EUROPEAN MEDICINES AGENCY
SCIENCE MEDICINES HEALTH

European Network of Centres for Pharmacoepidemiology and Pharmacovigilance

EMA/95098/2010 Rev.8

# The European Network of Centres for Pharmacoepidemiology and Pharmacovigilance (ENCePP) Guide on Methodological Standards in Pharmacoepidemiology

(Revision 8)



| | |
|---|---|
| **KEYWORDS** | *methodological standards, pharmacoepidemiology, pharmacovigilance, ENCePP, research, guidance* |



**This document should be cited as follows**: The European Network of Centres for Pharmacoepidemiology and Pharmacovigilance (ENCePP). Guide on Methodological Standards in Pharmacoepidemiology (Revision 8). EMA/95098/2010.

Available at http://www.encepp.eu/standards_and_guidance

**Official address** Domenico Scarlattilaan 6 ● 1083 HS Amsterdam ● The Netherlands
**Address for visits and deliveries** Refer to www.ema.europa.eu/how-to-find-us
**Send us a question** Go to www.ema.europa.eu/contact **Telephone** +31 (0)88 781 6000 An agency of the European Union

# Table of contents

# Foreword to 8th Revision

**ENCePP Guide supports strong observational research for the COVID-19 pandemic**

The rapid progression of the COVID-19 pandemic has generated several hypotheses on the safety and effectiveness of therapeutic interventions, such as repurposed medicines. The need for quick answers triggered the initiation of observational studies carried-out with fast data collection, analysis and reporting. In a pandemic situation, the same methodological standards as those applied in any other circumstance should nevertheless be used to provide valid and reliable evidence supporting rapid treatment decisions by clinicians and regulators. Adherence to existing guidance on the appropriate design and conduct of pharmacoepidemiologic studies is therefore of utmost importance. ENCePP believes that this 8th Revision of the Guide on Methodological Standards in Pharmacoepidemiology should be the backdrop against which observational studies related to the COVID-19 pandemic should be conducted.

Pottegård et al. provide methodological considerations for the conduct of pharmacoepidemiological studies in relation to the COVID-19 pandemic across eight domains. The ENCePP Guide addresses each of these domains: (1) timeliness of evidence generation, including the need to prioritise some questions over others in the acute phase of the pandemic (addressed in Chapter 2); (2) the need to align observational and interventional research on efficacy (Chapter 10.1); (3) the specific challenges related to "real-time epidemiology" during an ongoing pandemic (Chapters 4.1, 4.2 and 4.3); (4) what design to use to answer a specific question (Chapter 5); (5) considerations on the definition of exposures (Chapter 5.1); (6) what covariates to collect (Chapter 5.1); (7) considerations on the definition of outcomes (Chapter 5.1); and (8) the need for transparent reporting (Chapter 8).

The methodological challenges described by Pottegård et al. are illustrated by studies that examined the differences in the incidence and severity of the SARS-CoV -2 virus infection between patients receiving angiotensin converting enzyme inhibitors (ACEi) or angiotensin receptor blockers (ARB) and those not receiving ACEi/ARB, which may help to inform hypertension treatment decisions. In these studies, the subset of patients being tested was not a random and unbiased sample of the total population, which may lead to the selection bias described in Chapter 5.2. Patients with symptoms and with comorbidities, including hypertension, may have been more likely to be tested, therefore influencing the testing probability of patients exposed to ACEi/ARB. Although some of these studies have adjusted for relevant comorbidities and other potential confounding variables, some factors that may have affected the testing probability and the risk of being tested positive may have been unmeasured or uncontrolled for. Amongst the studies referenced in an EMA Press Release (June 2020), five large and well-conducted studies were considered adequate to assess this risk (de Abajo et al., Gnavi et al., Mancia et al., Mehta et al. and Reynolds et al.).

The risk of a severe outcome of infection in patients exposed to ACEi/ARB has been measured in several studies by mortality, admission to intensive care unit (ICU) or need for respiratory ventilation in hospitalised patients. A common issue in these studies is the potential for selection bias, due to factors associated with hospitalisation or admission to ICU, and for time-related bias due to misclassification of the observation time in different treatment groups. Another frequent limitation is the presence of unmeasured or unadjusted confounding factors, for example factors associated to the prescription of ACEi or ARB, or to higher risk of patient hospitalisation or patient treatment with ventilation. Several studies attempted to adjust for confounding by indication but potential issues have included the limited number of variables introduced in the model and the timing of the assessment of the relevant variables. Chapters 5.2 on bias and 5.3 on methods to address bias offer guidance to identify and address these sources of errors in observational studies. The study by Yung et al. is

considered to be a good example of a study providing valid results between ACEi/ARB and the risk of mortality in patients hospitalised with SARS-CoV -2 virus infection.

In the context of a pandemic where rapid answers to research questions are needed, combining data across different databases affords insight into the generalisability of the results and may improve precision if outcomes or exposure of interest are rare or when there is interest in subgroup effects. It may also inform on specific patterns of drug utilisation. Chapter 4.6. on research networks for multi-database studies describes the different models that can be applied for combining data or results from multiple databases. An example of a collaboration in the context of the COVID-19 pandemic is the study published by Lane et al. using data from 14 multinational sources of claims data or electronic medical records. This study followed the model of a general common data model (CDM) described in Chapter 4.6.2.5.

During the COVID-19 pandemic, many studies were published on the effects of a specific drug or drug class. Systematic reviews and meta-analyses were subsequently performed to provide summaries of their results but some of them lacked the methodological rigor needed for the selection and review of studies and the appropriate statistical methods to pool estimates from individual studies if a meta-analysis was conducted. Annex 1 of this Guide provides a Guidance on conducting systematic reviews and meta-analyses of completed comparative pharmacoepidemiologic studies of safety outcomes, and it may serve as a helpful tool to generate valid conclusions from systematic reviews and meta-analyses.

The impact of the COVID-19 pandemic has accelerated the development of vaccines. By June 2020, the first COVID-19 vaccine candidates have entered human clinical testing and might be available rapidly with potentially remaining questions and the need for close monitoring. Chapter 10.2. on vaccine safety and effectiveness helps developing such studies.

As stated by Watson et al.in relation to one of the published studies, lack of transparency and uncertainties about research standards applied raise doubts about published results. Morales et al. supported the reproducibility of their study by publishing the study protocol in the EU PAS Register ahead of time, providing a start-to-finish executable code, facilitating the sharing and exploration of the complete result set with an interactive web application and asking clinicians and epidemiologists to perform a blinded evaluation of propensity score diagnostics for the treatment comparisons.

# 1. Introduction

The choice of epidemiological methods to answer a research question should be based on principles and methodological standards supporting the validity of the study results. There are many textbooks describing methodological standards in pharmacoepidemiology but they cannot incorporate all new developments. ENCePP therefore considered there was a need for a regularly updated resource providing recommendations on the practical implementation of pharmacoepidemiological principles and innovative methods, based on published guidance and illustrative examples.

This Guide aims to offer a dynamic and publicly available web resource for methodological English language guidance in pharmacoepidemiology. It provides links to selected published articles and guidelines that illustrate important principles of pharmacoepidemiological research. For each topic covered, recommendations are provided with direct electronic access to textbooks, reference documents and examples selected by experts from ENCePP. Where relevant, gaps in existing guidance are addressed with what ENCePP considers as being good practice. The Guide is updated annually by a structured review in order to maintain its dynamic nature. It may also be amended as necessary in response to comments received. For this purpose, any comment and additional relevant guidance document may be forwarded to ENCePP_Secretariat@ema.europa.eu.

The Guide does not discuss general methods of pharmacoepidemiology as they are already covered in existing textbooks. For example, it does not describe traditional study designs like the cross-sectional, cohort and case-control designs but it discusses important aspects of more recent designs such as the self-controlled case series (SCCS) design. Chapters 2 to 9 describe methods that may be used in studies with different goals, be they safety, effectiveness, drug utilisation, health technology assessment or any other objective. For some specific topics, specific recommendations and references are provided where they differ from general principles. Chapter 10 provides more extensive guidance on three specific topics, comparative effectiveness research, vaccine safety and effectiveness and pharmacogenetic studies.

Annex 1 has been developed separately by an ENCePP working group and provides methodological guidance addressing the conduct of systematic reviews and meta-analyses of drug safety endpoints.

Annex 2 has been developed by an ENCePP Special Interest Group and provides recommendations on methods for measuring the impact of pharmacovigilance activities on patients and public health.

General guidance on the conduct of pharmacoepidemiology studies can be found in the ISPE Good Pharmacoepidemiology Practices (GPP) and the IEA Good Epidemiology Practice (GEP). The GPP guidance is especially useful for its recommendations on aspects rarely covered by guidelines, such as data quality issues and archiving. The Guidelines and recommendations for ensuring Good Epidemiological Practice (GEP): a guideline developed by the German Society for Epidemiology (Eur J Epidemiol. 2019;34(3):301-17) provides detailed recommendations addressed to everyone involved in the planning, preparation, execution, analysis, and evaluation of epidemiological research and is relevant for the evaluation of medicines.

The Guideline of good pharmacovigilance practices (GVP) Module VIII - Post-authorisation safety studies provides a general guidance on the development, conduct and reporting of post-authorisation safety studies (PASS) conducted by marketing authorisation holders voluntarily or pursuant to the EU legislation (Directive 2001/83/EC). It also describes the criteria applicable in the European Union (EU) to define a post-authorisation study as non-interventional, but investigators should be aware that implementation of these criteria may vary at national level in different EU countries. The Scientific guidance on post-authorisation efficacy studies provides general scientific guidance in the context of EU regulatory decision-making with regard to the need for such studies and methodological considerations.

The terms "Real-world data" (RWD) and "Real-world evidence" (RWE) are increasingly used in the regulatory setting to denote the secondary use of observational data and pharmacoepidemiological methods for regulatory decision-making. The article Real-World Data for Regulatory Decision Making: Challenges and Possible Solutions for Europe (Clin Pharmacol Ther. 2019;106(1):36-9) defines RWD as "routinely collected data relating to a patient's health status or the delivery of health care from a variety of sources other than traditional clinical trials", and RWE as "the information derived from the analysis of RWD". The article introduces the OPTIMAL framework to describe the operational, technical and methodological challenges for the acceptability of RWE for regulatory purposes and presents possible solutions to address these challenges, including the use of best methodological standards in statistics and epidemiology. The FDA's Real-World Evidence website also provides definitions and links to a set of useful guidelines on the submission and use of RWD and RWE to support decision-making.

Textbooks on standard methods in pharmacoepidemiology that are considered useful are listed below. The list is not exhaustive, and researchers may find other textbooks more appropriate to their specific needs. Others are cited in specific chapters.

- Modern Epidemiology, Fourth Edition (K. Rothman, S. Greenland, T. Lash. Lippincott Williams & Wilkins, 2020) is a comprehensive textbook on methods in epidemiology.

- Epidemiology: Study Design and Data Analysis, Third Edition (M. Woodward, Chapman & Hall, 2014) focuses on the quantitative aspects of epidemiological research.
- A Dictionary of Epidemiology, Sixth Edition (M Porta, Editor. Oxford University Press, 2014), sponsored by the International Epidemiological Association (IEA), provides a definition and concise explanation of epidemiologic terms and is a key to understanding epidemiological concepts.
- Clinical epidemiology: practice and methods, Second Edition (PS Parfrey, BJ Barret, Human Press, 2015) focuses on the diagnosis, prognosis and management of human disease using appropriate research design, measurement and evaluation.
- Pharmacoepidemiology, Sixth Edition (B. Strom, S.E. Kimmel, S. Hennessy, Wiley, 2019) provides a comprehensive guidance on pharmacoepidemiology addressing data sources, applications and methodologies.
- Pharmacoepidemiology and Therapeutic Risk Management, First Edition (A.G. Hartzema, H.H. Tilson and K.A. Chan, Editors, Harvey Whitney Books Company, 2008) illustrates practical issues with a large number of real life examples in addition to a general review of drug-specific methodologies.
- Practical Statistics for Medical Research, Second Edition (D. Altman. Chapman & Hall, 2020) presents a problem-based statistical text for medical researchers.
- Drug Utilization Research. Methods and Applications (M Elseviers, B Wettermark, AB Almarsdóttir, et al. Editors. Wiley Blackwell, 2016) provides a comprehensive manual of methodology and applications of drug utilisation research.
- Mann's Pharmacovigilance, Third Edition (EB Andrews, N Moore, Editors, Wiley-Blackwell, 2014) is a reference for the science of detection, assessment, understanding and prevention of the adverse effects of medicines, including vaccines and biologics.
- Post-Authorization Safety Studies of Medicinal Products. The PASS Book, 1st Edition (Ayad Ali, Abraham Hartzema, Ed., Academic Press, 2018) covers the use of observational studies in post-marketing drug safety assessment, presents various types of post-authorisation safety studies and discusses challenges and solutions in the design and conduct of these studies.

# 2. Formulating the research question

Generating evidence involves three steps: asking the right research questions, finding or collecting fit-for-purpose data, and conducting the appropriate analyses. The first step in any research is to formulate the research question clearly and accurately. The research question should stem from the problem or gap in knowledge to be addressed and should be supported by a theoretical framework. The research question should be formulated in collaboration with the primary end-users of the study results and state who will be these end-users, e.g. patients, health care professionals, regulators or public health authorities, health technology assessment organisations, payers, pharmaceutical company or research community. It should state the 'why' (main justification for starting the research), the 'what' (exposure and endpoints), the 'who' (target population), the 'how' (main study design) and the 'when' (time period of the study) of the research in a way that helps understanding the choice of study objectives and methods. It should make it clear whether a hypothesis will be tested and, in this case, whether the hypothesis is pre-specified or data driven.

Previous findings are useful for the methodological planning of the current study and may support the background, research question, hypotheses and design of the proposed study. They may also serve to determine the expected effect size and, if available in the target population, to characterise risk factors

for the event, to identify relevant outcomes and measures and to assess the feasibility of the proposed study. A critical and thorough review of the literature forms the basis for the theoretical framework of the research question and should usually be included in the background section of a protocol. Such a review aims to evaluate current evidence around the question at hand and identify gaps in knowledge that a study is intended to fill.

How to formulate research recommendations (BMJ 2006;333:804-6) proposes the EPICOT format with 5 core elements for research recommendations on the effects of treatments: Evidence (source of the current evidence), Population (population characterised by any diagnosis, disease stage, comorbidity, risk factor, sex, age, ethnic group, specific inclusion or exclusion criteria, clinical setting), Intervention (type, frequency, dose, duration, prognostic factor), Comparison (placebo, routine care, alternative treatment/management), Outcome (which clinical or patient related outcomes will the researcher need to measure, improve, influence or accomplish; which methods of measurement should be used), and Time stamp (date of literature search or recommendation). This format was adopted by the European network of Health Technology Assessment (EUnetHTA) in its Position paper on how to formulate research recommendations. Chapter 5.7 and Annex 1 present methods for reviewing and synthesising findings from the literature through the means of systematic review and meta-analysis.

Research questions relevant to regulatory authorities and health technology assessment bodies regarding the utilisation, safety, efficacy and impact of medicines are detailed the European Public Assessment Report (EPAR) available for each centrally authorised product on the EMA website, with general pharmacovigilance related aspects being described in Modules of the Good Pharmacovigilance Practices (GVP), and The criteria to select and prioritise health technologies for additional evidence generation document.

When the study data source is not well characterised or known, a feasibility study should be considered. The aim of a feasibility study is not to answer the research question directly but to determine whether the data source could answer it within the expected timelines and what is the required statistical power for the proposed study design. Feasibility studies can provide information on the number of people with a specific exposure or outcome, the availability of covariates and the follow up period needed. A feasibility study can also provide insights into the potential difficulties which may be encountered in the conduct of the study or which may introduce bias. Importance of feasibility assessments before implementing non-interventional pharmacoepidemiologic studies of vaccines: lessons learned and recommendations for future studies (Pharmacoepidemiol Drug Saf. 2016;25(12):1397-406) illustrates a vaccine manufacturer's pragmatic approach for conducting feasibility assessments for post-authorisation studies required to address regulatory requests. The ISPE Good pharmacoepidemiology practice (GPP) explains how a data collection method or data source can answer a research question with justifications coming from the feasibility study when relevant.

# 3. Development of the study protocol

The study protocol is the core document of a study that should be drafted as one of the first steps in any research project once the research question has been clearly defined. The final version must precisely describe everything being done in the study to ensure reproducibility of the study. The protocol should be amended as needed and amendments should be justified.

For PASS described in the GVP Module VIII - Post-authorisation safety studies, the Commission Implementing Regulation (EU) No 520/2012 provides legal definitions of the start of data collection (the date from which information on the first study subject is first recorded in the study dataset, or, in the case of secondary use of data, the date from which data extraction starts) and of the end of data collection (the date from which the analytical dataset is completely available). These dates provide

time references for the commencement of the study and the submission of the final study report to the competent authorities. It also provides the format of protocols, abstracts and final study reports for imposed PASS. Based on these formats, the EMA published detailed templates for the protocol and final study report which it recommends to be used for all PASS, including meta-analyses and systematic reviews. The ISPE Guidelines for Good Pharmacoepidemiology Practices (GPP) provides guidance on what is expected from a pharmacoepidemiology study protocol and on the different aspects to be covered. It states that the protocol should include a description of the data quality and integrity, including abstraction of original documents, extent of source data verification, and validation of endpoints. The FDA's Best Practices for Conducting and Reporting Pharmacoepidemiologic Safety Studies Using Electronic Health Care Data Sets includes a description of the elements that should be addressed in the protocols of such studies, including the choice of data sources and study population, the study design and the analyses. The ENCePP Checklist for Study Protocols seeks to stimulate researchers to consider important epidemiological aspects when designing a pharmacoepidemiological study and writing a study protocol. The Agency for Healthcare Research and Quality (AHRQ) published Developing a Protocol for Observational Comparative Effectiveness Research: A User's Guide including best practice principles and checklists on a wide range of topics that are also applicable to observational studies outside the scope of comparative effectiveness research.

For studies involving human patients, consent form and special ethical guidelines apply, see Chapter 9.2.

GVP Module VIII - Post-authorisation safety studies provides a structure for study protocols, which should cover at least the following aspects:

- The research question the study is designed to answer, which might be purely descriptive, exploratory or explanatory (hypothesis driven). The protocol should include a background description that explains the origin (scientific, regulatory, etc.) and current knowledge on the research question. It will also explain the context of the research question, including what data are currently available and how these data can or cannot contribute to answering the question. The context will also be defined in terms of what information sources can be used to generate appropriate data and how the proposed study methodology will be shaped around these.
- The main study objective and possible secondary objectives, which are operational definitions of the research question. In defining secondary objectives, consideration could be given to time and cost, which may impose constraints and choices, for example in terms of sample size, duration of follow-up or data collection.
- The source and study population to be used to answer the research question. The protocol should describe whether this population is already identified, and whether data are already available (secondary data collection) or whether it needs to be recruited de novo (primary data collection). The boundaries of the desired population will be defined, including inclusion/exclusion criteria, timelines (such as index dates for inclusion in the study) and any exposure or events defining the population. Exposure of interest that needs to be pre-specified and defined, including duration and intensity of exposure, source of data and methods of ascertainment.
- Outcomes of interest that need to be pre-specified and defined, including data sources, operational definitions and methods of ascertainment such as data elements in field studies or appropriate codes in database studies.
- Adverse events/reactions that will or will not be collected and reported and the procedures put in place for this purpose. In the EU, the collection and reporting of adverse events or reactions by companies sponsoring a post-authorisation study should follow the recommendations specified in Module VI of the Guideline on good pharmacovigilance practice (GVP) - Management and reporting

of adverse reactions to medicinal products. If the study qualifies as an interventional trial, the reporting criteria laid down in Directive 2001/20/EC and Volume 10 of the Rules Governing Medicinal Products in the European Union should be followed.

- The covariates and potential confounders that need to be pre-specified and defined, including how they will be measured.
- The statistical plan for the analysis of the resulting data, including statistical methods and software, adjustment strategies, and how the results are going to be presented.
- The identification and way of minimisation of potential biases.
- Major assumptions, critical uncertainties and challenges in the design, conduct and interpretation of the results of the study given the research question and the data used.
- Ethical considerations, as described in Chapter 9.

Various data collection forms including the Case Report Form (CRF), list of disease codes or descriptions of the data elements may be appended to the protocol, providing an exact representation of how the data will be collected. The study protocols could include a section specifying ways in which the CRF will be piloted, tested and finalised. Amendments of final CRFs should be justified. For field studies, physician or patient forms would be included depending on the data collection methodology. Other forms may be included as needed, such as patient information, consent form or patient-oriented summaries.

Registration of the study protocol before the start of data collection provides information to other researchers about the ongoing study, improves transparency and, especially for studies based on secondary use of data, provides assurance that the stated hypotheses have not been influenced by the results. The EU PAS Register is a public register open to everyone for the registration of non-interventional studies.

# 4. Approaches to data collection

There are two main approaches for data collection: collection of data specifically for a particular study ('primary data collection') or use of data already collected for another purpose, e.g. as part of administrative records of patient health care ('secondary use of data). The distinction between primary data collection and secondary use of data is important for marketing authorisation holders as it implies different regulatory requirements for the collection and reporting of suspected adverse reactions, as described in Module VI of the Guideline on good pharmacovigilance practice (GVP) - Management and reporting of adverse reactions to medicinal products.

Secondary use of data has become a common approach used in pharmacoepidemiology due to the increasing availability of electronic healthcare records, administrative claims data and other already existing data sources (see Chapter 4.2 Secondary use of data) and due to its increased efficiency and lower cost. In addition, networking between centres active in pharmacoepidemiology and pharmacovigilance is rapidly changing the landscape of drug safety research in Europe, both in terms of networks of data and networks of researchers who can contribute to a particular study with a particular data source (see Chapter 4.6 Research Networks).

## *4.1. Primary data collection*

The methodological aspects of primary data collection studies are well covered in the textbooks and guidelines referred to in the Introduction chapter. Annex 1 of Module VIII of the Good

pharmacovigilance practice provides examples of different study designs based on prospective primary data collection such as cross-sectional study, prospective cohort study, active surveillance. Surveys and randomised controlled trials are also presented below as examples of primary data collection.

Studies using hospital or community-based primary data collection have allowed the evaluation of drug-disease associations for rare complex conditions that require very large source populations and in-depth case assessment by clinical experts. Classic examples are Appetite-Suppressant Drugs and the Risk of Primary Pulmonary Hypertension (N Engl J Med 1996;335:609-16), The design of a study of the drug etiology of agranulocytosis and aplastic anemia (Eur J Clin Pharmacol 1983;24:833-6) and Medication Use and the Risk of Stevens–Johnson Syndrome or Toxic Epidermal Necrolysis (N Engl J Med 1995;333:1600-8). For some conditions, case-control surveillance networks have been developed and used for selected studies and for signal generation and clarification, e.g. Signal generation and clarification: use of case-control data (Pharmacoepidemiol Drug Saf 2001;10:197-203).

### 4.1.1. Surveys

A survey is the collection of data on knowledge, attitudes, behaviour, practices, opinions, beliefs or feelings of selected groups of individuals, by asking them in person, on paper, by phone or online from some sampling frame. They generally have a cross-sectional design, but repeated measures overtime may apply for trends assessment.

Surveys have been used for a long time in fields such as marketing, social science and epidemiology. General guidance on constructing and testing the survey questionnaire, modes of data collection, sampling frames and ways to achieve representativeness can be found in general texts (*Survey Sampling* (L. Kish, Wiley, 1995) and *Survey Methodology* (R.M. Groves, F.J. Fowler, M.P. Couper et al., 2nd Edition, Wiley 2009). The book *Quality of Life: the assessment, analysis and interpretation of patient-related outcomes* (P.M. Fayers, D. Machin, 2nd Edition, Wiley, 2007) offers a comprehensive review of the theory and practice of developing, testing and analysing quality of life questionnaires in different settings.

Surveys have an important role in the evaluation of the effectiveness of risk minimisation measures (RMM) or of a risk evaluation and mitigation strategy (REMS) (see chapter 5.9). The application of methods described in these aforementioned textbooks needs adaptation for surveys to evaluate the effectiveness of RMM or REMS. For example, the extensive methods for questionnaire development of quality of life scales (construct, criterion and content validity, inter-rater and test-retest reliability, sensitivity and responsiveness) are not appropriate to questionnaires in RM which are often used only once. The EMA and FDA issued guidance documents on the conduct of surveys for RM which, together, encompass the selection of risk minimisation measures, study design, instrument development, data collection, processing and data analysis and presentation of results. This guidance include the EMA Guideline on good pharmacovigilance practices (GVP) Module XVI (2017), the FDA draft guidance for industry REMS Assessment: Planning and Reporting on REMS (2019) and the FDA Guidance on Survey Methodologies to Assess REMS Goals That Relate to Knowledge (2019).

A checklist to assess the quality of studies evaluating RM programs is provided in The RIMES Statement: A Checklist to Assess the Quality of Studies Evaluating Risk Minimization Programs for Medicinal Products (Drug Saf 2018;41(4): 389-401). The article Are Risk Minimization Measures for Approved Drugs in Europe Effective? A Systematic Review (Expert Opin Drug Saf 2019;18(5):443-54) highlights the need for improvement in the methods and presentation of results and for more hybrid designs that link survey data with health and safety outcomes as requested by regulators. This article also reports on low response rates found in many studies, allowing for the possibility of important bias. The response rate should therefore be reported in a standardised way in surveys to allow comparisons.

Standard Definitions. Final Dispositions of Case Codes and Outcome Rates for Surveys (2016) of the American Association for Public Opinion Research provides standard definitions which can be adapted to RM surveys and the FDA Guidance on Survey Methodologies to Assess REMS Goals That Relate to Knowledge (2019) provides guidance for RM surveys.

The increasing use of online RMM require that survey methods adapt but should not sacrifice representativeness by accessing only populations which visit these websites. They should provide evidence that the results using these sampling methods are not biased. Similarly, the increasing use of health care professional and patient panels needs to ensure that survey methods do not sacrifice representativeness by accessing only self-selected participants in these panels and should provide evidence that the results are not biased by using these convenient sampling frames.

### 4.1.2. Randomised clinical trials

Randomised clinical trials is an experimental design that involves primary data collection. There are numerous textbooks and publications on methodological and operational aspects of clinical trials and they are not covered here. An essential guideline on clinical trials is the European Medicines Agency (EMA) Guideline for good clinical practice E6(R2), which specifies obligations for the conduct of clinical trials to ensure that the data generated in the trial are valid. From a legal perspective, the Volume 10 of the Rules Governing Medicinal Products in the European Union contains all guidance and legislation relevant for conduct of clinical trials. A number of documents are under revision.

The way clinical trials are conducted in the European Union (EU) will undergo a major change when the Clinical Trial Regulation (Regulation (EU) No 536/2014) will fully come into effect and will replace the existing Directive 2001/20/EC.

Hybrid data collection as used in pragmatic trials, large simple trials and randomised database studies are described in Chapter 5.6.

## *4.2. Secondary use of data*

Secondary use of data refers to the utilisation of data already gathered for another purpose (e.g. electronic and non-electronic healthcare data). These can be further linked to prospectively collected data including medical and non-medical data. The last decades have witnessed the development of key data resources, expertise and methodology that have allowed use of such data for pharmacoepidemiology. The ENCePP Inventory of Data Sources contains information on existing European databases. However, this field is continuously evolving and it is recommended to look for recently published reviews and lists of databases.

A comprehensive description of the main features and applications of frequently used electronic healthcare databases for pharmacoepidemiology research in the United States and in Europe appears in the book *Pharmacoepidemiology* (B. Strom, S.E. Kimmel, S. Hennessy. 6th Edition, Wiley, 2019, Chapters 11 - 14). The limitations of using electronic healthcare databases should be acknowledged, as detailed in A review of uses of healthcare utilisation databases for epidemiologic research on therapeutics (J Clin Epidemiol 2005; 58: 23-337).

The primary purpose of the ISPE-endorsed Guidelines for Good Database Selection and use in Pharmacoepidemiology Research (Pharmacoepidemiol Drug Saf 2012;21:1-10) is to assist in the selection and use of data resources in pharmacoepidemiology by highlighting potential limitations and recommending correct procedures. This guideline refers to the secondary use of databases containing routinely collected healthcare information such as electronic medical records and claims databases and does not include spontaneous reporting databases. It is a simple, well-structured guideline that will

help investigators to select the most suitable databases to address specific research question and helps database custodians to describe their database in a useful manner. An entire section is dedicated to the use of multi-database studies. The document also contains references to data quality and validation procedures, data processing/transformation, privacy and security.

The Guidelines and recommendations for ensuring Good Epidemiological Practice (GEP): a guideline developed by the German Society for Epidemiology (European Journal of Epidemiology 2019;34(3):301-17) provide detailed recommendations on all aspects of the design and conduct of epidemiological studies, and many of these recommendations address aspects to be considered when making secondary use of data. The FDA's Best Practices for Conducting and Reporting Pharmacoepidemiologic Safety Studies Using Electronic Health Care Data Sets provides criteria for best practice that apply to design, analysis, conduct and documentation. It emphasizes that investigators should understand the potential limitations of electronic healthcare data systems, make provisions for their appropriate use and refer to validation studies of safety outcomes of interest in the proposed study and captured in the database. Guidance for conduction studies within electronic healthcare databases can also be found in the International Society for Pharmacoepidemiology Guidelines for Good Pharmacoepidemiology Practices (ISPE GPP), in particular sections IV-B (Study conduct, Data collection). This guidance emphasizes the importance of patient data protection.

The concepts of "Real-world data" (RWD) and "Real-world evidence" (RWE) are increasingly used in the regulatory setting to denote the secondary use of observational data and pharmacoepidemiological methods for regulatory decision-making. The article Real-World Data for Regulatory Decision Making: Challenges and Possible Solutions for Europe. (Clin Pharmacol Ther. 2019;106(1):36-9) describes the operational, technical and methodological challenges for the acceptability of real-world data for regulatory purposes and presents possible solutions to address these challenges. The FDA's Real-World Evidence website also provides definitions and links to a set of useful guidelines on the submission and use of real-world data, including electronic health care databases, to support decision-making. The Joint ISPE-ISPOR Special Task Force Report on Good Practices for Real-World Data Studies of Treatment and/or Comparative Effectiveness recommends good research practices for designing and analysing retrospective databases for comparative effectiveness research (CER). and reviews methodological issues and possible solutions for CER studies based on secondary data analysis (see also Chapter 10.1 on comparative effectiveness research). Many of the principles are applicable to studies with other objectives than CER, but aspects of pharmacoepidemiological studies based on secondary use of data, such as data quality, ethical issues, data ownership and privacy, are not covered.

The majority of the examples and methods covered in Chapter 5 are based on studies and methodologic developments in secondary data collection, since this is the most frequent approach used in pharmacoepidemiology.  Several potential issues need to be considered in the use of electronic healthcare data for pharmacoepidemiological studies as they may affect the validity of the results. They include completeness of data capture, bias in the assessment of exposure, outcome and covariates, variability between data sources and the impact of changes over time in data, access methodology and the healthcare system.

Chapter 4.6. deals with models of studies conducted across multiple data sources.

## *4.3. Patient registries*

### 4.3.1. Definitions

A patient registry is an organised system that uses observational methods to collect uniform data on specified outcomes in a population defined by a particular disease, condition or exposure. A registry-based study is an investigation of a research question using a patient registry infrastructure for patient recruitment and data collection. The term 'registry' is sometimes used incorrectly to designate a cohort study with primary data collection or a list of all patients meeting the eligibility criteria for a study (the term 'patient log' or 'patient log-list' could be used for the latter purpose).

A patient registry should be considered as an infrastructure for the standardised recording of data from routine clinical practice on individual patients identified by a characteristic or an event, for example the diagnosis of a disease (*disease registry)*, the occurrence of a condition (e.g., *pregnancy registry*), a birth defect (e.g. *birth defect registry*), a molecular or a genomic feature or any other patient characteristics, or an encounter with a particular healthcare service. The term *product registry* is sometimes used for a system where data are collected on patients exposed to a particular medicinal product, single substance or therapeutic class in order to evaluate their use or their effects, but such system should rather be considered a clinical trial or a non-interventional study as data is collected for a specific pre-planned analysis purpose in line with performing a trial/study.

### 4.3.2. Conceptual differences between a registry and a study

As illustrated in Imposed registries within the European postmarketing surveillance system (Pharmacoepidemiol Drug Saf 2018; 27(7):823-826), there are methodological differences between registries and registry-based studies.

Patient registries are often integrated into routine clinical practice with systematic and sometimes automated data capture in electronic healthcare records. Whilst the duration of a registry is normally open-ended, that of a registry-based study is dictated by the time needed to define and collect data relevant for the specific study objectives. Studies may also require introduction of specific procedures, questionnaires or data collection tools. Studies are set up and managed based on a limited number of endpoints and a specific protocol, whereas patient registries should focus on system specifications in order to ensure continuous, efficient and collaborative data collection, safe data hosting and availability of retrievable, interoperable and re-usable data.

A registry can be used as a source of patients for studies based on either primary data collection (where the events of interest for the study are collected directly from the patients, caregivers, healthcare professionals or other persons involved in the patient care) or secondary use of data already collected (where the study uses data collected for another purpose, analogously to the use of electronic healthcare records). For this purpose, registry data can be enriched with additional information on outcomes, lifestyle data, immunisation or mortality information obtained from linkage to the existing databases such as national cancer registries, prescription databases or mortality records.

### 4.3.3. Methodological aspects

To support better use of existing registries for the benefit-risk evaluation of medicines, the EU regulatory network developed the Patient registries initiative. As part of this initiative, the European Medicines Agency organised several workshops on disease-specific registries. The reports of these workshops describe regulators' expectation on common data elements to be collected and best

practices on topics such as governance, data quality control, data sharing or reporting of safety data. The ENCePP Resource database of data sources is also used to support an inventory of existing disease registries.

The EMA's Scientific Advice Working Party issued two Qualification Opinions for two registry platforms, the ECFSPR and the EBMT, with an evaluation of their potential use as data sources for registry-based studies. Although they apply only to two registry platforms, these opinions provide a good indication of the key methodological components expected by regulators for using a disease registry for such studies.

The US Agency for Health Care Research and Quality (AHRQ) published a comprehensive document on 'good registry practices' entitled Registries for Evaluating Patient Outcomes: A User's Guide, 3rd Edition, which provides methodological guidance on planning, design, implementation, analysis, interpretation and evaluation of the quality of a registry. There is a dedicated section for linkage of registries to other data sources. The EU PARENT Joint Action developed Methodological guidelines and recommendations for efficient and rational governance of patient registries to facilitate cross-border use of registries.

Results obtained from analyses of registry data may be affected by the same biases as those of studies described in Chapter 5.2 Bias and confounding. Registry-based studies are sensitive to selection bias. This is due to the fact that factors that may influence the enlistment of patients in a registry may be numerous (including clinical, demographic and socio-economic factors) and difficult to predict and identify, potentially resulting in a biased sample of the patient population in case the recruitment has not been exhaustive. In addition, registry-based studies may also introduce selection bias in the recruitment or selection of registered patient for the specific study, as well as in the differential completeness of follow-up and data collection. It is therefore important to systematically compare the characteristics of the study population with those of the source population.

As illustrated in The randomized registry trial--the next disruptive technology in clinical research? (N Engl J Med 2013; 369: 1579-81) and Registry-based randomized controlled trials: what are the advantages, challenges and areas for future research? (J Clin Epidemiol 2016;80:16-24), the randomised registry-based trial may support enhanced generalisability of findings, rapid consecutive enrollment, and the potential completeness of follow-up for the reference population, when compared with conventional randomized effectiveness trials, but several challenges need to be considered (see also Chapter 5.6.3).

### 4.3.4. Population registries

In European Nordic countries, a comprehensive registration of data for a high proportion or all of the population allows linkage between government-administered patient registries that may include hospital encounters, diagnoses and procedures, such as the Norwegian Patient Registry, the Danish National Patient Registry or the Swedish National Patient Register. They may however lack information on lifestyle factors, patient-related outcomes and laboratory data. A Review of 103 Swedish Healthcare Quality Registries (J Intern Med 2015; 277(1): 94–136) describes additional healthcare quality registries focusing on specific disorders initiated in Sweden mostly by physicians with data on aspects of disease management, self-reported quality of life, lifestyle, and general health status, providing an important source for research.

### 4.3.5. Registries which capture special populations

Special populations can be identified based on age (e.g., paediatric or elderly), pregnancy status, renal or hepatic function, race, or genetic differences. Some registries are focused on these particular

populations. Examples of these are the birth registries in Nordic countries and registries for rare diseases. The European Platform on Rare Diseases Registration (EU RD Platform) serves as platform for information on registries for rare diseases and has developed a set of common data elements for the European Reference Network and other rare disease registries.

The FDA's Draft Postapproval Pregnancy Safety Studies Guidance for Industry (May 2019) include recommendations for designing a pregnancy registry with a description of research methods and elements to be addressed. The Systematic overview of data sources for Drug Safety in pregnancy research provides an inventory of pregnancy exposure registries and alternative data sources on safety of prenatal drug exposure and discusses their strengths and limitations. Example of population-based registers allowing to assess outcome of drug exposure during pregnancy are the European network of registries for the epidemiologic surveillance of congenital anomalies EUROCAT, and the pan-Nordic registries which record drug use during pregnancy as illustrated in Selective serotonin reuptake inhibitors and venlafaxine in early pregnancy and risk of birth defects: population based cohort study and sibling design (BMJ 2015;350:h1798).

For paediatric populations, specific and detailed information as neonatal age (e.g. in days), pharmacokinetic parameters and organ maturation need to be considered and is usually missing from the classical datasources, therefore paediatric specific registries are important. The CHMP Guideline on Conduct of Pharmacovigilance for Medicines Used by the Paediatric Population provides further relevant information. An example of registry which focuses on paediatric patients is Pharmachild, which captures children with juvenile idiopathic arthritis undergoing treatment with methotrexate or biologic agents.

Other registries that focus on special populations (e.g., the UK Renal Registry) can be found in the ENCePP Inventory of data sources.

### 4.3.6. Disease registries in regulatory practice and health technology assessment

The article Patient Registries: An Underused Resource for Medicines Evaluation: Operational proposals for increasing the use of patient registries in regulatory assessments (Drug Saf. 2019;42(11):1343-1351) proposes sets of measures to improve use of registries in relation to: (1) nature of the data collected and registry quality assurance processes; (2) registry governance, informed consent, data protection and sharing; and (3) stakeholder communication and planning of benefit-risk assessments. Appendix 1 of Module VIII of the Good pharmacovigilance practice discusses the use of registries for conducting post authorisation studies. The use of registries to support the post-authorisation collection of data on effectiveness and safety of medicinal products in the routine treatment of diseases is also discussed in the EMA Scientific guidance on post-authorisation efficacy studies. Use of existing disease registries is recommended as they allow continued assessment of disease outcomes and a comparison of different treatment options using a similar methodology. Data of existing registries could be supplemented with additional data collection or linkage to external data sources.

When efficacy has been demonstrated in RCTs, registry-based studies may also be useful to study aspects related to long term effectiveness and safety in heterogeneous populations, study effect modifiers such as doses that have been prescribed by physicians and that may differ from those used in RCTs, and study patient sub-groups defined by variables such as age, co-morbidities, use of concomitant medication or genetic factors, or other factors that might influence effectiveness or safety.

Incorporating data from clinical practice into the drug development process is a growing interest from health technology assessment (HTA) bodies and payers since reimbursement decisions can benefit from better estimation and prediction of effectiveness of treatments at the time of product launch. An

example of where registries can provide clinical practice data is the building of predictive models that incorporate data from both RCTs and registries to generalise results observed in RCTs to a real-world setting. In this context, the EUnetHTA Joint Action 3 project has issued the Registry Evaluation and Quality Standards Tool (REQueST) aiming to guide the evaluation of registries for effective usage in HTA.

## *4.4. Spontaneous reports*

Spontaneous reports of adverse drug effects remain a cornerstone of pharmacovigilance and are collected from a variety of sources, including healthcare providers, national authorities, pharmaceutical companies, medical literature and more recently directly from patients. EudraVigilance is the European Union data processing network and management system for reporting and evaluation of suspected adverse drug reactions (ADRs). The Global Individual Case Safety Reports Database System (VigiBase) pools reports of suspected ADRs from the members of the WHO programme for international drug monitoring. These systems deal with the electronic exchange of Individual Case Safety Reports (ICSRs), the early detection of possible safety signals and the continuous monitoring and evaluation of potential safety issues in relation to reported ADRs. The report Characterization of databases (DB) used for signal detection (SD) of the PROTECT project shows the heterogeneity of spontaneous databases and the lack of comparability of SD methods employed. This heterogeneity is an important consideration when assessing the performance of SD algorithms.

The strength of spontaneous reporting systems is that they cover all types of legal drugs used in any setting. In addition to this, the reporting systems are built to obtain information specifically on potential adverse drug reactions and the data collection concentrates on variables relevant to this objective and directs reporters towards careful coding and communication of all aspects of an ADR. The increase in systematic collection of ICSRs in large electronic databases has allowed the application of data mining and statistical techniques for the detection of safety signals. There are known limitations of spontaneous ADR reporting systems, which include limitations embedded in the concept of voluntary reporting, whereby known or unknown external factors may influence the reporting rate and data quality. ICSRs may be limited in their utility by a lack of data for an accurate quantification of the frequency of events or the identification of possible risk factors for their occurrence. For these reasons, the concept is now well accepted that any signal from spontaneous reports needs to be verified clinically before further communication.

One challenge in spontaneous report databases is report duplication. Duplicates are separate and unlinked records that refer to one and the same case of a suspected ADR and may mislead clinical assessment or distort statistical screening. They are generally detected by individual case review of all reports or by computerised duplicate detection algorithms. In Performance of probabilistic method to detect duplicate individual case safety reports (Drug Saf 2014;37(4):249-58) a probabilistic method highlighted duplicates that had been missed by a rule-based method and also improved the accuracy of manual review. In the study, however, a demonstration of the performance of de-duplication methods to improve signal detection is lacking. The FDA have also implemented probabilistic duplicate detection in the FAERS and VAERS databases. A novel feature is an attempt to use narrative text analysed via NLP methods as demonstrated in Using Probabilistic Record Linkage of Structured and Unstructured Data to Identify Duplicate Cases in Spontaneous Adverse Event Reporting Systems (Drug Saf 2017;40(7):571–58).

Validation of statistical signal detection procedures in EudraVigilance post-authorisation data: a retrospective evaluation of the potential for earlier signalling (Drug Saf 2010;33: 475 – 87) has shown that the statistical methods applied in EudraVigilance can provide significantly early warning in a large

proportion of Drug Safety problems. Nonetheless, this approach should supplement, rather than replace, other pharmacovigilance methods.

Chapters IV and V of the Report of the CIOMS Working Group VIII 'Practical aspects of Signal detection in Pharmacovigilance' present sources and limitations of spontaneously-reported drug-safety information and databases that support signal detection. Appendix 3 of the report provides a list of international and national spontaneous reporting system databases.

## *4.5. Social media*

### 4.5.1. Definition

Technological advances have dramatically increased the range of data sources that can be used to complement traditional ones and may provide compelling insights into effectiveness and safety of interventions. Such data include digital media that exist in a computer-readable format as websites, web pages, blogs, vlogs, social networking sites, internet forums, chat rooms, health portals. A recent addition to this list is represented by the biomedical data collected through wearable technology (e.g., heart rate, physical activity and sleep pattern, dietary patterns). This data is unsolicited and generated in real time.

Social media is considered as a sub-set of digital media. The European Commission's Digital Single Market Glossary defines social media as "*a group of Internet-based applications that build on the ideological and technological foundations of Web 2.0 and that allow the creation and exchange of user-generated content. It employs mobile and web-based technologies to create highly interactive platforms via which individuals and communities share, co-create, discuss, and modify user-generated content.*"

### 4.5.2. Use in pharmacovigilance

Social media has been used to provide insights into the patient's perception of the effectiveness of drugs and for the collection of patient reported outcomes, as discussed in Web-based patient-reported outcomes in Drug Safety and risk management: challenges and opportunities? (Drug Saf 2012;35(6):437-46).

The IMI WEB-RADR European collaborative project explored different aspects related to the use of social media data as a basis for pharmacovigilance and summarised its recommendations in Recommendations for the Use of Social Media in Pharmacovigilance: Lessons From IMI WEB-RADR (Drug Saf 2019;42(12):1393-1407). The French Vigi4Med project, which evaluated the use of social media, mainly web forums, for pharmacovigilance activities,has published a set of recommendation in Use of Social Media for Pharmacovigilance Activities: Key Findings and Recommendations from the Vigi4Med Project (Drug Saf. 2020;10.1007/s40264-020-00951-2 [published online ahead of print, 2020 Jun 16]).

One possible use of social media would be source of information for signal detection or assessment. Studies including Using Social Media Data in Routine Pharmacovigilance: A Pilot Study to Identify Safety Signals and Patient Perspectives (Pharm Med 2017;31(3): 167-74) and Assessment of the Utility of Social Media for Broad-Ranging Statistical Signal Detection in Pharmacovigilance: Results from the WEB-RADR Project (Drug Saf 2018;41(12):1355–1369) have evaluated whether analysis of social media data (specifically Facebook and Twitter posts) could identify pharmacovigilance signals early, but in their respective settings, found that this was not the case.

Using Social Media Data in Routine Pharmacovigilance: A Pilot Study to Identify Safety Signals and Patient Perspectives (Pharm Med 2017;31(3): 167-74) also tried to determine the quantity of posts with resemblance to adverse events and the types and characteristics of products that would benefit from social media analysis. It concludes that, although analysis of data from social media did not identify new safety signals, it can provide unique insight into the patient perspective.

From a regulatory perspective, social media is a source of potential reports of suspected adverse drug reactions and marketing authorisation holders are legally obliged to screen web sites under their management and assess whether reports of adverse reactions qualify for spontaneous reporting (see Good Pharmacovigilance practice Module VI- (Rev. 2), Chapter VI.B.1.1.4). Principles for continuous monitoring of the safety of medicines without overburdening established pharmacovigilance systems and a regulatory framework on the use of social media in pharmacovigilance have been proposed in Establishing a Framework for the Use of Social Media in Pharmacovigilance in Europe (Drug Saf. 2019;42(8):921-30).

