The European Network of Centres for Pharmacoepidemiology and Pharmacovigilance (ENCePP)

Guide on Methodological Standards in Pharmacoepidemiology (Revision 3)

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 3). EMA/95098/2010. Available at http://www.encepp.eu/standards_and_guidances
Table of Contents

1. General aspects of study protocol ............................................................ 4
2. Research question ................................................................................... 5
3. Approaches to data collection ................................................................. 6
   3.1. Primary data collection ........................................................................ 6
   3.2. Secondary use of data ......................................................................... 7
   3.3. Research Networks .............................................................................. 9
   3.4. Spontaneous report database ............................................................... 11
4. Study design and methods ..................................................................... 12
   4.1. General considerations ....................................................................... 12
   4.2. Challenges and lessons learned ......................................................... 13
       4.2.1. Drug exposure/outcome/covariate definition and validation .......... 13
       4.2.1.1. Assessment of outcomes .......................................................... 14
       4.2.1.2. Validation ............................................................................... 14
       4.2.1.3. Drug consumption databases ............................................... 15
       4.2.2. Bias and confounding ................................................................. 15
       4.2.2.1. Choice of exposure risk-windows ......................................... 15
       4.2.2.2. Time-related bias .................................................................. 15
       4.2.2.3. Confounding by indication ...................................................... 16
       4.2.2.4. Protopathic bias .................................................................... 17
       4.2.2.5. Unmeasured confounding ....................................................... 17
       4.2.3. Methods to handle bias and confounding .................................... 18
           4.2.3.1. New-user designs .............................................................. 18
           4.2.3.2. Case-only designs .............................................................. 18
           4.2.3.3. Disease risk scores ............................................................ 19
           4.2.3.4. Propensity scores .............................................................. 20
           4.2.3.5. Instrumental variables ....................................................... 21
           4.2.3.6. Handling time-dependent confounding in the analysis .......... 22
       4.2.4. Effect modification ...................................................................... 22
   4.3. Ecological analyses and case-population studies ................................. 23
   4.4. Hybrid studies ................................................................................. 24
       4.4.1. Large simple trials ................................................................. 24
       4.4.2. Randomised database studies .................................................... 25
   4.5. Systematic review and meta-analysis .................................................. 25
   4.6. Signal detection methodology and application .................................... 27
5. Statistical and epidemiological analysis plan ......................................... 28
   5.1. General considerations ................................................................... 28
   5.2. Statistical plan ................................................................................... 29
   5.3. Handling of missing data ................................................................. 30
6. Quality control and quality assurance .................................................. 31
7. Communication ...................................................................................... 32
   7.1. Principles of communication ........................................................... 32
   7.2. Guidelines on communication ......................................................... 33
1. General aspects of study protocol

The study protocol is the core document of a study. A protocol should be drafted as one of the first steps in any research project. It must precisely describe everything being done in the study, so that the study can be reproduced. It should be amended and updated as needed. Amendments should be justified.

The EU Guidance on the format and content of the protocol of non-interventional post-authorisation safety studies (PASS) provides a template for protocols and may be applied to all non-interventional PASS, including meta-analyses and systematic reviews. Chapter II of the ISPE GPP provides guidance on what is expected of 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, for example, 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 all the elements that should be addressed and included in the protocols of such studies, including on the choice of data sources and study population, the study design and the analyses. The ENCePP Checklist for Study Protocols also 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 whose best practice principles and checklists on a wide range of topics are applicable to observational studies outside the scope of comparative effectiveness research. It should be borne in mind that, as stated in the GPP, adherence to guidelines will not guarantee valid research.

The protocol 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 of the research question. It will also explain the context of the research question, including what data are currently available and how this 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 populations to be used to answer the research question. The protocol should describe whether this population is already identified, and whether data are already available (allowing a secondary data collection from a database) or whether it needs to be recruited de novo. The limits of the desired population will be defined, including inclusion/exclusion criteria, timelines (such as index dates for inclusion in the study) and any exposure criteria and events defining cases and exposed study groups.

- Exposures of interest that need to be pre-specified and defined, including duration of exposure or follow-up, visits or time-dependent appraisals and details of which data are collected when, using what methods.

- 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.
• 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 (see section 6 of this Guide).

• The identification of possible 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.

