pregnant of their first child on pregnancy outcomes are presented in chapter 7. In chapter 8 record linkage is used to analyze non-response within a large cohort study (the ABCD-study) by linking it to the integrated national register of all pregnancies. This thesis ends with a general discussion.
Introduction

References