### 4.5.3. Challenges

While offering the promise of new research models and approaches, the rapidly evolving social media environment presents many challenges including the need for strong and systematic processes for selection, validation and study implementation. Articles which detail associated challenges are: Evaluating Social Media Networks in Medicines Safety Surveillance: Two Case Studies (Drug Saf 2015; 38(10): 921-30.) and Social media and pharmacovigilance: A review of the opportunities and challenges (Br J Clin Pharmacol 2015; 80(4): 910-20).

There is currently no defined strategy or framework in place in order to meet the standards around data validity, generalisability for this type of data, and their regulatory acceptance may therefore be lower than for traditional sources. However, more tools and solutions for analysing unstructured data are becoming available, especially for pharmacoepidemiology and Drug Safety research, as in Deep learning for pharmacovigilance: recurrent neural network architectures for labeling adverse drug reactions in Twitter posts (J Am Med Inform Assoc 2017 Feb 22) and Social Media Listening for Routine Post-Marketing Safety Surveillance (Drug Saf 2016;39(5):443-54). However, the recognition and disambiguation of references to drugs and adverse events in free text remains a challenge and performance evaluations need to be critically assessed as discussed in Prospective Evaluation of Adverse Event Recognition Systems in Twitter: Results from the Web-RADR Project (Drug Saf 2020;10.1007/s40264-020-00942-3 [published online ahead of print, 2020 May 14]).

### 4.5.4. Data protection

The EU General Data Protection Regulation (GDPR) introduces EU-wide legislation on personal data and security. It specifies that the impact of data protection at the time of study design concept should be assessed and reviewed periodically. Other technical documents may also be applicable such as Smartphone Secure Development Guidelines (2011) published by the European Network and Information Security Agency (ENISA), which advises on design and technical solutions. The principles of these security measures are found in the European Data Protection Supervisor (EDPS) opinion on mobile health (Opinion 1/2015 Mobile Health-Reconciling technological innovation with data protection).

## *4.6. Research networks for multi-database studies*

### 4.6.1. General considerations

Pooling data across different databases affords insight into the generalisability of the results and may improve precision. A growing number of studies use data from networks of databases, often from different countries. Some of these networks are based on long-term contracts with selected partners and are very well structured (such as Sentinel, the Vaccine Safety Datalink (VSD) or the Canadian Network for Observational Drug Effect Studies (CNODES)), but others are looser collaborations based on an open community principle (e.g. Observational Health Data Sciences and Informatics (OHDSI)). In Europe, collaborations for multi-database studies have been strongly encouraged by the Drug Safety research funded by the European Commission (EC) and public-private partnerships such as the Innovative Medicines Initiative (IMI). This funding resulted in the conduct of groundwork necessary to overcome the hurdles of data sharing across countries for specific projects (e.g. PROTECT, ADVANCE, EMIF, EHDEN) or for specific post-authorisation studies.

In this chapter, networking is used to mean collaboration between investigators for sharing expertise and resources. The ENCePP Database of Research Resources may facilitate such networking by providing an inventory of research centres and data sources that can collaborate on specific pharmacoepidemiology and pharmacovigilance studies in Europe. It allows the identification of centres and data sets by country, type of research and other relevant fields.

The use of research networks in drug safety analyses is well established and a significant body of practical experience exists. By contrast, no consensus exists on the use of such networks, or indeed of single sources of observational data, in estimating effectiveness. In particular, the use in support of licensing applications will require evaluations of the reliability of results and the verifiability of research processes that are currently at an early stage. Specific advice on effectiveness can only be given once this work has been done and incorporated into regulatory guidelines. Hence this discussion currently relates only to product safety (see Assessing strength of evidence for regulatory decision making in licensing: What proof do we need for observational studies of effectiveness?; Pharmacoepidemiol. Drug Saf. 2020 Apr 16).

From a methodological point of view, research networks have many advantages over single database studies:

- Increase the *size* of study populations which facilitates research on rare events, drugs used in specialised setting (see Ability of primary care health databases to assess medicinal products discussed by the European Union Pharmacovigilance Risk Assessment Committee. Clin. Pharmacol. Ther. 2020 Apr;107(4):957-965), or when the interest is in subgroup effects.
- In case of primary data collection, *shorten the time* needed for obtaining the desired sample size and speed-up investigation of drug safety issues or other outcomes.
- Benefit from the *heterogeneity* of treatment options across countries, which allows studying the effect of different drugs used for the same indication or specific patterns of utilisation.
- May provide additional knowledge on the *generalisability* of results and on the *consistency of information*, for instance whether a safety issue exists in several countries. Possible inconsistencies might be caused by different biases or truly different effects in the databases revealing causes of differential drug effects, and these might be investigated.

- Involve experts from various countries addressing case definitions, terminologies, coding in databases and research practices provides opportunities to increase *consistency of results* of observational studies.
- Allow pooling data or results and increase *the amount of information* gathered for a specific issue addressed in different databases.

The article Different strategies to execute multi-database studies for medicines surveillance in real world setting: a reflection on the European model (Clin. Pharmacol. Ther. 2020 Apr 3) describes different models applied for combining data or results from multiple databases. A common characteristic of all models is the fact that data partners maintain physical and operational control over electronic data in their existing environment and therefore the data extraction is always done locally. Differences however exist in the following areas: use of a common protocol; use of a common data model (CDM); and where and how the data analysis is done.

Use of a common data model (CDM) implies that local formats are translated into a predefined, common data structure, which allows launching a similar data extraction and analysis script across several databases. Sometimes the CDM imposes a common terminology as well, as in the case of the OMOP CDM. The CDM can be systematically applied on the entire database (generalised CDM) or on the subset of data needed for a specific study (study specific CDM). In the EU, study specific CDMs have generated results in several projects and studies and initial steps have been taken to create generalised CDMs, but experience based on real-life studies is still limited. An example is the study Safety of hydroxychloroquine, alone and in combination with azithromycin, in light of rapid wide-spread use for COVID-19: a multinational, network cohort and self-controlled case series study.

## 4.6.2. Models of studies using multiple data sources

Five models of studies are presented, classified according to specific choices in the steps needed to execute a study: protocol development and agreement (whether separate or common); where the data are extracted and analysed (locally or centrally); how the data are extracted and analysed (using individual or common programs); and use of a CDM and which type (study specific or general) (see Table 1).

### *4.6.2.1. Meta-analysis: separate protocols, local and individual data extraction and analysis, no CDM*

The traditional mode to combine data from multiple data sources is when data extraction and analysis are performed independently at each centre based on separate protocols. This is usually followed by meta-analysis of the different estimates obtained (see Chapter 5.7).

This type of model may be viewed as a baseline situation which a research network will try to improve. Moreover, meta-analysis should be used in all models of studies presented, as there is always the possibility that different data sources provides different results and hence explicitly looking for such variation should always be considered. If all the data sources can be accessed, explaining variations in term of covariates should also be attempted. This is coherent with the recommendations from Multi-centre, multi-database studies with common protocols: lessons learnt from the IMI PROTECT project (Pharmacoepidemiol. Drug Saf. 2016;25(S1):156-165) that states that a priori pooling of data from several databases may disguise heterogeneity that may provide useful information on the safety issue under investigation. On the other hand, parallel analysis of databases allows exploring reasons for heterogeneity through extensive sensitivity analyses. This approach eventually increases consistency in findings from observational drug effect studies or reveal causes of differential drug effects.

#### *4.6.2.2. Local analysis: common protocol, local and individual data extraction and analysis, no CDM*

In this option, data are extracted and analysed locally, with site-specific programs that are developed by each centre, on the basis of a common protocol. Definitions of exposure, outcomes and covariates, analytical programmes and reporting formats are standardised according to a common protocol and the results of each analysis, either at a patient level or in an aggregated format depending on the governance of the network, are shared and pooled together through meta-analysis.

This approach allows assessment of database or population characteristics and their impact on estimates but reduces variability of results determined by differences in design. Examples of research networks that use the common protocol approach are PROTECT (as described in Improving Consistency and Understanding of Discrepancies of Findings from Pharmacoepidemiological Studies: the IMI PROTECT Project. (Pharmacoepidemiol Drug Saf 2016;25(S1): 1-165) and the Canadian Network for Observational Drug Effect Studies (CNODES). The latter is experimenting with a CDM as explained in Building a framework for the evaluation of knowledge translation for the Canadian Network for Observational Drug Effect Studies (Pharmacoepidemiol. Drug Saf. 2020;29 (S1),8-25)

This approach requires very detailed common protocols and data specifications that reduce variability in interpretations by researchers.

#### *4.6.2.3. Sharing of raw data: common protocol, local and individual data extraction, central analysis, no CDM*

In this approach, a mutually agreed protocol is agreed by the study partners. Data intended to be used for the study are locally extracted with site-specific programs, transferred without analysis and conversion to a CDM, and pooled and analyzed at the central partner receiving them.

Examples for this approach are when databases are very similar in structure and content as is the case for some Nordic registries, or on the Italian regional databases. Examples of such models are Selective serotonin reuptake inhibitors during pregnancy and risk of persistent pulmonary hypertension in the newborn: population based cohort study from the five Nordic Countries (BMJ 2012;344:d8012) and All-cause mortality and antipsychotic use among elderly persons with high baseline cardiovascular and cerebrovascular risk: a multi-center retrospective cohort study in Italy (Expert Opin. Drug Metab. Toxicol. 2019;15:179-88).

The central analysis allows removing an additional source of variability linked to the statistical programing and analysis.

#### *4.6.2.4. Study specific CDM: common protocol, local and individual data extraction, local and common analysis, study specific CDM*

In this approach, a mutually agreed protocol is agreed by the study partners and data intended to be used for the study are locally extracted and loaded into a CDM; data in the CDM are then processed locally in all the sites with one common program. The output of the common program is transferred to a specific partner. The output to be shared may be an analytical dataset or study estimates, depending on the governance of the network.

Examples of research networks that used this approach by employing a study-specific CDM with transmission of anonymised patient-level data (allowing a detailed characterisation of each database) are EU-ADR (as explained in Combining multiple healthcare databases for postmarketing drug and vaccine safety surveillance: why and how?, J Intern Med 2014;275(6):551-61), SOS, ARITMO, SAFEGUARD, GRIP, EMIF, EUROmediCAT and ADVANCE. In all these projects, a basic and simple CDM was utilised and R, SAS, STATA or Jerboa scripts have been used to create and share common

analytics. Diagnosis codes for case finding can be mapped across terminologies by using the Codemapper, developed in the ADVANCE project, as explained in CodeMapper: semiautomatic coding of case definitions (Pharmacoepidemiol Drug Saf 2017;26(8):998-1005).

An approach to quantify the impact of different case finding algorithms, called the component strategy, was developed in the EMIF and ADVANCE projects and could also be compatible with the simple and generalised common data model (see Identifying Cases of Type 2 Diabetes in Heterogeneous Data Sources: Strategy from the EMIF Project. PLoS One 2016;11(8):e0160648).

#### *4.6.2.5. General CDM: common protocol, local and common data extraction and analysis, general CDM*

In this approach, the local databases are transformed into a CDM prior to and independent of any study protocol. When a study is required, a protocol is agreed by the study partners and a centrally developed analysis program is created that runs locally on each database to extract and analyse the data. The output of the common programs shared may be an analytical dataset or study estimates, depending on the governance of the network.

Two examples of research networks which use a generalised CDM are the Sentinel Initiative (as described in The U.S. Food and Drug Administration's Mini-Sentinel Program, Pharmacoepidemiol Drug Saf 2012;21(S1):1–303) and OHDSI. The main advantage of a general CDM is that it can be used for virtually any study involving that database. OHDSI is based on the Observational Medical Outcomes Partnership (OMOP) CDM which is now used by many organisations and has been tested for its suitability for safety studies (see for example Validation of a common data model for active safety surveillance research. J Am Med Inform Assoc 2012;19(1):54–60 and Can We Rely on Results From IQVIA Medical Research Data UK Converted to the Observational Medical Outcome Partnership Common Data Model?: A Validation Study Based on Prescribing Codeine in Children (Clin Pharmacol Ther 2020;107(4):915-25)). Conversion into the OMOP CDM, requires formal mapping of database items to standardised concepts. This is resource intensive and will need to be updated every time the databases is refreshed. An example of a study performed with the OMOP CDM in Europe is Safety of hydroxychloroquine, alone and in combination with azithromycin, in light of rapid wide-spread use for COVID-19: a multinational, network cohort and self-controlled case series study.

In a Comparative Assessment of Observational Medical Outcomes Partnership and Mini-Sentinel Common Data Models and Analytics: Implications for Active Drug Safety Surveillance (Drug Saf 2015;38(8):749-65), it is suggested that slight conceptual differences between the Sentinel and the OMOP models do not significantly impact on identifying known safety associations. Differences in risk estimations can be primarily attributed to the choices and implementation of the analytic approach.

Table 1: Models of studies using multiple data sources: key characteristics following the steps needed to execute a study

| | Before study start | After study start | | | | |
|---|---|---|---|---|---|---|
| Model | Conversion to general CDM | Common protocol development | Data extraction | Conversion to study specific CDM | Data analysis | What is shared |
| Meta-analysis | | Separate | Local - individual programs | | Local - individual programs | Study estimates |
| Local analyses | | Mutually developed | Local - individual programs | | Local - individual programs | Analytical dataset or study estimates |
| Sharing of raw data | | Mutually developed | Local - individual programs | | Local - individual programs | Extracted raw data |
| Study specific CDM | * | Unilaterally or mutually developed | Local - individual programs | Local - individual programs | Local - individual programs | Analytical dataset or study estimates |
| General CDM | Local - individual programs | Unilaterally or mutually developed | Local - common programs | | Local - individual programs | Analytical dataset or study estimates |

*Programs to transforms sets of variables into a CDM for a specific study are archived and can be re-used for future studies

### 4.6.3. Challenges of different models

The different models presented above present several challenges:

*Related to the scientific content*

- Differences in the underlying health care systems
- Different mechanisms of data generation and collection
- Mapping of differing disease coding systems (e.g., the International Classification of Disease, 10th Revision (ICD-10), Read codes, the International Classification of Primary Care (ICPC-2)) and narrative medical information in different languages
- Validation of study variables and access to source documents for validation

*Related to the organisation of the network*

- Differences in culture and experience between academia, public institutions and private partners
- Differences in the type and quality of information contained within each mapped database
- Different ethical and governance requirements in each country regarding processing of anonymised or pseudo-anonymised healthcare data
- Choice of data sharing model and access rights of partners
- Issues linked to intellectual property and authorship.
- Sustainability and funding mechanisms.

Each model has strengths and weaknesses in facing the above challenges, as illustrated in Data Extraction and Management in Networks of Observational Health Care Databases for Scientific Research: A Comparison of EU-ADR, OMOP, Mini-Sentinel and MATRICE Strategies (eGEMs 2016;4(1):2). In particular, a central analysis or a CDM provide protection from problems related to variation in how protocols are implemented as individual analysts might implement protocols differently (as described in Quantifying how small variations in design elements affect risk in an incident cohort study in claims; Pharmacoepidemiol. Drug Saf. 2020;29(1):84-93). Experience has shown that many of these difficulties can be overcome by full involvement and good communication between partners, and a project agreement between network members defining roles and responsibilities and addressing issues of intellectual property and authorship. Several of the networks have made their code, products data models and analytics software publicly available as OHDSI, Sentinel, ADVANCE.

Timeliness or speed for running studies is important in order to meet short regulatory timelines in circumstances where prompt decisions are needed. Solutions need therefore to be further developed and introduced to be able to run multi-database studies with shorter timelines. Independently from the model used, major factors that should be considered in speeding up studies include having work independent of any particular study already done. This includes factors such as: prespecified agreements on data access and processes for protocol development and study management, identification and characterisation of a large set of databases, creation of common definitions for variables that seem likely to occur in studies, and a common analytical systems where the most typical and routine analyses are already defined (this latter point is made easier with the use of CDMs, especially general ones, with standardised analytics and tools that can be re-used to support faster analysis).

# 5. Study design and methods

## *5.1. Definition and validation of drug exposure, outcomes and covariates*

Historically, pharmacoepidemiology studies relied on patient-supplied information or searches through paper-based health records. The rapid increase in access to electronic healthcare records and large administrative databases has changed the way exposures and outcomes are defined, measured and validated. All variables should be defined with care taking into account the fact that information is often recorded for purposes other than pharmacoepidemiology.

### 5.1.1. Assessment of exposure

In pharmacoepidemiology studies, exposure data originate mainly from four data sources: data on prescribing (e.g. CPRD primary care data), data on dispensing (e.g. PHARMO outpatient pharmacy database), data on payment for medication (namely claims data, e.g. IMS LifeLink PharMetrics Plus) and data collected in surveys. The population included in these data sources follows a process of attrition: drugs that are prescribed are not necessarily dispensed, and drugs that are dispensed are not necessarily ingested. In Primary non-adherence in general practice: a Danish register study (Eur J Clin Pharmacol 2014;70(6):757-63), 9.3% of all prescriptions for new therapies were never redeemed at the pharmacy, with different percentages per therapeutic and patient groups. The attrition from dispensing to ingestion is even more difficult to measure, as it is compounded by uncertainties about which dispensed drugs are actually taken by the patients and the patients' ability to provide an accurate account of their intake. In addition, paediatric adherence is dependent on parents' accurate recollection and recording.

Exposure definitions can include simple dichotomous variables (e.g. ever exposed vs. never exposed) or be more detailed, including estimates of duration, exposure windows (e.g. current vs. past exposure) or dosage (e.g. current dosage, cumulative dosage over time). Consideration should be given to the level of detail available from the data sources on the timing of exposure, including the quantity prescribed, dispensed or ingested and the capture of dosage instructions. This will vary across data sources and exposures (e.g. estimating anticonvulsant ingestion is typically easier than estimating rescue medication for asthma attacks). Discussions with clinicians regarding sensible assumptions will be informative for the variable definition.

The Methodology chapter of the book *Drug Utilization Research. Methods and Applications* (M. Elseviers, B. Wettermark, A.B. Almarsdottir et al. Ed. Wiley Blackwell, 2016) discusses different methods for data collection on drug utilisation.

### 5.1.2. Assessment of outcomes

A case definition compatible with the data source should be developed for each outcome of a study at the design stage. This description should include how events will be identified and classified as cases, whether cases will include prevalent as well as incident cases, exacerbations and second episodes (as differentiated from repeat codes) and all other inclusion or exclusion criteria. The reason for the data collection and the nature of the healthcare system that generated the data should also be described as they can impact on the quality of the available information and the presence of potential biases. Published case definitions of outcomes, such as those developed by the Brighton Collaboration in the context of vaccinations, are useful but are not necessarily compatible with the information available in the observational data sources. For example, information on the duration of symptoms may not be available.

Search criteria to identify outcomes should be defined and the list of codes and any used algorithm should be provided. Generation of code lists requires expertise in both the coding system and the disease area. Researchers should consult clinicians who are familiar with the coding practice within the studied field. Suggested methodologies are available for some coding systems (see Creating medical and drug code lists to identify cases in primary care databases. Pharmacoepidemiol Drug Saf 2009;18(8):704-7). Coding systems used in some commonly used databases are updated regularly so sustainability issues in prospective studies should be addressed at the protocol stage. Moreover, great care should be given when re-using a code list from another study as code lists depend on the study objective and methods. Public repository of codes as Clinicalcodes.org is available and researchers are also encouraged to make their own set of coding available.

In some circumstances, chart review or free text entries in electronic format linked to coded entries can be useful for outcome identification. Such identification may involve an algorithm with use of multiple code lists (for example disease plus therapy codes) or an endpoint committee to adjudicate available information against a case definition. In some cases, initial plausibility checks or subsequent medical chart review will be necessary. When databases contain prescription data only, drug exposure may be used as a proxy for an outcome, or linkage to different databases is required.

### 5.1.3. Assessment of covariates

In pharmacoepidemiology studies, covariates are used for selecting and matching study subjects, comparing characteristics of the cohorts, developing propensity scores, creating stratification variables, evaluating effect modifiers and adjusting for confounders. Reliable assessment of covariates is therefore essential for the validity of results. Patient characteristics and other key covariates that could be confounding variables need to be evaluated using all available data. A given database may or may not be suitable for studying a research question depending on the availability of information on these covariates.

Some patient characteristics and covariates vary with time and accurate assessment is therefore time dependent. The timing of assessment of the covariates is an important factor for the correct classification of the subjects and should be clearly specified in the protocol. Capturing covariates can be done at one or multiple points during the study period. In the later scenario, the variable will be modeled as time-dependent variable.

Assessment of covariates can be done using different periods of time (look-back periods or run-in periods). Fixed look-back periods (for example 6 months or 1 year) are sometimes used when there are changes in coding methods or in practices or when using the entire medical history of a patient is not feasible. Estimation using all available covariates information versus a fixed look-back window for dichotomous covariates (Pharmacoepidemiol Drug Saf 2013; 22(5):542-50) establishes that defining covariates based on all available historical data, rather than on data observed over a commonly shared fixed historical window will result in estimates with less bias. However, this approach may not always be applicable, for example when data from paediatric and adult periods are combined because covariates may significantly differ between paediatric and adult populations (e.g. height and weight).

### 5.1.4. Validation

In healthcare databases, the correct assessment of drug exposure, outcome and covariate is crucial to avoid misclassification. Validity of diagnostic coding within the General Practice Research Database: a systematic review (Br J Gen Pract 2010;60:e128-36), the book *Pharmacoepidemiology* (B. Strom, S.E. Kimmel, S. Hennessy. 5th Edition, Wiley, 2012) and Mini-Sentinel's systematic reviews of validated

methods for identifying health outcomes using administrative and claims data: methods and lessons learned (Pharmacoepidemiol Drug Saf 2012; Suppl 1:82-9) provide examples.

Potential misclassification of exposure, outcome and other variables should be measured and removed or reduced. Misclassification by exposure should be measured by validating each comparison group. External validation against chart review or physician/patient questionnaire is possible in some instances but the questionnaires cannot always be considered as 'gold standard'. While the positive predicted value is more easily measured than the negative predictive value, a low specificity is more damageable than a low sensitivity when considering bias in relative risk estimates (see A review of uses of health care utilization databases for epidemiologic research on therapeutics, J Clin Epidemiol 2005;58(4):323-37). When validation of the variable is complete, the study point estimate should be adjusted accordingly (see Use of the Positive Predictive Value to Correct for Disease Misclassification in Epidemiologic Studies, Am J Epidemiol 1993;138 (11):1007–15 and Sentinel Quantitative Bias Analysis Methodology Development: Sequential Bias Adjustment for Outcome Misclassification, 2017).

For databases routinely used in research, documented validation of key variables may have been done previously by the data provider or other researchers. Any extrapolation of previous validation study should however consider the effect of any differences in prevalence and inclusion and exclusion criteria, the distribution and analysis of risk factors as well as subsequent changes to health care, procedures and coding, as illustrated in Basic Methods for Sensitivity Analysis of Biases, (Int J Epidemiol 1996; 25(6): 1107-16). The accurate date of onset is particularly important for studies relying upon timing of exposure and outcome such as in the self-controlled designs. A comparison of data from registries with clinical or administrative records can also validate individual records on a specific outcome.

Linkage validation can be used when another database is used for the validation through linkage methods (see Using linked electronic data to validate algorithms for health outcomes in administrative databases, J Comp Eff Res 2015; 4:359-66). In some situations, there is no access to a resource to provide data for comparison. In this case, indirect validation may be an option, as explained in the book *Applying quantitative bias analysis to epidemiologic data* (Lash T, Fox MP, Fink AK Springer-Verlag, New-York, 2009).

Structural validation of the database with internal logic checks can also be performed to verify the completeness and accuracy of variables. For example, one can investigate whether an outcome was followed by (or proceeded from) appropriate exposure or procedures or if a certain variable has values within a known reasonable range.

## *5.2. Bias (systematic error)*

### 5.2.1. Selection bias

Selection bias means the selective recruitment into the study of subjects that are not representative of the exposure or outcome pattern in the source population. Examples of common selection bias are referral bias and self-selection bias (Strom BL, Kimmel SE, Hennessy S. *Pharmacoepidemiology*, 5th Edition, Wiley, 2012). Other forms of selection biases are presented below.

*Prevalence bias*

The practice of including prevalent users in observational studies, i.e. patients already taking a therapy for some time before study follow-up began, can cause two types of bias. Firstly, prevalent users are 'survivors' (healthy-users) of the early period of pharmacotherapy, which can introduce substantial selection bias if the risk varies with time, as seen in the association between contraceptive intake and

venous thrombosis which was initially overestimated due to the heathy-users bias. (see The Transnational Study on Oral Contraceptives and the Health of Young Women. Methods, results, new analyses and the healthy user effect, Hum Reprod Update 1999;5(6)707-20). Secondly, covariates for drug use at study entry are often influenced by the previous intake of the drug.

### 5.2.2. Information bias

Information bias arises when incorrect information about either exposure or outcome or any covariates is collected in the study. It can be either non-differential when it does occur randomly across exposed/non-exposed participants or differential when it is influenced by the disease or exposure status.

Non-differential misclassification bias drives the risk estimate towards the null value, while differential bias can drive the risk estimate in either direction. Examples of non-differential misclassification bias are recall bias (e.g., in case controls studies cases and controls can have different recall of their past exposures) and surveillance or detection bias.

*Protopathic bias*

Protopathic bias arises when the initiation of a drug (exposure) occurs in response to a symptom of the (at this point undiagnosed) disease under study (outcome). For example, use of analgesics in response to pain caused by an undiagnosed tumour might lead to the erroneous conclusion that the analgesic caused the tumour. Protopathic bias thus reflects a reversal of cause and effect (see Bias: Considerations for research practice. Am J Health Syst Pharm 2008;65:2159-68). This is particularly a problem in studies of drug-cancer associations and other outcomes with long latencies. It may be handled by including a time-lag, (i.e. by disregarding all exposure during a specified period of time before the index date).

Protopathic bias has also been described as a selection bias and it should not be confused with confounding by indication (see Confounding by Indication: An Example of Variation in the Use of Epidemiologic Terminology, Am J Epidemiol. 1999;149(11):981-3).

*Surveillance bias (or detection bias)*

Surveillance or detection bias arises when patients in one exposure group have a higher probability of having the study outcome detected, due to increased surveillance, screening or testing of the outcome itself, or of an associated symptom. For example, post-menopausal exposure to estrogen is associated with an increased risk of bleeding that can trigger screening for endometrial cancers, leading to a higher probability of early stage endometrial cancers being detected. Any association between estrogen exposure and endometrial cancer potentially overestimates risk, because unexposed patients with sub-clinical cancers would have a lower probability of their cancer being diagnosed or recorded. This is discussed in Alternative analytic methods for case-control studies of estrogens and endometrial cancer (N Engl J Med 1978;299(20):1089-94).

This non-random type of misclassification bias can be reduced by selecting an unexposed comparator group with a similar likelihood of screening or testing, selecting outcomes that are likely to be diagnosed equally in both exposure groups, or by adjusting for the surveillance rate in the analysis. These issues and recommendations are outlined in Surveillance Bias in Outcomes Reporting (JAMA 2011;305(23):2462-3).

*Immortal time bias*

Immortal time bias refers to a period of cohort follow-up time during which death (or an outcome that determines end of follow-up) cannot occur. (K. Rothman, S. Greenland, T. Lash. *Modern Epidemiology*, 3rd Edition, Lippincott Williams & Wilkins, 2008 p. 106-7).

Immortal time bias can arise when the period between cohort entry and date of first exposure to a drug, during which the event of interest has not occurred, is either misclassified or simply excluded and not accounted for in the analysis. Immortal time bias in observational studies of drug effects (Pharmacoepidemiol Drug Saf 2007(3);16:241-9) demonstrates how several observational studies used a flawed approach to design and data analysis, leading to immortal time bias, which can generate an illusion of treatment effectiveness. This is frequently found in studies that compare groups of 'users' against 'non-users'. Observational studies with surprisingly beneficial drug effects should therefore be re-assessed to account for this bias.

Immortal Time Bias in Pharmacoepidemiology (Am J Epidemiol 2008(4);167:492-9) describes various cohort study designs leading to this bias, quantifies its magnitude under different survival distributions and illustrates it with data from a cohort of lung cancer patients. For time-based, event-based and exposure-based cohort definitions, the bias in the rate ratio resulting from misclassified or excluded immortal time increases proportionately to the duration of immortal time. It is asserted that immortal time bias arises by conditioning on future exposure and that it can be avoided by analyzing the data as if the exposures and outcomes were included as they developed, without ever looking into the future.

Survival bias associated with time-to-treatment initiation in drug effectiveness evaluation: a comparison of methods (Am J Epidemiol 2005(10);162:1016-23) describes five different approaches to deal with immortal time bias. The use of a time-dependent approach had several advantages: no subjects are excluded from the analysis and the study allows effect estimation at any point in time after discharge. However, changes of exposure might be predictive of the study endpoint and need adjustment for time-varying confounders using complex methods. Problem of immortal time bias in cohort studies: example using statins for preventing progression of diabetes (BMJ 2010; 340:b5087) describes how immortal time in observational studies can bias the results in favor of the treatment group and how they can be identified and avoided. It is recommended that all cohort studies should be assessed for the presence of immortal time bias using appropriate validity criteria. However, Re. 'Immortal time bias in pharmacoepidemiology' (Am J Epidemiol 2009;170(5): 667-8) argues that sound efforts at minimising the influence of more common biases should not be sacrificed to that of avoiding immortal time bias.

*Other forms of time-related bias*

In many database studies, exposure status during hospitalisations is unknown. Exposure misclassification bias may occur with a direction depending on whether exposure to drugs prescribed preceding hospitalisations are continued or discontinued and if days of hospitalisation are considered as gaps of exposure, especially when several exposure categories are assigned, such as current, recent and past. The differential bias arising from the lack of information on (or lack of consideration of) hospitalisations that occur during the observation period (called 'immeasurable time bias' in Immeasurable time bias in observational studies of drug effects on mortality. Am J Epidemiol 2008;168(3):329-35) can be particularly problematic when studying serious chronic diseases that require extensive medication use and multiple hospitalisations.

In the case of case control studies assessing chronic diseases with multiple hospitalisations and in-patient treatment (such as the use of inhaled corticosteroids and death in chronic obstructive pulmonary disease patients), no clearly valid approach to data analysis can fully circumvent this bias. However, sensitivity analyses such as restricting the analysis to non-hospitalised patients or providing estimates weighted by exposable time may provide additional information on the potential impact of

this bias, as shown in Immeasurable time bias in observational studies of drug effects on mortality. (Am J Epidemiol 2008;168(3):329-35).

In cohort studies where a first-line therapy (such as metformin) has been compared with second- or third-line therapies, patients are unlikely to be at the same stage of the disease (e.g. diabetes), which can induce confounding of the association with an outcome (e.g. cancer incidence) by disease duration. An outcome related to the first-line therapy may also be attributed to the second-line therapy if it occurs after a long period of exposure. Such situation requires matching on disease duration and consideration of latency time windows in the analysis (example drawn from Metformin and the Risk of Cancer. Time-related biases in observational studies. Diabetes Care 2012;35(12):2665-73).

### 5.2.3. Confounding

Confounding occurs when the estimate of measure of association is distorted by the presence of another risk factor. For a variable to be a confounder, it must be associated with both the exposure and the outcome, without being in the causal pathway.

*Confounding by indication*

Confounding by indication refers to a determinant of the outcome parameter that is present in people at perceived high risk or poor prognosis and is an indication for intervention. This means that differences in care between the exposed and non-exposed, for example, may partly originate from differences in indication for medical intervention such as the presence of risk factors for particular health problems. Other names for this type of confounding are 'channelling'. Confounding by severity is a type of confounding by indication, where not only the disease but its severity acts as confounders (see Confounding by Indication: An Example of Variation in the Use of Epidemiologic Terminology, Am J Epidemiol. 1999;149(11):981-3).

This type of confounding has frequently been reported in studies evaluating the efficacy of pharmaceutical interventions and is almost always encountered in various extents in pharmacoepidemiological studies. A good example can be found in Confounding and indication for treatment in evaluation of drug treatment for hypertension (BMJ 1997;315:1151-4).

The article Confounding by indication: the case of the calcium channel blockers (Pharmacoepidemiol Drug Saf 2000;9(1):37-41) demonstrates that studies with potential confounding by indication can benefit from appropriate analytic methods, including separating the effects of a drug taken at different times (see Chapter 5.3).

With the more recent application of pharmacoepidemiological methods to assess effectiveness, confounding by indication is a greater challenge and the article Approaches to combat with confounding by indication in observational studies of intended drug effects (Pharmacoepidemiol Drug Saf 2003;12(7):551-8) focusses on its possible reduction in studies of intended effects. An extensive review of these and other methodological approaches discussing their strengths and limitations is discussed in Methods to assess intended effects of drug treatment in observational studies are reviewed (J Clin Epidemiol 2004;57(12):1223-31).

*Unmeasured confounding*

Complete adjustment for confounders would require detailed information on clinical parameters, lifestyle or over-the-counter medications, which are often not measured in electronic healthcare records, causing residual confounding bias. Using directed acyclic graphs to detect limitations of traditional regression in longitudinal studies (Int J Public Health 2010;55(6):701-3) reviews

confounding and intermediate effects in longitudinal data and introduces causal graphs to understand the relationships between the variables in an epidemiological study.

Unmeasured confounding can be adjusted for only through randomisation. When this is not possible, as most often in pharmacoepidemiological studies, the potential impact of residual confounding on the results should be estimated and considered in the discussion.

Sensitivity analysis and external adjustment for unmeasured confounders in epidemiologic database studies of therapeutics (Pharmacoepidemiol Drug Saf 2006;15(5):291-303) provides a systematic approach to sensitivity analyses to investigate the impact of residual confounding in pharmacoepidemiological studies that use healthcare utilisation databases. In this article, four basic approaches to sensitivity analysis were identified: (1) sensitivity analyses based on an array of informed assumptions; (2) analyses to identify the strength of residual confounding that would be necessary to explain an observed drug-outcome association; (3) external adjustment of a drug-outcome association given additional information on single binary confounders from survey data using algebraic solutions; (4) external adjustment considering the joint distribution of multiple confounders of any distribution from external sources of information using propensity score calibration. The paper concludes that sensitivity analyses and external adjustments can improve our understanding of the effects of drugs in epidemiological database studies. With the availability of easy-to-apply spreadsheets (e.g. at https://www.drugepi.org/dope/software#Sensitivity), sensitivity analyses should be used more frequently, substituting qualitative discussions of residual confounding.

The impact of residual and unmeasured confounding in epidemiologic studies: a simulation study (Am J Epidemiol 2007;166(6):646–55) considers the extent and patterns of bias in estimates of exposure-outcome associations that can result from residual or unmeasured confounding, when there is no true association between the exposure and the outcome.. Another important finding of this study was that when confounding factors (measured or unmeasured) are interrelated (e.g. in situations of confounding by indication), adjustment for a few factors can almost completely eliminate confounding.

## *5.3. Methods to address bias*

### 5.3.1. The target trial approach

The target trial approach and its emulation by an observational study was initially introduced in 1989 (The clinical trial as a paradigm for epidemiologic research. J Clin Epidemiol 1989;42(6):491-6) and later extended to pharmacoepidemiology as a conceptual framework helping researchers to identify and avoid potential biases (Using Big Data to Emulate a Target Trial When a Randomized Trial Is Not Available. Am. J. Epidemiol 2016;183(8) 758-64). The underlying idea is to first imagine a hypothetical randomised trial ("target trial") that would answer the research question instead of starting to design a study around the limitations of the available observational data. In the first step, the target trial is described with regards to the eligibility criteria, the treatment strategies, the assignment procedure, the follow-up period, the outcome, the causal contrasts and the analysis plan. In the second step, the researcher specifies how the observational data is used to emulate the target trial, e.g. how time zero is defined, and the trade-offs needed to conduct the observational study, e.g. regarding eligibility criteria, interventions, confounders and outcomes. The explicit description of the target trial and the specification of how this trial is emulated with observational data lead to study designs and analytic approaches that prevent common biases such as immortal time bias or prevalent user bias. It also facilitates a systematic methodological evaluation and comparison of observational studies (Specifying a target trial prevents immortal time bias and other self-inflicted injuries in observational analyses. J Clin Epidemiol. 2016;79: 70-5). How to estimate the effect of treatment duration on survival outcomes

using observational data (BMJ 2018;360: k182) proposes methods for overcoming bias with this approach when quantifying the effect of treatment duration.

### 5.3.2. Methods to address selection bias

New user (incident user) designs restrict the study population to persons who are observed at the start of treatment. New user design helps mitigate selection bias by preventing 'depletion of susceptibles' – an unwanted exclusion from a safety assessment of persons discontinuing treatments following early adverse reactions. It also helps alleviate healthy user bias for preventive treatments in some circumstances (see Healthy User and Related Biases in Observational Studies of Preventive Interventions: A Primer for Physicians. J Gen Intern Med 2011, 26(5):546-50). The article Evaluating medication effects outside of clinical trials: new-user designs (Am J Epidemiol 2003;158 (9):915–20) defines new-user designs and explains how they can be implemented in case-control studies. One should also be aware of the difference between a new user (which requires absence of prior use of a given drug/drug class during a prespecified washout period) and a treatment-naïve user (which requires absence of prior treatment for a given indication). A treatment-naïve status may not be ascertainable in left-truncated data.

The active comparator new user design (see Chapter 5.3.4.2) would ideally compare two treatments that are marketed contemporaneously. However, a more common situation is where a recently marketed drug is compared with an older established alternative. For such situations, the article Prevalent new-user cohort designs for comparative drug effect studies by time-conditional propensity scores (Pharmacoepidemiol Drug Saf 2017;26(4):459-68) introduces a cohort design allowing identification of matched subjects using the comparator drug at the same point in the course of disease as the (newly marketed) drug of interest. The design utilises time-based and prescription-based exposure sets to compute time-dependent propensity scores of initiating the new drug

The use of case only study designs can also reduce selection bias if the statistical assumptions of the method are fulfilled (see Chapter 5.3.4.1).

### 5.3.3. Methods to address misclassification bias

Misclassification can occur in exposure, outcome or covariate variables. Outcome misclassification occurs when a non-case is classified as a case (false positive error) or a case is classified as a non-case (false negative error). Errors are quantified as estimates of positive predictive value, negative predictive value, sensitivity and specificity. Most database studies will be subject to outcome misclassification to some degree, although case adjudication against an established case definition or a reference standard can remove false positives, and false negatives can be mitigated if a broad search algorithm is used. The influence of misclassification on the point estimate should be quantified or, if this is not possible, its impact on the interpretation of the results should be discussed. Exposure misclassification may also occur and one should avoid the epidemiologic 'mantra' about non-differential misclassification of exposure producing conservative estimates. It holds true, on the average, for dichotomous exposures that have an effect, but does not necessarily apply to any given estimate (Proper interpretation of non-differential misclassification effects: expectations vs observations. Int J Epidemiol 2005;34(3):680-7).

Good practices for quantitative bias analysis (Int J Epidemiol 2014;43(6):1969-85) advocates explicit and quantitative assessment of misclassification bias, including guidance on which biases to assess in each situation, what level of sophistication to use, and how to present the results. When outcome status is misclassified, relative measures of association are unbiased if specificity of ascertainment is high.

In Use of the Positive Predictive Value to Correct for Disease Misclassification in Epidemiologic Studies (Am J Epidemiol 1993;138(11):1007-15), Brenner and Gefeller propose a method based on estimates of the positive predictive value which requires validation of a sample of patients with the outcome only, while assuming that sensitivity is non-differential. A web application allows correction of risk ratio or cumulative incidence point estimates and confidence intervals for bias due to outcome misclassification based on Brenner and Gefeller's methodology. The article Basic methods for sensitivity analysis of biases (Int J Epidemiol 1996;25(6):1107-16) provides different examples of methods for examining the sensitivity of study results to biases, with a focus on methods that can be implemented without computer programming.

## 5.3.4. Methods to address confounding

### *5.3.4.1. Case-only designs*

Case-only designs reduce confounding by using the exposure history of each case as its own control and thereby eliminate confounding by characteristics that are constant over time, such as sex, socio-economic factors, genetics and chronic diseases. A review of case only designs is available in Use of self-controlled designs in pharmacoepidemiology (J Intern Med 2014; 275(6): 581-9).

A simple form of a case-only design is the symmetry analysis (initially described as prescription sequence symmetry analysis), introduced as a screening tool in Evidence of depression provoked by cardiovascular medication: a prescription sequence symmetry analysis (Epidemiology 1996;7(5):478-84).

The case-crossover design compares the risk of exposure in a time period prior to an outcome with that in an earlier reference time-period, or set of time periods, to examine the effect of transient exposures on acute events (see The Case-Crossover Design: A Method for Studying Transient Effects on the Risk of Acute Events, Am J Epidemiol 1991;133(2):144-53). The case-time-control designs are a modification of case-crossover designs which use exposure history data from a traditional control group to estimate and adjust for the bias from temporal changes in prescribing (The case-time-control design, Epidemiology 1995;6(3):248-53). However, if not well matched, the case-time-control group may reintroduce selection bias (Confounding and exposure trends in case-crossover and case-time-control designs (Epidemiology 1996;7(3):231-9). Methods have been suggested to overcome the exposure-trend bias while controlling for time-invariant confounders (see Future cases as present controls to adjust for exposure trend bias in case-only studies, Epidemiology 2011;22(4):568-74. Persistent User Bias in Case-Crossover Studies in Pharmacoepidemiology (Am J Epidemiol 2016; 184(10):761-9) demonstrates that case-crossover studies of drugs that may be used indefinitely are biased upward. This bias is alleviated, but not removed completely, by using a control group.

In the self-controlled case series (SCCS) design, the observation period following each exposure for each case is divided into risk period(s) (e.g. number of days immediately following each exposure) and a control period (observed time outside this risk period). Incidence rates within the risk period after exposure are compared with incidence rates within the control period. The Tutorial in biostatistics: the self-controlled case series method (Stat Med 2006; 25(10):1768-97) explains how to fit SCCS models using standard statistical packages. The bias introduced by inaccurate specification of the risk window is discussed and a data-based approach for identifying the optimal risk windows is proposed in Identifying optimal risk windows for self-controlled case series studies of vaccine safety (Stat Med 2011; 30(7):742-52). The SCCS also assumes that the event itself does not affect the chance of being exposed. The pseudo-likelihood method developed to address this possible issue is described in Cases series analysis for censored, perturbed, or curtailed post-event exposures (Biostatistics 2009;10(1):3-16). Use of the self-controlled case-series method in vaccine safety studies: review and

recommendations for best practice (Epidemiol Infect 2011;139(12):1805-17) assesses how the SCCS method has been used across 40 vaccine studies, highlights good practice and gives guidance on how the method should be used and reported. Using several methods of analysis is recommended, as it can reinforce conclusions or shed light on possible sources of bias when these differ for different study designs.

When should case-only designs be used for safety monitoring of medical products? (Pharmacoepidemiol Drug Saf 2012;21(Suppl. 1):50-61) compares the SCCS and case-crossover methods as to their use, strength and major difference (directionality). It concludes that case-only analyses of intermittent users complement the cohort analyses of prolonged users because their different biases compensate for one another. It also provides recommendations on when case-only designs should and should not be used for Drug Safety monitoring. Empirical performance of the self-controlled case series design: lessons for developing a risk identification and analysis system (Drug Saf 2013;36(Suppl. 1):S83-S93) evaluates the performance of the SCCS design using 399 drug-health outcome pairs in 5 observational databases and 6 simulated datasets. Four outcomes and five design choices were assessed. Within-person study designs had lower precision and greater susceptibility to bias because of trends in exposure than cohort and nested case-control designs (J Clin Epidemiol 2012;65(4):384-93) compares cohort, case-control, case-cross-over and SCCS designs to explore the association between thiazolidinediones and the risks of heart failure and fracture and anticonvulsants and the risk of fracture. Bias was removed when follow-up was sampled both before and after the outcome, or when a case-time-control design was used.

#### *5.3.4.2. Use of an active comparator*

The main purpose of using an active comparator is to reduce confounding by indication or by severity. Its use is optimal in the context of the new user design, whereby comparison is between patients with the same indication initiating different treatments (The active comparator, new user study design in pharmacoepidemiology: historical foundations and contemporary application, Curr Epidemiol Rep 2015;2(4):221-8). An active comparator should be chosen to represent the counterfactual risk of a given outcome with a different treatment, i.e. it should have a known and positive safety profile with respect to the events of interest and ideally represent the background risk in the diseased (for example, safety of antiepileptics in pregnancy in relation to risk of congenital malformations could be compared against that of lamotrigine, which is not known to be teratogenic). The paper Using Big Data to Emulate a Target Trial When a Randomized Trial Is Not Available (Am J Epidemiol. 2016;183(8):758-64) channels counterfactual theory for comparing the effects of treatment strategies helping avoid common methodologic pitfalls. The C-Word: Scientific Euphemisms Do Not Improve Causal Inference From Observational Data (Am J Public Health 2018;108(5):616-19) highlights the need to be explicit about the causal objective of a study to help for the emulation of a particular target trial and support the choice of confounding adjustment variables.

With newly marketed medicines, no active comparator with ideal comparability of patients' characteristics may be available because prescribing newly marketed medicines may be driven to a greater extent by patients' prognostic characteristics (early users may be either sicker or healthier than all patients with the indication) and by reimbursement considerations than prescribing of established medicines. This is described for comparative effectiveness studies in Assessing the comparative effectiveness of newly marketed medications: methodological challenges and implications for drug development (Clin Pharmacol Ther 2011;90(6):777-90) and in Newly marketed medications present unique challenges for nonrandomized comparative effectiveness analyses. (J Comp Eff Res 2012;1(2):109-11). Other challenges include treatment effect heterogeneity as patient characteristics of users evolve over time, and low precision owing to slow drug uptake.