• Adverse events/reactions that will or will not be collected and reported and the procedures put in place for this purpose.

• Ethical considerations, as described in the section on governance of the current document.

• The various data collection forms including the Case Report Form (CRF) or descriptions of the data elements may be appended to the protocol, providing an exact representation of the data collection. 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 data collection methodology. Other forms may be included as needed, such as patient information, patient-oriented summaries, etc.

2. Research question

The research question and the associated objectives address the knowledge or information to be gained from the study. It is important that current knowledge gaps are properly identified. Existing guidance on this aspect includes the EU Guidance on the format and content of the protocol of non-interventional post-authorisation safety studies, the ISPE GPP and the ENCePP Checklist for Study Protocols. These guidance documents emphasise that it should be clearly explained why the study is to be conducted (e.g. to answer an important public health concern, to confirm or further characterise a risk identified in a Risk Management Plan, to assess a new or emerging safety issue or to determine health outcomes or the benefit/risk profile). It also should be clear whether the results that will be reported represent pre-formed hypotheses or research questions, or are data driven. If there is no pre-formed hypothesis, this should be clearly stated.

A critical and thorough review of the literature usually forms the basis for the background description and theoretical framework of the research question and should be included in a protocol. Such review aims at evaluating the pertinent information and identifying gaps in knowledge. The review should include findings of relevant animal and human experiments, clinical studies and previous epidemiological studies. The findings of similar studies should be mentioned and gaps in knowledge that the study is intended to fill should be described.

In addition, previous findings are useful for the methodological planning of the current study. They may be used to discuss how they support the background, significance, research question, hypotheses, and/or design of the proposed study. They may also serve to determine the expected magnitude of the event(s) under study and, if available, in the target population, to characterise the various risk factors for the event, to identify the outcomes and measures that have been used in previous studies and to assess the feasibility of the proposed study. Several methods for reviewing and synthesising findings from the literature exist, including narrative review, for which guidance is available in Writing narrative literature reviews (Rev of Gen Psychol 1997;1(3):11-320).
3. Approaches to data collection

There are two main approaches for data collection. One is primary collection of data specifically for the study. Another option is to use data already collected for another purpose, eg. as part of administrative records of patient health care ("secondary data collection"). Module VIII Post-authorisation safety studies of the Good pharmacovigilance practice distinguishes between studies that involve either approach.

Increasingly often, a combination of approaches is used. In addition, networking among centres active in pharmacoepidemiology and pharmacovigilance is rapidly changing the landscape of drug safety research in Europe, both in terms of data networks and networks of interested researchers who may have data sources that they can contribute to a particular study.

3.1. Primary data collection

Despite some limitations, primary data collection has an important role in pharmacoepidemiology. Case-control 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 base populations and in-depth case assessment by clinical experts. Good 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).

General guidance on proper conduct of prospective patient-based studies can be found in the ISPE Good Pharmacoepidemiology Practices (GPP) and the IEA Good Epidemiology Practice (GEP). The GPP is especially useful for its recommendations on aspects rarely covered by guidelines, such as data quality issues and archiving. Both guidelines address the importance of patient data protection and the ethical principles of research using patient healthcare and personal data.

Patient registries are sometimes requested by regulators at the time of authorisation of a medicinal product in order to determine clinical effectiveness and monitor safety. A registry should be considered a structure within which studies can be performed, i.e. a data source where entry is defined either by diagnosis of a disease (disease registry) or prescription of a drug (exposure registry). AHRQ has published a comprehensive document on ‘good registry practices’ entitled Registries for Evaluating Patient Outcomes: A User's Guide, 3rd Edition, which guides the planning, design, implementation, analysis, interpretation, and evaluation of the quality of a registry. A section also covers linking of registries to other data sources. The FDA’s Guidance for Industry-Establishing Pregnancy Exposure Registries provides best practice for designing a pregnancy registry, with a description of research methods and elements for 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.

