

# Use of Electronic Medical Records in the Epidemiological Research

Silvia Dombrowski  
Karel Kostev



**Cuvillier Verlag Göttingen**  
Internationaler wissenschaftlicher Fachverlag



## Use of Electronic Medical Records in the Epidemiological Research





# Use of Electronic Medical Records in the Epidemiological Research

Silvia Dombrowski, Karel Kostev



### **Bibliographical information held by the German National Library**

The German National Library has listed this book in the Deutsche Nationalbibliografie (German national bibliography); detailed bibliographic information is available online at <http://dnb.d-nb.de>.

1<sup>st</sup> edition - Göttingen: Cuvillier, 2017

© CUVILLIER VERLAG, Göttingen, Germany 2017

Nonnenstieg 8, 37075 Göttingen, Germany

Telephone: +49 (0)551-54724-0

Telefax: +49 (0)551-54724-21

[www.cuvillier.de](http://www.cuvillier.de)

All rights reserved. This publication may not be reproduced by photomechanical means (photocopying, microfiche), in whole or in part, without the prior express permission of the publisher.

1<sup>st</sup> edition, 2017

This publication is printed on acid-free paper.

ISBN 978-3-7369-9500-0

eISBN 978-3-7369-8500-1



# Table of contents

Introduction.....	1
Chapter 1. Diabetes .....	5
Chapter 2. Cancer .....	30
Chapter 3. Cardiovascular diseases .....	42
Chapter 4. Osteoporosis .....	46
Chapter 5. Dementia .....	52
Chapter 6. Depression .....	58
Chapter 7. Neurologic diseases .....	66
Chapter 8. Obstetrics and Gynecology .....	70
Chapter 9. Children's Health .....	78
Chapter 10. Allergy.....	82
Chapter 11. Limitations .....	86



# Introduction

Medical record databases are longitudinal patient record databases that are used by health care providers in caring for their patients and that are anonymized for the purpose of research. Data from such retrospective databases allow for investigations into specific subpopulations – e.g., groups with specific diagnoses – thanks to their size and duration of observation. Regarding the quality of such data, it has been proven that carefully planned observational studies can produce results comparable to those of randomized controlled trials.

Patient and health care databases are available in many countries and are often based on routinely collected diagnosis and prescription data. Over time, patient data from such databases have been linked with each other via pseudonyms and then analyzed. In Germany, examples of these databases include not only several statutory health insurance (SHI) databases but also commercial databases like the QuintilesIMS Disease Analyzer database. These databases can be used to evaluate important questions concerning health services, such as whether therapy regimens being applied reflect the current state of scientific knowledge or whether supply shortages, surpluses, or mismatches occur. Using these databases, numerous studies have been conducted to analyze the duration, adverse effects, success, costs, and courses of and compliance with therapies and therapy changes. These studies also play an important role in drug safety and risk prevention. A sufficiently valid database is required in order to be able to guarantee the scientific relevance of epidemiological studies.

Data from German SHI bodies have been identified as an important data source for pharmacoepidemiological studies (Hoffmann, 2009), but so far, only a few German data sources have been presented transparently to the scientific community (Pigeot & Ahrens, 2008). Andersohn and Walker were able to show the good overall agreement between the SHI



database and the German population in terms of morbidity, mortality, and drug usage. The demographic structure of insurants was slightly different than that among the German population, with the database population being younger and with eastern parts of Germany being underrepresented. There was a high persistence of insurants with the database over time, indicating suitability of the data source for longitudinal epidemiological analyses (Andersohn & Walker, 2016).

QuintilesIMS Disease Analyzer is one of the major European patient databases. It contains data from Germany, the UK, and France and allows for anonymous access to a selected panel of physicians' practices and patients. The data are generated directly from the computers in the physicians' practices via standardized interfaces and provide daily routine information regarding patients' diseases and therapies. A practice transmits patient data stored in the physician's computer to IMS on a monthly basis. Before transmission, the data are encrypted for data protection purposes and contain in similar scope and detail the information in the files of patients in the doctor's practice. Patients and practices can be analyzed in a cross-sectional and longitudinal fashion. In Germany, the database contains data from more than 2,000 practices and more than 20 million patients. In addition to data from general practitioners and specialists in internal medicine, data for various specialist groups are also recorded in Germany. The database includes only anonymized data in compliance with the regulations of the applicable data protection laws.

The sampling method for the Disease Analyzer database is based on summary statistics from all physicians in Germany published every year by the German Medical Association. The statistical unit of IMS uses these statistics to determine the panel design according to the following categories: specialist group, German federal state, community size category, and age of physician.

This panel design forms the basis for the acquisition of the practices processed in the Disease Analyzer. The acquisition of and support for the practices is performed by cooperating software companies using a standardized interface that enables the practices to collect the required data and send them to IMS in an anonymized form. To account for natural fluctuations in the practices and an annual check of the summary statistics by the German Medical Association, the panel design is adjusted each year. Whenever a practice ends its collaboration with IMS, it is replaced by a new one. Altogether, eleven specialist fields are taken into account in the random sampling plan. For this purpose, the field of internal medicine has been subdivided into five subgroups. Furthermore, the field of neurology also includes pediatric and adolescent psychiatrists.

The sampling plan is subdivided into eight regions, which are summaries of the 16 German federal states. This stratification results in 176 cells derived from the summary statistics with regard to specialist fields and proportional to the summary statistics with regard to the German federal states. Within each specialist field, at least 30 doctors must be sampled. Within each region, a minimum of seven physicians must be sampled within each specialist field to allow for estimates at the specialist field level for each region (Ogdie et al., 2012).

The main strength of studies based on the Disease Analyzer database is the large number of patients available for analysis. Another strength is the use of real-world data in primary care practices where diagnoses are continuously documented, allowing for an unbiased exposure assessment without recall bias.

The Disease Analyzer database has been the basis of a large number of peer-reviewed scientific publications in the fields of epidemiology, health economics, pharmacovigilance, compliance/persistence, pharmaceutical guidelines, prescribing behavior, and drug application. This book