#### *5.3.4.3. Disease risk scores*

An approach to controlling for a large number of confounding variables is to summarise them in a single multivariable confounder score. Stratification by a multivariate confounder score (Am J Epidemiol 1976;104(6):609-20) shows how control for confounding may be based on stratification by the score. An example is a disease risk score (DRS) that estimates the probability or rate of disease occurrence conditional on being unexposed. The association between exposure and disease is then estimated with adjustment for the disease risk score in place of the individual covariates.

DRSs are however difficult to estimate if outcomes are rare. Use of disease risk scores in pharmacoepidemiologic studies (Stat Methods Med Res 2009;18(1):67-80) includes a detailed description of their construction and use, a summary of simulation studies comparing their performance to traditional models, a comparison of their utility with that of propensity scores, and some further topics for future research. Disease risk score as a confounder summary method: systematic review and recommendations (Pharmacoepidemiol Drug Saf 2013;22(2);122-29), examines trends in the use and application of DRS as a confounder summary method and shows that large variation exists with differences in terminology and methods used.

In Role of disease risk scores in comparative effectiveness research with emerging therapies (Pharmacoepidemiol Drug Saf 2012;21 Suppl 2:138–47), it is argued that DRS may have a place when studying drugs that are recently introduced to the market. In such situations, as characteristics of users change rapidly, exposure propensity scores may prove highly unstable. DRSs based mostly on biological associations would be more stable. However, DRS models are still sensitive to misspecification as discussed in Adjusting for Confounding in Early Postlaunch Settings: Going Beyond Logistic Regression Models (Epidemiology 2016;27(1):133-42).

#### *5.3.4.4. Propensity scores*

Databases used in pharmacoepidemiological studies often include records of prescribed medications and encounters with medical care providers, from which one can construct surrogate measures for both drug exposure and covariates that are potential confounders. It is often possible to track day-by-day changes in these variables. However, while this information can be critical for study success, its volume can pose challenges for statistical analysis.

A propensity score (PS) is analogous to the disease risk score in that it combines a large number of possible confounders into a single variable (the score). The exposure propensity score (EPS) is the conditional probability of exposure to a treatment given observed covariates. In a cohort study, matching or stratifying treated and comparison subjects on EPS tends to balance all of the observed covariates. However, unlike random assignment of treatments, the propensity score may not balance unobserved covariates. Invited Commentary: Propensity Scores (Am J Epidemiol. 1999;150(4):327–33) reviews the uses and limitations of propensity scores and provide a brief outline of the associated statistical theory. The authors present results of adjustment by matching or stratification on the propensity score.

High-dimensional Propensity Score Adjustment in Studies of Treatment Effects Using Healthcare Claims Data (Epidemiol. 2009;20(4):512-22) discusses the high dimensional propensity score (hd-PS) model approach. It attempts to empirically identify large numbers of potential confounders in healthcare databases and, by doing so, to extract more information on confounders and proxies. Covariate selection in high-dimensional propensity score analyses of treatment effects in small samples (Am J Epidemiol 2011;173(12):1404-13) evaluates the relative performance of hd-PS in smaller samples. Confounding adjustment via a semi-automated high-dimensional propensity score algorithm: an application to electronic medical records (Pharmacoepidemiol Drug Saf 2012;20(8):849-57) evaluates the use of hd-PS in a primary care electronic medical record database. In addition, the article Using high-dimensional propensity scores to automate confounding control in a distributed medical product

safety surveillance system (Pharmacoepidemiol Drug Saf 2012;21(S1):41-9) summarises the application of this method for automating confounding control in sequential cohort studies as applied to safety monitoring systems using healthcare databases and also discusses the strengths and limitations of hd-PS.

Most cohort studies match patients 1:1 on the propensity score. Increasing the matching ratio may increase precision but also bias. One-to-many propensity score matching in cohort studies (Pharmacoepidemiol Drug Saf 2012;21(S2):69-80) tests several methods for 1:n propensity score matching in simulation and empirical studies and recommends using a variable ratio that increases precision at a small cost of bias. Matching by propensity score in cohort studies with three treatment groups (Epidemiology 2013;24(3):401-9) develops and tests a 1:1:1 propensity score matching approach offering a way to compare three treatment options.

The use of several measures of balance for developing an optimal propensity score model is described in Measuring balance and model selection in propensity score methods (Pharmacoepidemiol Drug Saf 2011;20(11):1115-29) and further evaluated in Propensity score balance measures in pharmacoepidemiology: a simulation study (Pharmacoepidemiol Drug Saf 2014;23(8):802-11). In most situations, the standardised difference performs best and is easy to calculate (see Balance measures for propensity score methods: a clinical example on beta-agonist use and the risk of myocardial infarction (Pharmacoepidemiol Drug Saf 2011;20(11):1130-7) and Reporting of covariate selection and balance assessment in propensity score analysis is suboptimal: a systematic review (J Clin Epidemiol 2015;68(2):112-21)). Metrics for covariate balance in cohort studies of causal effects (Stat Med 2013;33:1685-99) shows in a simulation study that the c-statistics of the PS model after matching and the general weighted difference perform as well as the standardized difference and are preferred when an overall summary measure of balance is requested. Treatment effects in the presence of unmeasured confounding: dealing with observations in the tails of the propensity score distribution--a simulation study (Am J Epidemiol. 2010;172(7):843-54) demonstrates how 'trimming' of the propensity score eliminates subjects who are treated contrary to prediction and their exposed/unexposed counterparts, thereby reducing bias by unmeasured confounders.

Performance of propensity score calibration--a simulation study (Am J Epidemiol 2007;165(10):1110-8) introduces 'propensity score calibration' (PSC). This technique combines propensity score matching methods with measurement error regression models to address confounding by variables unobserved in the main study. This is done by using additional covariate measurements observed in a validation study, which is often a subset of the main study.

Although in most situations propensity score models, with the possible exception of hd-PS, do not have any advantages over conventional multivariate modelling in terms of adjustment for identified confounders, several other benefits may be derived. Propensity score methods may help to gain insight into determinants of treatment including age, frailty and comorbidity and to identify individuals treated against expectation. A statistical advantage of PS analyses is that if exposure is not infrequent it is possible to adjust for a large number of covariates even if outcomes are rare, a situation often encountered in Drug Safety research. Furthermore, assessment of the PS distribution may reveal non-positivity. An important limitation of PS is that it is not directly amenable for case-control studies. A critical assessment of propensity scores is provided in Propensity scores: from naive enthusiasm to intuitive understanding (Stat Methods Med Res 2012;21(3):273-93). Semiautomated and machine-learning based approaches to propensity score methods are currently being developed (Automated data-adaptive analytics for electronic healthcare data to study causal treatment effects (Clin Epidemiol 2018;10:771-88).

### *5.3.4.5. Instrumental variables*

Instrumental variable (IV) analysis is an approach to address uncontrolled confounding in comparative studies. An introduction to instrumental variables for epidemiologists (Int J Epidemiol 2000;29(4):722-9) presents those developments, illustrated by an application of IV methods to non-parametric adjustment for non-compliance in randomised trials. The author mentions a number of caveats but concludes that IV corrections can be valuable in many situations. IV analysis in comparative safety and effectiveness research is reviewed in Instrumental variable methods in comparative safety and effectiveness research (Pharmacoepidemiol Drug Saf 2010; 19(6):537-54). A review of IV analysis for observational comparative effectiveness studies suggested that in the large majority of studies, in which IV analysis was applied, one of the assumption could be violated (Potential bias of instrumental variable analyses for observational comparative effectiveness research, Ann Intern Med. 2014;161(2):131-8).

A proposal for reporting instrumental variable analyses has been suggested in Commentary: how to report instrumental variable analyses (suggestions welcome) (Epidemiology 2013;24(3):370-4). In particular the type of treatment effect (average treatment effect/homogeneity condition or local average treatment effect/monotonicity condition) and the testing of critical assumptions for valid IV analyses should be reported. In support of these guidelines, the standardized difference has been proposed to falsify the assumption that confounders are not related to the instrumental variable (Quantitative falsification of instrumental variables assumption using balance measures, Epidemiology 2014;25(5):770-2).

The complexity of the issues associated with confounding by indication, channelling and selective prescribing is explored in Evaluating short-term drug effects using a physician-specific prescribing preference as an instrumental variable (Epidemiology 2006;17(3):268-75). A conventional, adjusted multivariable analysis showed a higher risk of gastrointestinal toxicity for selective COX-2-inhibitors than for traditional NSAIDs, which was at odds with results from clinical trials. However, a physician-level instrumental variable approach (a time-varying estimate of a physician's relative preference for a given drug, where at least two therapeutic alternatives exist) yielded evidence of a protective effect due to COX-2 exposure, particularly for shorter term drug exposures. Despite the potential benefits of physician-level IVs their performance can vary across databases and strongly depends on the definition of IV used as discussed in Evaluating different physician's prescribing preference based instrumental variables in two primary care databases: a study of inhaled long-acting beta2-agonist use and the risk of myocardial infarction (Pharmacoepidemiol Drug Saf 2016;25 Suppl 1:132-41).

Instrumental variable methods in comparative safety and effectiveness research (Pharmacoepidemiol Drug Saf 2010;19(6):537–54) is a practical guidance on IV analyses in pharmacoepidemiology. Instrumental variable methods for causal inference (Stat Med 2014;33(13):2297-340) is a tutorial, including statistical code for performing IV analysis.

An important limitation of IV analysis is that weak instruments (small association between IV and exposure) lead to decreased statistical efficiency and biased IV estimates as detailed in Instrumental variables: application and limitations (Epidemiology 2006;17:260-7). For example, in the above mentioned study on non-selective NSAIDs and COX-2-inhibitors, the confidence intervals for IV estimates were in the order of five times wider than with conventional analysis. Performance of instrumental variable methods in cohort and nested case-control studies: a simulation study (Pharmacoepidemiol Drug Saf 2014; 2014;23(2):165-77) demonstrated that a stronger IV-exposure association is needed in nested case-control studies compared to cohort studies in order to achieve the same bias reduction. Increasing the number of controls reduces this bias from IV analysis with relatively weak instruments.

Selecting on treatment: a pervasive form of bias in instrumental variable analyses (Am J Epidemiol 2015;181(3):191-7) warns against bias in IV analysis by including only a subset of possible treatment options.

#### *5.3.4.6. Prior event rate ratios*

Another method proposed to control for unmeasured confounding is the Prior Event Rate Ratio (PERR) adjustment method, in which the effect of exposure is estimated using the ratio of rate ratios (RRs) from periods before and after initiation of a drug exposure, as discussed in Replicated studies of two randomized trials of angiotensin converting enzyme inhibitors: further empiric validation of the 'prior event rate ratio' to adjust for unmeasured confounding by indication (Pharmacoepidemiol Drug Saf 2008;17(7):671-685). For example, when a new drug is launched, direct estimation of the drugs effect observed in the period after launch is potentially confounded. Differences in event rates in the period before the launch between future users and future non-users may provide a measure of the amount of confounding present. By dividing the effect estimate from the period after launch by the effect obtained in the period before launch, the confounding in the second period can be adjusted for. This method requires that confounding effects are constant over time, that there is no confounder-by-treatment interaction, and outcomes are non-lethal events.

Performance of prior event rate ratio adjustment method in pharmacoepidemiology: a simulation study (Pharmacoepidemiol Drug Saf 2015(5);24:468-477) discusses that the PERR adjustment method can help to reduce bias as a result of unmeasured confounding in certain situations but that theoretical justification of assumptions should be provided.

#### *5.3.4.7. Handling time-dependent confounding in the analysis*

Methods for dealing with time-dependent confounding (Stat Med. 2013;32(9):1584-618) provides an overview of how time-dependent confounding can be handled in the analysis of a study. It provides an in-depth discussion of marginal structural models and g-computation.

Beyond the G-estimation and the Marginal Structural Model (MSM) described below, traditional and efficient approaches to deal with time dependent variables should be considered in the design of the study, such as nested case control studies with assessment of time varying exposure windows.

G-estimation is a method for estimating the joint effects of time-varying treatments using ideas from instrumental variables methods. G-estimation of Causal Effects: Isolated Systolic Hypertension and Cardiovascular Death in the Framingham Heart Study (Am J Epidemiol 1998;148(4):390-401) demonstrates how the G-estimation procedure allows for appropriate adjustment of the effect of a time-varying exposure in the presence of time-dependent confounders that are themselves influenced by the exposure.

The use of Marginal Structural Models can be an alternative to G-estimation. Marginal Structural Models and Causal Inference in Epidemiology (Epidemiology 2000;11(5):550-60) introduces a class of causal models that allow for improved adjustment for confounding in situations of time-dependent confounding.

MSMs have two major advantages over G-estimation. Even if it is useful for survival time outcomes, continuous measured outcomes and Poisson count outcomes, logistic G-estimation cannot be conveniently used to estimate the effect of treatment on dichotomous outcomes unless the outcome is rare. The second major advantage of MSMs is that they resemble standard models, whereas G-estimation does not (see Marginal Structural Models to Estimate the Causal Effect of Zidovudine on the Survival of HIV-Positive Men. Epidemiology 2000;11(5):561-70).

Effect of highly active antiretroviral therapy on time to acquired immunodeficiency syndrome or death using marginal structural models (Am J Epidemiol 2003;158(7):687-94) provides a clear example in which standard Cox analysis failed to detect a clinically meaningful net benefit of treatment because it does not appropriately adjust for time-dependent covariates that are simultaneously confounders and intermediate variables. This net benefit was shown using a marginal structural survival model. In Time-dependent propensity score and collider-stratification bias: an example of beta2-agonist use and the risk of coronary heart disease (Eur J Epidemiol 2013;28(4):291-9), various methods to control for time-dependent confounding are compared in an empirical study on the association between inhaled beta-2-agonists and the risk of coronary heart disease. MSMs resulted in slightly reduced associations compared to standard Cox-regression.

#### *5.3.4.8. The trend-in-trend design*

The Trend-in-trend Research Design for Causal Inference (Epidemiology 2017;28: 529-36) presents a semi-ecological design, whereby trends in exposure and in outcome rates are compared in subsets of the population that have different rates of uptake for the drug in question. These subsets are identified through PS modelling. There is a formal framework for transforming the observed trends into an effect estimate. Simulation and empirical studies showed the design to be less statistically efficient than a cohort study, but more resistant to confounding. The trend-in-trend method may be useful in settings where there is a strong time trend in exposure, such as a newly approved drug.

### 5.3.5. Positive and negative control exposures and outcomes

One may test the validity of putative causal associations by using control exposures or outcomes. Well-chosen positive and negative controls help convince investigator that the data at hand correctly detect existing associations or correctly demonstrate lack of association when none is expected. Positive controls turning out as negative and negative as positive may signal presence of a bias, as illustrated in a study demonstrating health adherer bias by showing that adherence to statins was associated with decreased risks of biologically implausible outcomes (Statin adherence and risk of accidents: a cautionary tale, Circulation 2009;119(15):2051-7). The general principle, with additional examples, is described in Control Outcomes and Exposures for Improving Internal Validity of Nonrandomized Studies (Health Serv Res 2015;50(5):1432-51).

Selecting drug-event combinations as reliable controls poses a challenge: it is difficult to establish for negative controls proof of absence of an association, and it is still more problematic to select positive controls because it is desirable not only to establish an association but also an accurate estimate of the effect size. This has led to attempts to establish libraries of controls that can be used to characterise the performance of different observational datasets in detecting various types of association using a number of different study designs. Although this kind of controls may be questioned according to Evidence of Misclassification of Drug-Event Associations Classified as Gold Standard 'Negative Controls' by the Observational Medical Outcomes Partnership (OMOP) (Drug Saf 2016;39(5):421-32), the approach of calibrating the performance of epidemiological methods prior to performing a study holds the promise of providing a trustworthy framework for interpretation of the results, as shown by Interpreting observational studies: Why empirical calibration is needed to correct p-values (Stat Med. 2014;33(2):209-18), Robust empirical calibration of p-values using observational data (Stat Med 2016;35(22):3883-8) and Empirical confidence interval calibration for population-level effect estimation studies in observational healthcare data (Proc Natl Acad Sci USA 2018;115 (11): 571-7).

### 5.3.6. Triangulation

Triangulation is not a separate methodological approach, but rather a framework, formally described in Triangulation in aetiological epidemiology (Int J Epidemiol 2016;45(6):1866-86). Triangulation is defined as "the practice of obtaining more reliable answers to research questions through integrating results from several different approaches, where each approach has different key sources of potential bias that are unrelated to each other." In some ways, the paper formalises approaches already used in many nonrandomised pharmacoepidemiologic studies, including control exposures and outcomes, sensitivity analyses, comparing results from different population and different study designs – all within the same study and while explicitly specifying the direction of bias in each approach. Triangulation was used (without using the explicit term) in Associations of maternal antidepressant use during the first trimester of pregnancy with preterm birth, small for gestational age, autism spectrum disorder, and attention-deficit/hyperactivity disorder in offspring (JAMA 2017;317(15):1553-62), whereby, within the same study, the authors used negative controls (paternal exposure to antidepressants), and assess the association using different study design and study population (sibling design).

### 5.3.7. Interrupted time series analysis

In evaluating effectiveness of population-level interventions that are implemented at a specific point in time (clear before-after periods, such as policy effect date, regulatory action date) interrupted time series (ITS) studies are becoming the standard approach. The ITS analysis establishes the expected pre-intervention trend for an outcome of interest. The counterfactual scenario in the absence of the intervention serves as the comparator, the expected trend that provides a comparison for the evaluation of the impact of the intervention by examining any change occurring following the intervention period (Interrupted time series regression for the evaluation of public health interventions: a tutorial. Int J Epidemiol. 2017; 46(1):348-55). ITS is a quasi-experimental design and has been described as the "next best" approach for dealing with interventions in the absence of randomisation. ITS analysis requires several assumptions and its implementation is technically sophisticated, as explained in Regression based quasi-experimental approach when randomisation is not an option: Interrupted time series analysis (BMJ 2015; 350:h2750). The use of ITS regression in impact research is illustrated in Annex 2 'Guidance on methods for pharmacovigilance impact research' of this Guide.

## *5.4. Effect measure modification and interaction*

Effect measure modification and interaction are often encountered in epidemiological research and it is important to recognize their occurrence. The difference between these terms is rather subtle and has been described in On the distinction between interaction and effect modification (Epidemiology 2009;20(6):863–71). Effect measure modification occurs when the measure of an effect changes over values of some other variable (which does not necessarily need to be a causal factor). Interaction occurs when two exposures contribute to the causal effect of interest, and they are both causal factors. Interaction is generally studied in order to clarify etiology while effect modification is used to identify populations that are particularly susceptible to the exposure of interest.

To check the presence of an effect measure modifier, one can stratify the study population by a certain variable, e.g. by gender, and compare the effects in these subgroups. It is recommended to perform a formal statistical test to assess if there are statistically significant differences between subgroups for the effects (see CONSORT 2010 Explanation and Elaboration: Updated guidelines for reporting parallel group randomised trials, J Clin Epidemiol 2010;63(8):e1-37 and Interaction revisited: the difference

between two estimates, BMJ 2003;326(7382):219). The study report should explain which method was used to examine these differences and specify which subgroup analyses were predefined in the study protocol and which ones were performed while analysing the data (Strengthening the Reporting of Observational Studies in Epidemiology (STROBE): explanation and elaboration. Epidemiology 2007;18(6):805-35).

The presence of effect measure modification depends on which measure is used in the study (absolute or relative) and can be measured in two ways: on an additive scale (based on risk differences [RD]), or on a multiplicative scale (based on relative risks [RR]). From the perspective of public health and clinical decision making, the additive scale is usually considered the most appropriate. An example of potential effect modifier in studies assessing the risk of occurrence of events associated with recent drug use is the past use of the same drug. This is shown in Evidence of the depletion of susceptibles effect in non-experimental pharmacoepidemiologic research (J Clin Epidemiol 1994;47(7):731-7) in the context of a hospital-based case-control study on NSAIDs and the risk of upper gastrointestinal bleeding.

For the evaluation of interaction, the standard measure is the relative excess risk due to interaction (RERI), as explained in the textbook Modern Epidemiology (K. Rothman, S. Greenland, T. Lash. 3rd Edition, Lippincott Williams & Wilkins, 2008). Other measures of interaction include the attributable proportion (A) and the synergy index (S). According to Exploring interaction effects in small samples increases rates of false-positive and false-negative findings: results from a systematic review and simulation study (J Clin Epidemiol 2014; 67(7):821-9), with sufficient sample size, most interaction tests perform similarly with regard to type 1 error rates and power.

Due to confusion about these terms, it is important that effect measure modification and interaction analysis are presented in a way that is easy to interpret and allows readers to reproduce the analysis. For recommendations regarding reporting, Strengthening the Reporting of Observational Studies in Epidemiology (STROBE): explanation and elaboration (Epidemiology 2007;18(6):805-35) and Recommendations for presenting analyses of effect modification and interaction (Int J Epidemiol 2012;41(2):514-20) are useful resources. They recommend to present the results as follows:

- Separate effects (rate ratios, odds ratios or risk differences, with confidence intervals) of the exposure of interest (e.g. drug), of the effect modifier (e.g. gender) and of their joint effect using one single reference category (preferably the stratum with the lowest risk of the outcome) as suggested in Estimating measures of interaction on an additive scale for preventive exposures (Eur J Epidemiol 2011;26(6):433-8), as this gives enough information to the reader to calculate effect modification on an additive and multiplicative scale;
- Effects of the exposure within strata of the potential effect modifier;
- Measures of effect modification on both additive (e.g. RERI) and multiplicative (e.g. S) scales including confidence intervals;
- List of the confounders for which the association between exposure and outcome was adjusted for.

## *5.5. Ecological analyses and case-population studies*

Ecological analyses are not hypothesis testing but hypothesis generating studies. As illustrated in Control without separate controls: evaluation of vaccine safety using case-only methods (Vaccine 2004; 22(15-16):2064-70), ecological analyses assume that a strong correlation between the trend in an indicator of an exposure (vaccine coverage in this example) and the trend in incidence of a disease (trends calculated over time or across geographical regions) is consistent with a causal relationship. Such comparisons at the population level may only generate hypotheses as they do not allow

controlling for time-related confounding variables, such as age and seasonal factors. Moreover, they do not establish that the vaccine effect occurred in the vaccinated individuals.

Case-population studies are a form of ecological studies where cases are compared to an aggregated comparator consisting of population data. The case-population study design: an analysis of its application in pharmacovigilance (Drug Saf 2011;34(10):861-8) explains its design and its application in pharmacovigilance for signal generation and drug surveillance. The design is also explained in Chapter 2: *Study designs in drug utilization research* of the textbook *Drug Utilization Research - Methods and Applications* (M Elseviers, B Wettermark, AB Almarsdóttir, et al. Editors. Wiley Blackwell, 2016). An example is a multinational case-population study aiming to estimate population rates of a suspected adverse event using national sales data (see Transplantation for Acute Liver Failure in Patients Exposed to NSAIDs or Paracetamol (Drug Saf 2013;36(2):135–44). Based on the same study, Choice of the denominator in case population studies: event rates for registration for liver transplantation after exposure to NSAIDs in the SALT study in France (Pharmacoepidemiol Drug Saf 2013;22(2):160-7) compared sales data and healthcare insurance data as denominators to estimate population exposure and found large differences in the event rates. Choosing the wrong denominator in case population studies might generate erroneous results. The choice of the right denominator depends not only on a valid data source but will also depend on the hazard function of the adverse event.

A pragmatic attitude towards case-population studies is recommended: in situations where nation-wide or region-wide electronic health records (EHR) are available and allow assessing the outcomes and confounders with sufficient validity, a case-population approach is neither necessary nor desirable, as one can perform a population-based cohort or case-control study with adequate control for confounding. In situations where outcomes are difficult to ascertain in EHR or where such databases do not exist, the case-population design might give an approximation of the absolute and relative risk when both events and exposures are rare. This is limited by the ecological nature of the reference data that restricts the ability to control for confounding.

## *5.6. Pragmatic trials and large simple trials*

### 5.6.1. Pragmatic trials

RCTs are considered the gold standard for demonstrating the efficacy of medicinal products and for obtaining an initial estimate of the risk of adverse outcomes. However, they are not necessarily indicative of the benefits, risks or comparative effectiveness of an intervention when used in clinical practice. The IMI GetReal Glossary defines a pragmatic clinical trial as 'a study comparing several health interventions among a randomised, diverse population representing clinical practice, and measuring a broad range of health outcomes'. Pragmatic clinical trials are focused on evaluating benefits and risks of treatments in patient populations and settings that are more representative of routine clinical practice. To ensure generalisability, pragmatic trials should represent the patients to whom the treatment will be applied, for instance, inclusion criteria may be broader (e.g. allowing co-morbidity, co-medication, wider age range), and the follow-up may be minimised and allow for treatment switching. Monitoring safety in a phase III real-world effectiveness trial: use of novel methodology in the Salford Lung Study (Pharmacoepidemiol Drug Saf 2017;26(3):344-352) describes the model of a phase III pragmatic clinical trial where patients were enrolled through primary care practices using minimal exclusion criteria and without extensive diagnostic testing, and where potential safety events were captured through patients' electronic health records and triggered review by the specialist safety team.

Pragmatic explanatory continuum summary (PRECIS): a tool to help trial designers (CMAJ 2009; 180(10): E45-E57) is a tool to support pragmatic trial designs and help define and evaluate the degree of pragmatism. The PRECIS tool has been further refined and now comprises nine domains each scored on a 5 point Likert scale ranging from very explanatory to very pragmatic with an exclusive focus on the issue of applicability (The PRECIS-2 tool: designing trials that are fit for purpose. BMJ 2015;350: h2147). A checklist and additional guidance is also provided in Improving the reporting of pragmatic trials: an extension of the CONSORT statement (BMJ 2008; 337 (a2390): 1-8).

Based on the evidence that the current costs and complexity of conducting randomised trials lead to more restrictive eligibility criteria and short durations of trials, and therefore reduce the generalisability and reliability of the evidence about the efficacy and safety of new and existing interventions, the article The Magic of Randomization versus the Myth of Real-World Evidence (N Engl J Med. 2020;382(7):674-678) proposes measures to remove practical obstacles to the conduct of randomised trials of appropriate size.

### 5.6.2. Large simple trials

Large simple trials are pragmatic clinical trials with minimal data collection narrowly focused on clearly defined outcomes important to patients as well as clinicians. Their large sample size provides adequate statistical power to detect even small differences in effects. Additionally, large simple trials include a follow-up time that mimics routine clinical practice.

Large simple trials are particularly suited when an adverse event is very rare or has a delayed latency (with a large expected attrition rate), when the population exposed to the risk is heterogeneous (e.g. different indications and age groups), when several risks need to be assessed in the same trial or when many confounding factors need to be balanced between treatment groups. In these circumstances, the cost and complexity of a traditional RCT may outweigh its advantages and large simple trials can help keep the volume and complexity of data collection to a minimum.

Outcomes that are simple and objective can also be measured from the routine process of care using epidemiological follow-up methods, for example by using questionnaires or hospital discharge records. Examples of published large simple trials are An assessment of the safety of paediatric ibuprofen: a practitioner based randomised clinical trial (JAMA 1995;279:929-33) and Comparative mortality associated with ziprasidone and olanzapine in real-world use among 18,154 patients with schizophrenia: The Zodiac Observational Study of Cardiac Outcomes (ZODIAC) (Am J Psychiatry 2011;168(2):193-201).

Note that the use of the term 'simple' in the expression 'Large simple trials' refers to data structure and not to data collection. It is used in relation to situations in which a small number of outcomes are measured. The term may therefore not adequately reflect the complexity of the studies undertaken.

### 5.6.3. Randomised database studies

Randomised database studies can be considered a special form of a large simple trial where patients included in the trial are enrolled in a healthcare system with electronic records. Eligible patients may be identified and flagged automatically by the software, with the advantage of allowing comparison of included and non-included patients. Database screening or record linkage can be used to detect and measure outcomes of interest otherwise assessed through the normal process of care. Patient recruitment, informed consent and proper documentation of patient information are hurdles that still need to be addressed in accordance with the applicable legislation for RCTs. Randomised database studies attempt to combine the advantages of randomisation and observational database studies. These and other aspects of randomised database studies are discussed in The opportunities and

challenges of pragmatic point-of-care randomised trials using routinely collected electronic records: evaluations of two exemplar trials (Health Technol Assess. 2014;18(43):1-146) which illustrates the practical implementation of randomised studies in general practice databases.

There are few published examples of randomised database studies, but this design could become more common in the near future with the increasing computerisation of medical records. Pragmatic randomised trials using routine electronic health records: putting them to the test (BMJ 2012;344:e55) describes a project to implement randomised trials in the everyday clinical work of general practitioners, comparing treatments that are already in common use, and using routinely collected electronic healthcare records both to identify participants and to gather results.

A particular form of randomised databases studies is the registry-based randomised trial, which uses an existing registry as a platform for the identification of cases, their randomisation and their follow-up. The editorial The randomized registry trial - the next disruptive technology in clinical research? (N Engl J Med 2013; 369(17):1579-1581) introduces the concept. This hybrid design tries to achieve both internal and external validity by using a robust design (a RCT) in a data source with higher generalisability (registries). Other examples are the TASTE trial that followed patients in the long-term using data from a Scandinavian registry (Thrombus aspiration during ST-segment elevation myocardial infarction. N. Engl J Med. 2013;369(17):1587-97) and A registry-based randomized trial comparing radial and femoral approaches in women undergoing percutaneous coronary intervention: the SAFE-PCI for Women (Study of Access Site for Enhancement of PCI for Women) trial (JACC Cardiovasc Interv. 2014;7(8):857-67).

## *5.7. Systematic reviews and meta-analysis*

Identification and integration of evidence derived from results from several studies with the same or similar research objective can extend our understanding of the research question. A systematic literature review aims to collect in a systematic and explicit manner all empirical evidence that fits pre-specified eligibility criteria to answer a specific research question and to critically appraise relevant results. A meta-analysis involves the use of statistical techniques to integrate and summarise the results of identified studies. The focus of this activity may be to learn from the diversity of designs, results and associated gaps in knowledge as well as to obtain overall risk estimates. An example of a systematic review and meta-analysis of results of individual studies with potentially different design is given in Variability in risk of gastrointestinal complications with individual NSAIDs: results of a collaborative meta-analysis (BMJ 1996;312(7046):1563-6), which compared the relative risks of serious gastrointestinal complications reported with individual NSAIDs by conducting a systematic review of twelve hospital and community based case-control and cohort studies, and found a relation between use of the drugs and admission to hospital for haemorrhage or perforation.

Systematic review and meta-analysis of observational studies and other epidemiological sources are becoming as common as those of randomised clinical trials (RCTs). Challenges in systematic reviews that assess treatment harms (Ann Intern Med 2005;142:1090-9) explains the different reasons why both are important in providing relevant information and knowledge for pharmacovigilance. However, the method of analysis differs when meta-analyses pool evidence from observational studies compared to RCTs.

A detailed guidance on the methodological conduct of systematic reviews and meta-analysis is reported in Annex 1 of this guide. This guidance includes links to other relevant resources.

It should be noted that meta-analysis, even of RCTs, shares characteristics with observational research as subjective criteria are often involved in the selection of studies to include. Careful planning in design of a meta-analysis and pre-specification of selection criteria, outcomes and analytical methods before

review of any study results may thus contribute to the confidence placed in the results. A further useful reference is the CIOMS Working Group X Guideline on Evidence Synthesis and Meta-Analysis for Drug Safety (Geneva 2016).

## *5.8. Signal detection methodology and application*

### 5.8.1. General aspects of signal detection

A general overview of methods for signal detection and recommendations for their application are provided in the report of the CIOMS Working Group VIII Practical aspects of signal detection in pharmacovigilance and empirical results on various aspects of signal detection obtained from the IMI PROTECT project have been summarised in Good Signal Detection Practices: Evidence from IMI PROTECT. (Drug Saf. 2016;39(6):469-90).

The EU Guideline on good pharmacovigilance practices (GVP) Module IX (Rev 1)- Signal Management defines signal management as the set of activities performed to determine whether, based on an examination of individual case safety reports (ICSRs), aggregated data from active surveillance systems or studies, literature information or other data sources, there are new risks associated with an active substance or a medicinal product or whether risks have changed. Signal management covers all steps from detecting signals (signal detection), through their validation and confirmation, analysis, prioritisation and assessment to recommending action, as well as the tracking of the steps taken and of any recommendations made.

The FDA's Guidance for Industry-Good Pharmacovigilance Practices and Pharmacoepidemiologic Assessment provides best practice for documenting, assessing and reporting individual case safety reports and case series and for identifying, evaluating, investigating and interpreting safety signals, including recommendations on data mining techniques and use of pharmacoepidemiological studies.

### 5.8.2. Methods of statistical signal detection

Quantitative analysis of spontaneous adverse drug reaction reports is routinely used in drug safety research. Quantitative signal detection using spontaneous ADR reporting (Pharmacoepidemiol Drug Saf 2009;18(6):427-36) describes the core concepts behind the most common methods, the proportional reporting ratio (PRR), reporting odds ratio (ROR), information component (IC) and empirical Bayes geometric mean (EBGM). The authors also discuss the role of Bayesian shrinkage in screening spontaneous reports and the importance of changes over time in screening the properties of the measures. Additionally, they discuss major areas of controversy (such as stratification and evaluation and implementation of methods) and give some suggestions as to where emerging research is likely to lead. Data mining for signals in spontaneous reporting databases: proceed with caution (Pharmacoepidemiol Drug Saf 2007;16(4):359–65) reviews data mining methodologies and their limitations and provides useful points to consider before incorporating data mining as a routine component of any pharmacovigilance program.

The revised guidance on Screening for adverse reactions in EudraVigilance describes methods (statistical and clinical information based) for screening adverse reactions and used by the European Medicines Agency, national competent authorities and Marketing Authorisation Holders. For the methods recommended, it addresses elements of their interpretation, their potential advantages and limitations and the evidence behind. Areas of uncertainty, that require resolution before firm recommendations can be made, are also mentioned.

Methods such as multiple logistic regression (that may use propensity score-adjustment) have the theoretical capability to reduce masking and confounding by co-medication and underlying disease.

The letter Logistic regression in signal detection: another piece added to the puzzle (Clin Pharmacol Ther 2013;94(3):312) highlights the variability of results obtained in different studies based on this method and the daunting computational task it requires. More work is needed on its value for pharmacovigilance in the real-world setting.

A more recent proposal involves a broadening of the basis for computational screening of individual case safety reports, by considering multiple aspects of the strength of evidence in a predictive model. This approach combines disproportionality analysis with features such as the number of well-documented reports, the number of recent reports and geographical spread of the case series (Improved statistical signal detection in pharmacovigilance by combining multiple strength-of-evidence aspects in vigiRank. Drug Saf 2014;37(8):617–28). In a similar spirit, logistic regression has been proposed to combine a disproportionality measure with a measure of unexpectedness for the time-to-onset distribution (Use of logistic regression to combine two causality criteria for signal detection in vaccine spontaneous report data, Drug Saf. 2014;37(12):1047-57). In A prediction model-based algorithm for computer-assisted database screening of adverse drug reactions in the Netherlands (Pharmacoepidemiol Drug Saf. 2018;27(2):199-205), five relevant characteristics (number of reports, disproportionality, Naranjo score, proportion of MAH reports, proportion of HCP reports) were chosen as potential predictors in the model and tested against the presence in the SPC of each unique drug-ADR association at the time of the analysis. All candidate predictors were included into the final model with an increased screening efficiency. The authors comment that the choice of candidate predictors may depend on each spontaneous report databases but that the method of generating a prediction model-based priority list of signals could be useful in other databases.

Disproportionality methods are usually calculated on the cumulative data and therefore do not provide a direct insight into temporal changes in frequency of reports. Methodologies to monitor changes in the frequency of reporting over time have been developed with the focus to enhance pharmacovigilance when databases are small, when drugs have established safety profiles and/or when product quality defects, medication errors and cases of abuse or misuse are of concern. Automated method for detecting increases in frequency of spontaneous adverse event reports over time (J Biopharm Stat. 2013; 23(1):161-77) presents a regression method with both smooth trend and seasonal components, while An algorithm to detect unexpected increases in frequency of reports of adverse events in EudraVigilance (Pharmacoepidemiol Drug Saf 2018;27(1):38-45) presents the testing of a model based on a negative binomial time-series regression model on thirteen historical concerns. Additionally, a modification of the Information Component to screen for spatial-temporal disproportionality is described in Using VigiBase to Identify Substandard Medicines: Detection Capacity and Key Prerequisites (Drug Saf 2015; 38(4): 373–382). Despite the promising results of these methods, and even if theoretically they seem appealing, limited work has been performed to assess their effectiveness. Thus, these methods should be implemented with quality control measures to ensure acceptable performance.

As understanding increases regarding the mechanisms at a molecular level that are involved in adverse effects of drugs it would be expected that this information will inform efforts to predict and detect drug safety problems. Such modeling is currently at an early stage, as presented in Data-driven prediction of drug effects and interactions (Sci Transl Med. 2012 14;4(125):125ra31), but should be a major focus of drug safety research activity. An example of an application of this concept is illustrated in the paper Cheminformatics-aided pharmacovigilance: application to Stevens-Johnson Syndrome (J Am Med Inform Assoc 2016; 23(5): 968–78) where the authors apply a Quantitative Structure-Activity Relationship (QSAR) model to predict the drugs associated with Stevens Johnson syndrome in a pharmacovigilance database.

### 5.8.3. Performance comparison of signal detection methods

The role of data mining in pharmacovigilance (Expert Opin Drug Saf 2005;4(5):929-48) explains how signal detection algorithms work and addresses questions regarding their validation, comparative performance, limitations and potential for use and misuse in pharmacovigilance.

An empirical evaluation of several disproportionality methods in a number of different spontaneous reporting databases is given in Comparison of statistical detection methods within and across spontaneous reporting databases (Drug Saf 2015;38(6);577-87).

Performance of pharmacovigilance signal detection algorithms for the FDA adverse event reporting system (Clin Pharmacol Ther 2013;93(6):539-46) describes the performance of signal-detection algorithms for spontaneous reports in the US FDA adverse event reporting system against a benchmark constructed by the Observational Medical Outcomes Partnership OMOP. It concludes that logistic regression performs better than traditional disproportionality analysis. Other studies have addressed similar or related questions, for examples Large-scale regression-based pattern discovery: The example of screening the WHO global Drug Safety database (Stat. Anal. Data Min 2010;3(4):197–208), Are all quantitative postmarketing signal detection methods equal? Performance characteristics of logistic regression and Multi-item Gamma Poisson Shrinker (Pharmacoepidemiol Drug Saf. 2012; 21(6):622–630 and Data-driven prediction of drug effects and interactions (Sci. Transl. Med. 2012; 4(125):125ra31).

### 5.8.4. Stratification and sub-group analyses

Many statistical signal detection algorithms disregard the underlying diversity and give equal weight to reports on all patients when computing the expected number of reports for a drug-event pair. This may give them vulnerability to confounding and distortions due to effect modification, and could result in true signals being masked or false associations being flagged as potential signals. Stratification and/or subgroup analyses might address these issues, and whereas stratification is implemented in some standard software packages, routine use of subgroup analyses is less common. Performance of stratified and subgrouped disproportionality analyses in spontaneous databases (Drug Saf 2016; 39(4):355-64) performed a comparison across a range of spontaneous report databases and covariates and found subgroup analyses to improve first pass signal detection, whereas stratification did not; subgroup analyses by patient age and country of origin were found to bring the greatest value.

### 5.8.5. Masking

Masking is a statistical issue by which true signals of disproportionate reporting are hidden by the presence of other products in the database. While it is not currently perfectly understood, publications have described methods assessing the extent and impact of the masking effect of measures of disproportionality. They include A conceptual approach to the masking effect of measures of disproportionality (Pharmacoepidemiol Drug Saf. 2014;23(2):208-17), with an application described in Assessing the extent and impact of the masking effect of disproportionality analyses on two spontaneous reporting systems databases (Pharmacoepidemiol Drug Saf. 2014;23(2):195-207), Outlier removal to uncover patterns in adverse drug reaction surveillance - a simple unmasking strategy (Pharmacoepidemiol Drug Saf 2013;22(10):1119-29) and A potential event-competition bias in safety signal detection: results from a spontaneous reporting research database in France (Drug Saf 2013;36(7):565-72). The value of these methods in practice needs to be further investigated.

### 5.8.6. Complementary role of databases

A time-consuming step in signal detection of adverse reactions is the determination of whether an effect is already recorded in the product information. A database which can be searched for this information allows filtering or flagging reaction monitoring reports for signals related to unlisted reactions, thus improving considerably the efficiency of the signal detection process by restricting attention to those drugs and adverse event not already considered causally related. In research, it permits an evaluation of the effect of background restriction on the performance of statistical signal detection. An example of such database is the PROTECT Database of adverse drug reactions (EU SPC ADR database), a structured Excel database of all adverse drug reactions (ADRs) listed in Chapter 4.8 of the Summary of Product Characteristics (SPC) of medicinal products authorised in the European Union (EU) according to the centralised procedure, based exclusively on the Medical Dictionary for Regulatory Activities (MedDRA) terminology.

Other large observational databases such as claims and electronic medical records databases are potentially useful as part of a larger signal detection and refinement strategy. Modern methods of pharmacovigilance: detecting adverse effects of drugs (Clin Med 2009;9(5):486-9) describes the strengths and weaknesses of different data sources for signal detection (spontaneous reports, electronic patient records and cohort-event monitoring). A number of studies have considered the use of observational data in electronic systems that complement existing methods of safety surveillance e.g. the PROTECT, OHDSI and Sentinel projects. Toward multimodal signal detection of adverse drug reactions (J Biomed Inform. 2017;76:41-9) concludes that utilising and jointly analysing multiple data sources may lead to improved signal detection but development of this approach requires a deeper understanding the data sources used, additional benchmarks and further research on methods to generate and synthesise signals.

## *5.9. Methods for pharmacovigilance impact research*

Assessment of the impact of pharmacovigilance actions at the population level is an area currently under-investigated but with increasing importance for regulators. Impact research identifies the net impact of a regulatory intervention by measuring both the intended outcomes and the unintended consequences of a regulatory intervention, such as stopping a useful medication or switching to alternatives. A detailed guidance on the methodological conduct of impact studies is provided in Annex 2 of this Guide, together with a comprehensive reference list.

Although it uses existing datasources and methods, the area of impact research has some distinctive characteristics that are worth discussing.

To measure the impact of pharmacovigilance activities, process or outcome indicators can be used depending on the type of intervention, target population, drug or disease characteristics. Determining and measuring the right outcomes can be challenging. It may be further complicated by unavailability of data and may therefore require use of surrogate outcomes. Data sources for the analysis include both primary data and secondary use of data, the latter being used more frequently as they reflect routine clinical practice (real world population). However, secondary use of data that is originally collected for other purposes as such presents limitations, especially in terms of missing relevant data.

If the date or time period of the intervention is known, a before/after time series is a design frequently used allowing to analyse changes of trends in incidence or prevalence of an outcome before and after the intervention occurred. Changes may be affected by simultaneously occurring interventions or events and the use of comparator groups that did not receive the intervention may facilitate the interpretation of any associations found.

The analytical methods will depend on the study design and type of data collection. Interrupted time series (ITS) regression is a strong analytical tool for before/after time series, especially if autocorrelation and adjusting for seasonality are taken into account, and the time point (or period) of the intervention is known. For adequate power, sufficient time points before and after the intervention are required. Joinpoint regression models calculating time points of trend line changes offer an alternative if the date of the intervention is unknown.

Specific analytical approaches are needed to measure unintended effects of pharmacovigilance activities which may not be expected at the design stage, for example switching to alternative medicines following product withdrawal or restriction, and determine the net attributable impact on patient outcomes.

Future challenges include the identification of long-term consequences of regulatory actions and the definition of thresholds for successful risk minimisation activities.

# 6. The statistical analysis plan

## *6.1. General considerations*

There is a considerable body of literature explaining statistical methods for observational studies but very little addressing the statistical analysis plan. A clear guide to general principles and the need for a plan is given in *Design of Observational Studies* (P.R. Rosenbaum, Springer Series in Statistics, 2010. Chapter 18), which also gives useful advice on how to test complex hypotheses in a way that minimises the chances of drawing incorrect conclusions.

Planning analyses for randomised clinical trials is covered in a number of publications. These often give checklists of the component parts of an analysis plan and much of this applies equally to non-randomised designs. A good reference in this respect is the International Conference on Harmonisation of Technical Requirements for Registration of Pharmaceuticals for Human Use (ICH): ICH E9 'Statistical Principles for Clinical Trials' and its addendum on estimands and sensitivity analysis in clinical trials (ICH E9(R1)).

While specific guidance on the statistical analysis plan for epidemiological studies is sparse, the following principles will apply to most of the studies.

- A study is generally designed with the objective of addressing a set of research questions. However, the initial product of a study is a set of numerical and categorical observations that do not usually provide a direct answer to the questions that the study is designed to address. The statistical analysis plan details the mathematical transformations that will be performed on the observed data in the study and the patterns of results that will be interpreted as supporting answers to the questions. An important part of the statistical analysis plan will explain how problems in the data will be handled in such calculations, for example missing or partial data.
- The statistical analysis plan should be sufficiently detailed so that it can be followed and reproduced by any competent analyst. Thus, it should provide clear and complete templates for each analysis.
- Pre-specification of statistical and epidemiological analyses can be challenging for data that are not collected specifically to answer the study questions. This is often the case in observational studies, where secondary data are used. However, thoughtful specification of the way missing values will be handled or the use of a small part of the data as a pilot set to guide analysis can be useful techniques to overcome such problems. A feature common to most studies is that some not pre-specified analyses will be performed in response to observations in the data to help interpretation

of results. It is important to distinguish between such data-driven analyses and the pre-specified findings. Post-hoc modifications to the analysis strategy should be noted and explained. The statistical analysis plan provides a confirmation of this process.

- Strong emphasis will be given in studies using observational data to measures taken to control and quantify levels of bias. Thus, part of the analysis plan will be devoted to converting scientific understanding of the causal relationships between the exposures and outcomes that are the primary focus of the study and other variables that are available in the dataset into a credible mathematical model. It is also advisable to include appropriate negative controls – (exposure, outcome) pairs that are strongly believed not to be causally related for which a similar model is considered reasonable – in the analysis as these may indicate uncontrolled confounding.