Surveys are increasingly used in pharmacoepidemiology, especially in the areas of disease epidemiology and risk minimisation evaluation. They require a sampling strategy that allows for external validity and maximised response rates. Useful textbooks on these aspects are Survey Sampling (L. Kish, Wiley, 1995) and Survey Methodology (R.M. Groves, F.J. Fowler, M.P. Couper, J.M. Lepkowski, E. Singer, R. Tourangeau, 2nd Edition, Wiley 2009). Depending of the purpose of the survey, questionnaires are often used. They should be validated based on accepted measures including, if appropriate, construct, criterion and content validity, inter-rater and test-retest reliability, sensitivity and responsiveness. Although primarily focused on quality of life research, the book Quality of Life: the assessment, analysis and

Randomised clinical trials (RCTs) are a form of primary data collection. There are numerous textbooks and publications on methodological and operational aspects of clinical trials; they are not covered here. An essential guideline on clinical trials is the European Medicines Agency (EMA) Note for Guidance on Good Clinical Practice, which specifies obligations for the conduct of clinical trials to ensure that the data generated in the trial are valid.

3.2. Secondary use of data

The use of already available electronic patient healthcare data for research has had a marked impact on pharmacoepidemiology research. The last two decades have witnessed the development of key data resources, expertise and methodology that have allowed the conduct of landmark studies in the field. Electronic medical records and record linkage of administrative health records are the main types of databases from a data structure and origin perspective. Examples of the first and second are the CPRD in the UK and the national or regional databases in the Nordic countries, Italy, Netherlands and other countries, respectively. The ENCePP Inventory of Databases contains key information on the databases that are registered in ENCePP. Section 3.3 of this Guide also describes databases and healthcare records used by research networks.


The primary purpose of the ISPE-endorsed Guidelines for Good Database Selection and use in Pharmacoepidemiology Research (Pharmacoepidemial 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 tested procedures. Although it refers in the title and objective to data resources or databases, it mainly refers to databases of routinely collected healthcare information and does not include spontaneous report databases. It is a simple, well-structured guideline that will help investigators when selecting databases for their research. If used, it will help database custodians to describe their database in a useful manner. A section is entirely dedicated to the use of multi-site studies. The entire document contains references to data quality and data processing/transformation issues and there are sections dedicated to quality and validation procedures. There are also separate sections on privacy and security.

The Working Group for the Survey and Utilisation of Secondary Data (AGENS) with representatives from the German Society for Social Medicine and Prevention (DGSPM) and the German Society for Epidemiology (DGEpi) developed a Good Practice in Secondary Data Analysis Version 2 aiming to establish a standard for planning, conducting and analysing studies on the basis of secondary data. It is also aimed to be used as the basis for contracts between data owners (so-called primary users) and secondary users. It is divided in 11 sections addressing, among other aspects, the study protocol, quality assurance and data protection.
The FDA’s Best Practices for Conducting and Reporting Pharmacoepidemiologic Safety Studies Using Electronic Health Care Data Sets provides criteria of best practice that apply to the design, analysis, conduct and documentation of pharmacoepidemiological safety studies using electronic healthcare data. It emphasizes that investigators should understand 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.

General guidance for studies including those conducted with electronic healthcare databases can also be found in the ISPE GPP, in particular sections IV-B (Study conduct, Data collection). This guidance emphasises the paramount importance of patient data protection.

The International Society for Pharmacoeconomics and Outcome Research (ISPOR) established a task force to recommend good research practices for designing and analysing retrospective databases for comparative effectiveness research (CER). The Task Force has subsequently published three articles (Part I, Part II and Part III) that review methodological issues and possible solutions for CER studies based on secondary data analysis. (see also section 9.1 on comparative effectiveness research ). Many of the principles are applicable to studies with other objectives than CER, but important aspects of pharmacoepidemiological studies based on secondary use of data, such as data quality, ethical issues, data ownership and privacy, are not covered.

The use of technology including administrative databases for pharmacoepidemiological research has limitations, including the following issues:

- **Consistency and totality of data capture i.e. does the database reliably capture all of the patient’s healthcare interactions or are there known gaps in coverage, capture, longitudinality or eligibility?** Researchers using claims data rarely have the opportunity to carry out quality assurance of the whole data set. Descriptive analyses of the integrity of a US Medicaid Claims Database (Pharmacoepidemiol Drug Saf 2003;12:103–11) concludes that performing such analyses can reveal important limitations of the data and whenever possible, researchers should examine the ‘parent’ data set for apparent irregularities.

- **Bias in assessment of drug exposure from an administrative database.** The relevance of these biases for quality control in more clinical databases is explored in European Surveillance of Antimicrobial Consumption (ESAC): Data Collection Performance and Methodological Approach (Br J Clin Pharmacol 2004;58: 419-28). This article describes a retrospective data collection effort (1997–2001) through an international network of surveillance systems, aimed at collecting publicly available, comparable and reliable data on antibiotic use in Europe. The data collected were screened for bias, using a checklist focusing on detection bias in sample and census data, errors in assigning medicinal product packages to the Anatomical Therapeutic Chemical Classification System, errors in calculations of Defined Daily Doses per package, bias by over-the-counter sales and parallel trade, and bias in ambulatory/hospital care mix. The authors conclude that methodological rigour is needed to assure data validity and to ensure reliable cross-national comparison.

- **Validity of the data and the definitions used, which is not simply about source record validation of a particular endpoint.** There are many possible ways to define endpoints and researchers may only seek to validate their choice. Validation and validity of diagnoses in the General Practice Research Database (GPRD): a systematic review (Br J Clin Pharmacol 2010;69:4-14) investigated the range of methods used to validate diagnoses in a primary care database and concluded that a number of methods had been used to assess validity and that overall, estimates of validity were high. The quality of reporting of the validations was however often inadequate to permit a clear interpretation. Not all methods provided a quantitative estimate of validity and most methods considered only the positive predictive value of a set of diagnostic codes in a highly selected group of cases.
Discordance between data sources, such as in Discordance of databases designed for claims payment versus clinical information systems: implications for outcomes research (Ann Intern Med 1993; 119: 844-50), which was a comparative study of a clinical versus an insurance claims database for predictors of prognosis in patients with ischaemic heart disease. A finding was that claims data failed to identify more than half of the patients with prognostically important conditions when compared with the clinical information system.

Another example of the hazards of using large linked databases is provided in Vaccine safety surveillance using large linked databases: opportunities, hazards and proposed guidelines (Expert Rev Vaccines 2003; 2(1):21-9). In general it is clear that the quality of pharmacoepidemiological studies that rely heavily on clinical databases from medical practice could be greatly enhanced by stimulating the quality of medical registration in electronic health records, through the provision of elaborate end-user terminologies and classification aides at the point-of-care.

Quality control and assurance are further addressed in section 6 of the Guide.

### 3.3. Research Networks

Collaborations for multinational studies are not new and have been strongly encouraged over the last years by the drug safety research funded by the European Commission (EC). The funding resulted in the conduct of groundwork necessary to overcome the hurdles of data sharing across countries.

Networking implies collaboration between investigators, which is based on trust and willingness to share and to maximise the advantage of bundling expertise. The ENCePP Database of Research Resources may facilitate such collaborations by providing an inventory of research centres and data sources available for 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. In addition, an important component of ENCePP is the potential for meta-analyses to maximise the information gathered for an issue that is addressed in different databases. In the US, the HMO Research Network, the Vaccine Safety Datalink (VSD) and Mini-Sentinel are examples of consortia involving health maintenance organisations that have formal, recognised research capabilities.

From a methodological point of view, data networks have many advantages:

- By increasing the size of study populations, networks may shorten the time needed for obtaining the desired sample size. Hence, networks can facilitate research on rare events and accelerate investigation of drug safety issues.
- Heterogeneity of drug exposure across countries allows studying the effect of individual drugs.
- Multinational studies may provide additional knowledge on whether a drug safety issue exists in several countries, on the consistency of information and on the impact of biases on estimates.
- Involvement of experts from various countries addressing case definitions, terminologies, coding in databases and research practices provides opportunities to increase consistency of results of observational studies.
- Requirement to share data forces harmonisation of data elaboration and transparency in analyses, and benchmarking of data management.

Different models have been applied for combining data from various countries ranging from a very disparate to a more integrated approach:

- Meta-analysis of results of individual studies with potentially different design e.g. Variability in risk of gastrointestinal complications with individual NSAIDs: results of a collaborative meta-analysis (BMJ 1996;312:1563-6), 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.

- Pooling of results from common protocol studies conducted in different databases, allowing assessment of database/population characteristics and of choices of study design and analysis as determinants of variability of results (e.g. Pharmacoepidemiological Research on Outcomes of Therapeutics by a European Consortium (PROTECT) project).