A particular concern in retrospective studies is that decisions about the analysis should be made blinded to any knowledge of the results. This should be a consideration in the study design, particularly when feasibility studies are to be performed to inform the design phase. Feasibility studies should be independent of the main study results.

## *6.2. Timing of the statistical analysis plan*

The study protocol will have specified the questions to be addressed in the study and will contain a generic description of the study type and the statistical techniques. However, the statistical analysis plan is likely to be the document in which the statistics to be calculated and tabular and graphical presentations are fully described. Since the decision criteria for the study are specified in terms of the observed values of these detailed statistics, it is worth formulating the statistical analysis plan at an early stage and, in particular, before any informal inspection of aspects of the data or results that might influence opinions regarding the study hypotheses. Ideally the statistical analysis plan will be developed as soon as the protocol is finalised.

## *6.3. Decision criteria*

If decisions are to be made based on the results of the study, a section of the statistical analysis plan should explain the different outcomes that might be selected for each decision, which statistics influence the decision making process and which values of the statistics will be considered to support each outcome. Often the statistical analysis will employ standard routines incorporated in statistical packages that have outputs seen as implicit decision criteria – for instance p values or confidence intervals. However, different applications of the study may require lower or higher strength of evidence – for instance policy recommendations regarding drug licensing may require a lower chance of false positive decisions than the classical one when deciding whether further investigation is needed for a product safety issue. Hence consideration of decision-making criteria with explicit reference to the type of decision to be made is beneficial.

## *6.4. Statistical analysis plan structure*

The statistical and epidemiological analysis plan is usually structured to reflect the protocol and will address, where relevant, the following points:

1. A description of the study data sources, linkage methods, and study design including intended study population, inclusion and exclusion criteria and study period with discussion of strengths and weaknesses.
2. Formal definitions of exposure including transformations to determine duration and quantity of exposure.

3. Definition of follow-up and censoring if applicable.
4. Formal definitions of any outcomes, for example 'fatal myocardial infarction' that might be defined as 'death within 30 days of a myocardial infarction'. Outcome variables based on historical data may involve complex transformations to approximate clinical variables not explicitly measured in the dataset used. These transformations should be discriminated from those made to improve the fit of a statistical model. In either case the rationale should be given. In the latter case this will include which tests of fit will be used and under what conditions a transformation will be used.
5. Formal definitions for other variables – e.g. thresholds for abnormal levels of blood parameters. When values of variables for a subject vary with time, care should be given to explaining how the values will be determined at each time point and recorded in the dataset for use in a statistical model.
6. The effect measures and statistical methods used to address each primary and secondary objective.
7. Blinding evaluators to exposure variables in order to avoid making subjective judgments about the study.
8. Methods of dealing with confounding, and assessing bias such as:
    8.1. Which confounders will be considered and how they will be defined
    8.2. Adjustment for confounders in statistical models
    8.3. Restriction in analysis
    8.4. Matching, including propensity-score matching
    8.5. Self-controlled study designs
    8.6. Statistical approach for any selection of a subset of confounders
    8.7. Methods for assessing the level of confounding adjustment achieved
    8.8. Sensitivity analysis for residual confounding
    8.9. How negative controls will be selected for the model
9. Handling of missing data, including:
    9.1. How missing data will be reported;
    9.2. Methods of imputation;
    9.3. Sensitivity analyses for handling missing data;
    9.4. How censored data will be treated and rationale
10. Fit of the model – if considered for a predictive model, including:
    10.1. Criteria for assessing fit;
    10.2. Alternative models in the event of clear lack of it.
11. Interim analyses – if considered:
    11.1. Criteria, circumstances and possible drawbacks for performing an interim analysis and possible actions (including stopping rules) that can be taken on the basis of such an analysis
12. How the achieved patient population will be characterised:

12.1. Description of target population;

12.2. Description of the analysis population if different, e.g. after propensity score matching or in instrumental variable analyses.

13. Treatment of multiplicity issues not elsewhere covered.

14. Sample size considerations should be presented, making explicit the data source from which the expected variation of relevant quantities and the clinically relevant differences are derived. It should be noted that in observational studies on data that already exist and where no additional data can be collected, sample size is not preclusive and the ethical injunction against 'underpowered' studies has no obvious force provided the results, in particular the 'absence of effect' and 'insufficient evidence', are properly presented and interpreted.

### *6.5. Handling of missing data*

Missing data occur when no data value is stored for the variable in the current observation. Missing data are a common occurrence and can have a significant effect on the conclusions that can be drawn from the data. There are different patterns of missing data: completely at random, at random or not at random.

The book *Statistical analysis with missing data* (Little RJA, Rubin DB. 2nd ed.,Wiley 2002) describes many aspects of the handling of missing data. The section 'Handling of missing values' in Rothman's *Modern Epidemiology*, 3rd ed. (K. Rothman, S. Greenland, T. Lash. Lippincott Williams & Wilkins, 2008) is a summary of the state of the art, focused on practical issues for epidemiologists. Ways of dealing with such data include complete subject analysis (subjects with missing values are deleted from the analyses) and imputation methods (missing data are predicted based on the observed values and the pattern of missingness). A method commonly used in epidemiology is to create a category of the variable, or an indicator, for the missing values. This practice can be invalid even if the data are missing completely at random and should be avoided (see Indicator and Stratification Methods for Missing Explanatory Variables in Multiple Linear Regression. J Am Stat Assoc 1996;91(433):222-30).

A concise review of methods to handle missing data is provided in the section 'Missing data' of the *Encyclopedia of Epidemiologic Methods* (Gail MH, Benichou J, Editors. Wiley 2000). Identifying the pattern of missing data is important as some methods for handling missing data assume a defined pattern of missingness. Biased results may be obtained if it is incorrectly assumed that data are missing at random. In general, it is desirable to show that conclusions drawn from the data are not sensitive to the particular strategy used to handle missing values. To investigate this, it may be helpful to repeat the analysis with a variety of approaches.

Other useful references on handling of missing data include the books *Multiple Imputation for Nonresponse in Surveys* (Rubin DB, Wiley, 2004) and *Analysis of Incomplete Multivariate Data* (Schafer JL, Chapman & Hall/CRC, 1997), and the articles Using the outcome for imputation of missing predictor values was preferred (J Clin Epi 2006;59(10):1092-101), Recovery of information from multiple imputation: a simulation study (Emerg Themes Epidemiol 2012;9(1):3) and Evaluation of two-fold fully conditional specification multiple imputation for longitudinal electronic health record data (Stat Med. 2014;33(21):3725-37).

# 7. Quality management

Quality in research ultimately impacts on regulatory practice, medicines development and public health. Quality is a measure of excellence and quality management includes all the activities that

organisations use to direct, control and coordinate quality (International Standards Organization, ISO 9000). Quality management principles as described in ISO Quality management principles are also applicable to pharmacoepidemiological research. The book *Total Quality Management-Key Concepts and Case Studies* (D.R. Kiran, BSP Books, Elsevier, 2016) deals with the management principles and practices that govern the quality function and presents all the aspects of quality control and management both in practice.

Quality management consists in four main activities: quality planning, quality assurance, quality control and quality improvement. Quality planning is defined as a set of activities whose purpose is to define quality system policies, objectives, and requirements, and to explain how these will be applied and achieved. Quality assurance defines the standards to be followed in order to meet the quality requirements for a product or service, whereas quality control ensures that these defined standards are followed at every step. Quality improvement refers to enhancing an organisation's ability to meet quality requirements.

Quality control should be designed as a study and involve identifying the study's objective, determining the relevant data to collect, choosing appropriate instruments to collect the data, analysing the data, recommending appropriate actions, implementing them, and evaluating the implementation to be used effectively in order to act strategically.

Rules, procedures, roles and responsibilities of quality assurance and quality control for clinical trials and biomedical research are well defined and described in many documents, such as Chapter 11 of the book *Principles of Good Clinical Practice* (M.J. McGraw, A.N. George, S.P. Shearn, eds., Pharmaceutical Press, London, 2010), the ICH E6 (R2) Good clinical practice, the European Forum for Good Clinical Practice (EFCGP) Guidelines, the Imperial College Academic Health Science Centre (AHSC)'s Quality Control and Quality Assurance SOP, the article Quality by Design in Clinical Trials: A Collaborative Pilot With FDA (Therapeutic Innovation & Regulatory Science 2013; 47;161-6), or the article Guidelines for Quality Assurance in Multicenter Trials: A Position Paper (Control Clin Trials 1998;19(5);477- 493).

For post-authorisation safety studies, resources include: Commission Implementing Regulation (EU) No 520/2012, GVP Module I, FDA's Best Practices for Conducting and Reporting Pharmacoepide Safety Studies Using Electronic Health Care Data Sets, the ISPE GPP and the Guidelines and recommendations for ensuring Good Epidemiological Practice (GEP): a guideline developed by the German Society for Epidemiology (European Journal of Epidemiology 2019;34(3):301-17).

The article Quality Assurance and Quality Control in Longitudinal Studies (Epidemiol Rev 1998, 20(1); 71-80) provides a comprehensive overview of components of QA and QC in multi-centre cohort studies with primary data collection. Such studies typically involve collection of an extensive amount of data for processing over an extended period of time and at several centres, with quality depending on a variety of factors relating to study personnel and equipment. Consequently, the QC process in such studies should be considered an integral part of the design of the study and a condition for the validity of its results. Quality assurance in non-interventional studies (Ger Med Sci 2009;7:Doc 29: 1-14) proposes measures of quality assurance that can be applied at different stages of non-interventional studies without compromising the character of non-intervention. Chapter 11 'Data Collection and Quality Assurance' of the AHRQ Registries for Evaluating Patient Outcomes: A User's Guide, 3rd Edition, reviews key areas of data collection, cleaning, storing, and quality assurance for registries, with practical examples.

For clinical trials which become more complex and more globally driven and produce a huge amount of data that has grown exponentially, the proposed system for managing quality implies a risk-based approach, *i.e.*, Risk-Based Quality Management (RBQM). The RBQM is incorporated as Good Clinical Practice expectation in the ICH E8 (R1).

The following articles are practical examples of quality aspects implementation in pharmacovigilance and pharmacoepidemiological as well as other biomedical studies:

- Training, quality assurance, and assessment of medical record abstraction in a multisite study (Am J Epidemiol 2003;157:546-51) describes a practical approach to assurance of good quality control in a large multi-site study.
- Interviewer variability – quality aspects in a case–control study (Eur J Epidemiol 2006;21(4);267-77) describes the procedures used to reduce interviewer variability, including procedures of quality assurance (i.e. education and training of interviewers and data validity checks) and quality control (i.e. a classification test, annual test interviews, expert case validation and database validation).
- Establishment of the nationwide Norwegian Prescription Database (NorPD) – new opportunities for research in pharmacoepidemiology in Norway (Norsk epidemiologi 2008;18(2):129-36) describes the quality checks applied to the database.
- Validation and validity of diagnoses in the General Practice Research Database (GPRD): a systematic review (Br J Clin Pharmacol 2010;69:4-14) assesses the quality of the methods used to validate diagnoses in the GPRD. The article contains methodological and reporting recommendations to further strengthen the use of the GPRD in research that are potentially applicable to other databases.
- EuroDURG Quality Indicator Meeting (DURQUIM) presents a report of a meeting which recommended indicators of prescribing quality in drug utilisation research [report published in full in Indicators of prescribing quality in drug utilisation research: report of a European meeting (DURQUIM, 13-15 May 2004) (Eur J Clin Pharmacol 2005;60(11):831-4)].
- Data quality management in pharmacovigilance (Drug Saf 2004;27(12):857-70) focusses on the intial three steps of data processing cycle (collection and data entry; storage and maintenance; selection, retrieval and manipulation), the different quality dimensions associated with these steps together with examples relevant to pharmacovigilance data.
- Quality assessment of structure and language elements of written responses given by seven Scandinavian drug information centres (Eur J Clin Pharmacol 2017;73(5):623-631) deals with the identification of structure and language elements affecting the quality of responses from Scandinavian drug information centres that have been evaluated by internal and external, medical and language experts.
- Total quality management in the health-care context: integrating the literature and directing future research (Risk Manag Healthc Policy 2019;12:167–77) focusses through systematic literature review on the synergistic integration of predictors and elements that determine the success of total quality management (TQM) implementations in hospitals.

# 8. Dissemination and communication of study results

Aspects of dissemination and communication of study results include, but are not limited to, reports to health authorities and study sponsors, presentations in scientific fora, scientific publications, patient focused communications and websites.

The Declaration of Helsinki provides overarching guidance on the registration, publication and dissemination of research results. Every research study involving human subjects must be registered in a publicly accessible database before recruitment of the first subject. A means to achieve this with pharmacoepidemiology and pharmacovigilance studies is through registration of protocols and reports of studies in the European Union electronic Register of Post-Authorisation Studies (EU PAS Register),

ideally before they start, and protocols and study results should be made public. This is compulsory only for study imposed by regulators.

Authorship should conform to the guidelines established by the International Committee of Medical Journal Editors (ICJME) 'Recommendations for the Conduct, Reporting, Editing, and Publication of Scholarly work in Medical Journals'.

Important specific points relating to reporting of study results that are common to the various guidelines cited below are that:

- Sources of research funding should always be disclosed whether in oral or written presentation of results.
- A dissemination and communication strategy should be pre-defined as part of the funding contract for a given study.
- All results with a scientific or public health impact must be reported to relevant authorities and made publicly available without undue delay.
- Quantitative measures of association should be reported rather than just results of statistical testing.

The ISPE GPP contain a section on communication (section V) which includes a statement that there is an ethical obligation to disseminate findings of potential scientific or public health importance and that research sponsors (government agencies, private sector, etc.) shall be informed of study results in a manner that complies with local regulatory requirements. The European Medicines Agency (EMA) Guidance for the format and content of the protocol of non-interventional post-authorisation safety studies states that plans for disseminating and communicating study results are to be described as part of study planning activities.

The EMA Guidance for the format and content of the final study report of non-interventional post-authorisation safety studies (PASS) provides a template for final study reports that may be applied to any non-interventional PASS, including meta-analyses and systematic reviews. The FDA's Best Practices for Conducting and Reporting Pharmacoepidemiologic Safety Studies Using Electronic Health Care Data Sets includes a description of all the elements that should be addressed and included in the final study report of such studies.

The Enhancing the Quality and Transparency of Health Research (EQUATOR) network is an international initiative that aims to enhance the reliability and value of the published health research literature. A catalogue of reporting guidelines for health research (Eur J Clin Invest 2010;40(1):35-53) presents a collection of tools and guidelines available on the EQUATOR website relating to resources, education and training to facilitate good research reporting and the development, dissemination and implementation of robust reporting guidelines to increase the accuracy and transparency of health research reporting.

The Strengthening the Reporting of Observational studies in Epidemiology (STROBE) Statement Guidelines for reporting observational studies has established recommendations for improving the quality of reporting of observational studies and seeks to ensure a clear presentation of what was planned, done, and found. Of note, the aim of these guidelines was not to require the reporting of observational research in a rigid format, but to address what should be the essential information contained in a publication on an observational study.

The REporting of studies Conducted using Observational Routinely-collected health Data (RECORD) Statement (PLoS Med. 2015;12(10):e1001885) was created as an extension to the STROBE statement to address reporting items specific to observational studies using routinely collected health data.

RECORD makes additional recommendations on the reporting of methods of selection of study populations, exposures, outcomes and covariates (including codes or algorithms used), whether validation has been conducted, the level of access to databases used, and data linkages that were required to conduct the study. The RECORD-PE statement (BMJ 2018;363:k3532) aims to extend existing STROBE and RECORD guidelines providing guidance for the reporting of pharmacoepidemiological studies using routinely collected data.

The joint ISPE-ISPOR Special Task Force on Real World Evidence in Health Care Decision Making developed a guidance on Reporting to Improve Reproducibility and Facilitate Validity Assessment for Healthcare Database Studies (Pharmacoepidemiol Drug Saf. 2017;26(9):1018-32), with the objective "to catalogue scientific decisions underpinning study execution that should be reported to facilitate replication and enable assessment of validity of studies conducted in large healthcare databases." A key recommendation is that "A substantial improvement in reproducibility, rigor and confidence in real world evidence generated from healthcare databases could be achieved with greater transparency about operational study parameters used to create analytic datasets from longitudinal healthcare databases".

The Good ReseArch for Comparative Effectiveness (GRACE) guidance includes recommendations on reporting comparative effectiveness studies. The STARD guidelines (BMJ Open 2016;14;6(11):e012799) focus on reporting diagnostic accuracy studies.

The Meta-analysis of Observational Studies in Epidemiology (MOOSE) group (JAMA 2000;283(15):2008-15) has developed a consensus statement and recommendations for reporting meta-analyses of observational studies. It is equivalent to the STROBE Statement and the Consolidated Standards of Reporting Trials Consolidated Standards for Reporting Trials (CONSORT) 2010 Statement for RCTs, in focusing primarily on communication and list the minimum requirements for adequate reporting. The authors recommend a broad inclusion of studies and conduct of post-hoc sensitivity testing on the dependence of the results on factors such as quality of underlying papers, design, accounting for confounders, etc. The authors comment on the particular problems in merging observational studies with highly variable sets of confounders that were or were not controlled for, but they do not suggest any solution or give any references to possible ways to address it. As pragmatic trials increase in our field, another CONSORT extension focused on this type of studies, Improving the reporting of pragmatic trials: an extension of the CONSORT Statement (BMJ 2008;337:a2390) might be also relevant.

The Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement (BMJ 2009;339:b2535) is an evidence-based minimum set of items for reporting in systematic reviews and meta-analyses. While focused on randomised trials, PRISMA can also be used as a basis for reporting systematic reviews of other types of research, particularly evaluations of interventions. PRISMA may also be useful for critical appraisal of published systematic reviews, although it is not designed as a quality assessment instrument.

Module VI of the Guideline on good pharmacovigilance practices (GVP) addresses the legal requirements which are applicable regards submission of individual reports of suspected adverse reactions associated with medicinal products authorised in the European Union. The Guidelines for Submitting Adverse Event Reports for Publication (Pharmacoepidemiol Drug Saf 2007;16(5): 581–7) also list key elements that have to be included when publishing a report of one or more adverse events. These guidelines have been endorsed by the International Society for Pharmacoepidemiology (ISPE) and the International Society of Pharmacovigilance (ISoP) and are available on their web sites.

Several of the guidelines listed above include recommendations on the minimum information to be provided on data sources used in pharmacoepidemiology. The name, type, content and validity of the

database and the original reasons why the data were collected should be reported with the extent of access, data cleaning methods and details of linkage (see The REporting of studies Conducted using Observational Routinely-collected health Data (RECORD) Statement. PLoS Med. 2015;12(10):e1001885). To allow reproducibility reporting of extraction date, data version, data sampling strategy and the years of source data used for the study are required, as described in Reporting to Improve Reproducibility and Facilitate Validity Assessment for Healthcare Database Studies, V1.0. (Pharmacoepidemiol Drug Saf. 2017;26(9):1018-1032).

Additional guidance on reporting of study results is provided in the ENCePP Checklist for Study Protocols and Code of Conduct and the IEA GEP guideline that are reviewed elsewhere in this Guide.

# 9. Data protection and ethical aspects

## *9.1. Patient and data protection*

In Europe, the EU Charter of Fundamental Rights stipulates that EU citizens have the right to protection of their personal data. Additionally, EU and national legislation addresses patient data access, data linkage and consent issues, including duty of confidentiality. Therefore, while individual data custodians may have differing requirements related to what approvals are needed before their data can be released for a particular study, all studies conducted in Europe must meet all applicable legislation.

The General Data Protection Regulation (GDPR) Regulation (EU) 2016/679 on the protection of individuals with regard to the processing of personal data and on free movement of such data came into force on 24 May 2016. It repeals Directive 95/46/EC and is aimed at making data protection fit for the digital age.

Regulation (EC) No 45/2001 sets forth the rules applicable to the processing of personal data by EU institutions and bodies. On 10 January 2017, a proposal was put forward to amend those rules to bring them in line with the GDPR.

For interventional research, Directive 2001/20/EC and the Guidelines for Good Clinical Practice (Commission Directive 2005/28/EC) apply. Directive 2001/20 EC will be repealed when the Clinical Trials Regulation (Regulation (EU) No 536/2014) comes into application. It will also apply to trials authorised under the previous legislation if they are still ongoing three years after the Regulation has come into operation. In addition, marketing authorisation holders (MAHs) and investigators must follow relevant national guidance of those Member States where the study is being conducted.

Article 36 of the Commission Implementing Regulation (EU) No. 520/2012 specifies that for post-authorisation safety studies (PASS) imposed as an obligation, MAHs shall ensure that all study information is handled and stored in a way that ensure the confidentiality of the study records. The GVP Module VIII - Post-authorisation safety studies recommends that these provisions should also be applied to PASS that are voluntarily initiated, managed or financed by a MAH.

The ISPE Good pharmacoepidemiology practice provides recommendations on the protection of human subjects and refers to the ISPE guidelines on Data Privacy, Medical Record Confidentiality, and Research in the Interest of Public Health. It also recommends that the plans for protecting human subjects should be described in a stand-alone section of the study protocol.

## 9.2. Scientific integrity and ethical conduct

Principles of scientific integrity and ethical conduct are paramount in any medical research. The Declaration of Helsinki provides ethical principles addressed primarily to physicians involved in medical research involving human subjects, including research on identifiable human material and data and is the main document on human research ethics. The ENCePP Code of Conduct offers standards for scientific independence and transparency of research in pharmacoepidemiology and pharmacovigilance and promotes best practice for the interactions between investigators and study funders in critical areas such as planning, conduct and reporting of studies. As a core transparency measure, it recommends that the protocols of all pharmacoepidemiology and pharmacovigilance studies should be registered in the European Union electronic Register of Post-Authorisation Studies (EU PAS Register), ideally before they start. The Code also recommends that study findings should be published irrespective of positive or negative results.

Guided by three core values (best science, strengthening public health and transparency), the ADVANCE Code of Conduct for Collaborative Vaccine Studies (Vaccine 2017;35(15):1844-55) includes recommendations about 10 topics: Scientific integrity, Scientific independence, Transparency, Conflicts of interest, Study protocol, Study report, Publication, Subject privacy, Sharing of study data, Research contract. Each topic includes a definition, a set of recommendations and a list of additional reading. The concept of the study team is introduced as a key component of the ADVANCE Code of Conduct with a core set of roles and responsibilities. It also provides direct access to a comprehensive list of relevant guidelines.

The Good Pharmacoepidemiology Practices (GPP) of the International Society for Pharmacoepidemiology (ISPE) proposes practices and procedures that should be considered to help ensure the quality and integrity of pharmacoepidemiological research, including detailed guidance for protocol development, roles and responsibilities, study conduct, communication, reporting of adverse events and archiving. The Good Epidemiology Practice (GEP) of the International Epidemiological Association addresses four general ethical principles for research (Autonomy, Beneficence, Non-maleficence and Justice) and proposes rules for good research behaviour in relation to working with personal data, data documentation, publication, the exercise of judgment and scientific misconduct.

The CIOMS International Ethical Guidelines for Health-related Research Involving Humans (Geneva: 2016) provides detailed commentary on how universal ethical principles should be applied, with particular attention to conducting research in low-resource settings. It includes 25 guidelines addressing different topics, settings and population groups concerned by health-related research.

The Recommendations for the Conduct, Reporting, Editing, and Publication of Scholarly work in Medical Journals by the International Committee of Medical Journal Editors (ICJME) include clear statements on ethical principles related to publication in biomedical journals. Authorship and contributorship, editorship, peer review, conflicts of interest, privacy and confidentiality and protection of human subjects and animals in research are addressed.

The Agency for Healthcare Research and Quality (AHRQ) published Registries to Evaluate Patient Outcomes: a User's guide, Third Edition, 2014, which is a reference for establishing, maintaining and evaluating the success of registries created to collect data about patient outcomes. Section II: 'Legal and Ethical Considerations for Registries' is a specific chapter dedicated to ethics, data ownership, and privacy. The concepts within are focused on US law.

# 10. Specific topics

## *10.1. Comparative effectiveness research*

**Note: Chapter 10.1 has not been updated for revision 8**

### 10.1.1. Introduction

Comparative effectiveness research (CER) is designed to inform health-care decisions at the level of both policy and the individual by comparing the benefits and harms of therapeutic strategies available in routine practice, for the prevention, the diagnosis or the treatment of a given health condition. The interventions under comparison may be related to similar treatments, such as competing drugs, or different approaches, such as surgical procedures and drug therapy. The comparison may focus only on the relative medical benefits and risks of the different options or it may weigh both their costs and their benefits. The methods of comparative effectiveness research (Annu Rev Public Health 2012;33:425-45) defines the key elements of CER as (a) head-to-head comparison of active treatments, (b) study populations typical of day-to-day clinical practice, and (c) a focus on evidence to inform health care tailored to the characteristics of individual patients. In What is Comparative Effectiveness Research, the AHRQ highlights that CER requires the development, expansion and use of a variety of data sources and methods to conduct timely and relevant research and disseminate the results in a form that is quickly usable. The evidence may come from a review and synthesis of available evidence from existing clinical trials or observational studies or from the conduct of studies that generate new evidence. In Developing a Protocol for Observational Comparative Effectiveness Research: A User's Guide, AHRQ also highlights that CER is still a relatively new field of enquiry that has its origin across multiple disciplines and is likely to evolve and be refined over time.

Among resources for keeping up with the evolution in this field, the US National Library of Medicine provides a web site for queries on CER.

The terminology 'Relative effectiveness assessment (REA)' is also used when comparing multiple technologies or a new technology against standard of care, while 'rapid' REA refers to performing an assessment within a limited timeframe in the case of a new marketing authorisation or a new indication granted for an approved medicine (What is a rapid review? A methodological exploration of rapid reviews in Health Technology Assessments. Int J Evid Based Healthc. 2012;10(4):397-410).

### 10.1.2. General aspects

Several initiatives have promoted the conduct of CER and REA and proposed general methodological guidance to help in the design and analysis of such studies.

The Methodological Guidelines for Rapid Relative Effectiveness Assessment of Pharmaceuticals developed by EUnetHTA cover a broad spectrum of issues on REA. They address methodological challenges that are encountered by health technology assessors while performing rapid REA and provide and discuss practical recommendations on definitions to be used and how to extract, assess and present relevant information in assessment reports. Specific topics covered include the choice of comparators, strengths and limitations of various data sources and methods, internal and external validity of studies, the selection and assessment of endpoints (including composite and surrogate endpoints and Health Related Quality of Life [HRQoL]) and the evaluation of relative safety.

AHRQ's Developing a Protocol for Observational Comparative Effectiveness Research: A User's Guide identifies minimal standards and best practices for observational CER. It provides principles on a wide range of topics for designing research and developing protocols, with relevant questions to be

addressed and checklists of key elements to be considered. The GRACE Principles provide guidance on the evaluation of the quality of observational CER studies to help decision-makers in recognizing high-quality studies and researchers in design and conduct high quality studies. A checklist to evaluate the quality of observational CER studies is also provided. the International Society for Pharmacoeconomics and Outcomes Research (ISPOR) addressed several key issues of CER in three publications: Part I includes the selection of study design and data sources and the reporting and interpretation of results in the light of policy questions; Part II relates to the validity and generalisability of study results, with an overview of potential threats to validity; Part III includes approaches to reducing such threats and, in particular, to controlling of confounding. The Patient-Centered Outcomes Research Institute (PCORI) Methodology Standards document provides standards for patient-centred outcome research that aims to improve the way research questions are selected, formulated and addressed, and findings reported. The PCORI group has recently published how stakeholders may be involved in PCORI research, Stakeholder-Driven Comparative Effectiveness Research (JAMA 2015; 314: 2235-2236). In a Journal of Clinical Epidemiology series of articles, the Grading of Recommendations Assessment, Development, and Evaluation (GRADE) working group offers a structured process for rating quality of evidence and grading strength of recommendations in systematic reviews, health technology assessment and clinical practice guidelines. The GRADE group recommends individuals new to GRADE to first read the 6-part 2008 BMJ series.

A guideline on methods for performing systematic reviews of existing comparative effectiveness research has been published by the AHRQ (Methods Guide for Effectiveness and Comparative Effectiveness Reviews).

The RWE Navigator website has been developed by the IMI GetReal consortium to provide recommendations on the use of real-world evidence for decision-making on effectiveness and relative effectiveness of medicinal products. It discusses important topics such as the sources of real-world data, study designs, approaches to summarising and synthesising the evidence, modelling of effectiveness and methods to adjust for bias and governance aspects. It also presents a glossary of terms and case studies relevant for RWD research, with a focus on effectiveness research.

## 10.1.3. Prominent issues in CER

### *10.1.3.1. Randomised clinical trials vs. observational studies*

While RCTs are considered to provide the most robust evidence of the efficacy of therapeutic options, they are affected by well-recognised qualitative and quantitative limitations that may not reflect how the drug of interest will perform in real-life. Moreover, relatively few RCTs are traditionally designed using an alternative therapeutic strategy as a comparator, which limits the utility of the resulting data in establishing recommendations for treatment choices. For these reasons, other research methodologies such as pragmatic trials and observational studies may complement traditional explanatory RCTs in CER.

Explanatory and Pragmatic Attitudes in Therapeutic Trials (J Chron Dis 1967; republished in J Clin Epidemiol 2009;62(5):499-505) distinguishes between two approaches in designing clinical trials: the 'explanatory' approach, which seeks to understand differences between the effects of treatments administered in experimental conditions, and the 'pragmatic' approach which seeks to answer the practical question of choosing the best treatment administered in normal conditions of use. The two approaches affect the definition of the treatments, the assessment of results, the choice of subjects and the way in which the treatments are compared. A pragmatic-explanatory continuum indicator summary (PRECIS): a tool to help trial designers (CMAJ 2009; 180 (10):E47-57) quantifies distinguishing characteristics between pragmatic and explanatory trials and has been updated in The

Precis-2 tool: designing trials that are fit for purpose (BMJ 2015; 350: h2147). A checklist of eight items for the reporting of pragmatic trials was also developed as an extension of the CONSORT statement to facilitate the use of results from such trials in decisions about health-care (Improving the reporting of pragmatic trials: an extension of the CONSORT statement. BMJ 2008;337 (a2390):1-8).

The article Why we need observational studies to evaluate effectiveness of health care (BMJ 1996;312(7040):1215-18) documents situations in the field of health care intervention assessment where observational studies are needed because randomised trials are either unnecessary, inappropriate, impossible or inadequate. In a review of five interventions, Randomized, controlled trials, observational studies, and the hierarchy of research designs (N Engl J Med 2000;342(25):1887-92) found that the results of well-designed observational studies (with either a cohort or case-control design) did not systematically overestimate the magnitude of treatment effects. In defense of Pharmacoepidemiology-Embracing the Yin and Yang of Drug Research (N Engl J Med 2007;357(22):2219-21) shows that strengths and weaknesses of RCTs and observational studies make both designs necessary in the study of drug effects. However, When are observational studies as credible as randomised trials? (Lancet 2004;363(9422):1728-31) explains that observational studies are suitable for the study of adverse (non-predictable) effects of drugs but should not be used for intended effects of drugs because of the potential for selection bias.

With regard to the selection and assessment of endpoints for CER, the COMET (Core Outcome Measures in Effectiveness Trials) Initiative aims at developing agreed minimum standardized sets of outcomes ('core outcome sets', COS) to be assessed and reported in effectiveness trials of a specific condition as discussed in Choosing Important Health Outcomes for Comparative Effectiveness Research: An Updated Review and User Survey (PLoS One 2016 ;11(1):e0146444.).

#### *10.1.3.2. Use of electronic healthcare databases*

A review of uses of health care utilization databases for epidemiologic research on therapeutics (J Clin Epidemiol 2005;58(4):323-37) considers the application of health care utilisation databases to epidemiology and health services research, with particular reference to the study of medications. Information on relevant covariates and in particular on confounding factors may not be available or adequately measured in electronic healthcare databases. To overcome this limit, CER studies have integrated information from health databases with information collected ad hoc from study subjects. Enhancing electronic health record measurement of depression severity and suicide ideation: a Distributed Ambulatory Research in Therapeutics Network (DARTNet) study (J Am Board Fam Med. 2012;25(5):582-93) shows the value of adding direct measurements and pharmacy claims data to data from electronic healthcare records participating in Assessing medication exposures and outcomes in the frail elderly: assessing research challenges in nursing home pharmacotherapy (Med Care 2010;48(6 Suppl):S23-31) describe how merging longitudinal electronic clinical and functional data from nursing home sources with Medicare and Medicaid claims data can support unique study designs in CER but pose many challenging design and analytic issues. Pragmatic randomised trials using routine electronic health records: putting them to the test (BMJ 2012;344:e55) discusses opportunities for using electronic healthcare records for conducting pragmatic trials.

A model based on counterfactual theory for CER using large administrative healthcare databases has been suggested, in which causal inference from observational studies based on large administrative health databases is viewed as an emulation of a randomized trial. This 'target trial' is made explicit and design and analytic approaches are reviewed in Using Big Data to Emulate a Target Trial When a Randomized Trial Is Not Available (Am J Epidemiol (2016) 183 (8): 758-764).

#### *10.1.3.3. Bias and confounding in observational CER*

Methodological issues and principles of Chapter 5 of the ENCePP Guide are applicable to CER as well and the textbooks cited in that chapter are recommended for consultation.

Methods to assess intended effects of drug treatment in observational studies are reviewed (J Clin Epidemiol 2004;57(12):1223-31) provides an overview of methods that seek to adjust for confounding in observational studies when assessing intended drug effects. Developments in post-marketing comparative effectiveness research (Clin Pharmacol Ther 2007;82(2):143-56) also reviews the roles of propensity scores (PS), instrumental variables and sensitivity analyses to reduce measured and unmeasured confounding in CER. Use of propensity scores and disease risk scores in the context of observational health-care programme research is described in Summary Variables in Observational Research: Propensity Scores and Disease Risk Scores. More recently, high-dimensional propensity score has been suggested as a method to further improve control for confounding as these variables may collectively be proxies for unobserved factors.

Results presented in High-dimensional propensity score adjustment in studies of treatment effects using health care claims data (Epidemiology 2009;20(4):512-22) show that in a selected empirical evaluation, high-dimensional propensity score improved confounding control compared to conventional PS adjustment when benchmarked against results from randomized controlled trials. See Chapter 5.3.4 of the Guide for an in-depth discussion of propensity scores. Several methods can be considered to handle cofounders in non-experimental CER (Confounding adjustment in comparative effectiveness research conducted within distributed research networks (Med Care 2013 ; 51(8 Suppl 3) : S4-S10); Disease Risk Score (DRS) as a Confounder Summary Method: Systematic Review and Recommendations (Pharmacoepidemiol Drug Saf 2013; 22(2): 122–129). Strategies for selecting variables for adjustment in non-experimental CER have also been proposed (Pharmacoepidemiol Drug Saf 2013; 22(11): 1139–1145).

A reason for discrepancies between results of randomised trials and observational studies may be the use of prevalent drug users in the latter. Evaluating medication effects outside of clinical trials: new-user designs (Am J Epidemiol 2003;158(9):915-20) explains the biases introduced by use of prevalent drug users and how a new-user (or incident user) design eliminate these biases by restricting analyses to persons under observation at the start of the current course of treatment. The Incident User Design in Comparative Effectiveness Research (Pharmacoepidemiol Drug Saf 2013; 22(1): 1–6) reviews published CER case studies in which investigators had used the incident user design, discusses its strengths (reduced bias) and weakness (reduced precision of comparative effectiveness estimates) and provides recommendations to investigators considering to use this design. The value of incident user design and its exceptions have been reviewed.

## *10.2. Vaccine safety and effectiveness*

## 10.2.1. Vaccine safety

#### *10.2.1.1. General aspects*

A reference to be consulted for vaccine safety assessment is the ADVANCE Report on appraisal of vaccine safety methods. Together with a large number of references, it provides a brief description of a very wide range of direct and indirect methods of risk assessment for vaccines (listed in the Table of Contents) and evaluates them based on nine criteria related to five domains: Effect Measure, Statistical Criteria, Timeliness, Restriction and Robustness, and Operational Criteria. It also emphasises the specificities of safety assessment for vaccines and how they differ from other medicines, evaluates study designs, discusses perspectives of different stakeholders on risk assessment, describes

experiences from other projects and systems, and provides recommendations. This document is highly relevant for all the topics covered in this chapter on vaccine safety.

Specific aspects related to vaccine safety are discussed in several other documents.

- The Report of the CIOMS/WHO Working Group on Definition and Application of Terms for Vaccine Pharmacovigilance (2012) provides definitions and explanatory notes for the terms 'vaccine pharmacovigilance', 'vaccination failure' and 'adverse event following immunisation (AEFI)'.
- The CIOMS Guide to Active Vaccine Safety Surveillance (2017) describes the process of determining whether active vaccine safety surveillance is necessary, more specifically in the context of resource-limited countries, and, if so, of choosing the best type of active safety surveillance and considering key implementation issues.
- The CIOMS Guide to Vaccine Safety Communication (2018) provides an overview of strategic communication issues faced by regulators, those responsible for vaccination policies and other stakeholders in introducing current or new vaccines in populations. Building upon existing recommendations, it provides a guide for vaccine risk safety communication approaches.
- The Brighton Collaboration provides resources to facilitate and harmonise collection, analysis and presentation of vaccine safety data, including case definitions specifically developed for pharmacoepidemiological research, an electronic tool to help the classification of reported signs and symptoms, template protocols and guidelines.
- Module 4 (Surveillance) of the e-learning training course Vaccine Safety Basics of the World Health Organization (WHO) describes phamacovigilance principles, causality assessment procedures, surveillance systems and places safety in the context of the benefit/risk profile of the vaccine. For example the systematic review Maternal Influenza Vaccination and Risk for Congenital Malformations: A Systematic Review and Meta-analysis (Obstet Gynecol 2015;126(5):1075-84) on influenza vaccination in pregnancy and the risk of congenital anomalies in newborns did not find an association, adding to the evidence base in favor of influenza vaccination in pregnancy ().
- Recommendations on vaccine-specific aspects of the EU pharmacovigilance system, including on risk management, signal detection and post-authorisation safety studies (PASS) are presented in the Module P.I: Vaccines for prophylaxis against infectious diseases of the Good pharmacovigilance practices (GVP).

#### *10.2.1.2. Signal detection and validation*

Aside from a qualitative analysis of spontaneous case reports or case series, quantitative methods such as disproportionality analyses and observed vs. expected (O/E) analyses are routinely employed in signal detection for vaccines. Several documents discuss the merits and review the methods of these approaches.

*Disproportionality analyses*

GVP Module P.I: Vaccines for prophylaxis against infectious diseases describes issues to be considered when applying methods for disproportionality analyses for vaccines, including the choice of the comparator group and the use of stratification. Effects of stratification on data mining in the US Vaccine Adverse Event Reporting System (VAERS) (Drug Saf 2008;31(8):667-74) demonstrates that stratification can reveal and reduce confounding and unmask some vaccine-event pairs not found by crude analyses. However, Stratification for Spontaneous Report Databases (Drug Saf 2008;31(11):1049-52) highlights that extensive use of stratification in signal detection algorithms should be avoided as it can mask true signals. Vaccine-Based Subgroup Analysis in VigiBase: Effect on Sensitivity in Paediatric Signal Detection (Drug Saf 2012;35(4)335-46) further examines the effects of

subgroup analyses based on the relative distribution of vaccine/non-vaccine reports in paediatric ADR data.

The article Optimization of a quantitative signal detection algorithm for spontaneous reports of adverse events post immunization (Pharmacoepidemiol Drug Saf 2013; 22(5): 477–87) explores various ways of improving performance of signal detection algorithms when looking for vaccines.

The article Adverse events associated with pandemic influenza vaccines: comparison of the results of a follow-up study with those coming from spontaneous reporting (Vaccine 2011;29(3):519-22) reported a more complete pattern of reactions when using two complementary methods for first characterisation of the post-marketing safety profile of a new vaccine, which may impact on signal detection.

In Review of the initial post-marketing safety surveillance for the recombinant zoster vaccine (Vaccine 2020;38(18):3489-500), the time-to-onset distribution of zoster vaccine-adverse event pairs is used to generate a quantitative signal of unexpected temporal relationship; the method was applied when the distribution within 60 days post-vaccination is significantly different from the same adverse event reported with comparators or from the reported distribution of other vaccine- adverse event pairs.

*Observed-to-expected analyses*

When prompt decision-making about a safety concern is required and there is insufficient time to review individual cases, GVP Module P.I: Vaccines for prophylaxis against infectious diseases suggests the conduct of O/E analyses for signal validation and preliminary signal evaluation. The module discusses key requirements of O/E analyses: the observed number of cases detected in a passive or active surveillance systems, near real-time exposure data, appropriately stratified background incidence rates (to calculate the expected number of cases) and sensitivity analyses around these measures. O/E analyses for vaccines are further discussed in Pharmacoepidemiological considerations in observed-to-expected analyses for vaccines (Pharmacoepidemiol Drug Saf 2016;25(2): 215-22) and are also addressed in the review Near real-time vaccine safety surveillance using electronic health records - a systematic review of the application of statistical methods (Pharmacoepidemiol Drug Saf 2016;25(3):225-37).

O/E analyses require several assumptions and each one is associated with uncertainties. How to deal with these uncertainties is addressed in Pharmacoepidemiological considerations in observed-to-expected analyses for vaccines (Pharmacoepidemiol Drug Saf. 2016;25(2):215–222).

Simple 'snapshot' O/E analyses require near-real-time exposure data, appropriately stratified background incidence rates (to calculate the expected number of cases) and sensitivity analyses around these measures, and they may not be appropriate for continuous monitoring due to inflation of type 1 error rates when multiple tests are performed. Safety monitoring of Influenza A/H1N1 pandemic vaccines in EudraVigilance (Vaccine 2011;29(26):4378-87) illustrates that simple 'snapshot' O/E analyses are affected by uncertainties regarding the numbers of vaccinated individuals and age-specific background incidence rates.

The articles Human papilloma virus immunization in adolescents and young adults: a cohort study to illustrate what events might be mistaken for adverse reactions (Pediatr Infect Dis J 2007;26(11):979-84), Health problems most commonly diagnosed among young female patients during visits to general practitioners and gynecologists in France before the initiation of the human papillomavirus vaccination program (Pharmacoepidemiol Drug Saf 2012; 21(3):261-80) and Use of population based background rates of disease to assess vaccine safety in childhood and mass immunisation in Denmark: nationwide population based cohort study (BMJ 2012;345:e5823) illustrate the importance of collecting background rates by estimating risks of coincident associations of emergency consultations, hospitalisations and outpatients consultations with vaccination. Rates of selected disease events for

several countries also vary by age, sex, method of ascertainment and geography, as shown in Importance of background rates of disease in assessment of vaccine safety during mass immunisation with pandemic H1N1 influenza vaccines (Lancet 2009; 374(9707):2115-22). Moreover, Guillain-Barré syndrome and influenza vaccines: A meta-analysis (Vaccine 2015; 33(31):3773-8) suggests that a trend observed between different geographical areas would be consistent with a different susceptibility of developing a particular adverse reaction among different populations. In addition, comparisons with background rates may be invalid if conditions are unmasked at vaccination visits (see Human papillomavirus vaccination of adult women and risk of autoimmune and neurological diseases (J Intern Med. 2018;283(2):154-165)).

*Sequential methods*

Sequential methods, as described in Early detection of adverse drug events within population-based health networks: application of sequential methods (Pharmacoepidemiol Drug Saf 2007;16(12):1275-1284), allow O/E analyses to be performed on a routine (e.g. weekly) basis using cumulative data with adjustment for multiplicity. Such methods are routinely used for near-real time surveillance in the Vaccine Safety Datalink (VSD) (Near real-time surveillance for influenza vaccine safety: proof-of-concept in the Vaccine Safety Datalink Project. Am J Epidemiol 2010;171(2):177-88). Potential issues are described in Challenges in the design and analysis of sequentially monitored postmarket safety surveillance evaluations using electronic observational health care data (Pharmacoepidemiol Drug Saf 2012;21(S1):62-71). A review of signals detected over 3 years with these methods in Vaccine Safety Datalink concluded that care with data quality, outcome definitions, comparison groups and length of surveillance is required to enable detection of true safety problems while controlling error rates (Active surveillance for adverse events: the experience of the Vaccine Safety Datalink Project (Pediatrics 2011;127(S1):S54-S64)). Sequential methods are, therefore, more robust but also more complex to perform, understand and communicate to a non-statistical audience.

A new self-controlled case series method for analyzing spontaneous reports of adverse events after vaccination (Am J Epidemiol 2013;178(9):1496-504) extends the self-controlled case series approach to explore and quantify vaccine safety signals from spontaneous reports. It uses parametric and nonparametric versions with different assumptions to account for the specific features of the data (e.g., large amount of underreporting and variation of reporting with time since vaccination). The method should be seen as a signal strengthening approach for quickly exploring a signal based on spontaneous reports prior to a pharmacoepidemiologic study, if any. The method was used to document the risk of intussusception after rotavirus vaccines (see Intussusception after Rotavirus Vaccination -- Spontaneous Reports; N Engl J Med 2011; 365:2139) and the risk of Kawasaki disease following pneumococcal vaccination (see Kawasaki disease and 13-valent pneumococcal conjugate vaccination among young children: A self-controlled risk interval and cohort study with null results; PLoS Med. 2019;16(7):e100284).

#### *10.2.1.3. Hypothesis testing studies*

A complete review of study designs and methods from hypothesis testing studies in the field of vaccine safety is included in the ADVANCE Report on appraisal of vaccine safety methods.

Traditional study designs such as cohort and case-control studies may be difficult to implement for vaccines where studies involve populations with high vaccine coverage rates, an appropriate unvaccinated group is lacking or adequate information on covariates at the individual level is not available. Frequent sources of confounding to be considered are socioeconomic status, underlying health status and other factors influencing the probability of being vaccinated. Control without separate controls: evaluation of vaccine safety using case-only methods (Vaccine 2004; 22(15-16):2064-70)

describes and illustrates epidemiological methods that are useful in such situations. They are mostly case-only design described in Chapter 5.3.2 of the Guide:

- The case-crossover design was primarily developed to investigate the association between a vaccine and an adverse event. In this design, control information for each case is based on own past exposure experience and a person can 'crossover' between two or more exposure levels. It is a retrospective design that requires the strong assumption that the underlying probability of vaccination should be the same in all defined time intervals, but this is unlikely to hold for paediatric vaccines administered according to strict schedules or for seasonally administered vaccines.
- The self-controlled case series (SCCS) design can be both prospective and retrospective and aims to estimate a relative incidence, which compares the incidence of acute adverse events within periods of hypothesised excess risk due to exposure with incidence during all other times (baseline risk).
- The case-coverage design uses exposure information on cases and population data on vaccination coverage to serve as control. It requires reliable and detailed vaccine coverage data corresponding to the population from which cases are drawn. This will allow control of confounding by stratified analysis. During vaccine introduction, it is also particularly important to address selection bias introduced by awareness of possible occurrence of a specific outcome. An example of a study using a case-coverage method is Risk of narcolepsy in children and young people receiving AS03 adjuvanted pandemic A/H1N1 2009 influenza vaccine: retrospective analysis (BMJ 2013; 346:f794). (But: compares odds of exposure in cases to odds of exposure in the general population, similar to screening method in vaccine effectiveness: limited control for residual confounding, and selection bias as doesn't consider propensity to seek care/vaccination and underlying medical conditions / low comparability between cases and controls).

The study Control without separate controls: evaluation of vaccine safety using case-only methods (Vaccine 2004; 22(15-6):2064-70) concludes that properly designed and analysed epidemiological studies using only cases, especially the SCCS method, may provide stronger evidence than large cohort studies as they control completely for fixed individual-level confounders (such as demographics, genetics and social deprivation) and typically have similar, sometimes better, power. Three factors are however critical in making optimal use of such methods: access to good data on cases, computerised vaccination records with the ability to link them to cases and availability of appropriate analysis techniques.

Several studies on vaccines have compared traditional and case-only study designs:

- Epidemiological designs for vaccine safety assessment: methods and pitfalls (Biologicals 2012;40(5):389-92) used three study designs (cohort, case-control and self-controlled case series) to illustrate the issues that may arise when designing an epidemiological study, such as understanding the vaccine safety question, case definition and finding, limitations of data sources, uncontrolled confounding, and pitfalls that apply to the individual designs.
- Comparison of epidemiologic methods for active surveillance of vaccine safety (Vaccine 2008; 26(26):3341-3345) performed a simulation study to compare four designs (matched-cohort, vaccinated-only (risk interval) cohort, case-control and self-controlled case series) in the context of vaccine safety surveillance. The cohort study design allowed for the most rapid signal detection, the least false-positive error and highest statistical power in performing sequential analysis. The authors highlight, however, that the chief limitation of this simulation is the exclusion of confounding effects and the lack of chart review, which is a time and resource intensive requirement.

- Another simulation study (Four different study designs to evaluate vaccine safety were equally validated with contrasting limitations. J Clin Epidemiol 2006; 59(8):808-818) compared four study designs (cohort, case-control, risk-interval and SCCS) with the conclusion that all the methods were valid designs, with contrasting strengths and weaknesses. The SCCS method, in particular, proved to be an efficient and valid alternative to the cohort method.
- Hepatitis B vaccination and first central nervous system demyelinating events: Reanalysis of a case-control study using the self-controlled case series method. Vaccine 2007;25(31):5938-43) describes how the SCCS found similar results as the case-control study but with greater precision as it used cases without matched controls excluded from the case-control analysis. This is at the cost of the assumption that exposures are independent of earlier events. The authors recommended that, if case-control studies of vaccination and adverse events are undertaken, parallel case-series analyses should also be conducted, where appropriate.

While the SCCS is suited to secondary use of data, it may not be appropriate in situations where primary data collection is needed (e.g. a pandemic) since follow-up time needs to be accrued. In such instances, the Self-controlled Risk Interval (SCRI) method can be used to shorten the observation time (see The risk of Guillain-Barre Syndrome associated with influenza A (H1N1) 2009 monovalent vaccine and 2009-2010 seasonal influenza vaccines: Results from self-controlled analyses. Pharmacoepidemiol Drug Saf 2012;21(5):546-52), historical background rates can be used for an O/E analysis (see Near real-time surveillance for influenza vaccine safety: proof-of-concept in the Vaccine Safety Datalink Project. Am J Epidemiol 2010;171(2):177-88) or a classical case-control study can be performed, as used in Guillain-Barré syndrome and adjuvanted pandemic influenza A (H1N1) 2009 vaccine: multinational case-control study in Europe. BMJ 2011;343:d3908).

Ecological analyses should not be considered hypothesis testing studies (see Chapter 5.5 of this Guide) but can be useful for hypothesis generating without requiring the implementation of specific data collection.

#### *10.2.1.4. Meta-analyses*

The guidance on conducting meta-analyses of completed comparative pharmacoepidemiological studies of safety outcomes (Annex 1 of the Guide) applies also to vaccines. A systematic review evaluating the potential for bias and the methodological quality of meta-analyses in vaccinology (Vaccine 2007; 25(52):8794-806) provides a comprehensive overview of the methodological quality and limitations of 121 meta-analyses of vaccine studies. Association between Guillain-Barré syndrome and influenza A (H1N1) 2009 monovalent inactivated vaccines in the USA: a meta-analysis (Lancet 2013;381(9876):1461-8) describes a self-controlled risk-interval design in a meta-analysis of six studies at the patient level with a reclassification of cases according to the Brighton Collaboration classification.

#### *10.2.1.5. Studies on vaccine safety in special populations*

The article Vaccine safety in special populations (Hum Vaccin 2011;7(2):269-71) highlights common methodological issues that may arise in evaluating vaccine safety in special populations, especially infants and children who often differ in important ways from healthy individuals and change rapidly during the first few years of life, and elderly patients.

Pregnancy registries include pregnant women followed until the end of pregnancy and provide information on pregnancy outcomes. Besides the difficulties of recruitment and retention of pregnant women, specific challenges of using pregnancy registries for observational studies on adverse effects of vaccines administered during pregnancy include the identification of relevant control groups for comparisons and completeness of information on pregnancy outcomes as embryonic and early foetal

loss are often not recognised or recorded and data on the gestational age at which these events occur are often missing. These studies may require linkage with data captured in birth defects registries, teratology information services or electronic health care records where mother-child linkage is possible. In addition, the likelihood of vaccination increases with gestational age whereas the likelihood of foetal death decreases. Assessing the effect of vaccine on spontaneous abortion using time-dependent covariates Cox models (Pharmacoepidemiol Drug Saf 2012;21(8):844-50) demonstrates that rates of spontaneous abortion can be severely underestimated without survival analysis techniques using time-dependent covariates to avoid immortal time bias and shows how to fit such models. Risk of miscarriage with bivalent vaccine against human papillomavirus (HPV) types 16 and 18: pooled analysis of two randomised controlled trials (BMJ 2010; 340:c712) explains methods to calculate rates of miscarriage, address the lack of knowledge of time of conception during which vaccination might confer risk and perform subgroup and sensitivity analyses.

In Harmonising Immunisation Safety Assessment in Pregnancy Part I (Vaccine 2016;34 (49): 5991-6110) and Part II (Vaccine 2017;35 (48), 6469-582), the Global Alignment of Immunization Safety Assessment in pregnancy (GAIA) project has provided a selection of case definitions and guidelines for the evaluation of pregnancy outcomes following immunization. The Systematic overview of data sources for Drug Safety in pregnancy research provides an inventory of pregnancy exposure registries and alternative data sources useful to assess the safety of prenatal vaccine exposure.

Few vaccine studies are performed in immunocompromised subjects. Influenza vaccination for immunocompromised patients: systematic review and meta-analysis by etiology (J Infect Dis 2012;206(8):1250-9) illustrates the importance of performing stratified analyses by aetiology of immunocompromise and possible limitations due to residual confounding, differences within and between etiological groups and small sample size in some etiological groups. Although not vaccine-related, Burden of herpes zoster in 16 selected immunocompromised populations in England: a cohort study in the Clinical Practice Research Datalink 2000–2012 (BMJ Open. 2018; 8(6): e020528) illustrates the difficulty met for defining an immunocompromised cohort and a relevant comparator cohort when making secondary use of a primary health care database.

There is an increasing interest in the influence of genetics on safety and efficacy outcomes of vaccinations. Understanding this influence may optimise the choice of vaccines and the vaccination schedule. Research in this field is illustrated by Effects of vaccines in patients with sickle cell disease: a systematic review protocol (BMJ Open 2018;8:e021140. doi:10.1136/bmjopen-2017-021140) and Adversomics: a new paradigm for vaccine safety and design (Expert Rev Vaccines. 2015 Jul; 14(7): 935–47).

## 10.2.2. Vaccine effectiveness

#### *10.2.2.1. General aspects*

The article A framework for research on vaccine effectiveness (Vaccine 2018;36(48): 7286-93) proposes standardised definitions, considers models of vaccine failure and provides methodological considerations for different designs. This article is useful to researchers who investigate the effectiveness of vaccines and vaccination programs and why they may fail.

World Health Organisation's Evaluation of influenza vaccine effectiveness: a guide to the design and interpretation of observational studies (2017) provides a complete overview of methods to study effectiveness of influenza vaccines which are also relevant for other vaccines. Study designs and methods for measuring vaccine effectiveness in the PRISM system are also explained in Exploring the Feasibility of Conducting Vaccine Effectiveness Studies in Sentinel's PRISM Program (2018).

The article Vaccine effects and impact of vaccination programmes in post-licensure studies (Vaccine 2013;31(48):5634-42) reviews effectiveness of vaccine and of vaccination programmes, proposes epidemiological measures of public health impact, describes relevant methods to measure these effects and discusses the assumptions and potential biases involved.

The ADVANCE Report on appraisal of vaccine safety methods describes a large number of methods to assess vaccine safety which are also relevant for effectiveness evaluation.

It is worth mentioning that there are few comparative effectiveness studies between vaccines (except of head-to-head immunogenicity studies), but comparative effectiveness has been used to compare formulations or doses of a same vaccine.

#### *10.2.2.2. Traditional cohort and case-control studies*

Generic protocols for retrospective case-control studies and retrospective cohort studies to assess the effectiveness of rotavirus and influenza vaccination in EU Member States based on computerised databases were published by the European Centre for Disease Prevention and Control (ECDC) and can be used as reference for other vaccines. They describe the information that should be collected by country and region in vaccine effectiveness studies and the data sources that may be available to identify virus-related outcomes a vaccine is intended to avert, including hospital registers, computerised primary care databases, specific surveillance systems (i.e. laboratory surveillance, hospital surveillance, primary care surveillance) and laboratory registers.

The case-control methodology is frequently used to evaluate vaccine effectiveness post-authorisation but the potential for bias and confounding in such studies are important limitations. The articles Case-control vaccine effectiveness studies: Preparation, design, and enrollment of cases and controls (Vaccine 2017; 35(25):3295-302) and Case-control vaccine effectiveness studies: Data collection, analysis and reporting results (Vaccine 2017; 35(25):3303-8) summarize the recommendations of an expert group regarding best practices for the design, analysis and reporting of case-control vaccine effectiveness studies.

Based on a meta-analysis comprising 49 cohort studies and 10 case-control studies, Efficacy and effectiveness of influenza vaccines in elderly people: a systematic review (Lancet 2005;366(9492):1165-74) highlights the heterogeneity of outcomes and study populations included in such studies and the high likelihood of selection bias.

Non-specific effects of vaccines, such as a decrease of mortality, have been claimed in observational studies but generally can be affected by bias and confounding. Epidemiological studies of the 'non-specific effects' of vaccines: I--data collection in observational studies (Trop Med Int Health 2009;14(9):969-76.) and Epidemiological studies of the non-specific effects of vaccines: II--methodological issues in the design and analysis of cohort studies (Trop Med Int Health 2009;14(9):977-85) provide recommendations for vaccine observational studies conducted in countries with high mortality; these recommendations have wider relevance. The study Observational studies of non-specific effects of Diphtheria-Tetanus-Pertussis vaccines in low-income countries: Assessing the potential impact of study characteristics, bias and confounding through meta-regression (Vaccine. 2019;37(1):34–40) uses meta-regression to analyse study characteristics significantly associated with increased relative risks of non-specific effects of DTP vaccines.

#### *10.2.2.3. Screening method*

The screening method estimates vaccine effectiveness by comparing vaccination coverage in positive (usually laboratory confirmed) cases of a disease (e.g. influenza) with the vaccination coverage in the population from which the cases are derived (e.g., the same age group). If representative data on cases and vaccination coverage are available, it can provide an inexpensive and ready-to-use method

that can be useful in providing early effectiveness estimates or identify changes in effectiveness over time. However, Application of the screening method to monitor influenza vaccine effectiveness among the elderly in Germany (BMC Infect Dis. 2015;15(1):137) emphasises that accurate and age-specific vaccine coverage rates are crucial to provide valid VE estimates. Since adjusting for important confounders and the assessment of product-specific VE is generally not possible, this method should be considered only a supplementary tool for assessing crude VE.

#### *10.2.2.4. Indirect cohort (Broome) method*

The indirect cohort method is a case-control type design which uses cases caused by non-vaccine serotypes as controls. Use of surveillance data to estimate the effectiveness of the 7-valent conjugate pneumococcal vaccine in children less than 5 years of age over a 9 year period (Vaccine 2012;30(27):4067-72) applied this method to evaluate the effectiveness of a pneumococcal conjugate vaccine against invasive pneumococcal disease (IPD) and compared the results to the effectiveness measured using a standard case-control study conducted during the same time period. The authors considered the method would be most useful shortly after vaccine introduction, and less useful in a setting of very high vaccine coverage and fewer vaccine-type cases. Using the Indirect Cohort Design to Estimate the Effectiveness of the Seven Valent Pneumococcal Conjugate Vaccine in England and Wales (PLoS One 6(12):e28435. doi:10.1371/journal.pone.0028435) describes how the method was used to estimate effectiveness of various numbers of doses as well as for each vaccine serotype.

#### *10.2.2.5. Density case-control design*

Effectiveness of live-attenuated Japanese encephalitis vaccine (SA14-14-2): a case-control study (Lancet 1996;347(9015):1583-6) describes a case control study of incident cases in which the control group consisted of all village-matched children of a given age who were at risk of developing disease at the time that the case occurred (density sampling). The effect measured is an incidence density rate ratio.

#### *10.2.2.6. Test negative design*

The test-negative design aims to reduce bias associated with confounding by health-care-seeking behavior and misclassification of cases. The article The test-negative design for estimating influenza vaccine effectiveness (Vaccine 2013;31(17):2165-8) explains the rationale, assumptions and analysis of the test-negative study. Study subjects were all persons who seek care for an acute respiratory illness and influenza VE was estimated from the ratio of the odds of vaccination among subjects testing positive for influenza to the odds of vaccination among subject testing negative. This design is less susceptible to bias due to misclassification of infection and the confounding by health care-seeking behaviour, at the cost of difficult-to-test assumptions. The article Theoretical basis of the test-negative study design for assessment of influenza vaccine effectiveness (Am J Epidemiol. 2016;184(5):345-53; see also the related Comments) uses directed acyclic graphs to characterize potential biases in studies using this design and shows how bias can be avoided or minimised and where bias may be introduced with particular design variations.

Effectiveness of rotavirus vaccines in preventing cases and hospitalizations due to rotavirus gastroenteritis in Navarre, Spain (Vaccine 2012;30(3):539-43) evaluates effectiveness using a test negative case-control design based on electronic clinical reports. Cases were children with confirmed rotavirus and controls were those who tested negative for rotavirus in all samples. The test-negative design was based on an assumption that the rate of gastroenteritis caused by pathogens other than rotavirus is the same in both vaccinated and unvaccinated persons. This approach may rule out differences in parental attitude when seeking medical care and of physician differences in making decisions about stool sampling or hospitalisation. A limitation is sensitivity of antigen detection which may underestimate vaccine effectiveness. In addition, if virus serotype is not available, it is not

possible to study the association between vaccine failure and a possible mismatch of vaccine strains and circulating strains of virus.

The article 2012/13 influenza vaccine effectiveness against hospitalised influenza A(H1N1)pdm09, A(H3N2) and B: estimates from a European network of hospitals (EuroSurveill 2015;20(2):pii=21011) illustrates a multicentre test-negative case-control study to estimate influenza VE in 18 hospitals. It is believed that confounding due to health-seeking behaviour is minimised since, in the study sites, all people needing hospitalisation are likely to be hospitalised. The study Trivalent inactivated seasonal influenza vaccine effectiveness for the prevention of laboratory-confirmed influenza in a Scottish population 2000 to 2009 (EuroSurveill 2015;20(8):pii=21043) applied this method using a Scotland-wide linkage of patient-level primary care, hospital and virological swab data over nine influenza seasons and discusses strengths and weaknesses of the design in this context.

#### *10.2.2.7. Case coverage design*

This design is described in Chapter 10.2.1.3.

#### *10.2.2.8. Impact assessment*

The impact of vaccination can be quantified in children in the age group targeted for the vaccine (overall effect) or in children of other age groups (indirect effect). The direct effect of a vaccine, however, needs to be defined by the protection it confers given a specific amount of exposure to infection and not just a comparable exposure. A generic study protocol to assess the impact of rotavirus vaccination in EU Member States has been published by the ECDC. It recommends the information that needs to be collected to compare the incidence/proportion of rotavirus cases in the period before and after the introduction of the vaccine. These generic protocols need to be adapted to each country/regions and specific situation. Direct and indirect effects in vaccine efficacy and effectiveness (Am J Epidemiol 1991; 133(4):323-31) describes how parameters intended to measure direct effects must be robust and interpretable in the midst of complex indirect effects of vaccine intervention programmes.

Impact of rotavirus vaccination in regions with low and moderate vaccine uptake in Germany (Hum Vaccin Immunother 2012; 8(10):1407-15) describes an impact assessment of rotavirus vaccination comparing the incidence rates of hospitalisations before, and in seasons after, vaccine introduction using data from national mandatory disease reporting system. First year experience of rotavirus immunisation programme in Finland (Vaccine 2012; 31(1):176-82) estimates the impact of a rotavirus immunisation programme on the total hospital inpatient and outpatient treated acute gastroenteritis burden and on severe rotavirus disease burden during the first year after introduction. The study may be considered as a vaccine-probe-study, where unspecific disease burden prevented by immunisation is assumed to be caused by the agent the vaccine is targeted against. The study Lack of impact of rotavirus vaccination on childhood seizure hospitalizations in England - An interrupted time series analysis (Vaccine 2018; 36(31):4589-92) discusses possible reasons for negative findings in this study although previous studies have established a protective vaccine association in this age group.

In a review of 65 included articles, Population-level impact and herd effects following the introduction of human papillomavirus vaccination programmes: updated systematic review and meta-analysis (Lancet. 2019;394(10197):497–509) compared the frequency (prevalence or incidence) of several HPV-related endpoints between the pre-vaccination and post-vaccination periods with stratification by sex, age, and years since introduction of HPV vaccination.

#### *10.2.2.9. Methods to study waning immunity*

The study of vaccine effectiveness against diseases where immunity wanes over time requires consideration of both the within-host dynamics of the pathogen and immune system as well as the

associated population-level transmission dynamics. Implications of vaccination and waning immunity (Proc Biol Sci 2009; 276(1664):2071-80) seeks to combine immunological and epidemiological models for measles infection to examine the interplay between disease incidence, waning immunity and boosting.

Besides a discussion on effectiveness of varicella vaccines over time, Global Varicella Vaccine Effectiveness: A Meta-analysis (Pediatrics 2016; 137(3):e20153741) reports low effectiveness in outbreak investigations and highlights the difficulties to reliably measure effectiveness in this situation where some confounders cannot controlled for, the force of infection may be high, the degree of exposure may be variable across study participants and measures may originate from settings where there is epidemiologic evidence of vaccine failure. More than a few estimates are therefore needed to accurately assess vaccine effectiveness and conclude in waning immunity.

#### *10.2.2.10. Misclassification in studies of vaccine effectiveness*

Like vaccine safety studies, studies of vaccine effectiveness rely on accurate identification of vaccination and cases of vaccine-preventable diseases but in practice diagnostic tests, clinical case definitions and vaccination records often present inaccuracies. For outcomes with a complex natural history, and particularly when using secondary data collection (where case finding may be difficult), such as neurological or potential immune mediated diseases, validation studies based on case validation may be needed in a first step. Bias due to differential and non-differential disease- and exposure misclassification in studies of vaccine effectiveness (PLoS One 2018;15;13(6):e0199180) explores through simulations the impact of non-differential and differential disease- and exposure-misclassification when estimating vaccine effectiveness using cohort, case-control, test-negative case-control and case-cohort designs.

Misclassification can lead to significant bias and its impact strongly depends on the vaccination scenarios. A web application designed in the ADVANCE project is publicly available to assess the potential (joint) impact of possibly differential disease- and exposure misclassification.

## *10.3. Design and analysis of pharmacogenetic studies*

**Note: Chapter 10.3 has not been updated for revision 8**

### 10.3.1. Introduction

Pharmacogenetics is defined as the study of genetic variation as a determinant of drug response. It can complement information on clinical factors and disease sub-phenotypes to optimise the prediction of treatment response and reduce the risk of adverse reactions.

Individual variation in the response to drugs is an important clinical issue and may range from a lack of therapeutic effect to serious adverse drug reactions. This heterogeneity of response has important policy implications if individual patients not responding to conventional agents are denied access to other agents based on clinical trial evidence and systematic reviews that show no overall benefit. While also clinical variables such as disease severity, age, concomitant drug use and illnesses are potentially important determinants of the response to drugs, heterogeneity in drug disposition (absorption, metabolism, distribution, and excretion) and targets (such as receptors and signal transduction modulators) may be an important cause of inter-individual variability in the therapeutic effects of drugs (see Pharmacogenomics: translating functional genomics into rational therapeutics. Science 1999;286(5439):487-91). Identification of variation in genes which modify the response to drugs provides the opportunity to optimise safety and effectiveness of the currently available drugs and

develop new drugs for paediatric and adult populations (see Drug discovery: a historical perspective. Science 2000;287(5460):1960-4).

It is important to note that genetic variants are not the only potentially useful biomarkers of drug effects but a first step in the chain of genomics [DNA variation, SNPs, Copy Number Variations, indels], epigenomics [methylation], transcriptomics [RNA transcription], and proteomics [protein function and structure].

### 10.3.2. Identification of genetic variants

Identification of genetic variation associated with important drug or therapy-related outcomes can follow two main approaches.

The first is the candidate gene approach in which as many as dozens to thousands of genetic variations within one or several genes, including a common form of variations known as single nucleotide polymorphisms (SNPs), are genotyped, including the coding and noncoding sequence. Generally they are chosen on the grounds of biological plausibility, which may have been proven before in previous studies, or of knowledge of functional genes known to be involved in pharmacokinetic and pharmacodynamics pathways or related to the disease or intermediate phenotype. Methodological and statistical issues in pharmacogenomics (J Pharm Pharmacol 2010;62(2):161-6) discusses pros and cons of a candidate gene approach and a genome-wide scan approach (see below), and A tutorial on statistical methods for population association studies (Nat Rev Genet 2006;7(10):781-91) gives an outline of key methods that can be used. The advantage of the candidate gene approach is that resources can be directed to several important genetic polymorphisms and the higher a priori chance of relevant drug-gene interactions. This approach, however, requires a priori information about the likelihood of the polymorphism, gene, or gene-product interacting with a drug or drug pathway. Moving towards individualized medicine with pharmacogenomics (Nature 2004;429:464-8) explains that lack or incompleteness of information on genes from previous studies may result in the failure in identifying every important genetic determinant in the genome.

The second approach is hypothesis-generating or hypothesis-agnostic, known as genome-wide, which identifies genetic variants across the whole genome. By comparing the frequency of genetic or SNP markers between drug responders and non-responders, or those with or without drug toxicity, important genetic determinants are identified. In this approach, no previous information or specific gene/variant hypothesis is needed. Because of the concept of linkage disequilibrium, whereby certain genetic determinants tend to be co-inherited together, it is possible that the genetic associations identified through a genome-wide approach may not be truly biologically functional polymorphisms, but instead may simply be a linkage-related marker of another genetic determinant that is the true biologically relevant genetic determinant. Thus, this approach is considered discovery in nature. It may detect the SNPs in genes, which were previously not considered as candidate genes, or even SNPs outside of the genes. Nonetheless, failure to cover all relevant genetic risk factors can still be a problem, though less than with the candidate gene approach. It is therefore important to conduct replication and validation studies (in vivo and in vitro) to ascertain the generalisability of findings to populations of individuals, to characterise the mechanistic basis of the effect of these genes on drug action, and to identify true biologic genetic determinants. This approach is useful for studying complex diseases where multiple genetic variations contribute to disease risk, but are applicable to disease and treatment outcomes.

Various genome-wide approaches are currently available including genome and exome sequencing, and application of various chips that type hundreds of thousands to billions of SNPs (e.g. exome chip). Finally, power is usually limited to detect only common variants with a large effect, and therefore large

sample sizes should be considered, e.g. through pooling of biobanks. An example of such pooling is the CHARGE Consortium wth its focus on cardiovascular diseases [The Cohorts for Heart and Aging Research in Genomic Epidemiology (CHARGE) Consortium as a model of collaborative science. Epidemiology 2013;24:346-8.]. It is important that findng are replicated in other cohorts and consortia, but also that other techniques are actively used to confirm or refute associations [epigenomics, transcriptomics, and proteomics].

### 10.3.3. Study designs

Several options are available for the design of pharmacogenetic studies. Firstly, RCTs, both pre- and post-authorisation, provide the opportunity to address several pharmacogenetic questions. Pharmacogenetics in randomized controlled trials: considerations for trial design (Pharmacogenomics 2011;12(10):1485-92) describes three different trial designs differing in the timing of randomization and genotyping, and Promises and challenges of pharmacogenetics: an overview of study design, methodological and statistical issues (JRSM Cardiovasc Dis 2012 5;1(1)) discusses outstanding methodological and statistical issues that may lead to heterogeneity among reported pharmacogenetic studies and how they may be addressed. Pharmacogenetic trials can be designed (or post hoc analysed) with the intention to study whether a subgroup of patients, defined by certain genetic characteristics, respond differently to the treatment under study. Alternatively, a trial can verify whether genotype-guided treatment is beneficial over standard care. Obvious limitations with regard to the assessment of rare adverse drug events are the large sample size required and its related high costs. In order to make a trial as efficient as possible in terms of time, money and/or sample size, it is possible to opt for an adaptive trial design, which allows prospectively planned modifications in design after patients have been enrolled in the study. Such a design uses accumulating data to decide how to modify aspects of the study during its progress, without undermining the validity and integrity of the trial. An additional benefit is that the expected number of patients exposed to an inferior/harmful treatment can be reduced (see Potential of adaptive clinical trial designs in pharmacogenetic research. Pharmacogenomics 2012;13(5):571-8).

Observational studies are the alternative and can be family-based (using twins or siblings) or population-based (using unrelated individuals). The main advantage of family-based studies is the avoidance of bias due to population stratification. A clear practical disadvantage for pharmacogenetic studies is the requirement to study families where patients have been treated with the same drugs (see Methodological quality of pharmacogenetic studies: issues of concern. Stat Med 2008;27(30):6547-69).

Population-based studies may be designed to assess drug-gene interactions as cohort (including exposure-only), case-cohort and case-control studies (including case-only, as described in Nontraditional epidemiologic approaches in the analysis of gene-environment interaction: case-control studies with no controls! Am J Epidemiol 1996;144(3):207-13). Sound pharmacoepidemiological principles as described in the current Guide also apply to observational pharmacogenetic studies. A specific type of confounding due to population stratification needs to be considered in pharmacogenetic studies, and, if present, needs to be dealt with. Its presence may be obvious where the study population includes more than one immediately recognisable ethnic group; however in other studies stratification may be more subtle. Population stratification can be detected by Pritchard and Rosenberg's method, which involves genotyping additional SNPs in other areas of the genome and testing for association between them and outcome. In genome-wide association studies, the data contained within the many SNPs typed can be used to assess population stratification without the need to undertake any further genotyping. Several methods have been suggested to control for population stratification such as genomic control, structure association and EIGENSTAT.

These methods are discussed in Methodological quality of pharmacogenetic studies: issues of concern (Stat Med 2008;27(30):6547-69) and Softwares and methods for estimating genetic ancestry in human populations (Hum Genomics 2013;7:1).

The main advantage of exposure-only and case-only designs is the smaller sample size that is required, at the cost of not being able to study the main effects of drug exposure (case-only) or genetic variant (exposure-only) on the outcome. Furthermore, interaction can be assessed only on a multiplicative scale, whereas from a public health perspective additive interactions are very relevant. However, up till now GWAs with gene*interactions have not been very rewarding because of the required huge power. An important condition that has to be fulfilled for case-only studies is that the exposure is independent of the genetic variant, e.g. prescribers are not aware of the genotype of a patient and do not take this into account, directly or indirectly (by observing clinical characteristics associated with the genetic variant). In the exposure-only design, the genetic variant should not be associated with the outcome, for example variants of genes coding for cytochrome p-450 enzymes. When these conditions are fulfilled and the main interest is in the drug-gene interaction, these designs may be an efficient option. In practice, case-control and case-only studies usually result in the same interaction effect as empirically assessed in Bias in the case-only design applied to studies of gene-environment and gene-gene interaction: a systematic review and meta-analysis (Int J Epidemiol 2011;40(5):1329-41). The assumption of independence of genetic and exposure factors can be verified among controls before proceeding to the case-only analysis. Further development of the case-only design for assessing gene-environment interaction: evaluation of and adjustment for bias (Int J Epidemiol 2004;33(5):1014-24) conducted sensitivity analyses to describe the circumstances in which controls can be used as proxy for the source population when evaluating gene-environment independence. The gene-environment association in controls will be a reasonably accurate reflection of that in the source population if baseline risk of disease is small (<1%) and the interaction and independent effects are moderate (i.e. risk ratio<2), or if the disease risk is low (e.g. <5%) in all strata of genotype and exposure. Furthermore, non-independence of gene-environment can be adjusted in multivariable models if non-independence can be measured in controls.

### 10.3.4. Data collection

The same principles and approaches to data collection as for other pharmacoepidemiological studies can be followed (see Chapter 3 of this Guide on Approaches to Data Collection). An efficient approach to data collection for pharmacogenetic studies is to combine secondary use of electronic health records with primary data collection (e.g. biological samples to extract DNA).

Examples are given by SLCO1B1 genetic variant associated with statin-induced myopathy: a proof-of-concept study using the clinical practice research datalink (Clin Pharmacol Ther 2013;94(6):695-701), Diuretic therapy, the alpha-adducin gene variant, and the risk of myocardial infarction or stroke in persons with treated hypertension (JAMA 2002;287(13):1680-9) and Interaction between the Gly460Trp alpha-adducin gene variant and diuretics on the risk of myocardial infarction (J Hypertens 2009 Jan;27(1):61-8). Another approach to enrich electronic health records with biological samples is record linkage to biobanks as illustrated in Genetic variation in the renin-angiotensin system modifies the beneficial effects of ACE inhibitors on the risk of diabetes mellitus among hypertensives (Hum Hypertens 2008;22(11):774-80). A third approach is to use active surveillance methods to fully characterise drug effects such that a rigorous phenotype can be developed prior to genetic analysis. This approach was followed in Adverse drug reaction active surveillance: developing a national network in Canada's children's hospitals (Pharmacoepidemiol Drug Saf 2009;18(8):713-21) and EUDRAGENE: European collaboration to establish a case-control DNA collection for studying the genetic basis of adverse drug reactions (Pharmacogenomics 2006;7(4):633-8).

### 10.3.5. Data analysis

The focus of data analysis should be on the measure of effect modification (see Chapter 4.2.4 of this Guide on Effect Modification). Attention should be given to whether the mode of inheritance (e.g. dominant, recessive or additive) is defined a priori based on prior knowledge from functional studies. However, investigators are usually naïve regarding the underlying mode of inheritance. A solution might be to undertake several analyses, each under a different assumption, though the approach to analysing data raises the problem of multiple testing (see Methodological quality of pharmacogenetic studies: issues of concern. Stat Med 2008;27(30):6547-69). The problem of multiple testing and the increased risk of type I error is in general a problem in pharmacogenetic studies evaluating multiple SNPs, multiple exposures and multiple interactions. The most common approach to correct for multiple testing is to use the Bonferroni correction. This correction may be considered too conservative and runs the risk of producing many pharmacogenetic studies with a null result. Other approaches to adjust for multiple testing include permutation testing and false discovery rate (FDR) control, which are less conservative. The FDR, described in Statistical significance for genomewide studies (Proc Natl Acad Sci USA 2003;100(16):9440-5), estimates the expected proportion of false-positives among associations that are declared significant, which is expressed as a q-value.

Alternative innovative methods are under development and may be used in the future, such as the systems biology approach, a Bayesian approach, or data mining (see Methodological and statistical issues in pharmacogenomics. J Pharm Pharmacol 2010;62(2):161-6).

Important complementary approaches include the conduct of individual patient data meta-analyses and/or replication studies to avoid the risk of false-positive findings.

An important step in analysis of genome-wide association studies data that needs to be considered is the conduct of rigorous quality control procedures before conducting the final association analyses. Relevant guidelines include Guideline for data analysis of genomewide association studies (Cancer Genomics Proteomics 2007;4(1):27-34) and Statistical Optimization of Pharmacogenomics Association Studies: Key Considerations from Study Design to Analysis (Curr Pharmacogenomics Person Med 2011;9(1):41-66).

### 10.3.6. Reporting

The guideline STrengthening the REporting of Genetic Association studies (STREGA)--an extension of the STROBE statement (Eur J Clin Invest 2009;39(4):247-66) should be followed for reporting findings of genetic studies.

### 10.3.7. Clinical practice guidelines

An important step towards the implementation of the use of genotype information to guide pharmacotherapy is the development of clinical practice guidelines. Several initiatives have been developed to provide these guidelines such as the Clinical Pharmacogenetics Implementation Consortium. Furthermore, several clinical practice recommendations have been published, for example Recommendations for HLA-B*15:02 and HLA-A*31:01 genetic testing to reduce the risk of carbamazepine-induced hypersensitivity reactions (Epilepsia 2014;55(4):496-506) or Clinical practice guideline: CYP2D6 genotyping for safe and efficacious codeine therapy (J Popul Ther Clin Pharmacol 2013;20(3):e369-96).

### 10.3.8. Resources

An important pharmacogenomics knowledge resource is available through PharmGKB that encompasses clinical information including dosing guidelines and drug labels, potentially clinically actionable gene-drug associations and genotype-phenotype relationships. PharmGKB collects curates and disseminates knowledge about the impact of human genetic variation on drug responses.

# References

**The hyperlinks in the document were last accessed online in June 2020.**


**Foreword**

de Abajo F, Rodríguez-Martín S, Lerma L et al. May 2020. Use of renin–angiotensin–aldosterone system inhibitors and risk of COVID-19 requiring admission to hospital: a case-population study. Lancet 2020; May 14; 395**.**

European Medicines Agency (EMA). Latest data support continued use of ACE inhibitors and ARB medicines during COVID-19 pandemic. Press Release 9 June 2020

Gnavi, Demaria M, Picariello R et al. Therapy with agents acting on the renin-angiotensin system and risk of SARS-CoV-2 infection. May 2020. Clinical Infectious Diseases, 2020

Jung S, Choi J-C, You SH et al. Association of renin-angiotensin-aldosterone system inhibitors with COVID-19-related outcomes in Korea: a nationwide population-based cohort study. Clinical Infectious Diseases, 2020; 22 May. ciaa624, https://doi.org/10.1093/cid/ciaa624

Lane JCE, Weaver J, Kostka K et al. Safety of hydroxychloroquine, alone and in combination with azithromycin, in light of rapid wide-spread use for COVID-19: a multinational, network cohort and self-controlled case series study. 2020, 31 May. doi: https://doi.org/10.1101/2020.04.08.20054551

Mancia G, Rea F, Ludergnani M et al . Renin–Angiotensin–Aldosterone System Blockers and the Risk of Covid-19. New Eng J Med 2020; May 1

Mehta N, Kalra A, Nowacki AS et al. Association of Use of Angiotensin-Converting Enzyme Inhibitors and Angiotensin II Receptor Blockers With Testing Positive for Coronavirus Disease 2019 (COVID-19). JAMA Cardiol .Published online May 5, 2020. doi:10.1001/jamacardio.2020.1855.

Morales DR, Conover MM, You SY et al. Renin-angiotensin system blockers and susceptibility to COVID-19: a multinational open science cohort study. https://doi.org/10.1101/2020.06.11.20125849

Pottegård A, Kurz X, Moore N, Christiansen CF, Klungel O. Considerations for pharmacoepidemiological analyses in the SARS-CoV-2 pandemic. Pharmacoepidemiol Drug Saf. 2020 May 5. doi: 10.1002/pds.5029

Reynolds, HR, Adhikari S, Pulgarin C et al. May 2020. Renin–Angiotensin–Aldosterone System Inhibitors and Risk of Covid-19. New Engl J Med 2020, May 1. DOI: 10.1056/NEJMoa2008975

Watson J et al. An open letter to Mehra et al and The Lancet. https://zenodo.org/record/3865253#.XtaoLjiP7ct

**Chapter 1 – Introduction**

Ali A, Hartzema A. (2018). Post-Authorization Safety Studies of Medicinal Products. The PASS Book (1st ed.). Academic Press.

Altman, D. (2020). Practical Statistics for Medical Research. Chapman & Hall.

Andrews AB, Moore N. (2014). Mann's Pharmacovigilance (3rd ed.). Wiley Blackwell.

Cave, A., Kurz, X., Arlett, P. (2019). Real-World Data for Regulatory Decision Making: Challenges and Possible Solutions for Europe. Clinical Pharmacology and Therapeutics. https://doi.org/10.1002/cpt.1426

Directive 2001/83/EC of the European Parliament and of the Council of 6 November 2001. (n.d.). Retrieved 14 May 2020, from https://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=CONSLEG:2001L0083:20070126:en:PDF

Elseviers M, Wettermark B, Almarsdóttir AB, et al. (2016). Drug Utilization Research: Methods and Applications. Wiley Blackwell.

Guideline on good pharmacovigilance practices (GVP) Module VIII – Post-authorisation safety studies (Rev 3). European Medicines Agency. https://www.ema.europa.eu/en/human-regulatory/post-authorisation/pharmacovigilance/good-pharmacovigilance-practices

Hartzema AG, Tilson HH, Chan KA. (2008). Pharmacoepidemiology and Therapeutic Risk Management (3rd ed.). Harvey Whitney Books Company.

Hoffmann, W., Latza, U., Baumeister, S. E., Brünger, M., Buttmann-Schweiger, N., Hardt, J., Hoffmann, V., Karch, A., Richter, A., Schmidt, C. O., Schmidtmann, I., Swart, E., & Berg, N. van den. (2019). Guidelines and recommendations for ensuring Good Epidemiological Practice (GEP): a guideline developed by the German Society for Epidemiology. European Journal of Epidemiology, 34(3), 301–317. https://doi.org/10.1007/s10654-019-00500-x

IEA Good Epidemiology Practices (GPP). (2007). [IEA Publications]. https://ieaweb.org/IEAWeb/Content/IEA_Publications.aspx

ISPE Guidelines for Good Pharmacoepidemiology Practices (GPP) (Rev 3). (2015). https://www.pharmacoepi.org/resources/policies/guidelines-08027/

Parfrey PS, Barret BJ. (2015). Clinical epidemiology: practice and methods (2nd ed.). Human Press.

Porta M, Last JM, Greenland. (2014). A Dictionary of Epidemiology (6th ed.). Oxford University Press.

Real-World Evidence. FDA. https://www.fda.gov/science-research/science-and-research-special-topics/real-world-evidence

Rothman K, Greenland S, Lash T. (2020). Modern Epidemiology (4th ed.). Lippincott Williams & Wilkins.

Scientific guidance on post-authorisation efficacy studies - EMA. (n.d.). Retrieved 14 May 2020, from https://www.ema.europa.eu/en/documents/scientific-guideline/scientific-guidance-post-authorisation-efficacy-studies-first-version_en.pdf

Strom B, Kimmel SE, Hennessy S. (2019). Pharmacoepidemiology (6th ed.). Wiley.

Woodward, M. (2014). Epidemiology: Study Design and Data Analysis (3rd ed.). Chapman & Hall.

**Chapter 2 - Formulating the research question**

Brown, P., Brunnhuber, K., Chalkidou, K., Chalmers, I., Clarke, M., Fenton, M., Forbes, C., Glanville, J., Hicks, N. J., Moody, J., Twaddle, S., Timimi, H., & Young, P. (2006). How to formulate research recommendations. BMJ : British Medical Journal, 333(7572), 804–806. https://doi.org/10.1136/bmj.38987.492014.94

European Medicines Agency. (n.d.). European Medicines Agency. Retrieved 14 May 2020, from https://www.ema.europa.eu/en

Guideline on good pharmacovigilance practices (GVP). European Medicines Agency. Retrieved 27 May 2020, from https://www.ema.europa.eu/en/human-regulatory/post-authorisation/pharmacovigilance/good-pharmacovigilance-practices

ISPE Guidelines for Good Pharmacoepidemiology Practices (GPP) (Rev 3). (2015). International Society for Pharmacoepidemiology. https://www.pharmacoepi.org/resources/policies/guidelines-08027/

Position paper on how to best formulate research recommendations - EUnetHTA. (n.d.). Retrieved 14 May 2020, from https://www.eunethta.eu/wp-content/uploads/2018/01/eunethta_position_paper_on_research_recommendations_0-1.pdf

The criteria to select and prioritise health technologies for additional evidence generation - EUnetHTA. (n.d.). Retrieved 14 May 2020, from https://eunethta.eu/wp-content/uploads/2018/01/Selection-prioritisation-criteria-1.pdf

Willame, C., Baril, L., van den Bosch, J., Ferreira, G. L. C., Williams, R., Rosillon, D., & Cohet, C. (2016). Importance of feasibility assessments before implementing non-interventional pharmacoepidemiologic studies of vaccines: lessons learned and recommendations for future studies. Pharmacoepidemiology and Drug Safety, 25(12), 1397–1406. https://doi.org/10.1002/pds.4081

**Chapter 3 - Development of the study protocol**

Agency for Healthcare Research and Quality (AHRQ). (n.d.). Retrieved 14 May 2020, from https://www.ahrq.gov/

Commission Implementing Regulation (EU) No 520/2012. (n.d.). Retrieved 14 May 2020, from https://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2012:159:0005:0025:EN:PDF

Developing a Protocol for Observational Comparative Effectiveness Research: A User's Guide - Effective Health Care Program - AHRQ. (2013). Agency for Healthcare Research and Quality. https://effectivehealthcare.ahrq.gov/products/observational-cer-protocol/research

Directive 2001/20/EC of the European Parliament and of the Council of 4 April 2001. (n.d.). Retrieved 14 May 2020, from https://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2001:121:0034:0044:en:PDF

ENCePP Checklist for Study Protocols. (n.d.). ENCePP. Retrieved 14 May 2020, from http://www.encepp.eu/standards_and_guidances/checkListProtocols.shtml

EU PAS Register. (n.d.). Retrieved 27 May 2020, from http://www.encepp.eu/encepp/studiesDatabase.jsp

Guidance for the format and content of the final study report of non-interventional post-authorisation safety studies - EMA. (n.d.). Retrieved 14 May 2020, from https://www.ema.europa.eu/en/documents/regulatory-procedural-guideline/guidance-format-content-final-study-report-non-interventional-post-authorisation-safety-studies_en.pdf

Guidance for the format and content of the protocol of non-interventional post-authorisation safety studies - EMA. (n.d.). Retrieved 14 May 2020, from https://www.ema.europa.eu/en/documents/other/guidance-format-content-protocol-non-interventional-post-authorisation-safety-studies_en.pdf

Guideline on good pharmacovigilance practices (GVP) Module VI – Collection, management and submission of reports of suspected adverse reactions to medicinal products (Rev 2) European Medicines Agency. Retrieved 26 May 2020, from https://www.ema.europa.eu/en/human-regulatory/post-authorisation/pharmacovigilance/good-pharmacovigilance-practices

Guideline on good pharmacovigilance practices (GVP) Module VIII – Post-authorisation safety studies (Rev 3) European Medicines Agency. Retrieved 15 May 2020, from https://www.ema.europa.eu/en/human-regulatory/post-authorisation/pharmacovigilance/good-pharmacovigilance-practices

ISPE Guidelines for Good Pharmacoepidemiology Practices (GPP) (Rev 3). (2015). International Society for Pharmacoepidemiology. https://www.pharmacoepi.org/resources/policies/guidelines-08027/

Research, C. for D. E. and R., Center for Biologics Evaluation and Research. (n.d.). FDA's Best Practices for Conducting and Reporting Pharmacoepidemiologic Safety Studies Using Electronic Healthcare Data Sets - FDA. Retrieved 14 May 2020, from https://www.fda.gov/regulatory-information/search-fda-guidance-documents/best-practices-conducting-and-reporting-pharmacoepidemiologic-safety-studies-using-electronic

The European Union electronic Register of Post-Authorisation Studies (EU PAS Register). (n.d.). ENCePP. Retrieved 14 May 2020, from http://www.encepp.eu/encepp/studiesDatabase.jsp

Volume 10 of the Rules Governing Medicinal Products in the European Union. Public Health - European Commission. Retrieved 14 May 2020, from https://ec.europa.eu/health/documents/eudralex/vol-10_en

## Chapter 4 - Approaches to data collection

Guideline on good pharmacovigilance practices (GVP) Module VI – Collection, management and submission of reports of suspected adverse reactions to medicinal products (Rev 2). European Medicines Agency. Retrieved 26 May 2020, from https://www.ema.europa.eu/en/human-regulatory/post-authorisation/pharmacovigilance/good-pharmacovigilance-practices

### 4.1 – Primary data collection

Abenhaim, L., Moride, Y., Brenot, F., Rich, S., Benichou, J., Kurz, X., Higenbottam, T., Oakley, C., Wouters, E., Aubier, M., Simonneau, G., & Bégaud, B. (1996). Appetite-Suppressant Drugs and the Risk of Primary Pulmonary Hypertension. New England Journal of Medicine, 335(9), 609–616. https://doi.org/10.1056/NEJM199608293350901

Clinical trials - Regulation EU No 536/2014. (n.d.). Public Health - European Commission. Retrieved 18 May 2020, from https://ec.europa.eu/health/human-use/clinical-trials/regulation_en

Artime, E., Qizilbash, N.,Garrido-Estepa, M, Vora, P, Soriano-Gabarró, M, Asiimwe, A, Pocock, S. (2019). Are Risk Minimization Measures for Approved Drugs in Europe Effective? A Systematic Review. Expert Opinion on Drug Safety. https://doi.org/10.1080/14740338.2019.1612875

EudraLex - Volume 10 - Clinical trials guidelines. (n.d.). Public Health - European Commission. Retrieved 18 May 2020, from https://ec.europa.eu/health/documents/eudralex/vol-10_en

Fayers PM, Machin D. (2007). Quality of Life: the assessment, analysis and interpretation of patient-related outcomes (2nd ed.). Wiley.