- The five Nordic countries with similar healthcare systems and databases have developed a collaborative cross-national pharmacoepidemiological network which covers the entire population of 25 million inhabitants (The Nordic countries as a cohort for pharmacoepidemiological research. Basic Clin Pharmacol & Toxicol 2010;106:86–94). This network has been used for analytical pharmacoepidemiological studies linking drug exposure to other health registries (for example in 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).

- Distributed data approach in which data partners maintain physical and operational control over electronic data in their existing environment (e.g. Mini-Sentinel project and PRISM, its extension for vaccines). A common data model allows to standardise administrative and clinical information across data partners, execute standardised programs and share the output of these programs in a summary form. Methods are available to allow multivariate adjusted analyses in federated databases without violating patient privacy (see Multivariate-adjusted pharmacoepidemiologic analyses of confidential information pooled from multiple healthcare utilisation databases. Pharmacoepidemiol Drug Saf 2010;19:848-57).

- Pooling of aggregated data (person-time based or person-level based) extracted locally from databases or electronic health records using a common data model and common software, and transmitted electronically to a central data warehouse for further analysis (see Combining electronic healthcare databases in Europe to allow for large-scale drug safety monitoring: the EU-ADR Project. Pharmacoepidemiol Drug Saf 2011;20(1):1-11).

These different models have different strengths and weaknesses and present different challenges. These may include:

- Differences in culture and experience between academia, public institutions and private partners.

- Different ethical and governance requirements in each country regarding processing of anonymised or pseudo-anonymised healthcare data.

- Mapping of differing disease coding systems (for examples, the International Classification of Disease, 10th Revision (ICD-10), Read codes in the United Kingdom and the International Classification of Primary Care (ICPC-2)) and languages of narrative medical information.

- Choice of data sharing model and access rights of partners.

- Validation of diagnoses and access to source documents for validation.

- Issues linked to intellectual property and authorship.

- Sustainability and funding mechanisms, especially when private funding (e.g. from pharmaceutical companies) is involved and when the study receives funding from several sponsors.

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. Technical solutions also
exist for data sharing and mapping of terminologies. A distributed data model and a JAVA (freely available)-based data elaboration software was developed by the EU-ADR project and shown to be feasible and able to deal effectively with ethical and governance issues in multinational drug safety studies, as illustrated in studies on Guillain-Barré syndrome and narcolepsy conducted by the VAESCO consortium in the context of H1N1 vaccination. A central datawarehouse that is remotely accessible is used in several EC funded projects (e.g. SOS, ARITMO, SAFEGUARD, GRiP). It increases collaboration and transparency and allows for rotating principal investigators.

Many of the current pharmacoepidemiology research networks in the EU have operated mainly with EC funds and under EC grant agreements. The coming years should demonstrate whether and how the expertise and infrastructures involved could be maintained and used in the conduct of post-authorisation studies.

3.4. **Spontaneous report database**

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 (ICSR), 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)](http://www.ema.europa.eu) 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 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 imbedded in the concept of voluntary reporting, whereby known or unknown external factors may influence the reporting rate and data quality. ADRs 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 or validated in a clinical context 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. [Performance of probabilistic method to detect duplicate individual case safety reports](http://www.drugsafetyjournal.com) (Drug Saf 2014;37(4):249-58) showed that a probabilistic method highlighted duplicates that had been missed by a rule-based method and also improved the accuracy of manual review. A demonstration of the performance of de-duplication methods to improve signal detection is however lacking.

[Validation of statistical signal detection procedures in EudraVigilance post-authorisation data: a retrospective evaluation of the potential for earlier signalling](http://www.drugsafetyjournal.com) (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. Its Appendix 3 provides a list of international and national spontaneous reporting system databases.

4. Study design and methods

There exists a number of evolving methodological challenges that recur in pharmacoepidemiological research, that are still in development or have not yet been adequately covered by recommendations. The following sections present such methodological challenges.

4.1. General considerations

The choice of study design and methods is a crucial part in every pharmacoepidemiological study and starts with the formulation of a relevant research question (whether non-steroidal anti-inflammatory drugs [NSAIDs] increase the risk of gastro-intestinal bleeding is cited throughout the present document as an illustrative working example). The study design and methods should follow the research question and are naturally interrelated.