FDA - Guidence on Survey Methodologies to Assess REMS Goals That Relate to Knowledge. (2019). Center for Drug Evaluation and Research Center for Biologics Evaluation and Research.

https://www.fda.gov/regulatory-information/search-fda-guidance-documents/survey-methodologies-assess-rems-goals-relate-knowledge

FDA - REMS Assessment: Planning and Reporting. (n.d.). Retrieved 4 June 2020, from https://www.fda.gov/media/119790/download

Guideline on good pharmacovigilance practices (GVP) Module VIII – Post-authorisation safety studies (Rev 3). European Medicines Agency. Retrieved 18 May 2020, from https://www.ema.europa.eu/en/human-regulatory/post-authorisation/pharmacovigilance/good-pharmacovigilance-practices

Guideline on good pharmacovigilance practices (GVP) Module XVI – Risk minimisation measures: selection of tools and effectiveness indicators (Rev 2). European Medicines Agency. Retrieved 4 June 2020, from https://www.ema.europa.eu/en/human-regulatory/post-authorisation/pharmacovigilance/good-pharmacovigilance-practices

Groves RM, Fowler FJ, Couper MP et al. (2009). Survey Methodology (2nd ed.). Wiley.

ICH E6 (R2) Good clinical practice European Medicines Agency. Retrieved 18 May 2020, from https://www.ema.europa.eu/en/ich-e6-r2-good-clinical-practice

Kaufman, D. W., Rosenberg, L., & Mitchell, A. A. (2001). Signal generation and clarification: use of case-control data. Pharmacoepidemiology and Drug Safety, 10(3), 197–203. https://doi.org/10.1002/pds.571

Kish L. (1995). Survey Sampling. Wiley.

My, S., A, R., P, B., Pgm, M., S, F., E, F., & Eh, M. (2018). The RIMES Statement: A Checklist to Assess the Quality of Studies Evaluating Risk Minimization Programs for Medicinal Products. Drug Safety. https://doi.org/10.1007/s40264-017-0619-x

Roujeau, J.-C., Kelly, J. P., Naldi, L., Rzany, B., Stern, R. S., Anderson, T., Auquier, A., Bastuji-Garin, S., Correia, O., Locati, F., Mockenhaupt, M., Paoletti, C., Shapiro, S., Shear, N., Schöpf, E., & Kaufman, D. W. (1995). Medication Use and the Risk of Stevens–Johnson Syndrome or Toxic Epidermal Necrolysis. New England Journal of Medicine, 333(24), 1600–1608. https://doi.org/10.1056/NEJM199512143332404

Shapiro, S. (1983). The design of a study of the drug etiology of agranulocytosis and aplastic anemia. European Journal of Clinical Pharmacology, 24(6), 833–836. https://doi.org/10.1007/BF00607096

Standard Definitions Final Dispositions of Case Codes and Outcome Rates for Surveys - AAPOR. (n.d.). Retrieved 4 June 2020, from https://www.aapor.org/AAPOR_Main/media/publications/Standard-Definitions20169theditionfinal.pdf

**4.2 - Secondary use of data**

Cave, A, Kurz, X, Arlett, P. (2019). Real-World Data for Regulatory Decision Making: Challenges and Possible Solutions for Europe. Clinical Pharmacology and Therapeutics. https://doi.org/10.1002/cpt.1426

Berger, M. L., Sox, H., Willke, R. J., Brixner, D. L., Eichler, H.-G., Goettsch, W., Madigan, D., Makady, A., Schneeweiss, S., Tarricone, R., Wang, S. V., Watkins, J., & Mullins, C. D. (2017). Good Practices for Real-World Data Studies of Treatment and/or Comparative Effectiveness: Recommendations from the Joint ISPOR-ISPE Special Task Force on Real-World Evidence in Health Care Decision Making. Value in Health, 20(8), 1003-1008. https://doi.org/10.1016/j.jval.2017.08.3019

ENCePP Resources Database. (n.d.). ENCePP. Retrieved 10 June 2020, from http://www.encepp.eu/encepp/resourcesDatabase.jsp

FDA's Best Practices for Conducting and Reporting Pharmacoepidemiologic Safety Studies Using Electronic Healthcare Data Sets. Center for Drug Evaluation and Research Center for Biologics Evaluation and Research. https://www.fda.gov/regulatory-information/search-fda-guidance-documents/best-practices-conducting-and-reporting-pharmacoepidemiologic-safety-studies-using-electronic

FDA Real-World Evidence. FDA. https://www.fda.gov/science-research/science-and-research-special-topics/real-world-evidence

Hall, Gc., Sauer, B., Bourke, A., Brown, Js., Reynolds, Mw., LoCasale, L. (2012). Guidelines for Good Database Selection and Use in Pharmacoepidemiology Research. Pharmacoepidemiology and Drug Safety. https://doi.org/10.1002/pds.2229

Hoffmann, W., Latza, U., Baumeister, S. E., Brünger, M., Buttmann-Schweiger, N., Hardt, J., Hoffmann, V., Karch, A., Richter, A., Schmidt, C. O., Schmidtmann, I., Swart, E., & Berg, N. van den. (2019). Guidelines and recommendations for ensuring Good Epidemiological Practice (GEP): a guideline developed by the German Society for Epidemiology. European Journal of Epidemiology, 34(3), 301–317. https://doi.org/10.1007/s10654-019-00500-x

ISPE Guidelines for Good Pharmacoepidemiology Practices (GPP). (2015). International Society for Pharmacoepidemiology. https://www.pharmacoepi.org/resources/policies/guidelines-08027/

Schneeweiss, S., Avorn, J. (2005). A Review of Uses of Health Care Utilization Databases for Epidemiologic Research on Therapeutics. Journal of Clinical Epidemiology. https://doi.org/10.1016/j.jclinepi.2004.10.012

Strom B, Kimmel SE, Hennessy S. (2019). Pharmacoepidemiology (6th ed.). Wiley.

**4.3 - Patient registries**

CHMP Guideline on Conduct of Pharmacovigilance for Medicines Used by the Paediatric Population - EMA. (n.d.). Retrieved 15 May 2020, from https://www.ema.europa.eu/en/documents/scientific-guideline/draft-guideline-conduct-pharmacovigilance-medicines-used-paediatric-population_en.pdf

Emilsson, L., Lindahl, B., Köster, M., Lambe, M., & Ludvigsson, J. F. (2015). Review of 103 Swedish Healthcare Quality Registries. Journal of Internal Medicine, 277(1), 94–136. https://doi.org/10.1111/joim.12303

ENCePP Resources Database. (n.d.). ENCePP. Retrieved 15 May 2020, from http://www.encepp.eu/encepp/resourcesDatabase.jsp

EUnetHTA Joint Action 3 (2016 - 2020). (n.d.). EUnetHTA. Retrieved 15 May 2020, from https://eunethta.eu/ja3-archive/

EUROCAT. (n.d.). Retrieved 15 May 2020, from http://www.eurocat-network.eu/

FDA Guidance for Industry Establishing Pregnancy Exposure Registries. (n.d.). Retrieved 15 May 2020, from https://www.fda.gov/media/75607/download

Furu K, Kieler H, Haglund B, Engeland A, Selmer R, Stephansson O, Valdimarsdottir UA, Zoega H, Artama M, Gissler M, Malm H, Nørgaard M. (2015). Selective serotonin reuptake inhibitors and venlafaxine in early pregnancy and risk of birth defects: population based cohort study and sibling design. BMJ : British Medical Journal, 350, h1798. http://www.bmj.com/content/bmj/350/bmj.h1798.full.pdf

Guideline on good pharmacovigilance practices (GVP) Module VIII – Post-authorisation safety studies (Rev 3). (2018, September 17). European Medicines Agency. https://www.ema.europa.eu/en/human-regulatory/post-authorisation/pharmacovigilance/good-pharmacovigilance-practices

JRC EU RD Platform releases the Set of Common Data Elements for RD Registration. (n.d.). ERA-Net E-Rare. Retrieved 15 May 2020, from http://www.erare.eu/news/jrc-eu-rd-platform-releases-set-common-data-elements-rd-registration

Lauer, M. S., & D'Agostino, R. B. (2013). The randomized registry trial--the next disruptive technology in clinical research? The New England Journal of Medicine, 369(17), 1579–1581. https://doi.org/10.1056/NEJMp1310102

Li, G., Sajobi, T. T., Menon, B. K., Korngut, L., Lowerison, M., James, M., Wilton, S. B., Williamson, T., Gill, S., Drogos, L. L., Smith, E. E., Vohra, S., Hill, M. D., Thabane, L., & 2016 Symposium on Registry-Based Randomized Controlled Trials in Calgary. (2016). Registry-based randomized controlled trials-what are the advantages, challenges, and areas for future research? Journal of Clinical Epidemiology, 80, 16–24. https://doi.org/10.1016/j.jclinepi.2016.08.003

McGettigan, P., Alonso Olmo, C., Plueschke, K., Castillon, M., Nogueras Zondag, D., Bahri, P., Kurz, X., & Mol, P. G. M. (2019). Patient Registries: An Underused Resource for Medicines Evaluation : Operational proposals for increasing the use of patient registries in regulatory assessments. Drug Safety, 42(11), 1343–1351. https://doi.org/10.1007/s40264-019-00848-9

Norwegian Patient Registry. (n.d.). Retrieved 15 May 2020, from https://helsedirektoratet.no/english/norwegian-patient-registry

Pacurariu, A., Plueschke, K., Olmo, C. A., & Kurz, X. (2018). Imposed registries within the European postmarketing surveillance system: Extended analysis and lessons learned for regulators. Pharmacoepidemiology and Drug Safety, 27(7), 823–826. https://doi.org/10.1002/pds.4449

PARENT - Methodological guidelines and recommendations for efficient and rational governance of patient registries. (n.d.). Retrieved 11 June 2020, from https://ec.europa.eu/health/sites/health/files/ehealth/docs/patient_registries_guidelines_en.pdf

Patient registries. (n.d.). European Medicines Agency. Retrieved 15 May 2020, from https://www.ema.europa.eu/en/human-regulatory/post-authorisation/patient-registries

PharmaChild PRINTO. (n.d.). Paediatric Rheumatology INternational Trials Organization. Retrieved 15 May 2020, from https://www.printo.it/projects/ongoing/15

Qualification Opinion - The European Cystic Fibrosis Society Patient Registry (ECFSPR) - EMA. (n.d.). Retrieved 15 May 2020, from https://www.ema.europa.eu/en/documents/regulatory-procedural-guideline/qualification-opinion-european-cystic-fibrosis-society-patient-registry-ecfspr_en.pdf

Qualification opinion on Cellular therapy module of the European Society for Blood & Marrow Transplantation (EBMT) Registry - EMA. (n.d.). Retrieved 15 May 2020, from https://www.ema.europa.eu/en/documents/scientific-guideline/qualification-opinion-cellular-therapy-module-european-society-blood-marrow-transplantation-ebmt_en.pdf

Registries for Evaluating Patient Outcomes: A User's Guide: 3rd Edition Addendum (Overview). (2018). Agency for Healthcare Research and Quality. https://effectivehealthcare.ahrq.gov/products/registries-guide-3rd-edition-addendum/overview

Registry Evaluation and Quality Standards Tool and its vision paper (REQueST). (n.d.). EUnetHTA. Retrieved 15 May 2020, from https://eunethta.eu/request-tool-and-its-vision-paper/

Research, C. for D. E. and. (2020). Draft Postapproval Pregnancy Safety Studies Guidance for Industry - FDA. FDA. https://www.fda.gov/regulatory-information/search-fda-guidance-documents/postapproval-pregnancy-safety-studies-guidance-industry

Schmidt, M., Schmidt, S. A. J., Sandegaard, J. L., Ehrenstein, V., Pedersen, L., & Sørensen, H. T. (2015). The Danish National Patient Registry: a review of content, data quality, and research potential. Clinical Epidemiology, 7, 449–490. https://doi.org/10.2147/CLEP.S91125

Scientific guidance on post-authorisation efficacy studies - EMA. (n.d.). Retrieved 11 June 2020, from https://www.ema.europa.eu/en/documents/scientific-guideline/scientific-guidance-post-authorisation-efficacy-studies-first-version_en.pdf

Swedish National Patient Register. (n.d.). SITENAME. Retrieved 15 May 2020, from https://www.socialstyrelsen.se/en/statistics-and-data/registers/register-information/the-national-patient-register/

Systematic overview of data sources for drug safety in pregnancy research. (n.d.). Retrieved 15 May 2020, from http://www.encepp.eu/structure/documents/Data_sources_for_medicines_in_pregnancy_research.pdf

UK Renal Registry. (n.d.). UK Renal Registry. Retrieved 19 June 2020, from https://www.renalreg.org/

**4.4 - Spontaneous reports**

Alvarez, Y, Hidalgo, A., Maignen, F., Slattery, J (2010, June 1). Validation of Statistical Signal Detection Procedures in Eudravigilance Post-Authorization Data: A Retrospective Evaluation of the Potential for Earlier Signalling. Drug Safety. https://doi.org/10.2165/11534410-000000000-00000

EudraVigilance. European Medicines Agency. Retrieved 18 May 2020, from https://www.ema.europa.eu/en/human-regulatory/research-development/pharmacovigilance/eudravigilance

Kreimeyer, K., Menschik, D., Winiecki, S., Paul, W., Barash, F., Woo, E. J., Alimchandani, M., Arya, D., Zinderman, C., Forshee, R., & Botsis, T. (2017). Using Probabilistic Record Linkage of Structured and Unstructured Data to Identify Duplicate Cases in Spontaneous Adverse Event Reporting Systems. Drug Safety, 40(7), 571–582. https://doi.org/10.1007/s40264-017-0523-4

Practical Aspects of Signal Detection in Pharmacovigilance: Report of CIOMS Working Group VIII • COUNCIL FOR INTERNATIONAL ORGANIZATIONS OF MEDICAL SCIENCES. (n.d.). COUNCIL FOR INTERNATIONAL ORGANIZATIONS OF MEDICAL SCIENCES. Retrieved 20 May 2020, from https://cioms.ch/publications/product/practical-aspects-of-signal-detection-in-pharmacovigilance-report-of-cioms-working-group-viii/

Tregunno, Pm., Fink, Db., Fernandez-Fernandez, C., Lázaro-Bengoa, E, Norén, Gn. (2014). Performance of Probabilistic Method to Detect Duplicate Individual Case Safety Reports. Drug Safety. https://doi.org/10.1007/s40264-014-0146-y

WHO-UMC - The Global Individual Case Safety Reports Database System (VigiBase). (n.d.). Retrieved 18 May 2020, from https://www.who-umc.org/vigibase/vigibase/

Wisniewski AFZ, Juhlin K, Vestergaard Laursen M, Macia MM, Manlik K, Pinkston VK, Seabroke S, Slattery J (PROTECT project). (n.d.). Characterization of databases (DB) used for signal detection (SD). http://www.imi-protect.eu/documents/WisniewskietalCharacterisationofdatabasesusedorsignaldetectionposterfinalICPE2012.pdf

### 4.5 - Social media


Audeh, B., Bellet, F., Beyens, M.-N., Louët, A. L.-L., & Bousquet, C. (2020). Use of Social Media for Pharmacovigilance Activities: Key Findings and Recommendations from the Vigi4Med Project. Drug Safety, 1–17. https://doi.org/10.1007/s40264-020-00951-2

Banerjee, Ak. et al. (2012). Web-based Patient-Reported Outcomes in Drug Safety and Risk Management: Challenges and Opportunities? Drug Safety. https://doi.org/10.2165/11632390-000000000-00000

Bhattacharya, M., Snyder, S., Malin, M., Truffa, M. M., Marinic, S., Engelmann, R., & Raheja, R. R. (2017). Using Social Media Data in Routine Pharmacovigilance: A Pilot Study to Identify Safety Signals and Patient Perspectives. Pharmaceutical Medicine, 31(3), 167–174. https://doi.org/10.1007/s40290-017-0186-6

Brosch, S. et al. (2019). Establishing a Framework for the Use of Social Media in Pharmacovigilance in Europe. Drug Safety. https://doi.org/10.1007/s40264-019-00811-8

Caster, O., Dietrich, J., Kürzinger, M.-L., Lerch, M., Maskell, S., Norén, G. N., Tcherny-Lessenot, S., Vroman, B., Wisniewski, A., & Stekelenborg, J. van. (2018). Assessment of the Utility of Social Media for Broad-Ranging Statistical Signal Detection in Pharmacovigilance: Results from the WEB-RADR Project. Drug Safety, 41(12), 1355–1369. https://doi.org/10.1007/s40264-018-0699-2

Cocos, A. et al. (2017). Deep Learning for Pharmacovigilance: Recurrent Neural Network Architectures for Labeling Adverse Drug Reactions in Twitter Posts. Journal of the American Medical Informatics Association : JAMIA. https://doi.org/10.1093/jamia/ocw180

Coloma, P. M., Becker, B., Sturkenboom, M. C. J. M., van Mulligen, E. M., & Kors, J. A. (2015). Evaluating Social Media Networks in Medicines Safety Surveillance: Two Case Studies. Drug Safety, 38(10), 921–930. https://doi.org/10.1007/s40264-015-0333-5

EDPS Opinion 1/2015 Mobile Health-Reconciling technological innovation with data protection. (n.d.). Retrieved 5 June 2020, from https://edps.europa.eu/sites/edp/files/publication/15-05-21_mhealth_en_0.pdf

ENISA - European Union Agency for Cybersecurity. (n.d.). [Plone Site]. Retrieved 5 June 2020, from https://www.enisa.europa.eu

ENISA - Smartphone Secure Development Guidelines. (n.d.). [Report/Study]. ENISA - European Union Agency for Cybersecurity. Retrieved 5 June 2020, from https://www.enisa.europa.eu/publications/smartphone-secure-development-guidelines-2016

European Commission's Digital Single Market Glossary. (n.d.). Shaping Europe's Digital Future - European Commission. Retrieved 5 June 2020, from https://ec.europa.eu/digital-single-market/en/glossary

Gattepaille, L. M., Vidlin, S. H., Bergvall, T., Pierce, C. E., & Ellenius, J. (2020). Prospective Evaluation of Adverse Event Recognition Systems in Twitter: Results from the Web-RADR Project. Drug Safety, 1–12. https://doi.org/10.1007/s40264-020-00942-3

Guideline on good pharmacovigilance practices (GVP) Module VI – Collection, management and submission of reports of suspected adverse reactions to medicinal products (Rev 2). European Medicines Agency. https://www.ema.europa.eu/en/human-regulatory/post-authorisation/pharmacovigilance/good-pharmacovigilance-practices

Powell, Ge. et al. (2016). Social Media Listening for Routine Post-Marketing Safety Surveillance. Drug Safety. https://doi.org/10.1007/s40264-015-0385-6

General Data Protection Regulation (GDPR) Compliance Guidelines. (n.d.). GDPR.Eu. Retrieved 5 June 2020, from https://gdpr.eu/

Sloane, R. et al. (2015). Social Media and Pharmacovigilance: A Review of the Opportunities and Challenges. British Journal of Clinical Pharmacology. https://doi.org/10.1111/bcp.12717

Stekelenborg, J. van, Ellenius, J., Maskell, S., Bergvall, T., Caster, O., Dasgupta, N., Dietrich, J., Gama, S., Lewis, D., Newbould, V., Brosch, S., Pierce, C. E., Powell, G., Ptaszyńska-Neophytou, A., Wiśniewski, A. F. Z., Tregunno, P., Norén, G. N., & Pirmohamed, M. (2019). Recommendations for the Use of Social Media in Pharmacovigilance: Lessons from IMI WEB-RADR. Drug Safety, 42(12), 1393–1407. https://doi.org/10.1007/s40264-019-00858-7

**4.6 - Research networks for multi-database studies**

ADVANCE - VAccine monitoring Collaboration for Europe. (n.d.). VAC4EU. Retrieved 20 May 2020, from https://vac4eu.org/

Arrhythmogenic potential of drugs | ARITMO Project | FP7 | CORDIS | European Commission. (n.d.). Retrieved 16 June 2020, from https://cordis.europa.eu/project/id/241679

Becker, Bfh et al. (2017). CodeMapper: Semiautomatic Coding of Case Definitions. A Contribution From the ADVANCE Project. Pharmacoepidemiology and Drug Safety. https://doi.org/10.1002/pds.4245

CNODES - Canadian Network for Observational Drug Effects Studies. (n.d.). Retrieved 20 May 2020, from https://www.cnodes.ca/

EHDEN - European Health Data Evidence Network. (n.d.). Ehden.Eu. Retrieved 20 May 2020, from https://www.ehden.eu/

EMIF. (n.d.). Retrieved 20 May 2020, from http://www.emif.eu/

ENCePP Resources Database. (n.d.). ENCePP. Retrieved 20 May 2020, from http://www.encepp.eu/encepp/resourcesDatabase.jsp

EU-ADR Project | FP7 | CORDIS | European Commission. (n.d.). Retrieved 16 June 2020, from https://cordis.europa.eu/project/id/215847

EUROmediCAT. (n.d.). Retrieved 20 May 2020, from http://www.euromedicat.eu/

Candore, G., Hedenmalm, K., Slattery, J., Cave, A., Kurz, X., Arlett, P. (2020). Can We Rely on Results From IQVIA Medical Research Data UK Converted to the Observational Medical Outcome Partnership Common Data Model?: A Validation Study Based on Prescribing Codeine in Children. Clinical Pharmacology and Therapeutics. https://doi.org/10.1002/cpt.1785

Trifirò, G., Coloma, Pm., Rijnbeek, Pr., Romio, S., Mosseveld, B., Weibel, D., Bonhoeffer, J., Schuemie, M., van der Lis, J, Sturkenboom, M. (2014). Combining Multiple Healthcare Databases for Postmarketing Drug and Vaccine Safety Surveillance: Why and How? Journal of Internal Medicine. https://doi.org/10.1111/joim.12159

Gini, R., Schuemie, M., Brown, J., Ryan, P., Vacchi, E., Coppola, M., Cazzola, W., Coloma, P., Berni, R., Diallo, G., Oliveira, J. L., Avillach, P., Trifirò, G., Rijnbeek, P., Bellentani, M., Lei, J. van D., Klazinga, N., & Sturkenboom, M. (2016). Data Extraction And Management In Networks Of Observational Health Care Databases For Scientific Research: A Comparison Among EU-ADR, OMOP, Mini-Sentinel And

MATRICE Strategies. EGEMs (Generating Evidence & Methods to Improve Patient Outcomes), 4(1). https://doi.org/10.13063/2327-9214.1189

GRIP - Global Research in Paediatrics - CORDIS | European Commission. (n.d.). Retrieved 16 June 2020, from https://cordis.europa.eu/project/id/261060

Improving Consistency and Understanding of Discrepancies of Findings from Pharmacoepidemiological Studies: the IMI PROTECT Project: Pharmacoepidemiology and Drug Safety: Vol 25, No S1. (n.d.). Wiley Online Library. Retrieved 20 May 2020, from https://onlinelibrary.wiley.com/toc/10991557/25/S1

Innovative Medicines Initiative (IMI). (n.d.). IMI Innovative Medicines Initiative. Retrieved 20 May 2020, from http://www.imi.europa.eu/

Sketris, Is et al. (2020). Building a Framework for the Evaluation of Knowledge Translation for the Canadian Network for Observational Drug Effect Studies. Pharmacoepidemiology and Drug Safety. https://doi.org/10.1002/pds.4738

Sultana, J et al. (2019). All-cause Mortality and Antipsychotic Use Among Elderly Persons With High Baseline Cardiovascular and Cerebrovascular Risk: A Multi-Center Retrospective Cohort Study in Italy. Expert Opinion on Drug Metabolism & Toxicology. https://doi.org/10.1080/17425255.2019.1561860

Kieler, H., Artama, M., Engeland, A., Ericsson, Ö., Furu, K., Gissler, M., Nielsen, R. B., Nørgaard, M., Stephansson, O., Valdimarsdottir, U., Zoega, H., & Haglund, B. (2012). Selective serotonin reuptake inhibitors during pregnancy and risk of persistent pulmonary hypertension in the newborn: population based cohort study from the five Nordic countries. BMJ, 344, d8012. https://doi.org/10.1136/bmj.d8012

Klungel Olaf H., Kurz Xavier, Groot Mark C. H., Schlienger Raymond G., Tcherny-Lessenot Stephanie, Grimaldi Lamiae, Ibáñez Luisa, Groenwold Rolf H. H., & Reynolds Robert F. (2016). Multi-centre, multi-database studies with common protocols: lessons learnt from the IMI PROTECT project. Pharmacoepidemiology and Drug Safety, 25(S1), 156–165. https://doi.org/10.1002/pds.3968

Lane, J. C. E., Weaver, J., Kostka, K., Duarte-Salles, T., Abrahao, M. T. F., Alghoul, H., Alser, O., Alshammari, T. M., Biedermann, P., Burn, E., Casajust, P., Conover, M., Culhane, A. C., Davydov, A., DuVall, S. L., Dymshyts, D., Bertolín, S. F., Fišter, K., Hardin, J., … Prieto-Alhambra, D. (2020). Safety of hydroxychloroquine, alone and in combination with azithromycin, in light of rapid wide-spread use for COVID-19: a multinational, network cohort and self-controlled case series study. MedRxiv, 2020.04.08.20054551. https://doi.org/10.1101/2020.04.08.20054551

OHDSI – Observational Health Data Sciences and Informatics. (n.d.). Retrieved 20 May 2020, from https://www.ohdsi.org/

OMOP Common Data Model – OHDSI. (n.d.). Retrieved 20 May 2020, from https://www.ohdsi.org/data-standardization/the-common-data-model/

Overhage, J. M., Ryan, P. B., Reich, C. G., Hartzema, A. G., & Stang, P. E. (2012). Validation of a common data model for active safety surveillance research. Journal of the American Medical Informatics Association : JAMIA, 19(1), 54–60. https://doi.org/10.1136/amiajnl-2011-000376

PROTECT. Retrieved 20 May 2020, from http://www.imi-protect.eu/about.shtml

Flyn, R et al. (2020). Ability of Primary Care Health Databases to Assess Medicinal Products Discussed by the European Union Pharmacovigilance Risk Assessment Committee. Clinical Pharmacology and Therapeutics. https://doi.org/10.1002/cpt.1775

Gini, R. et al. (2020). Different Strategies to Execute Multi-Database Studies for Medicines Surveillance in Real-World Setting: A Reflection on the European Model. Clinical Pharmacology and Therapeutics. https://doi.org/10.1002/cpt.1833

Izem, R. et al. (2020). Quantifying How Small Variations in Design Elements Affect Risk in an Incident Cohort Study in Claims. Pharmacoepidemiology and Drug Safety. https://doi.org/10.1002/pds.4892

Research, C. for D. E. and. (n.d.). FDA's Sentinel Initiative. FDA. Retrieved 20 May 2020, from https://www.fda.gov/safety/fdas-sentinel-initiative

Roberto, G., Leal, I., Sattar, N., Loomis, A. K., Avillach, P., Egger, P., Wijngaarden, R. van, Ansell, D., Reisberg, S., Tammesoo, M.-L., Alavere, H., Pasqua, A., Pedersen, L., Cunningham, J., Tramontan, L., Mayer, M. A., Herings, R., Coloma, P., Lapi, F., … Gini, R. (2016). Identifying Cases of Type 2 Diabetes in Heterogeneous Data Sources: Strategy from the EMIF Project. PLOS ONE, 11(8), e0160648. https://doi.org/10.1371/journal.pone.0160648

Safety Evaluation of Adverse Reactions in Diabetes | SAFEGUARD Project | FP7 | CORDIS | European Commission. (n.d.). Retrieved 16 June 2020, from https://cordis.europa.eu/project/id/282521

Slattery, J., Kurz, X. (2020). Assessing strength of evidence for regulatory decision making in licensing: What proof do we need for observational studies of effectiveness? Pharmacoepidemiology and Drug Safety, n/a(n/a). https://doi.org/10.1002/pds.5005

SOS NSAIDS PROJECT. (n.d.). Retrieved 26 May 2020, from https://www.sos-nsaids-project.org/

The U.S. Food and Drug Administration's Mini-Sentinel Program: Pharmacoepidemiology and Drug Safety: Vol 21, No S1. (n.d.). Wiley Online Library. Retrieved 20 May 2020, from https://onlinelibrary.wiley.com/toc/10991557/21/S1

Vaccine Safety Datalink (VSD). (n.d.). Retrieved 20 May 2020, from https://www.cdc.gov/vaccinesafety/ensuringsafety/monitoring/vsd/index.html

Xu, Y., Zhou, X., Suehs, B. T., Hartzema, A. G., Kahn, M. G., Moride, Y., Sauer, B. C., Liu, Q., Moll, K., Pasquale, M. K., Nair, V. P., & Bate, A. (2015). A Comparative Assessment of Observational Medical Outcomes Partnership and Mini-Sentinel Common Data Models and Analytics: Implications for Active Drug Safety Surveillance. Drug Safety, 38(8), 749–765. https://doi.org/10.1007/s40264-015-0297-5

## Chapter 5 - Study design and methods

### 5.1 - Definition and validation of drug exposure, outcomes and covariates

A review of uses of health care utilization databases for epidemiologic research on therapeutics. (2005). Journal of Clinical Epidemiology, 58(4), 323–337. https://doi.org/10.1016/j.jclinepi.2004.10.012

B. Strom, S.E. Kimmel, S. Hennessy. (2012). Pharmacoepidemiology (5th ed.). Wiley.

Brenner, H., & Gefeller, O. (1993). Use of the Positive Predictive Value to Correct for Disease Misclassification in Epidemiologic Studies. American Journal of Epidemiology, 138(11), 1007–1015. https://doi.org/10.1093/oxfordjournals.aje.a116805

Brighton Collaboration case definitions. (n.d.). Retrieved 19 June 2020, from https://www.brightoncollaboration.org/case-definitions

Brunelli, S. M., Gagne, J. J., Huybrechts, K. F., Wang, S. V., Patrick, A. R., Rothman, K. J., & Seeger, J. D. (2013). Estimation using all available covariate information versus a fixed look-back window for

dichotomous covariates. Pharmacoepidemiology and Drug Safety, 22(5), 542–550. https://doi.org/10.1002/pds.3434

Carnahan, R. M., & Moores, K. G. (2012). Mini-Sentinel's systematic reviews of validated methods for identifying health outcomes using administrative and claims data: methods and lessons learned. Pharmacoepidemiology and Drug Safety, 21(S1), 82–89. https://doi.org/10.1002/pds.2321

ClinicalCodes Repository. (n.d.). Retrieved 19 June 2020, from https://clinicalcodes.rss.mhs.man.ac.uk/

Davé, S., Petersen, I. (2009). Creating Medical and Drug Code Lists to Identify Cases in Primary Care Databases. Pharmacoepidemiology and Drug Safety. https://doi.org/10.1002/pds.1770

Greenland, S. (1996). Basic Methods for Sensitivity Analysis of Biases. International Journal of Epidemiology, 25(6), 1107–1116. https://doi.org/10.1093/ije/25.6.1107-a

Lash T, Fox MP, Fink AK. (2009). Applying quantitative bias analysis to epidemiologic data. Springer-Verlag.

Lee, Wj., Lee, Ta., Pickard, As, Shoaibi, A, Schumock, Gt. (2015, August). Using Linked Electronic Data to Validate Algorithms for Health Outcomes in Administrative Databases. Journal of Comparative Effectiveness Research. https://doi.org/10.2217/cer.15.14

M. Elseviers, B. Wettermark, A.B. Almarsdottir. (2016). Drug Utilization Research. Methods and Applications. Wiley Blackwell.

Khan, Nf., Harrison, Se., Rose, Pw. (2010). Validity of Diagnostic Coding Within the General Practice Research Database: A Systematic Review. The British Journal of General Practice : The Journal of the Royal College of General Practitioners. https://doi.org/10.3399/bjgp10X483562

Pottegard, A., Christensen, Rd, Houji, A., Christiansen, Cb, Paulsen, Ma, Thomsen, JI, Hallas, J. (2014). Primary Non-Adherence in General Practice: A Danish Register Study. European Journal of Clinical Pharmacology. https://doi.org/10.1007/s00228-014-1677-y

Quantitative Bias Analysis Methodology Development: Sequential Bias Adjustment for Outcome Misclassification. (n.d.). Retrieved 19 June 2020, from https://www.sentinelinitiative.org/sites/default/files/Methods/Sentinel_Methods_Sequential_bias.pdf

**5.2 - Bias (systematic error)**

Fewell, Z., Davey Smith, G., & Sterne, J. A. C. (2007). The Impact of Residual and Unmeasured Confounding in Epidemiologic Studies: A Simulation Study. American Journal of Epidemiology, 166(6), 646–655. https://doi.org/10.1093/aje/kwm165

Grobbee, D. E., & Hoes, A. W. (1997). Confounding and indication for treatment in evaluation of drug treatment for hypertension. BMJ, 315(7116), 1151–1154. https://doi.org/10.1136/bmj.315.7116.1151

Horwitz, R. I., & Feinstein, A. R. (1978). Alternative Analytic Methods for Case-Control Studies of Estrogens and Endometrial Cancer. New England Journal of Medicine, 299(20), 1089–1094. https://doi.org/10.1056/NEJM197811162992001

Joffe, M. M. (2000). Confounding by indication: the case of calcium channel blockers. Pharmacoepidemiology and Drug Safety, 9(1), 37–41. https://doi.org/10.1002/(SICI)1099-1557(200001/02)9:1<37::AID-PDS471>3.0.CO;2-U

Rothman, K., Greenland, S., Lash, T. (2008). Modern Epidemiology. Lippincott Williams & Wilkins.

Kiri, V. A., & MacKenzie, G. (2009). RE: "IMMORTAL TIME BIAS IN PHARMACOEPIDEMIOLOGY". American Journal of Epidemiology, 170(5), 667–668. https://doi.org/10.1093/aje/kwp239

Klungel, O. et al. Methods to assess intended effects of drug treatment in observational studies are reviewed. (2004). Journal of Clinical Epidemiology, 57(12), 1223–1231. https://doi.org/10.1016/j.jclinepi.2004.03.011

Lévesque, L. E., Hanley, J. A., Kezouh, A., & Suissa, S. (2010). Problem of immortal time bias in cohort studies: example using statins for preventing progression of diabetes. BMJ, 340. https://doi.org/10.1136/bmj.b5087

Ma, L. (1999). The Transnational Study on Oral Contraceptives and the Health of Young Women. Methods, Results, New Analyses and the Healthy User Effect. Human Reproduction Update. https://doi.org/10.1093/humupd/5.6.707

McMahon, A. D. (2003). Approaches to combat with confounding by indication in observational studies of intended drug effects. Pharmacoepidemiology and Drug Safety, 12(7), 551–558. https://doi.org/10.1002/pds.883

Moodie, E. E. M., & Stephens, D. A. (2010). Using Directed Acyclic Graphs to detect limitations of traditional regression in longitudinal studies. International Journal of Public Health, 55(6), 701–703. https://doi.org/10.1007/s00038-010-0184-x

Salas, M., Hotman, A., & Stricker, B. H. (1999). Confounding by Indication: An Example of Variation in the Use of Epidemiologic Terminology. American Journal of Epidemiology, 149(11), 981–983. https://doi.org/10.1093/oxfordjournals.aje.a009758

Schneeweiss, S. (2006). Sensitivity analysis and external adjustment for unmeasured confounders in epidemiologic database studies of therapeutics. Pharmacoepidemiology and Drug Safety, 15(5), 291–303. https://doi.org/10.1002/pds.1200

Sensitivity Analysis. (n.d.). Retrieved 23 June 2020, from https://www.drugepi.org/dope/software#Sensitivity

Strom, BL., Kimmel, SE., & Hennessy, S. (2012). Pharmacoepidemiology (5th ed.). Wiley.

Suissa, S. (2007). Immortal time bias in observational studies of drug effects. Pharmacoepidemiology and Drug Safety, 16(3), 241–249. https://doi.org/10.1002/pds.1357

Suissa, S. (2008). Immortal Time Bias in Pharmacoepidemiology. American Journal of Epidemiology, 167(4), 492–499. https://doi.org/10.1093/aje/kwm324

Suissa, S. (2008). Immeasurable Time Bias in Observational Studies of Drug Effects on Mortality. American Journal of Epidemiology, 168(3), 329–335. https://doi.org/10.1093/aje/kwn135

Suissa, S., & Azoulay, L. (2012). Metformin and the Risk of Cancer: Time-related biases in observational studies. Diabetes Care, 35(12), 2665–2673. https://doi.org/10.2337/dc12-0788

Surveillance Bias in Outcomes Reporting | Venous Thromboembolism | JAMA | JAMA Network. (n.d.). Retrieved 23 June 2020, from https://jamanetwork.com/journals/jama/fullarticle/900883

Gerhard, T. (2008). Bias: Considerations for Research Practice. American Journal of Health-System Pharmacy : AJHP : Official Journal of the American Society of Health-System Pharmacists. https://doi.org/10.2146/ajhp070369

Zhou, Z., Rahme, E., Abrahamowicz, M., & Pilote, L. (2005). Survival Bias Associated with Time-to-Treatment Initiation in Drug Effectiveness Evaluation: A Comparison of Methods. American Journal of Epidemiology, 162(10), 1016–1023. https://doi.org/10.1093/aje/kwi307

**5.3 - Methods to address bias**

Ali, Ms. et al. (2014). Propensity Score Balance Measures in Pharmacoepidemiology: A Simulation Study. Pharmacoepidemiology and Drug Safety. https://doi.org/10.1002/pds.3574

Ali, Ms. et al. (2015). Reporting of Covariate Selection and Balance Assessment in Propensity Score Analysis Is Suboptimal: A Systematic Review. Journal of Clinical Epidemiology. https://doi.org/10.1016/j.jclinepi.2014.08.011

Arbogast, P. G., & Ray, W. A. (2009). Use of disease risk scores in pharmacoepidemiologic studies. Statistical Methods in Medical Research, 18(1), 67–80. https://doi.org/10.1177/0962280208092347

Baiocchi, M. et al. (2014). Instrumental Variable Methods for Causal Inference. Statistics in Medicine. https://doi.org/10.1002/sim.6128

Belitser, Sv. et al.(2011). Measuring Balance and Model Selection in Propensity Score Methods. Pharmacoepidemiology and Drug Safety. https://doi.org/10.1002/pds.2188

Bernal, J. L., Cummins, S., & Gasparrini, A. (2017). Interrupted time series regression for the evaluation of public health interventions: a tutorial. International Journal of Epidemiology, 46(1), 348–355. https://doi.org/10.1093/ije/dyw098

Brenner, H., & Gefeller, O. (1993). Use of the Positive Predictive Value to Correct for Disease Misclassification in Epidemiologic Studies. American Journal of Epidemiology, 138(11), 1007–1015. https://doi.org/10.1093/oxfordjournals.aje.a116805

Brookhart, M. A., Rassen, J. A., & Schneeweiss, S. (2010). Instrumental variable methods in comparative safety and effectiveness research. Pharmacoepidemiology and Drug Safety, 19(6), 537–554. https://doi.org/10.1002/pds.1908

Brookhart, M. A., Wang, P., Solomon, D. H., & Schneeweiss, S. (2006). Evaluating short-term drug effects using a physician-specific prescribing preference as an instrumental variable. Epidemiology (Cambridge, Mass.), 17(3), 268–275. https://doi.org/10.1097/01.ede.0000193606.58671.c5

Cole, Sr. et al. (2003). Effect of Highly Active Antiretroviral Therapy on Time to Acquired Immunodeficiency Syndrome or Death Using Marginal Structural Models. American Journal of Epidemiology. https://doi.org/10.1093/aje/kwg206

Daniel, Rm. et al. (2013). Methods for Dealing With Time-Dependent Confounding. Statistics in Medicine. https://doi.org/10.1002/sim.5686

Dormuth, Cr. et al. (2009). Statin Adherence and Risk of Accidents: A Cautionary Tale. Circulation. https://doi.org/10.1161/CIRCULATIONAHA.108.824151

Dusetzina, Sb. et al. (2015). Control Outcomes and Exposures for Improving Internal Validity of Nonrandomized Studies. Health Services Research. https://doi.org/10.1111/1475-6773.12279

Farrington, Cp. et al. (2009). Case Series Analysis for Censored, Perturbed, or Curtailed Post-Event Exposures. Biostatistics (Oxford, England). https://doi.org/10.1093/biostatistics/kxn013

Franklin, Jm. et al. (2014). Metrics for Covariate Balance in Cohort Studies of Causal Effects. Statistics in Medicine. https://doi.org/10.1002/sim.6058

Garabedian, Lf. et al. (2014). Potential Bias of Instrumental Variable Analyses for Observational Comparative Effectiveness Research. Annals of Internal Medicine. https://doi.org/10.7326/M13-1887

Glynn, Rj. et al. (2012). Role of Disease Risk Scores in Comparative Effectiveness Research With Emerging Therapies. Pharmacoepidemiology and Drug Safety. https://doi.org/10.1002/pds.3231

Greenland, S. (1996a). Confounding and Exposure Trends in Case-Crossover and Case-Time-Control Designs. Epidemiology (Cambridge, Mass.). https://doi.org/10.1097/00001648-199605000-00003

Greenland, S. (1996b). Basic Methods for Sensitivity Analysis of Biases. International Journal of Epidemiology. https://pubmed.ncbi.nlm.nih.gov/9027513/

Greenland, S. (2000). An introduction to instrumental variables for epidemiologists. International Journal of Epidemiology, 29(4), 722–729. https://doi.org/10.1093/ije/29.4.722

Groenwold, Rh. et al. (2011). Balance Measures for Propensity Score Methods: A Clinical Example on Beta-Agonist Use and the Risk of Myocardial Infarction. Pharmacoepidemiology and Drug Safety. https://doi.org/10.1002/pds.2251

Hallas, J. (1996). Evidence of Depression Provoked by Cardiovascular Medication: A Prescription Sequence Symmetry Analysis. Epidemiology (Cambridge, Mass.). https://pubmed.ncbi.nlm.nih.gov/8862977/

Hallas, J. et al. (2014). Use of Self-Controlled Designs in Pharmacoepidemiology. Journal of Internal Medicine. https://doi.org/10.1111/joim.12186

Hallas, J. et al. (2016). Persistent User Bias in Case-Crossover Studies in Pharmacoepidemiology. American Journal of Epidemiology. https://doi.org/10.1093/aje/kww079

Hauben, M. et al. (2016). Evidence of Misclassification of Drug-Event Associations Classified as Gold Standard 'Negative Controls' by the Observational Medical Outcomes Partnership (OMOP). Drug Safety. https://doi.org/10.1007/s40264-016-0392-2

Hernán, M. A. (2018). The C-Word: Scientific Euphemisms Do Not Improve Causal Inference From Observational Data. American Journal of Public Health, 108(5), 616–619. https://doi.org/10.2105/AJPH.2018.304337

Hernán, Ma. et al. (2000). Marginal Structural Models to Estimate the Causal Effect of Zidovudine on the Survival of HIV-positive Men. Epidemiology (Cambridge, Mass.). https://doi.org/10.1097/00001648-200009000-00012

Hernán, Ma. et al. (2016). Specifying a Target Trial Prevents Immortal Time Bias and Other Self-Inflicted Injuries in Observational Analyses. Journal of Clinical Epidemiology. https://doi.org/10.1016/j.jclinepi.2016.04.014

Hernán, Ma. et al. (2016). Using Big Data to Emulate a Target Trial When a Randomized Trial Is Not Available. American Journal of Epidemiology. https://doi.org/10.1093/aje/kwv254

Hernán, Ma. et al. (2018). How to Estimate the Effect of Treatment Duration on Survival Outcomes Using Observational Data. BMJ (Clinical Research Ed.). https://doi.org/10.1136/bmj.k182

Joffe, Mm. et al. (1999). Invited Commentary: Propensity Scores. American Journal of Epidemiology. https://doi.org/10.1093/oxfordjournals.aje.a010011

Jurek, Am. (2005). Proper Interpretation of Non-Differential Misclassification Effects: Expectations vs Observations. International Journal of Epidemiology. https://doi.org/10.1093/ije/dyi060

Kontopantelis, E., Doran, T., Springate, D. A., Buchan, I., & Reeves, D. (2015). Regression based quasi-experimental approach when randomisation is not an option: interrupted time series analysis. The BMJ, 350. https://doi.org/10.1136/bmj.h2750

Lash, Tl. et al. (2014). Good Practices for Quantitative Bias Analysis. International Journal of Epidemiology. https://doi.org/10.1093/ije/dyu149

Lawlor, Da. et al. (2016). Triangulation in Aetiological Epidemiology. International Journal of Epidemiology. https://doi.org/10.1093/ije/dyw314

Lund, Jl. et al. (2015). The Active Comparator, New User Study Design in Pharmacoepidemiology: Historical Foundations and Contemporary Application. Current Epidemiology Reports. https://doi.org/10.1007/s40471-015-0053-5

Maclure, M. et al. (1991). The Case-Crossover Design: A Method for Studying Transient Effects on the Risk of Acute Events. American Journal of Epidemiology. https://doi.org/10.1093/oxfordjournals.aje.a115853

Maclure, Ma. et al. (2012). When Should Case-Only Designs Be Used for Safety Monitoring of Medical Products? Pharmacoepidemiology and Drug Safety. https://doi.org/10.1002/pds.2330

Martens, Ep. et al. (2006). Instrumental Variables: Application and Limitations. Epidemiology (Cambridge, Mass.). https://doi.org/10.1097/01.ede.0000215160.88317.cb

Miettinen, O. S. (1976). Stratification by a multivariate confounder score. American Journal of Epidemiology, 104(6), 609–620. https://doi.org/10.1093/oxfordjournals.aje.a112339

Miettinen, Os. et al. (1989). The Clinical Trial as a Paradigm for Epidemiologic Research. Journal of Clinical Epidemiology. https://doi.org/10.1016/0895-4356(89)90143-1

Nicholas, Jm. et al. (2012). Within-person Study Designs Had Lower Precision and Greater Susceptibility to Bias Because of Trends in Exposure Than Cohort and Nested Case-Control Designs. Journal of Clinical Epidemiology. https://doi.org/10.1016/j.jclinepi.2011.09.004

Outcome Misclassification. (n.d.). Retrieved 3 July 2020, from https://apps.p-95.com/ISPE/

Rassen, J. A., & Schneeweiss, S. (2012a). Using high-dimensional propensity scores to automate confounding control in a distributed medical product safety surveillance system. Pharmacoepidemiology and Drug Safety, 21(S1), 41–49. https://doi.org/10.1002/pds.2328

Rassen, J. A., Glynn, R. J., Brookhart, M. A., & Schneeweiss, S. (2011). Covariate Selection in High-Dimensional Propensity Score Analyses of Treatment Effects in Small Samples. American Journal of Epidemiology, 173(12), 1404–1413. https://doi.org/10.1093/aje/kwr001

Rassen, Ja. et al. (2012). Newly Marketed Medications Present Unique Challenges for Nonrandomized Comparative Effectiveness Analyses. Journal of Comparative Effectiveness Research. https://doi.org/10.2217/cer.12.12

Rassen, Ja. et al. (2012). One-to-many Propensity Score Matching in Cohort Studies. Pharmacoepidemiology and Drug Safety. https://doi.org/10.1002/pds.3263

Rassen, Ja. et al. (2013). Matching by Propensity Score in Cohort Studies With Three Treatment Groups. Epidemiology (Cambridge, Mass.). https://doi.org/10.1097/EDE.0b013e318289dedf

Ray, W. A. (2003). Evaluating Medication Effects Outside of Clinical Trials: New-User Designs. American Journal of Epidemiology, 158(9), 915–920. https://doi.org/10.1093/aje/kwg231

Robins, Jm. et al. (2000). Marginal Structural Models and Causal Inference in Epidemiology. Epidemiology (Cambridge, Mass.). https://doi.org/10.1097/00001648-200009000-00011

Sanni Ali, M. et al. (2013). Time-dependent Propensity Score and Collider-Stratification Bias: An Example of beta2-agonist Use and the Risk of Coronary Heart Disease. European Journal of Epidemiology. https://doi.org/10.1007/s10654-013-9766-2

Schmidt, Af. Et al. (2016). Adjusting for Confounding in Early Postlaunch Settings: Going Beyond Logistic Regression Models. Epidemiology (Cambridge, Mass.). https://doi.org/10.1097/EDE.0000000000000388

Schneeweiss, S. (2018). Automated data-adaptive analytics for electronic healthcare data to study causal treatment effects. Clinical Epidemiology, 10, 771–788. https://doi.org/10.2147/CLEP.S166545

Schneeweiss, S. et al. (2009). High-dimensional Propensity Score Adjustment in Studies of Treatment Effects Using Health Care Claims Data. Epidemiology (Cambridge, Mass.). https://doi.org/10.1097/EDE.0b013e3181a663cc

Schneeweiss, S. et al. (2011). Assessing the Comparative Effectiveness of Newly Marketed Medications: Methodological Challenges and Implications for Drug Development. Clinical Pharmacology and Therapeutics. https://doi.org/10.1038/clpt.2011.235

Schuemie, Mj. et al. (2014, January 30). Interpreting Observational Studies: Why Empirical Calibration Is Needed to Correct P-Values. Statistics in Medicine. https://doi.org/10.1002/sim.5925

Schuemie, Mj. et al. (2016). Robust Empirical Calibration of P-Values Using Observational Data. Statistics in Medicine. https://doi.org/10.1002/sim.6977

Schuemie, Mj. et al. (2018). Empirical Confidence Interval Calibration for Population-Level Effect Estimation Studies in Observational Healthcare Data. Proceedings of the National Academy of Sciences of the United States of America. https://doi.org/10.1073/pnas.1708282114

Shrank, Wh. et al. (2011). Healthy User and Related Biases in Observational Studies of Preventive Interventions: A Primer for Physicians. Journal of General Internal Medicine. https://doi.org/10.1007/s11606-010-1609-1

Stürmer, T. et al. (2007). Performance of Propensity Score Calibration--A Simulation Study. American Journal of Epidemiology. https://doi.org/10.1093/aje/kwm074

Stürmer, T. et al. (2010). Treatment Effects in the Presence of Unmeasured Confounding: Dealing With Observations in the Tails of the Propensity Score Distribution--A Simulation Study. American Journal of Epidemiology. https://doi.org/10.1093/aje/kwq198

Suchard, Ma. et al. (2013). Empirical Performance of the Self-Controlled Case Series Design: Lessons for Developing a Risk Identification and Analysis System. Drug Safety. https://doi.org/10.1007/s40264-013-0100-4

Suissa, S. (1995). The Case-Time-Control Design. Epidemiology (Cambridge, Mass.). https://doi.org/https://www.ncbi.nlm.nih.gov/pubmed/7619931

Suissa, S. et al. (2017). Prevalent New-User Cohort Designs for Comparative Drug Effect Studies by Time-Conditional Propensity Scores. Pharmacoepidemiology and Drug Safety. https://doi.org/10.1002/pds.4107

Sujan, Ac. et al. (2017). Associations of Maternal Antidepressant Use During the First Trimester of Pregnancy With Preterm Birth, Small for Gestational Age, Autism Spectrum Disorder, and Attention-Deficit/Hyperactivity Disorder in Offspring. JAMA. https://doi.org/10.1001/jama.2017.3413

Swanson, S. A., Robins, J. M., Miller, M., & Hernán, M. A. (2015). Selecting on Treatment: A Pervasive Form of Bias in Instrumental Variable Analyses. American Journal of Epidemiology, 181(3), 191–197. https://doi.org/10.1093/aje/kwu284

Swanson, Sa. et al. (2013). Commentary: How to Report Instrumental Variable Analyses (Suggestions Welcome). Epidemiology (Cambridge, Mass.). https://doi.org/10.1097/EDE.0b013e31828d0590

Tadrous, M., Gagne, J. J., Stürmer, T., & Cadarette, S. M. (2013). Disease Risk Score (DRS) as a Confounder Summary Method: Systematic Review and Recommendations. Pharmacoepidemiology and Drug Safety, 22(2), 122–129. https://doi.org/10.1002/pds.3377

Tannen, R. L., Weiner, M. G., & Xie, D. (2008). Replicated studies of two randomized trials of angiotensin- converting enzyme inhibitors: further empiric validation of the 'prior event rate ratio' to adjust for unmeasured confounding by indication. Pharmacoepidemiology and Drug Safety, 17(7), 671–685. https://doi.org/10.1002/pds.1584

Toh, S. et al. (2011). Confounding Adjustment via a Semi-Automated High-Dimensional Propensity Score Algorithm: An Application to Electronic Medical Records. Pharmacoepidemiology and Drug Safety. https://doi.org/10.1002/pds.2152

Uddin, M. J., Groenwold, R. H. H., de Boer, A., Afonso, A. S. M., Primatesta, P., Becker, C., Belitser, S. V., Hoes, A. W., Roes, K. C. B., & Klungel, O. H. (2016). Evaluating different physician's prescribing preference based instrumental variables in two primary care databases: a study of inhaled long-acting beta2-agonist use and the risk of myocardial infarction. Pharmacoepidemiology and Drug Safety, 25(S1), 132–141. https://doi.org/10.1002/pds.3860

Uddin, M. J., Groenwold, R. H. H., van Staa, T. P., de Boer, A., Belitser, S. V., Hoes, A. W., Roes, K. C. B., & Klungel, O. H. (2015). Performance of prior event rate ratio adjustment method in pharmacoepidemiology: a simulation study. Pharmacoepidemiology and Drug Safety, 24(5), 468–477. https://doi.org/10.1002/pds.3724

Uddin, Mj. et al. (2014). Performance of Instrumental Variable Methods in Cohort and Nested Case-Control Studies: A Simulation Study. Pharmacoepidemiology and Drug Safety. https://doi.org/10.1002/pds.3555

Wang, S. et al. (2011). Future Cases as Present Controls to Adjust for Exposure Trend Bias in Case-Only Studies. Epidemiology (Cambridge, Mass.). https://doi.org/10.1097/EDE.0b013e31821d09cd

Weldeselassie, Yg. et al. (2011). Use of the Self-Controlled Case-Series Method in Vaccine Safety Studies: Review and Recommendations for Best Practice. Epidemiology and Infection. https://doi.org/10.1017/S0950268811001531

Whitaker, Hj. et al. (2006). Tutorial in Biostatistics: The Self-Controlled Case Series Method. Statistics in Medicine. https://doi.org/10.1002/sim.2302

Williamson, E. et al. (2012). Propensity Scores: From Naive Enthusiasm to Intuitive Understanding. Statistical Methods in Medical Research. https://doi.org/10.1177/0962280210394483

Witteman, Jc. et al. (1998). G-estimation of Causal Effects: Isolated Systolic Hypertension and Cardiovascular Death in the Framingham Heart Study. American Journal of Epidemiology. https://doi.org/10.1093/oxfordjournals.aje.a009658

Xinyao, Ji. et al. (2017). The Trend-in-trend Research Design for Causal Inference. Epidemiology (Cambridge, Mass.), 28(4), 529–536. https://doi.org/10.1097/EDE.0000000000000579

Xu, S. et al. (2011). Identifying Optimal Risk Windows for Self-Controlled Case Series Studies of Vaccine Safety. Statistics in Medicine. https://doi.org/10.1002/sim.4125

### 5.4 - Effect measure modification and interaction

Altman, D. G., & Bland, J. M. (2003). Interaction revisited: the difference between two estimates. BMJ (Clinical Research Ed.), 326(7382), 219. https://doi.org/10.1136/bmj.326.7382.219

Knol, M. J., & VanderWeele, T. J. (2012). Recommendations for presenting analyses of effect modification and interaction. International Journal of Epidemiology, 41(2), 514–520. https://doi.org/10.1093/ije/dyr218

Knol, M. J., VanderWeele, T. J., Groenwold, R. H. H., Klungel, O. H., Rovers, M. M., & Grobbee, D. E. (2011). Estimating measures of interaction on an additive scale for preventive exposures. European Journal of Epidemiology, 26(6), 433–438.

Moher, D., Hopewell, S., Schulz, K. F., Montori, V., Gøtzsche, P. C., Devereaux, P. J., Elbourne, D., Egger, M., Altman, D. G., & Consolidated Standards of Reporting Trials Group. (2010). CONSORT 2010 Explanation and Elaboration: Updated guidelines for reporting parallel group randomised trials. Journal of Clinical Epidemiology, 63(8), e1-37. https://doi.org/10.1016/j.jclinepi.2010.03.004

Moride, Y., Abenhaim, L., Yola, M., & Lucein, A. (1994). Evidence of the depletion of susceptibles effect in non-experimental pharmacoepidemiologic research. Journal of Clinical Epidemiology, 47(7), 731–737. https://doi.org/10.1016/0895-4356(94)90170-8

Schmidt, A. et al. Exploring interaction effects in small samples increases rates of false-positive and false-negative findings: results from a systematic review and simulation study. (2014). Journal of Clinical Epidemiology, 67(7), 821–829. https://doi.org/10.1016/j.jclinepi.2014.02.008

Vandenbroucke, J. P., von Elm, E., Altman, D. G., Gøtzsche, P. C., Mulrow, C. D., Pocock, S. J., Poole, C., Schlesselman, J. J., Egger, M., & STROBE Initiative. (2007). Strengthening the Reporting of Observational Studies in Epidemiology (STROBE): explanation and elaboration. Epidemiology (Cambridge, Mass.), 18(6), 805–835. https://doi.org/10.1097/EDE.0b013e3181577511

VanderWeele, T. J. (2009). On the distinction between interaction and effect modification. Epidemiology (Cambridge, Mass.), 20(6), 863–871. https://doi.org/10.1097/EDE.0b013e3181ba333c

### 5.5 - Ecological analyses and case-population studies

Farrington, Cp. (2004). Control Without Separate Controls: Evaluation of Vaccine Safety Using Case-Only Methods. Vaccine. https://doi.org/10.1016/j.vaccine.2004.01.017

Gulmez, S. E., Larrey, D., Pageaux, G.-P., Lignot, S., Lassalle, R., Jové, J., Gatta, A., McCormick, P. A., Metselaar, H. J., Monteiro, E., Thorburn, D., Bernal, W., Zouboulis-Vafiadis, I., Vries, C. de, Perez-Gutthann, S., Sturkenboom, M., Bénichou, J., Montastruc, J.-L., Horsmans, Y., … Moore, N. (2013). Transplantation for Acute Liver Failure in Patients Exposed to NSAIDs or Paracetamol (Acetaminophen). Drug Safety, 36(2), 135–144. https://doi.org/10.1007/s40264-012-0013-7

Moore, N., Gulmez, S. E., Larrey, D., Pageaux, G.-P., Lignot, S., Lassalle, R., Jové, J., Pariente, A., Blin, P., Bénichou, J., & Bégaud, B. (2013). Choice of the denominator in case population studies: event rates for registration for liver transplantation after exposure to NSAIDs in the SALT study in France. Pharmacoepidemiology and Drug Safety, 22(2), 160–167. https://doi.org/10.1002/pds.3371

Théophile, H., Laporte, J.-R., Moore, N., Martin, K.-L., & Bégaud, B. (2011). The Case-Population Study Design. Drug Safety, 34(10), 861–868. https://doi.org/10.2165/11592140-000000000-00000

### 5.6 - Pragmatic trials and large simple trials

Collier, S., Harvey, C., Brewster, J., Bakerly, N. D., Elkhenini, H. F., Stanciu, R., Williams, C., Brereton, J., New, J. P., McCrae, J., McCorkindale, S., & Leather, D. (2016). Monitoring safety in a phase III real-world effectiveness trial: use of novel methodology in the Salford Lung Study. Pharmacoepidemiology and Drug Safety, 26(3), 344–352. https://doi.org/10.1002/pds.4118

Fröbert, O., Lagerqvist, B., Olivecrona, G. K., Omerovic, E., Gudnason, T., Maeng, M., Aasa, M., Angerås, O., Calais, F., Danielewicz, M., Erlinge, D., Hellsten, L., Jensen, U., Johansson, A. C., Kåregren, A., Nilsson, J., Robertson, L., Sandhall, L., Sjögren, I., … James, S. K. (2013). Thrombus Aspiration during ST-Segment Elevation Myocardial Infarction. The New England Journal of Medicine. https://doi.org/10.1056/NEJMoa1308789

IMI GetReal Project - Glossary of Definitions of Common Terms. (n.d.). Retrieved 18 May 2020, from https://www.imi-getreal.eu/Portals/1/Documents/01%20deliverables/D1.3%20-%20Revised%20GetReal%20glossary%20-%20FINAL%20updated%20version_25Oct16_webversion.pdf

Lauer, M. S., & D'Agostino, R. B. S. (2013). The Randomized Registry Trial — The Next Disruptive Technology in Clinical Research? The New England Journal of Medicine. https://doi.org/10.1056/NEJMp1310102

Loudon, K., Treweek, S., Sullivan, F., Donnan, P., Thorpe, K. E., & Zwarenstein, M. (2015). The PRECIS-2 tool: designing trials that are fit for purpose. BMJ, 350. https://doi.org/10.1136/bmj.h2147

Collins, R. et al. (2020). The Magic of Randomization Versus the Myth of Real-World Evidence. The New England Journal of Medicine. https://doi.org/10.1056/NEJMsb1901642

Lesko, Sm. et al. (1995). An Assessment of the Safety of Pediatric Ibuprofen. A Practitioner-Based Randomized Clinical Trial. JAMA. https://pubmed.ncbi.nlm.nih.gov/7884951/

Rao, S. et al. A Registry-Based Randomized Trial Comparing Radial and Femoral Approaches in Women Undergoing Percutaneous Coronary Intervention: The SAFE-PCI for Women (Study of Access Site for Enhancement of PCI for Women) Trial. (2014). JACC: Cardiovascular Interventions, 7(8), 857–867. https://doi.org/10.1016/j.jcin.2014.04.007

Staa, T.-P. van, Goldacre, B., Gulliford, M., Cassell, J., Pirmohamed, M., Taweel, A., Delaney, B., & Smeeth, L. (2012). Pragmatic randomised trials using routine electronic health records: putting them to the test. BMJ, 344. https://doi.org/10.1136/bmj.e55

Strom, B. L., Eng, S. M., Faich, G., Reynolds, R. F., D'Agostino, R. B., Ruskin, J., & Kane, J. M. (2011). Comparative Mortality Associated With Ziprasidone and Olanzapine in Real-World Use Among 18,154 Patients With Schizophrenia: The Ziprasidone Observational Study of Cardiac Outcomes (ZODIAC). American Journal of Psychiatry, 168(2), 193–201. https://doi.org/10.1176/appi.ajp.2010.08040484

Thorpe, K. E., Zwarenstein, M., Oxman, A. D., Treweek, S., Furberg, C. D., Altman, D. G., Tunis, S., Bergel, E., Harvey, I., Magid, D. J., & Chalkidou, K. (2009). A pragmatic–explanatory continuum indicator summary (PRECIS): a tool to help trial designers. CMAJ, 180(10), E47–E57. https://doi.org/10.1503/cmaj.090523

Van Staa, T-P. (2014). The opportunities and challenges of pragmatic point-of-care randomised trials using routinely collected electronic records: evaluations of two exemplar trials. Health Technology Assessment, 18(43). https://doi.org/10.3310/hta18430

Zwarenstein, M., Treweek, S., Gagnier, J. J., Altman, D. G., Tunis, S., Haynes, B., Oxman, A. D., & Moher, D. (2008). Improving the reporting of pragmatic trials: an extension of the CONSORT statement. BMJ, 337. https://doi.org/10.1136/bmj.a2390

### 5.7 – Systematic reviews and meta-analysis

Chou, R., & Helfand, M. (2005). Challenges in Systematic Reviews That Assess Treatment Harms. Annals of Internal Medicine, 142(12_Part_2), 1090–1099. https://doi.org/10.7326/0003-4819-142-12_Part_2-200506211-00009

CIOMS Working Group X - Guideline on Evidence Synthesis and Meta-Analysis for Drug Safety. (n.d.). COUNCIL FOR INTERNATIONAL ORGANIZATIONS OF MEDICAL SCIENCES - CIOMS. Retrieved 26 May 2020, from https://cioms.ch/working_groups/working-group-x/

Henry, D., Lim, L. L.-Y., Rodriguez, L. A. G., Gutthann, S. P., Carson, J. L., Griffin, M., Savage, R., Logan, R., Moride, Y., Hawkey, C., Hill, S., & Fries, J. T. (1996). Variability in risk of gastrointestinal complications with individual non-steroidal anti-inflammatory drugs: results of a collaborative meta-analysis. BMJ, 312(7046), 1563–1566. https://doi.org/10.1136/bmj.312.7046.1563

### 5.8 – Signal detection methodology and application

Wisniewski, Af. et al. (2016, June). Good Signal Detection Practices: Evidence From IMI PROTECT. Drug Safety. https://doi.org/10.1007/s40264-016-0405-1

Guideline on good pharmacovigilance practices (GVP) Module IX – Signal management (Rev 1) European Medicines Agency. https://www.ema.europa.eu/en/human-regulatory/post-authorisation/pharmacovigilance/good-pharmacovigilance-practices

Bate, A., & Evans, S. J. W. (2009). Quantitative signal detection using spontaneous ADR reporting. Pharmacoepidemiology and Drug Safety, 18(6), 427–436. https://doi.org/10.1002/pds.1742

Berlin, C., Blanch, C., Lewis, D. J., Maladorno, D. D., Michel, C., Petrin, M., Sarp, S., & Close, P. (2012). Are all quantitative postmarketing signal detection methods equal? Performance characteristics of logistic regression and Multi-item Gamma Poisson Shrinker. Pharmacoepidemiology and Drug Safety, 21(6), 622–630. https://doi.org/10.1002/pds.2247

Candore, G., Juhlin, K., Manlik, K., Thakrar, B., Quarcoo, N., Seabroke, S., Wisniewski, A., & Slattery, J. (2015). Comparison of Statistical Signal Detection Methods Within and Across Spontaneous Reporting Databases. Drug Safety, 38(6), 577–587. https://doi.org/10.1007/s40264-015-0289-5

Caster, O., Juhlin, K., Watson, S., & Norén, G. N. (2014). Improved Statistical Signal Detection in Pharmacovigilance by Combining Multiple Strength-of-Evidence Aspects in vigiRank. Drug Safety, 37(8), 617–628. https://doi.org/10.1007/s40264-014-0204-5

Caster, O., Norén, G. N., Madigan, D., & Bate, A. (2010). Large-scale regression-based pattern discovery: The example of screening the WHO global drug safety database. Statistical Analysis and Data Mining: The ASA Data Science Journal, 3(4), 197–208. https://doi.org/10.1002/sam.10078

Caster, O., Norén, G. N., Madigan, D., & Bate, A. (2013). Logistic Regression in Signal Detection: Another Piece Added to the Puzzle. Clinical Pharmacology & Therapeutics, 94(3), 312–312. https://doi.org/10.1038/clpt.2013.107

DuMouchel, W., Yuen, N., Payvandi, N., Booth, W., Rut, A., & Fram, D. (2013). Automated Method for Detecting Increases in Frequency of Spontaneous Adverse Event Reports over Time. Journal of Biopharmaceutical Statistics, 23(1), 161–177. https://doi.org/10.1080/10543406.2013.736809

Salvo, F. (2013). A Potential Event-Competition Bias in Safety Signal Detection: Results From a Spontaneous Reporting Research Database in France. Drug Safety. https://doi.org/10.1007/s40264-013-0063-5

Harpaz, R., DuMouchel, W., LePendu, P., Bauer-Mehren, A., Ryan, P., & Shah, N. H. (2013). Performance of Pharmacovigilance Signal Detection Algorithms for the FDA Adverse Event Reporting System. Clinical Pharmacology and Therapeutics, 93(6). https://doi.org/10.1038/clpt.2013.24

Hauben, M., Madigan, D., Gerrits, C. M., Walsh, L., & Puijenbroek, E. P. V. (2005). The role of data mining in pharmacovigilance. Expert Opinion on Drug Safety, 4(5), 929–948. https://doi.org/10.1517/14740338.4.5.929

Holle, L. V., & Bauchau, V. (2014). Use of Logistic Regression to Combine Two Causality Criteria for Signal Detection in Vaccine Spontaneous Report Data. Drug Safety, 37(12), 1047–1057. https://doi.org/10.1007/s40264-014-0237-9

Juhlin, K., Karimi, G., Andér, M., Camilli, S., Dheda, M., Har, T. S., Isahak, R., Lee, S.-J., Vaughan, S., Caduff, P., & Norén, G. N. (2015). Using VigiBase to Identify Substandard Medicines: Detection Capacity and Key Prerequisites. Drug Safety, 38(4), 373–382. https://doi.org/10.1007/s40264-015-0271-2

Juhlin, K. (2013). Outlier Removal to Uncover Patterns in Adverse Drug Reaction Surveillance - A Simple Unmasking Strategy. Pharmacoepidemiology and Drug Safety. https://doi.org/10.1002/pds.3474

Low, Y. S., Caster, O., Bergvall, T., Fourches, D., Zang, X., Norén, G. N., Rusyn, I., Edwards, R., & Tropsha, A. (2016). Cheminformatics-aided pharmacovigilance: application to Stevens-Johnson Syndrome. Journal of the American Medical Informatics Association, 23(5), 968–978. https://doi.org/10.1093/jamia/ocv127

Maignen, F., Hauben, M., Hung, E., Holle, L. V., & Dogne, J.-M. (2014). A conceptual approach to the masking effect of measures of disproportionality. Pharmacoepidemiology and Drug Safety, 23(2), 208–217. https://doi.org/10.1002/pds.3530

Maignen, F., Hauben, M., Hung, E., Van Holle, L., & Dogne, J.-M. (2014). Assessing the extent and impact of the masking effect of disproportionality analyses on two spontaneous reporting systems databases. Pharmacoepidemiology and Drug Safety, 23(2), 195–207. https://doi.org/10.1002/pds.3529

MedDRA - Medical Dictionary for Regulatory Activities. (n.d.). Retrieved 27 May 2020, from https://www.meddra.org/

Norén, G. N., & Edwards, I. R. (2009). Modern methods of pharmacovigilance: detecting adverse effects of drugs. Clinical Medicine, 9(5), 486–489. https://doi.org/10.7861/clinmedicine.9-5-486

OHDSI – Observational Health Data Sciences and Informatics. (n.d.). Retrieved 23 June 2020, from https://www.ohdsi.org/

OMOP Common Data Model – OHDSI. (n.d.). Retrieved 23 June 2020, from https://www.ohdsi.org/data-standardization/the-common-data-model/

Pinheiro, L. C., Candore, G., Zaccaria, C., Slattery, J., & Arlett, P. (2018). An algorithm to detect unexpected increases in frequency of reports of adverse events in EudraVigilance. Pharmacoepidemiology and Drug Safety, 27(1), 38–45. https://doi.org/10.1002/pds.4344

Practical Aspects of Signal Detection in Pharmacovigilance: Report of CIOMS Working Group VIII • COUNCIL FOR INTERNATIONAL ORGANIZATIONS OF MEDICAL SCIENCES. (n.d.). COUNCIL FOR INTERNATIONAL ORGANIZATIONS OF MEDICAL SCIENCES. Retrieved 26 May 2020, from https://cioms.ch/publications/product/practical-aspects-of-signal-detection-in-pharmacovigilance-report-of-cioms-working-group-viii/

PROTECT. Retrieved 27 May 2020, from http://www.imi-protect.eu/about.shtml

PROTECT - Database of adverse drug reactions listed in the SPC of centrally authorised medicinal products. (n.d.). Retrieved 27 May 2020, from http://www.imi-protect.eu/methodsRep.shtml

FDA's Sentinel Initiative. FDA. https://www.fda.gov/safety/fdas-sentinel-initiative

FDA's Good Pharmacovigilance Practices and Pharmacoepidemiologic Assessment. U.S. Food and Drug Administration. https://www.fda.gov/regulatory-information/search-fda-guidance-documents/good-pharmacovigilance-practices-and-pharmacoepidemiologic-assessment

Scholl, J. H. G., van Hunsel, F. P. A. M., Hak, E., & van Puijenbroek, E. P. (2018). A prediction model-based algorithm for computer-assisted database screening of adverse drug reactions in the Netherlands. Pharmacoepidemiology and Drug Safety, 27(2), 199–205. https://doi.org/10.1002/pds.4364

Screening for adverse reactions in EudraVigilance. (n.d.). Retrieved 26 May 2020, from https://www.ema.europa.eu/en/documents/other/screening-adverse-reactions-eudravigilance_en.pdf

Seabroke, S., Candore, G., Juhlin, K., Quarcoo, N., Wisniewski, A., Arani, R., Painter, J., Tregunno, P., Norén, G. N., & Slattery, J. (2016). Performance of Stratified and Subgrouped Disproportionality Analyses in Spontaneous Databases. Drug Safety, 39(4), 355–364. https://doi.org/10.1007/s40264-015-0388-3

Stephenson, W. P., & Hauben, M. (2007). Data mining for signals in spontaneous reporting databases: proceed with caution. Pharmacoepidemiology and Drug Safety, 16(4), 359–365. https://doi.org/10.1002/pds.1323

Tatonetti, N. P., Ye, P. P., Daneshjou, R., & Altman, R. B. (2012). Data-Driven Prediction of Drug Effects and Interactions. Science Translational Medicine, 4(125), 125ra31-125ra31. https://doi.org/10.1126/scitranslmed.3003377

Harpaz, R. et al. (2017). Toward Multimodal Signal Detection of Adverse Drug Reactions. Journal of Biomedical Informatics. https://doi.org/10.1016/j.jbi.2017.10.013

**5.9 - Methods for pharmacovigilance impact research**

No references.

**Chapter 6 – The statistical analysis plan**

ICH E9 statistical principles for clinical trials [Text]. European Medicines Agency. Retrieved 28 May 2020, from https://www.ema.europa.eu/en/ich-e9-statistical-principles-clinical-trials

Welch, Ca. et al. (2014). Evaluation of Two-Fold Fully Conditional Specification Multiple Imputation for Longitudinal Electronic Health Record Data. Statistics in Medicine. https://doi.org/10.1002/sim.6184

Gail MH, Benichou J. (2000). Encyclopedia of Epidemiologic Methods (2000th ed.). Wiley.

ICH - International Council for Harmonisation of Technical Requirements for Pharmaceuticals for Human Use. (n.d.). ICH. Retrieved 28 May 2020, from https://www.ich.org/

ICH E9(R1) - Addendum on estimands and sensitivity analysis in clinical trials. (n.d.). Retrieved 11 June 2020, from https://database.ich.org/sites/default/files/E9-R1_Step4_Guideline_2019_1203.pdf

Jones, M. P. (1996). Indicator and Stratification Methods for Missing Explanatory Variables in Multiple Linear Regression. Journal of the American Statistical Association, 91(433), 222–230. JSTOR. https://doi.org/10.2307/2291399

Lee, Kj., Carlin, Jb. (2012). Recovery of Information From Multiple Imputation: A Simulation Study. Emerging Themes in Epidemiology. https://doi.org/10.1186/1742-7622-9-3

Little RJA, Rubin DB. (2002). Statistical analysis with missing data (2nd ed.). Wiley.

Moons, K. et al. Using the outcome for imputation of missing predictor values was preferred. (2006). Journal of Clinical Epidemiology, 59(10), 1092–1101. https://doi.org/10.1016/j.jclinepi.2006.01.009

Rosenbaum, P. R. (2010). Design of Observational Studies. Springer.

Rothman K, Greenland S, Lash T. (2008). Modern Epidemiology (3rd ed.). Lippincott Williams & Wilkins.

Rubin DB. (2004). Multiple Imputation for Nonresponse in Surveys. Wiley.

Schafer JL. (1997). Analysis of Incomplete Multivariate Data. Chapman & Hall.

**Chapter 7 – Quality management**

AHSC's Quality Control and Quality Assurance SOP. (n.d.). Retrieved 28 May 2020, from https://www.imperial.ac.uk/media/imperial-college/research-and-innovation/joint-research-compliance-office/public/JRCO_SOP_025_QAQC_V7.0_25Oct2017.pdf

Guideline on good pharmacovigilance practices (GVP) Module I – Pharmacovigilance systems and their quality systems. European Medicines Agency. https://www.ema.europa.eu/en/human-regulatory/post-authorisation/pharmacovigilance/good-pharmacovigilance-practices

Commission Implementing Regulation (EU) No 520/2012. (n.d.). https://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2012:159:0005:0025:EN:PDF

Whitney, Cw. et al. (1998). Quality Assurance and Quality Control in Longitudinal Studies. Epidemiologic Reviews. https://doi.org/10.1093/oxfordjournals.epirev.a017973

Herrett, E. et al. (2010). Validation and Validity of Diagnoses in the General Practice Research Database: A Systematic Review. British Journal of Clinical Pharmacology. https://doi.org/10.1111/j.1365-2125.2009.03537.x

EURODURG - International Society for Pharmacoepidemiology. (n.d.). Retrieved 29 May 2020, from https://www.pharmacoepi.org/eurodurg/

European Forum for Good Clinical Practice (EFCGP) Guidelines. (n.d.). http://www.efgcp.eu/Publications.asp?Type=EFGCP%20publications&L1=13&L2=1

Furu, K. (2008). Establishment of the nationwide Norwegian Prescription Database (NorPD) – new opportunities for research in pharmacoepidemiology in Norway. Norsk Epidemiologi, 18(2). https://doi.org/10.5324/nje.v18i2.23

Guidelines for Quality Assurance in Multicenter Trials: A Position Paper. (1998). Controlled Clinical Trials, 19(5), 477–493. https://doi.org/10.1016/S0197-2456(98)00033-6

Hoffmann, W., Latza, U., Baumeister, S. E., Brünger, M., Buttmann-Schweiger, N., Hardt, J., Hoffmann, V., Karch, A., Richter, A., Schmidt, C. O., Schmidtmann, I., Swart, E., & Berg, N. van den. (2019). Guidelines and recommendations for ensuring Good Epidemiological Practice (GEP): a guideline

developed by the German Society for Epidemiology. European Journal of Epidemiology, 34(3), 301–317. https://doi.org/10.1007/s10654-019-00500-x

ICH E6 (R2) Good clinical practice. (n.d.). Retrieved 10 June 2020, from https://www.ema.europa.eu/en/ich-e6-r2-good-clinical-practice

ICH E8 (R1) General considerations for clinical studies. (2018, September 17). European Medicines Agency. https://www.ema.europa.eu/en/ich-e8-general-considerations-clinical-studies

ISO 9000:2015 - Quality management systems — Fundamentals and vocabulary. (n.d.). ISO. Retrieved 28 May 2020, from https://www.iso.org/cms/render/live/en/sites/isoorg/contents/data/standard/04/54/45481.html

ISO Quality management principles. (n.d.). Retrieved 28 May 2020, from https://www.iso.org/files/live/sites/isoorg/files/store/en/PUB100080.pdf

ISPE Guidelines for Good Pharmacoepidemiology Practices (GPP). (2015). International Society for Pharmacoepidemiology. https://www.pharmacoepi.org/resources/policies/guidelines-08027/

Hoven, Jl. et al. (2005). Indicators of Prescribing Quality in Drug Utilisation Research: Report of a European Meeting (DURQUIM, 13-15 May 2004). European Journal of Clinical Pharmacology. https://doi.org/10.1007/s00228-004-0845-x

Kiran, D. R. (2016). Total Quality Management - 1st Edition (1st ed.). BSP Books, Elsevier. https://www.elsevier.com/books/total-quality-management/kiran/978-0-12-811035-5

Blomgren, Kj. et al. (2006). Interviewer Variability - Quality Aspects in a Case-Control Study. European Journal of Epidemiology. https://doi.org/10.1007/s10654-006-0017-7

Reppe, La. et al. (2017). Quality Assessment of Structure and Language Elements of Written Responses Given by Seven Scandinavian Drug Information Centres. European Journal of Clinical Pharmacology. https://doi.org/10.1007/s00228-017-2209-3

Reisch, Lm. et al. (2003). Training, Quality Assurance, and Assessment of Medical Record Abstraction in a Multisite Study. American Journal of Epidemiology. https://doi.org/10.1093/aje/kwg016

Lindquist, M. (2004). Data Quality Management in Pharmacovigilance. Drug Safety. https://doi.org/10.2165/00002018-200427120-00003

McGraw, M. J., George, A. N., Shearn, S. P., Jr, T. F. H., & Hall, R. L. (2010). Principles of Good Clinical Practice. Pharmaceutical Press.

Alzoubi, Mm. et al. (2019). Total Quality Management in the Health-Care Context: Integrating the Literature and Directing Future Research. Risk Management and Healthcare Policy. https://doi.org/10.2147/RMHP.S197038

Registries for Evaluating Patient Outcomes: A User's Guide: 3rd Edition Addendum (Overview) | Effective Health Care Program. (n.d.). Agency for Healthcare Research and Quality. Retrieved 29 May 2020, from https://effectivehealthcare.ahrq.gov/products/registries-guide-3rd-edition-addendum/overview

Research, C. for D. E. and. (2013, May). FDA's Best Practices for Conducting and Reporting Pharmacoepidemiologic Safety Studies Using Electronic Healthcare Data Sets. U.S. Food and Drug Administration. https://www.fda.gov/regulatory-information/search-fda-guidance-documents/best-practices-conducting-and-reporting-pharmacoepidemiologic-safety-studies-using-electronic

Sprenger, K., Nickerson, D., Meeker-O'Connell, A., & Morrison, B. W. (2013). Quality by Design in Clinical Trials: A Collaborative Pilot With FDA. Therapeutic Innovation & Regulatory Science, 47(2), 161–166. https://doi.org/10.1177/0092861512458909

Theobald, K., Capan, M., Herbold, M., Schinzel, S., & Hundt, F. (2009). Quality assurance in non-interventional studies. GMS German Medical Science, 7. https://doi.org/10.3205/000088

**Chapter 8 – Dissemination and communication of study results**

Guideline on good pharmacovigilance practices (GVP) Module VI – Collection, management and submission of reports of suspected adverse reactions to medicinal products (Rev 2). European Medicines Agency. https://www.ema.europa.eu/en/human-regulatory/post-authorisation/pharmacovigilance/good-pharmacovigilance-practices

ENCePP Checklist for Study Protocols. (n.d.). ENCePP. Retrieved 11 June 2020, from http://www.encepp.eu/standards_and_guidances/checkListProtocols.shtml

ENCePP Code of Conduct. (n.d.). ENCePP. Retrieved 11 June 2020, from http://www.encepp.eu/code_of_conduct/index.shtml

Enhancing the QUAlity and Transparency Of Health Research - EQUATOR Network. (n.d.). EQUATOR Network. Retrieved 11 June 2020, from https://www.equator-network.org/

EU PAS Register. (n.d.). ENCePP. Retrieved 11 June 2020, from http://www.encepp.eu/encepp_studies/indexRegister.shtml

FDA. (2013). FDA's Best Practices for Conducting and Reporting Pharmacoepidemiologic Safety Studies Using Electronic Healthcare Data Sets. FDA. https://www.fda.gov/regulatory-information/search-fda-guidance-documents/best-practices-conducting-and-reporting-pharmacoepidemiologic-safety-studies-using-electronic

GRACE Principles - Good ReseArch for Comparative Effectiveness. (n.d.). GRACE - Good ReseArch for Comparative Effectiveness. Retrieved 11 June 2020, from https://www.graceprinciples.org/

Guidance for the format and content of the final study report of non-interventional post-authorisation safety studies - EMA. (n.d.). Retrieved 11 June 2020, from https://www.ema.europa.eu/en/documents/regulatory-procedural-guideline/guidance-format-content-final-study-report-non-interventional-post-authorisation-safety-studies_en.pdf

Guidance for the format and content of the protocol of non-interventional post-authorisation safety studies - EMA. (n.d.). Retrieved 11 June 2020, from https://www.ema.europa.eu/en/documents/other/guidance-format-content-protocol-non-interventional-post-authorisation-safety-studies_en.pdf

Simera, I. et al. (2010). A Catalogue of Reporting Guidelines for Health Research. European Journal of Clinical Investigation. https://doi.org/10.1111/j.1365-2362.2009.02234.x

ICMJE - International Committee of Medical Journal Editors. (n.d.). International Committee of Medical Journal Editors. Retrieved 11 June 2020, from http://www.icmje.org/

ICMJE - Recommendations for the Conduct, Reporting, Editing, and Publication of Scholarly work in Medical Journals. (n.d.). International Committee of Medical Journal Editors. Retrieved 11 June 2020, from http://www.icmje.org/recommendations/

IEA Publications - Good Epidemiology Practices - GEP (2007). (n.d.). IEA - International Epidemiological Association. Retrieved 11 June 2020, from https://ieaweb.org/IEAWeb/Content/IEA_Publications.aspx

ISPE Guidelines for Good Pharmacoepidemiology Practices (GPP). (2015). International Society for Pharmacoepidemiology (ISPE). https://www.pharmacoepi.org/resources/policies/guidelines-08027/

Cohen. Jf. et al. (2016). STARD 2015 Guidelines for Reporting Diagnostic Accuracy Studies: Explanation and Elaboration. BMJ Open. https://doi.org/10.1136/bmjopen-2016-012799

Langan, S. M., Schmidt, S. A., Wing, K., Ehrenstein, V., Nicholls, S. G., Filion, K. B., Klungel, O., Petersen, I., Sorensen, H. T., Dixon, W. G., Guttmann, A., Harron, K., Hemkens, L. G., Moher, D., Schneeweiss, S., Smeeth, L., Sturkenboom, M., Elm, E. von, Wang, S. V., & Benchimol, E. I. (2018). RECORD-PE statement - The reporting of studies conducted using observational routinely collected health data statement for pharmacoepidemiology. BMJ, 363. https://doi.org/10.1136/bmj.k3532

Moher, D., Liberati, A., Tetzlaff, J., & Altman, D. G. (2009). Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. BMJ, 339. https://doi.org/10.1136/bmj.b2535

Nicholls, S. G., Quach, P., Elm, E. von, Guttmann, A., Moher, D., Petersen, I., Sørensen, H. T., Smeeth, L., Langan, S. M., & Benchimol, E. I. (2015). The REporting of Studies Conducted Using Observational Routinely-Collected Health Data (RECORD) Statement: Methods for Arriving at Consensus and Developing Reporting Guidelines. PLOS ONE, 10(5), e0125620. https://doi.org/10.1371/journal.pone.0125620

Schulz, K. F., Altman, D. G., & Moher, D. (2010). CONSORT 2010 Statement: updated guidelines for reporting parallel group randomised trials. BMJ, 340. https://doi.org/10.1136/bmj.c332

STROBE - Strengthening the Reporting of Observational studies in Epidemiology Statement Guidelines for reporting observational studies. (n.d.). STROBE. Retrieved 11 June 2020, from https://strobe-statement.org/index.php?id=strobe-publications

Stroup, D. F., Berlin, J. A., Morton, S. C., Olkin, I., Williamson, G. D., Rennie, D., Moher, D., Becker, B. J., Sipe, T. A., Thacker, S. B., & Epidemiology (MOOSE)

Stroup, Df. at al. (2000). Meta-analysis of Observational Studies in Epidemiology: A Proposal for Reporting. JAMA, 283(15), 2008–2012. https://doi.org/10.1001/jama.283.15.2008

Wang, Sv. et al. (2017). Reporting to Improve Reproducibility and Facilitate Validity Assessment for Healthcare Database Studies V1.0. Pharmacoepidemiology and Drug Safety. https://doi.org/10.1002/pds.4295

WMA - The World Medical Association-WMA Declaration of Helsinki – Ethical Principles for Medical Research Involving Human Subjects. (2018). https://www.wma.net/policies-post/wma-declaration-of-helsinki-ethical-principles-for-medical-research-involving-human-subjects/

Kelly, Wn. et al. (2007). Guidelines for Submitting Adverse Event Reports for Publication. Pharmacoepidemiology and Drug Safety. https://doi.org/10.1002/pds.1399

Zwarenstein, M., Treweek, S., Gagnier, J. J., Altman, D. G., Tunis, S., Haynes, B., Oxman, A. D., & Moher, D. (2008). Improving the reporting of pragmatic trials: an extension of the CONSORT statement. BMJ, 337. https://doi.org/10.1136/bmj.a2390

**Chapter 9 – Data protection and ethical aspects**

### 9.1 – Patient and data protection

Commission Implementing Regulation (EU) No. 520/2012. (n.d.). Retrieved 29 May 2020, from https://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2012:159:0005:0025:EN:PDF

Data Privacy, Medical Record Confidentiality, and Research in the Interest of Public Health - International Society for Pharmacoepidemiology. (n.d.). International Society for Pharmacoepidemiology (ISPE). Retrieved 29 May 2020, from https://www.pharmacoepi.org/resources/policies/privacy/

Directive 95/46/EC on the protection of individuals with regard to the processing of personal data and on the free movement of such data. (n.d.). Retrieved 29 May 2020, from https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:31995L0046&from=en

Directive 2001/20/EC - Implementation of good clinical practice in the conduct of clinical trials on medicinal products for human use. (n.d.). Retrieved 29 May 2020, from https://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2001:121:0034:0044:en:PDF

Directive 2005/28/EC - Guidelines for good clinical practice. (n.d.). Retrieved 29 May 2020, from https://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2005:091:0013:0019:en:PDF

EU Charter of Fundamental Rights. (n.d.). [Text]. European Commission - European Commission. Retrieved 29 May 2020, from https://ec.europa.eu/info/aid-development-cooperation-fundamental-rights/your-rights-eu/eu-charter-fundamental-rights_en

Guideline on good pharmacovigilance practices (GVP) Module VIII – Post-authorisation safety studies (Rev 3). (2018, September 17). European Medicines Agency. https://www.ema.europa.eu/en/human-regulatory/post-authorisation/pharmacovigilance/good-pharmacovigilance-practices

ISPE Guidelines for Good Pharmacoepidemiology Practices (GPP) - International Society for Pharmacoepidemiology. (2015). International Society for Pharmacoepidemiology. https://www.pharmacoepi.org/resources/policies/guidelines-08027/

Proposal for a regulation on the protection of individuals with regard to the processing of personal data by the Union institutions, bodies, offices and agencies and on the free movement of such data, and repealing Regulation (EC) No 45/2001 and Decision No 1247/2002/EC. (2017). EUR-Lex. https://eur-lex.europa.eu/legal-content/EN/TXT/?qid=1511278760489&uri=CELEX:52017PC0008

Regulation (EC) No 45/2001 on the protection of individuals with regard to the processing of personal data by the Community institutions and bodies and on the free movement of such data. (n.d.). Retrieved 29 May 2020, from https://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2001:008:0001:0022:EN:PDF

Regulation (EU) 2016/679 - General Data Protection Regulation. (2016). EUR-Lex. https://eur-lex.europa.eu/legal-content/EN/TXT/?toc=OJ%3AL%3A2016%3A119%3ATOC&uri=uriserv%3AOJ.L_.2016.119.01.0001.01.ENG

Regulation (EU) No 536/2014 on clinical trials on medicinal products for human use, and repealing Directive 2001/20/EC. (n.d.). Retrieved 29 May 2020, from https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32014R0536&rid=1

### 9.2 – Scientific integrity and ethical conduct

Agency for Healthcare Research and Quality (AHRQ). (n.d.). Retrieved 4 June 2020, from https://www.ahrq.gov/

CIOMS International Ethical Guidelines for Health-related Research Involving Human. (n.d.). Retrieved 4 June 2020, from https://cioms.ch/wp-content/uploads/2017/01/WEB-CIOMS-EthicalGuidelines.pdf

ENCePP Code of Conduct. (n.d.). ENCePP. Retrieved 4 June 2020, from http://www.encepp.eu/code_of_conduct/

Good Epidemiology Practices - IEA. (2007). International Epidemiological Association. http://ieaweb.org/IEAWeb/Content/IEA_Publications.aspx

Guidelines for Good Pharmacoepidemiology Practices (GPP) - International Society for Pharmacoepidemiology (ISPE). (n.d.). Retrieved 4 June 2020, from https://www.pharmacoepi.org/resources/policies/guidelines-08027/

ICMJE - International Committee of Medical Journal Editors. (n.d.). Retrieved 4 June 2020, from http://www.icmje.org/

Recommendations for the Conduct, Reporting, Editing, and Publication of Scholarly work in Medical Journals (ICJME). (n.d.). Retrieved 4 June 2020, from http://www.icmje.org/recommendations/

Registries for Evaluating Patient Outcomes: A User's Guide: 3rd Edition - Effective Health Care Program. (2014). https://effectivehealthcare.ahrq.gov/products/registries-guide-3rd-edition/research/

The European Union electronic Register of Post-Authorisation Studies (EU PAS Register). (n.d.). ENCePP. Retrieved 4 June 2020, from http://www.encepp.eu/encepp_studies/indexRegister.shtml

WMA Declaration of Helsinki – Ethical Principles for Medical Research Involving Human Subjects. (2013). https://www.wma.net/policies-post/wma-declaration-of-helsinki-ethical-principles-for-medical-research-involving-human-subjects/

Kurz, X., Bauchau, V., Mahy, P., Glismann, S., Van der As, Lm., Simondon, F. (2017). The ADVANCE Code of Conduct for Collaborative Vaccine Studies. Vaccine. https://doi.org/10.1016/j.vaccine.2017.02.039

## Chapter 10 – Specific topics

### 10.1 – Comparative effectiveness research – *This chapter has not been updated for Revision 8*

Agency for Healthcare Research and Quality (AHRQ). (n.d.). Retrieved April 5, 2018, from https://www.ahrq.gov/

AHRQ Comparative Effectiveness Reviews. (2005). Agency for Healthcare Research and Quality (US).

AHRQ. Developing a Protocol for Observational Comparative Effectiveness Research: A User's Guide. (n.d.). Retrieved from https://www.effectivehealthcare.ahrq.gov/topics/observational-cer-protocol/research

AHRQ. Methods Guide for Effectiveness and Comparative Effectiveness Reviews. (n.d.). Retrieved from https://effectivehealthcare.ahrq.gov/topics/cer-methods-guide/overview

AHRQ. Summary Variables in Observational Research: Propensity Scores and Disease Risk Scores. Effective Health Care Program Research Report No. 33. May 2012. (n.d.). Retrieved from https://effectivehealthcare.ahrq.gov/topics/observational-research-scores/research

AHRQ. The Incident User Design in Comparative Effectiveness Research. Effective Health Care Program Research Report No. 33. May 2012. (n.d.). Retrieved from https://effectivehealthcare.ahrq.gov/topics/incident-user-design/research

Avorn J. (2007). In defense of Pharmacoepidemiology-Embracing the Yin and Yang of Drug Research. N Engl J Med, 357(22), 2219–2221. Retrieved from http://www.nejm.org/doi/pdf/10.1056/NEJMp0706892

Black, N. (1996). Why we need observational studies to evaluate the effectiveness of health care. BMJ (Clinical Research Ed.), 312(7040), 1215–1218. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/8634569

Choosing Important Health Outcomes for Comparative Effectiveness Research: An Updated Review and User Survey. (n.d.). Retrieved April 9, 2018, from http://journals.plos.org/plosone/article?id=10.1371/journal.pone.0146444

COMET - Core Outcome Measures in Effectiveness Trials Initiative. (n.d.). Retrieved April 9, 2018, from http://www.comet-initiative.org/

Crystal, S., Gaboda, D., Lucas, J., Gerhard, T., & Chakravarty, S. (2010). Assessing medication exposures and outcomes in the frail elderly: assessing research challenges in nursing home pharmacotherapy. Medical Care, 48(6 Suppl), S23-31. https://doi.org/10.1097/MLR.0b013e3181de9d10

DARTNet Institute. (n.d.). Retrieved April 9, 2018, from http://www.dartnet.info/

EUnetHTA. Methodological Guidelines for Rapid Relative Effectiveness Assessment of Pharmaceuticals. (n.d.). Retrieved April 9, 2018, from https://www.eunethta.eu/the-final-version-of-hta-core-model-and-the-methodological-guidelines-for-rapid-rea-of-pharmaceuticals-is-now-available/

European Network for Health Technology Assessment (EUnetHTA). (n.d.). Retrieved June 28, 2018, from https://www.eunethta.eu/

GRACE Principles - Good ReseArch for Comparative Effectiveness. (n.d.). Retrieved April 9, 2018, from https://www.graceprinciples.org/

GRADE Working Group. (n.d.). Retrieved June 28, 2018, from http://www.gradeworkinggroup.org/

GRADE working group. 6-part 2008 BMJ series. (n.d.). Retrieved from http://www.cochrane.de/grade-publications#bmj

GRADE working group. Journal of Clinical Epidemiology series of articles. (n.d.). Retrieved from https://www.jclinepi.com/content/jce-GRADE-Series

Grading of Recommendations Assessment, Development, and Evaluation (GRADE) working group. (n.d.). Retrieved April 9, 2018, from http://www.gradeworkinggroup.org/

Harker, J., & Kleijnen, J. (2012). What is a rapid review? A methodological exploration of rapid reviews in Health Technology Assessments. International Journal of Evidence-Based Healthcare, 10(4), 397–410. https://doi.org/10.1111/j.1744-1609.2012.00290.x

Hernán, M. A., & Robins, J. M. (2016). Using Big Data to Emulate a Target Trial When a Randomized Trial Is Not Available. American Journal of Epidemiology, 183(8), 758–764. https://doi.org/10.1093/aje/kwv254

IMI GetReal. (n.d.). Retrieved June 28, 2018, from http://www.imi-getreal.eu/

International Society for Pharmacoeconomics and Outcome Research (ISPOR). (n.d.). Retrieved April 5, 2018, from https://www.ispor.org/

ISPOR Good Research Practices for Retrospective Database Analysis Task Force Report—Part I. (n.d.). Retrieved from https://www.ispor.org/TaskForces/documents/RDPartI.pdf

ISPOR Good Research Practices for Retrospective Database Analysis Task Force Report—Part II. (n.d.). Retrieved from https://www.ispor.org/TaskForces/documents/RDPartII.pdf

ISPOR Good Research Practices for Retrospective Database Analysis Task Force Report—Part III. (n.d.). Retrieved from https://www.ispor.org/TaskForces/documents/RDPartIII.pdf

Johnson, E. S., Bartman, B. A., Briesacher, B. A., Fleming, N. S., Gerhard, T., Kornegay, C. J., … Schneeweiss, S. (2013). The incident user design in comparative effectiveness research. Pharmacoepidemiology and Drug Safety, 22(1), 1–6. https://doi.org/10.1002/pds.3334

Klungel, O. H., Martens, E. P., Psaty, B. M., Grobbee, D. E., Sullivan, S. D., Stricker, B. H. C., … de Boer, A. (2004). Methods to assess intended effects of drug treatment in observational studies are reviewed. Journal of Clinical Epidemiology, 57(12), 1223–1231. https://doi.org/10.1016/j.jclinepi.2004.03.011

Loudon, K., Treweek, S., Sullivan, F., Donnan, P., Thorpe, K. E., & Zwarenstein, M. (2015). The PRECIS-2 tool: designing trials that are fit for purpose. BMJ, 350, h2147. https://doi.org/10.1136/bmj.h2147

Patient-Centered Outcome Research Institute (PCORI). (n.d.). Retrieved April 9, 2018, from https://www.pcori.org/

PCORI. Methodology Report. (2014, July 30). Retrieved April 9, 2018, from https://www.pcori.org/research-results/about-our-research/research-methodology/pcori-methodology-report

Randomized, Controlled Trials, Observational Studies, and the Hierarchy of Research Designs | NEJM. (n.d.). Retrieved April 9, 2018, from http://www.nejm.org/doi/full/10.1056/NEJM200006223422507#t=article

Ray, W. A. (2003). Evaluating Medication Effects Outside of Clinical Trials: New-User Designs. American Journal of Epidemiology, 158(9), 915–920. https://doi.org/10.1093/aje/kwg231

Real World Evidence Navigator. Retrieved April 9, 2018, from https://rwe-navigator.eu/

S, S. (2007). Developments in Post-marketing Comparative Effectiveness Research. Clinical Pharmacology and Therapeutics, 82(2), 143–156. https://doi.org/10.1038/sj.clpt.6100249

Sauer, B. C., Brookhart, A., Roy, J., & Vanderweele, T. (2013). A Review of Covariate Selection for Nonexperimental Comparative Effectiveness Research. Pharmacoepidemiology and Drug Safety, 22(11), 1139–1145. https://doi.org/10.1002/pds.3506

Schneeweiss, S., & Avorn, J. (2005). A review of uses of health care utilization databases for epidemiologic research on therapeutics. Journal of Clinical Epidemiology, 58(4), 323–337. https://doi.org/10.1016/j.jclinepi.2004.10.012

Schneeweiss S, Rassen JA, Glynn RJ, Avorn J, Mogun H and Brookhart MA. (2009). High-dimensional propensity score adjustment in studies of treatment effects using health care claims data. Epidemiology, 20(4), 512–522. Retrieved from https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3077219/pdf/nihms282669.pdf

Schwartz, D., & Lellouch, J. (2009). Explanatory and pragmatic attitudes in therapeutical trials. Journal of Clinical Epidemiology, 62(5), 499–505. https://doi.org/10.1016/j.jclinepi.2009.01.012

Selby, J. V., Forsythe, L., & Sox, H. C. (2015). Stakeholder-Driven Comparative Effectiveness Research: An Update From PCORI. JAMA, 314(21), 2235–2236. https://doi.org/10.1001/jama.2015.15139

Sox, H. C., & Goodman, S. N. (2012). The methods of comparative effectiveness research. Annual Review of Public Health, 33, 425–445. https://doi.org/10.1146/annurev-publhealth-031811-124610

Staa, T.-P. van, Goldacre, B., Gulliford, M., Cassell, J., Pirmohamed, M., Taweel, A., … Smeeth, L. (2012). Pragmatic randomised trials using routine electronic health records: putting them to the test. BMJ, 344, e55. https://doi.org/10.1136/bmj.e55

Tadrous, M., Gagne, J. J., Stürmer, T., & Cadarette, S. M. (2013). Disease Risk Score (DRS) as a Confounder Summary Method: Systematic Review and Recommendations. Pharmacoepidemiology and Drug Safety, 22(2), 122–129. https://doi.org/10.1002/pds.3377

Thorpe, K. E., Zwarenstein, M., Oxman, A. D., Treweek, S., Furberg, C. D., Altman, D. G., … Chalkidou, K. (2009). A pragmatic–explanatory continuum indicator summary (PRECIS): a tool to help trial designers. Canadian Medical Association Journal, 180(10), E47–E57. https://doi.org/10.1503/cmaj.090523

Toh, S., Gagne, J. J., Rassen, J. A., Fireman, B. H., Kulldorff, M., & Brown, J. S. (2013). Confounding adjustment in comparative effectiveness research conducted within distributed research networks. Medical Care, 51(8 Suppl 3), S4-10. https://doi.org/10.1097/MLR.0b013e31829b1bb1

US National Library of Medicine: Queries on CER. (n.d.). [Other]. Retrieved April 9, 2018, from https://www.nlm.nih.gov/hsrinfo/cer.html

Valuck, R. J., Anderson, H. O., Libby, A. M., Brandt, E., Bryan, C., Allen, R. R., … Pace, W. D. (2012). Enhancing electronic health record measurement of depression severity and suicide ideation: a Distributed Ambulatory Research in Therapeutics Network (DARTNet) study. Journal of the American Board of Family Medicine: JABFM, 25(5), 582–593. https://doi.org/10.3122/jabfm.2012.05.110053

When are observational studies as credible as randomised trials? - The Lancet. (n.d.). Retrieved April 9, 2018, from http://www.thelancet.com/pdfs/journals/lancet/PIIS0140-6736(04)16261-2.pdf

Zwarenstein, M., Treweek, S., Gagnier, J. J., Altman, D. G., Tunis, S., Haynes, B., … Moher, D. (2008). Improving the reporting of pragmatic trials: an extension of the CONSORT statement. BMJ, 337, a2390. https://doi.org/10.1136/bmj.a2390

**10.2 – Vaccine effectiveness**

A framework for research on vaccine effectiveness. (2018). Vaccine, 36(48), 7286–7293. https://doi.org/10.1016/j.vaccine.2018.04.016

ADVANCE - Report on appraisal of vaccine safety methods. (n.d.). Retrieved 8 June 2020, from https://vac4eu.org/wp-content/uploads/2019/03/ADVANCE_D4-2_appraisal-safety-methods_final_PU.pdf

Andrews, N., Waight, P. A., Borrow, R., Ladhani, S., George, R. C., Slack, M. P. E., & Miller, E. (2011). Using the Indirect Cohort Design to Estimate the Effectiveness of the Seven Valent Pneumococcal Conjugate Vaccine in England and Wales. PLOS ONE, 6(12), e28435. https://doi.org/10.1371/journal.pone.0028435

Baker, M. A., Baer, B., Kulldorff, M., Zichittella, L., Reindel, R., DeLuccia, S., Lipowicz, H., Freitas, K., Jin, R., & Yih, W. K. (2019). Kawasaki disease and 13-valent pneumococcal conjugate vaccination among young children: A self-controlled risk interval and cohort study with null results. PLoS Medicine, 16(7). https://doi.org/10.1371/journal.pmed.1002844

Beck, Cr. et al. (2012). Influenza Vaccination for Immunocompromised Patients: Systematic Review and Meta-Analysis by Etiology. The Journal of Infectious Diseases. https://doi.org/10.1093/infdis/jis487

Biggart, R. et al. Lack of impact of rotavirus vaccination on childhood seizure hospitalizations in England – An interrupted time series analysis. (2018). Vaccine, 36(31), 4589–4592. https://doi.org/10.1016/j.vaccine.2018.06.029

Black, S. et al. (2009). Importance of Background Rates of Disease in Assessment of Vaccine Safety During Mass Immunisation With Pandemic H1N1 Influenza Vaccines. Lancet (London, England). https://doi.org/10.1016/S0140-6736(09)61877-8

Bollaerts, K. et al. (2019). Observational Studies of Non-Specific Effects of Diphtheria-Tetanus-Pertussis Vaccines in Low-Income Countries: Assessing the Potential Impact of Study Characteristics, Bias and Confounding Through Meta-Regression. Vaccine. https://doi.org/10.1016/j.vaccine.2018.11.049

Brighton Collaboration - Home Page. (n.d.). Brighton Collaboration. Retrieved 8 June 2020, from https://brightoncollaboration.us/

Brown, Js. et al. (2007). Early Detection of Adverse Drug Events Within Population-Based Health Networks: Application of Sequential Testing Methods. Pharmacoepidemiology and Drug Safety. https://doi.org/10.1002/pds.1509

Carvajal, A. et al. (2011). Adverse Events Associated With Pandemic Influenza Vaccines: Comparison of the Results of a Follow-Up Study With Those Coming From Spontaneous Reporting. Vaccine. https://doi.org/10.1016/j.vaccine.2010.10.067

Castilla, J. et al. (2012). Effectiveness of Rotavirus Vaccines in Preventing Cases and Hospitalizations Due to Rotavirus Gastroenteritis in Navarre, Spain. Vaccine. https://doi.org/10.1016/j.vaccine.2011.11.071

CIOMS Guide to Vaccine Safety Communication. (2018). Council for International Organizations of Medical Sciences. https://cioms.ch/publications/product/cioms-guide-vaccine-safety-communication/

de Bie, S. et al. (2012). Vaccine-based Subgroup Analysis in VigiBase: Effect on Sensitivity in Paediatric Signal Detection. Drug Safety. https://doi.org/10.2165/11598120-000000000-00000

De Serres, G. et al. (2012). Use of Surveillance Data to Estimate the Effectiveness of the 7-valent Conjugate Pneumococcal Vaccine in Children Less Than 5 Years of Age Over a 9 Year Period. Vaccine. https://doi.org/10.1016/j.vaccine.2012.04.017

De Smedt, T. et al. (2018). Bias Due to Differential and Non-Differential Disease- And Exposure Misclassification in Studies of Vaccine Effectiveness. PloS One. https://doi.org/10.1371/journal.pone.0199180

De Vito, C. et al. (2007). A Systematic Review Evaluating the Potential for Bias and the Methodological Quality of Meta-Analyses in Vaccinology. Vaccine. https://doi.org/10.1016/j.vaccine.2007.10.034

Dieleman, J., Romio, S., Johansen, K., Weibel, D., Bonhoeffer, J., & Sturkenboom, M. (2011). Guillain-Barré syndrome and adjuvanted pandemic influenza A (H1N1) 2009 vaccine: multinational case-control study in Europe. BMJ, 343. https://doi.org/10.1136/bmj.d3908

Drolet, M. et al. (2019). Population-level Impact and Herd Effects Following the Introduction of Human Papillomavirus Vaccination Programmes: Updated Systematic Review and Meta-Analysis. Lancet (London, England). https://doi.org/10.1016/S0140-6736(19)30298-3

Dudareva-Vizule, S. et al. (2012). Impact of Rotavirus Vaccination in Regions With Low and Moderate Vaccine Uptake in Germany. Human Vaccines & Immunotherapeutics. https://doi.org/10.4161/hv.21593

Effectiveness of rotavirus vaccination – Generic study protocol for retrospective case control studies based on computerised databases - ECDC. (n.d.). Retrieved 8 June 2020, from https://www.ecdc.europa.eu/sites/portal/files/media/en/publications/Publications/rotavirus-vaccination-case-control-april2013.pdf

Effectiveness of rotavirus vaccination – Generic study protocol for retrospective cohort studies based on computerised databases - ECDC. (n.d.). Retrieved 8 June 2020, from https://www.ecdc.europa.eu/sites/portal/files/media/en/publications/Publications/rotavirus-vaccination-cohort-study-april2013.pdf

Epidemiological Designs for Vaccine Safety Assessment: Methods and Pitfalls. Biologicals : Journal of the International Association of Biological Standardization. https://doi.org/10.1016/j.biologicals.2011.08.010

Escolano, S. (2013). A New Self-Controlled Case Series Method for Analyzing Spontaneous Reports of Adverse Events After Vaccination. American Journal of Epidemiology. https://doi.org/10.1093/aje/kwt128

Escolano, S., Farrington, C. P., Hill, C., & Tubert-Bitter, P. (2011). Intussusception after Rotavirus Vaccination — Spontaneous Reports. New England Journal of Medicine, 365(22), 2139–2139. https://doi.org/10.1056/NEJMc1107771

European Centre for Disease Prevention and Control. (n.d.). Retrieved 8 June 2020, from https://www.ecdc.europa.eu/en

Evans, S. J. W. (2008). Stratification for Spontaneous Report Databases. Drug Safety, 31(11), 1049–1052. https://doi.org/10.2165/00002018-200831110-00009

Exploring the Feasibility of Conducting Vaccine Effectiveness Studies - Sentinel (2018). (n.d.). Retrieved 8 June 2020, from https://www.sentinelinitiative.org/sites/default/files/Methods/Sentinel_PRISM_Vaccine_Effectiveness_White_Paper.pdf

Farrington, Cp. (2004). Control Without Separate Controls: Evaluation of Vaccine Safety Using Case-Only Methods. Vaccine. https://doi.org/10.1016/j.vaccine.2004.01.017

Farrington, Cp. et al. (2009). Epidemiological Studies of the Non-Specific Effects of Vaccines: II--methodological Issues in the Design and Analysis of Cohort Studies. Tropical Medicine & International Health : TM & IH. https://doi.org/10.1111/j.1365-3156.2009.02302.x

Fine, Pe. et al. (2009). Epidemiological Studies of the 'Non-Specific Effects' of Vaccines: I--data Collection in Observational Studies. Tropical Medicine & International Health : TM & IH. https://doi.org/10.1111/j.1365-3156.2009.02301.x

GAIA project. (2016). Harmonising Immunisation Safety Assessment in Pregnancy. Vaccine, 34(49). https://www.sciencedirect.com/journal/vaccine/vol/34/issue/49

GAIA project. (2017). Harmonising Immunisation Safety Assessment in Pregnancy - Part II. Vaccine, 35(48). https://www.sciencedirect.com/journal/vaccine/vol/35/issue/48/part/PA

Glanz, Jm. et al. (2006). Four Different Study Designs to Evaluate Vaccine Safety Were Equally Validated With Contrasting Limitations. Journal of Clinical Epidemiology. https://doi.org/10.1016/j.jclinepi.2005.11.012

Greene, Sk. et al. (2010). Near Real-Time Surveillance for Influenza Vaccine Safety: Proof-Of-Concept in the Vaccine Safety Datalink Project. American Journal of Epidemiology. https://doi.org/10.1093/aje/kwp345

Guideline on good pharmacovigilance practices (GVP) - Product- or Population-Specific Considerations I: Vaccines for prophylaxis against infectious diseases [Text]. European Medicines Agency. https://www.ema.europa.eu/en/human-regulatory/post-authorisation/pharmacovigilance/good-pharmacovigilance-practices

Halloran, M. E., Haber, M., Longini, I. M., & Struchiner, C. J. (1991). Direct and Indirect Effects in Vaccine Efficacy and Effectiveness. American Journal of Epidemiology, 133(4), 323–331. https://doi.org/10.1093/oxfordjournals.aje.a115884

Hanquet, G. et al. (2013). Vaccine Effects and Impact of Vaccination Programmes in Post-Licensure Studies. Vaccine. https://doi.org/10.1016/j.vaccine.2013.07.006

Heffernan, J. M., & Keeling, M. J. (2009). Implications of vaccination and waning immunity. Proceedings of the Royal Society B: Biological Sciences, 276(1664), 2071–2080. https://doi.org/10.1098/rspb.2009.0057

Heininger, U. et al. (2017). Guide to Active Vaccine Safety Surveillance: Report of CIOMS Working Group on Vaccine Safety - Executive Summary. Vaccine. https://doi.org/10.1016/j.vaccine.2017.06.033

Hennessy, S. et al. (1996). Effectiveness of Live-Attenuated Japanese Encephalitis Vaccine (SA14-14-2): A Case-Control Study. Lancet (London, England). https://doi.org/10.1016/s0140-6736(96)91075-2

Hocine, Mn. et al. (2007). Hepatitis B Vaccination and First Central Nervous System Demyelinating Events: Reanalysis of a Case-Control Study Using the Self-Controlled Case Series Method. Vaccine. https://doi.org/10.1016/j.vaccine.2007.05.032

Hviid, A. et al. (2018). Human Papillomavirus Vaccination of Adult Women and Risk of Autoimmune and Neurological Diseases. Journal of Internal Medicine. https://doi.org/10.1111/joim.12694

Impact of rotavirus vaccination – Generic study protocol - ECDC. (n.d.). Retrieved 8 June 2020, from https://www.ecdc.europa.eu/sites/portal/files/media/en/publications/Publications/Rotavirus-impact-vaccination-April-2013.pdf

Jackson, Ml et al. (2013). The Test-Negative Design for Estimating Influenza Vaccine Effectiveness. Vaccine. https://doi.org/10.1016/j.vaccine.2013.02.053

Jefferson, T. et al. (2005). Efficacy and Effectiveness of Influenza Vaccines in Elderly People: A Systematic Review. Lancet (London, England). https://doi.org/10.1016/S0140-6736(05)67339-4

Klein, N. P. (2011). Vaccine safety in special populations. Human Vaccines, 7(2), 269–271. https://doi.org/10.4161/hv.7.2.13860

Kurz, X. et al. (2011). Safety Monitoring of Influenza A/H1N1 Pandemic Vaccines in EudraVigilance. Vaccine. https://doi.org/10.1016/j.vaccine.2011.04.005

Leino, T. et al. (2012). First Year Experience of Rotavirus Immunisation Programme in Finland. Vaccine. https://doi.org/10.1016/j.vaccine.2012.10.068

Leite, A., Andrews, N. J., & Thomas, S. L. (2016). Near real-time vaccine safety surveillance using electronic health records—a systematic review of the application of statistical methods. Pharmacoepidemiology and Drug Safety, 25(3), 225–237. https://doi.org/10.1002/pds.3966

Mahaux, O., Bauchau, V., & Van Holle, L. (2016a). Pharmacoepidemiological considerations in observed-to-expected analyses for vaccines. Pharmacoepidemiology and Drug Safety, 25(2), 215–222. https://doi.org/10.1002/pds.3918

Mahaux, O., Bauchau, V., & Van Holle, L. (2016b). Pharmacoepidemiological considerations in observed-to-expected analyses for vaccines. Pharmacoepidemiology and Drug Safety, 25(2), 215–222. https://doi.org/10.1002/pds.3918

Marin, M., Marti, M., Kambhampati, A., Jeram, S. M., & Seward, J. F. (2016). Global Varicella Vaccine Effectiveness: A Meta-analysis. Pediatrics, 137(3). https://doi.org/10.1542/peds.2015-3741

Martín Arias, Lh. et al. (2015). Guillain-Barré Syndrome and Influenza Vaccines: A Meta-Analysis. Vaccine. https://doi.org/10.1016/j.vaccine.2015.05.013

McClure, Dl. et al. (2008). Comparison of Epidemiologic Methods for Active Surveillance of Vaccine Safety. Vaccine. https://doi.org/10.1016/j.vaccine.2008.03.074

Miller, E. et al. Risk of Narcolepsy in Children and Young People Receiving AS03 Adjuvanted Pandemic A/H1N1 2009 Influenza Vaccine: Retrospective Analysis - PubMed. (n.d.). https://doi.org/10.1136/bmj.f794

MODULE 4 – Surveillance - WHO Vaccine Safety Basics. (n.d.). WHO Vaccine Safety Basics. Retrieved 8 June 2020, from https://vaccine-safety-training.org/overview-and-outcomes-4.html

Nelson, Jc. et al. (2012). Challenges in the Design and Analysis of Sequentially Monitored Postmarket Safety Surveillance Evaluations Using Electronic Observational Health Care Data. Pharmacoepidemiology and Drug Safety. https://doi.org/10.1002/pds.2324

Organization, W. H. (2017). Evaluation of influenza vaccine effectiveness: a guide to the design and interpretation of observational studies. World Health Organization. https://apps.who.int/iris/handle/10665/255203

Polyzos, Ka. et al. (2015). Maternal Influenza Vaccination and Risk for Congenital Malformations: A Systematic Review and Meta-analysis. Obstetrics and Gynecology. https://doi.org/10.1097/AOG.0000000000001068

Rasmussen, T. A., Jørgensen, M. R. S., Bjerrum, S., Jensen-Fangel, S., Støvring, H., Østergaard, L., & Søgaard, O. S. (2012). Use of population based background rates of disease to assess vaccine safety in childhood and mass immunisation in Denmark: nationwide population based cohort study. The BMJ, 345. https://doi.org/10.1136/bmj.e5823

Remschmidt, C., Rieck, T., Bödeker, B., & Wichmann, O. (2015). Application of the screening method to monitor influenza vaccine effectiveness among the elderly in Germany. BMC Infectious Diseases, 15. https://doi.org/10.1186/s12879-015-0882-3

Report of the CIOMS/WHO Working Group on Definition and Application of Terms for Vaccine Pharmacovigilance (2012). (n.d.). Retrieved 8 June 2020, from https://www.who.int/vaccine_safety/initiative/tools/CIOMS_report_WG_vaccine.pdf

Rondy, M. et al. (2015). 2012/13 Influenza Vaccine Effectiveness Against Hospitalised Influenza A(H1N1)pdm09, A(H3N2) and B: Estimates From a European Network of Hospitals. Euro Surveillance : Bulletin Europeen Sur Les Maladies Transmissibles = European Communicable Disease Bulletin. https://doi.org/10.2807/1560-7917.es2015.20.2.21011

Salmon, Da. et al. (2013). Association Between Guillain-Barré Syndrome and Influenza A (H1N1) 2009 Monovalent Inactivated Vaccines in the USA: A Meta-Analysis. Lancet (London, England). https://doi.org/10.1016/S0140-6736(12)62189-8

Siegrist, Ca. et al. (2007). Human Papilloma Virus Immunization in Adolescent and Young Adults: A Cohort Study to Illustrate What Events Might Be Mistaken for Adverse Reactions. The Pediatric Infectious Disease Journal. https://doi.org/10.1097/INF.0b013e318149dfea

Simpson, Cr. et al. (2015). Trivalent Inactivated Seasonal Influenza Vaccine Effectiveness for the Prevention of Laboratory-Confirmed Influenza in a Scottish Population 2000 to 2009. Euro Surveillance : Bulletin Europeen Sur Les Maladies Transmissibles = European Communicable Disease Bulletin. https://doi.org/10.2807/1560-7917.es2015.20.8.21043

Sullivan, Sg. et al. (2016). Theoretical Basis of the Test-Negative Study Design for Assessment of Influenza Vaccine Effectiveness. American Journal of Epidemiology. https://doi.org/10.1093/aje/kww064

Systematic overview of data sources for drug safety in pregnancy research Consultancy. (n.d.). Retrieved 8 June 2020, from http://www.encepp.eu/structure/documents/Data_sources_for_medicines_in_pregnancy_research.pdf

Tavares-Da-Silva, F. et al. Review of the initial post-marketing safety surveillance for the recombinant zoster vaccine. (2020). Vaccine, 38(18), 3489–3500. https://doi.org/10.1016/j.vaccine.2019.11.058

Tokars, Ji. et al. (2012). The Risk of Guillain-Barré Syndrome Associated With Influenza A (H1N1) 2009 Monovalent Vaccine and 2009-2010 Seasonal Influenza Vaccines: Results From Self-Controlled Analyses. Pharmacoepidemiology and Drug Safety. https://doi.org/10.1002/pds.3220

Vaccine Effectiveness Simulation. (n.d.). Retrieved 8 June 2020, from https://apps.p-95.com/VEMisclassification/

Vaccine Safety Datalink (VSD). (2019, June 17). https://www.cdc.gov/vaccinesafety/ensuringsafety/monitoring/vsd/index.html

Van Ganse, E. et al. (2012). Health Problems Most Commonly Diagnosed Among Young Female Patients During Visits to General Practitioners and Gynecologists in France Before the Initiation of the Human Papillomavirus Vaccination Program. Pharmacoepidemiology and Drug Safety. https://doi.org/10.1002/pds.2243

Van Holle, L. et al. (2013). Optimization of a Quantitative Signal Detection Algorithm for Spontaneous Reports of Adverse Events Post Immunization. Pharmacoepidemiology and Drug Safety. https://doi.org/10.1002/pds.3392

Verani, Jr. et al. (2017). Case-control Vaccine Effectiveness Studies: Data Collection, Analysis and Reporting Results. Vaccine. https://doi.org/10.1016/j.vaccine.2017.04.035

Verani, Jr. et al. (2017). Case-control Vaccine Effectiveness Studies: Preparation, Design, and Enrollment of Cases and Controls. Vaccine. https://doi.org/10.1016/j.vaccine.2017.04.037

Wacholder, S. et al. (2010). Risk of Miscarriage With Bivalent Vaccine Against Human Papillomavirus (HPV) Types 16 and 18: Pooled Analysis of Two Randomised Controlled Trials. BMJ (Clinical Research Ed.). https://doi.org/10.1136/bmj.c712

Whitaker, J. A., Ovsyannikova, I. G., & Poland, G. A. (2015). Adversomics: a new paradigm for vaccine safety and design. Expert Review of Vaccines, 14(7), 935–947. https://doi.org/10.1586/14760584.2015.1038249

WHO Vaccine Safety Basics. (n.d.). Retrieved 8 June 2020, from https://vaccine-safety-training.org/home.html

Wiyeh, A. B., Abdullahi, L. H., Wonkam, A., Wiysonge, C. S., & Kaba, M. (2018). Effects of vaccines in patients with sickle cell disease: a systematic review protocol. BMJ Open, 8(3), e021140. https://doi.org/10.1136/bmjopen-2017-021140

Woo, Ej. et al. (2008). Effects of Stratification on Data Mining in the US Vaccine Adverse Event Reporting System (VAERS). Drug Safety. https://doi.org/10.2165/00002018-200831080-00003

Xu, R. et al. (2012). Assessing the Effect of Vaccine on Spontaneous Abortion Using Time-Dependent Covariates Cox Models. Pharmacoepidemiology and Drug Safety. https://doi.org/10.1002/pds.3301

Yanni, E. A., Ferreira, G., Guennec, M., El Hahi, Y., El Ghachi, A., Haguinet, F., Espie, E., & Bianco, V. (2018). Burden of herpes zoster in 16 selected immunocompromised populations in England: a cohort study in the Clinical Practice Research Datalink 2000–2012. BMJ Open, 8(6). https://doi.org/10.1136/bmjopen-2017-020528

Yih, Wk.et al. (2011). Active Surveillance for Adverse Events: The Experience of the Vaccine Safety Datalink Project. Pediatrics. https://doi.org/10.1542/peds.2010-1722I

### 10.3 – Design and analysis of pharmacogenetic studies – *This chapter has not been updated for Revision 8*

Amstutz, U., Shear, N. H., Rieder, M. J., Hwang, S., Fung, V., Nakamura, H., CPNDS clinical recommendation group. (2014). Recommendations for HLA-B*15:02 and HLA-A*31:01 genetic testing to reduce the risk of carbamazepine-induced hypersensitivity reactions. Epilepsia, 55(4), 496–506. https://doi.org/10.1111/epi.12564

Balding, D. J. (2006). A tutorial on statistical methods for population association studies. Nature Reviews. Genetics, 7(10), 781–791. https://doi.org/10.1038/nrg1916

Bozkurt, O., Verschuren, W. M. M., van Wieren-de Wijer, B. M. A., Knol, M. J., de Boer, A., Grobbee, D. E., Klungel, O. H. (2008). Genetic variation in the renin-angiotensin system modifies the beneficial effects of ACE inhibitors on the risk of diabetes mellitus among hypertensives. Journal of Human Hypertension, 22(11), 774–780. https://doi.org/10.1038/jhh.2008.62

Carleton, B., Poole, R., Smith, M., Leeder, J., Ghannadan, R., Ross, C., Hayden, M. (2009). Adverse drug reaction active surveillance: developing a national network in Canada's children's hospitals. Pharmacoepidemiology and Drug Safety, 18(8), 713–721. https://doi.org/10.1002/pds.1772

Carr, D. F., O'Meara, H., Jorgensen, A. L., Campbell, J., Hobbs, M., McCann, G., Pirmohamed, M. (2013). SLCO1B1 genetic variant associated with statin-induced myopathy: a proof-of-concept study using the clinical practice research datalink. Clinical Pharmacology and Therapeutics, 94(6), 695–701. https://doi.org/10.1038/clpt.2013.161

CPIC: Clinical Pharmacogenetics Implementation Consortium. (n.d.). Retrieved April 9, 2018, from https://cpicpgx.org/

Dennis, J., Hawken, S., Krewski, D., Birkett, N., Gheorghe, M., Frei, J., Little, J. (2011). Bias in the case-only design applied to studies of gene-environment and gene-gene interaction: a systematic review and meta-analysis. International Journal of Epidemiology, 40(5), 1329–1341. https://doi.org/10.1093/ije/dyr088

Drews, J. (2000). Drug discovery: a historical perspective. Science (New York, N.Y.), 287(5460), 1960–1964.

Evans, W. E., & Relling, M. V. (1999). Pharmacogenomics: translating functional genomics into rational therapeutics. Science (New York, N.Y.), 286(5439), 487–491.

Evans, W. E., & Relling, M. V. (2004). Moving towards individualized medicine with pharmacogenomics. Nature, 429(6990), 464–468. https://doi.org/10.1038/nature02626

Gatto, N. M., Campbell, U. B., Rundle, A. G., & Ahsan, H. (2004). Further development of the case-only design for assessing gene-environment interaction: evaluation of and adjustment for bias. International Journal of Epidemiology, 33(5), 1014–1024. https://doi.org/10.1093/ije/dyh306

Grady, B. J., & Ritchie, M. D. (2011). Statistical Optimization of Pharmacogenomics Association Studies: Key Considerations from Study Design to Analysis. Current Pharmacogenomics and Personalized Medicine, 9(1), 41–66. https://doi.org/10.2174/187569211794728805

Guideline for data analysis of genomewide association studies. (n.d.). Retrieved April 9, 2018, from https://www.ncbi.nlm.nih.gov/pubmed/17726238

Jorgensen, A. L., & Williamson, P. R. (2008). Methodological quality of pharmacogenetic studies: issues of concern. Statistics in Medicine, 27(30), 6547–6569. https://doi.org/10.1002/sim.3420

Khoury, M. J., & Flanders, W. D. (1996). Nontraditional epidemiologic approaches in the analysis of gene-environment interaction: case-control studies with no controls! American Journal of Epidemiology, 144(3), 207–213. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/8686689

Little, J., Higgins, J. P. T., Ioannidis, J. P. A., Moher, D., Gagnon, F., von Elm, E., Birkett, N. (2009). STrengthening the REporting of Genetic Association studies (STREGA)--an extension of the STROBE statement. European Journal of Clinical Investigation, 39(4), 247–266. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/19297801

Liu, Y., Nyunoya, T., Leng, S., Belinsky, S. A., Tesfaigzi, Y., & Bruse, S. (2013). Softwares and methods for estimating genetic ancestry in human populations. Human Genomics, 7, 1. https://doi.org/10.1186/1479-7364-7-1

Madadi, P., Amstutz, U., Rieder, M., Ito, S., Fung, V., Hwang, S., CPNDS Clinical Recommendations Group. (2013). Clinical practice guideline: CYP2D6 genotyping for safe and efficacious codeine therapy. Journal of Population Therapeutics and Clinical Pharmacology = Journal De La Therapeutique Des Populations Et De La Pharamcologie Clinique, 20(3), e369-396. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/24214521

Molokhia, M., & McKeigue, P. (2006). EUDRAGENE: European collaboration to establish a case-control DNA collection for studying the genetic basis of adverse drug reactions. Pharmacogenomics, 7(4), 633–638. https://doi.org/10.2217/14622416.7.4.633

Peters, B. J. M., Rodin, A. S., de Boer, A., & Maitland-van der Zee, A.-H. (2010). Methodological and statistical issues in pharmacogenomics. The Journal of Pharmacy and Pharmacology, 62(2), 161–166. https://doi.org/10.1211/jpp.62.02.0002

PharmGKB (The Pharmacogenomics Knowledgebase). (n.d.). Retrieved April 9, 2018, from https://www.pharmgkb.org/

Pritchard, J. K., Stephens, M., Rosenberg, N. A., & Donnelly, P. (2000). Association Mapping in Structured Populations. American Journal of Human Genetics, 67(1), 170–181. Retrieved from https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1287075/

Psaty, B. M., Smith, N. L., Heckbert, S. R., Vos, H. L., Lemaitre, R. N., Reiner, A. P., Rosendaal, F. R. (2002). Diuretic therapy, the alpha-adducin gene variant, and the risk of myocardial infarction or stroke in persons with treated hypertension. JAMA, 287(13), 1680–1689. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/11926892

Ross, S., Anand, S. S., Joseph, P., & Paré, G. (2012). Promises and challenges of pharmacogenetics: an overview of study design, methodological and statistical issues. JRSM Cardiovascular Disease, 1(1). https://doi.org/10.1258/cvd.2012.012001

Storey, J. D., & Tibshirani, R. (2003). Statistical significance for genomewide studies. Proceedings of the National Academy of Sciences of the United States of America, 100(16), 9440–9445. https://doi.org/10.1073/pnas.1530509100

van der Baan, F. H., Klungel, O. H., Egberts, A. C. G., Leufkens, H. G., Grobbee, D. E., Roes, K. C. B., & Knol, M. J. (2011). Pharmacogenetics in randomized controlled trials: considerations for trial design. Pharmacogenomics, 12(10), 1485–1492. https://doi.org/10.2217/pgs.11.95

van der Baan, F. H., Knol, M. J., Klungel, O. H., Egberts, A. C., Grobbee, D. E., & Roes, K. C. B. (2012). Potential of adaptive clinical trial designs in pharmacogenetic research. Pharmacogenomics, 13(5), 571–578. https://doi.org/10.2217/pgs.12.10

van Wieren-de Wijer, D. B. M. A., Maitland-van der Zee, A.-H., de Boer, A., Kroon, A. A., de Leeuw, P. W., Schiffers, P., Klungel, O. H. (2009). Interaction between the Gly460Trp alpha-adducin gene variant and diuretics on the risk of myocardial infarction. Journal of Hypertension, 27(1), 61–68. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/19145769

# ENCePP Guide, Revision 8 – Authorship

***Editors***

Valarie Strassmann for the ENCePP Working Group *Research Standards and Guidances*

***Design and layout***

Julianna Fogd

***Main authors***

**Foreword – ENCePP Guide supports strong observational research for the COVID-19 pandemic:** Alejandro Arana for the ENCePP WG1

**1. Introduction:** Xavier Kurz

**2. Formulating the research question:** Yola Moride, Massoud Toussi

**3. Development of the study protocol:** Annie Fourrier-Réglat

**4. Approaches to data collection**

**4.1. Primary data collection:** Susana Perez-Gutthann

**4.1.1. Surveys**: Massoud Toussi, Xavier Fournie, Nawab Qizilbash

**4.1.2. Randomised clinical trials:** Susana Perez-Gutthann

**4.2. Secondary data collection:** Susana Perez-Gutthann, Gianluca Trifirò

**4.3. Patient registries:** Xavier Kurz

**4.4. Spontaneous reports:** Niklas Norèn, Jim Slattery

**4.5. Social media:** Annie Fourrier-Reglat, Massoud Toussi, Victoria Newbould, Niklas Noren

**4.6. Research networks:** Gianmario Candore, Alexandra Pacurariu, Miriam Sturkenboom

**5. Study design and methods**

**5.1. Definition and validation of drug exposure, outcomes and covariates**: Alejandro Arana, Gillian Hall, Jesper Hallas

**5.2. Bias (systematic error)**: Olaf Klungel, Jesper Hallas, Vera Ehrenstein, Mirko Di Martino, Ursula Kirchmayer, Alexandra Pacurariu, Alejandro Arana

**5.3. Methods to address bias:** Jesper Hallas, Vera Ehrenstein, Gillian Hall, Tania Schink, Olaf Klungel

**5.4. Effect measure modification and interaction**: Olaf Klungel, Federica Pisa

**5.5. Ecological analyses and case-population studies:** Ursula Kirchmayer

**5.6. Pragmatic trials and large simple trials:** Olaf Klungel, Tjeerd van Staa

**5.7. Systematic reviews and meta-analysis:** Gillian Hall, Daniel Morales

**5.8. Signal detection methodology and application:** Niklas Norèn, Jim Slattery, Gianmario Candore

**5.9. Methods for pharmacovigilance impact research:** Helga Gardarsdottir, Thomas Goedecke for the *ENCePP Special Interest Group on Measuring the Impact of Pharmacovigilance Activities*

**6. The statistical analysis plan:** Jim Slattery

**7. Quality management:** Milena Jadrijević-Mladar Takač

**8. Dissemination and communication of study results:** Kevin Blake, Susana Perez-Gutthann, Gillian Hall, Tania Schink

**9. Data protection and ethical aspects**

**9.1. Patient and data protection:** Xavier Kurz, Kevin Blake

**9.2. Scientific integrity and ethical conduct:** Xavier Kurz, Kevin Blake

**10. Specific topics**

**10.1. Comparative effectiveness research:** Annie Fourrier-Reglat, Olaf Klungel, Federica Pisa

**10.2. Vaccine safety and effectiveness:** Xavier Kurz, Catherine Cohet

**10.3. Design and analysis of pharmacogenetic studies:** Olaf Klungel, Bruce Carleton, Bruno Stricker

**Annex 1** Nawab Qizilbash, Jim Slattery and Kevin Blake, Daniel Morales

**Annex 2** Helga Gardarsdottir, Thomas Goedecke, Alexandra Pacurariu, Daniel Morales and the *ENCePP Special Interest Group on Measuring the Impact of Pharmacovigilance Activities*

***Acknowledgements***

Peer review of one or several chapters was received from:

Alfonso Carvajal

Amalia Issa

Ana Marta Anes

Ann Daly

Annalisa Rubino

Brian Strom

Camilla Stephens

Consuelo Pedros

David Martin

Elisabetta Poluzzi

Frank DeStefano

Geoff Liu

Gillian Bartlet

Gonzalo Calvo

Gunilla Sjolin-Forsberg

Hervé Le Louet

Hubert Leufkens

Ivana Silva

Jan Bonhoeffer

John Parkinson

Kenneth J. Rothman

Lina Titievski

Marina Davoli

Mark de Groot

Martin Daumer

Michael Theodorakis

Munir Pirmohamed

Nada Božina

Nancy Dreyer

Paddy Farrington

Patrick Zuber

Peter Arlett

Phillip Bryan

Robert Vander Stichele

Rolf Groenwold

Samy Suissa

Sarah McFarlane

Sebastian Botzenhardt

Sebastian Schneeweiss

Stefanie Prilla

Stephen Evans

Teresa Herdeiro

Ulf Bergman

Viktor Kiri

Wei Zhou

Yvonne Lis