Pharmacoepidemiological studies involving multiple objectives are not so uncommon. One approach for a given study might be to consider the study population as a cohort in which to implement the most appropriate design for each objective, thus ensuring alignment of each objective to the best possible design and analysis.

In a descriptive study, the research question relates to describing a population with respect to pre-defined parameters. In analytical studies, the research question drives three key sequentially structured phases in the design and conduct of an epidemiological study:

- Relation of a parameter of incidence to a determinant or a set of determinants (e.g. the incidence rate ratio of gastro-intestinal bleeds among users and non-users of NSAIDs),
- Collection of data to empirically document this relation (e.g. collection from a database of exposure [use of NSAIDs] and outcomes data [gastro-intestinal bleeding] in a cohort of patients that are/have been NSAIDs users), and
- Analysis of data (from raw data to quantification of associations).

These three phases are not independent. A hypothesised relation may lead to an array of designs for data collection based, in this example, on different data sources available on use of NSAIDs (exposure) and occurrence of gastro-intestinal bleeds in patients (outcomes). Each design for data collection will be followed by only a few appropriate designs of data analysis. If applicable, the selection of appropriate electronic health data sources will be an important aspect of the design of data collection, but depending on the research question, other sources of data may be needed (e.g. some claims databases may not have a ‘reason for stopping’ a NSAID whereas another may have).

The choice of epidemiological methods to answer a research question is based on principles rather than on rules. These principles may provide opportunities for creativeness and new innovative methods, when appropriate and needed. However, there are certain ‘dos and don’ts’ and certain standards in order to assure validity and robustness of the study results.

Many research organisations (including those owning or hosting databases) have scientific review boards ensuring scientific standards are met. Some national competent authorities also have their own review board for registering/approving studies. In addition, it is good practice to invite experts to review the protocol and study report, and any publications and/or communications thereof. The role of scientific committees in governance is of particular importance in any study.
General aspects of study designs, their relevance to types of research question and issues relating to internal and external validity, including biases and confounding, are covered by many textbooks on epidemiology and pharmacoepidemiology. The following list proposes a list of textbooks recommended for consultation. Researchers may find other textbooks more appropriate to their specific needs.

- *Epidemiology: Principles and Methods 2nd Edition* (B. MacMahon, D. Trichopoulos. Lippincott Williams & Wilkins, 1996) offers an introductory understanding of epidemiological methods and processes, including on study designs and control for confounding.

- *Modern Epidemiology 3rd Edition* (K. Rothman, S. Greenland, T. Lash. Lippincott Williams & Wilkins, 2008) serves as a comprehensive textbook on methods in epidemiology. Chapter 8 deals with validity but rather than dichotomise validity into the two components, internal and external, details a view in which the essence of scientific generalisation is the formulation of abstract concepts relating the study factors.

- *Pharmacoepidemiology 5th Edition* (B. Strom, S.E. Kimmel, S. Hennessy. Wiley, 2012) provides a complete review of epidemiological methods applied to the study of drugs. In Chapter 41, it emphasises that, whatever the source of the data, the veracity of a study’s conclusion rests on the validity of the data.


- *Encyclopedia of Epidemiologic Methods* (M.H. Gail, J. Benichou, Editors. Wiley, 2000). This compilation of articles complements existing textbooks by providing a large coverage of specialised topics in epidemiological and statistical methods.


- *A Dictionary of Epidemiology 5th Edition* (M Porta, Editor. J.M. Last S. Greenland, Associate Editors. Oxford University Press, 2008), sponsored by the International Epidemiological Association (IEA), provides a definition and concise explanation of epidemiologic terms and is a key to understanding epidemiological concepts.


### 4.2. Challenges and lessons learned

#### 4.2.1. Drug exposure/outcome/covariate definition and validation

Historically, physicians relied on patient-supplied information on past drug use and illness to assist with the diagnosis of current disease. This reliance has been reduced with the rapid expansion of use and access to electronic health records.

Chapter 41 of Pharmacoepidemiology (B. Strom, S.E. Kimmel, S. Hennessy. 5th Edition, Wiley, 2012) includes a literature review of the studies that have evaluated the validity of drug, diagnosis and hospitalisation data and the factors that influence the accuracy of these data. The book presents information on the two primary information sources available for pharmacoepidemiology studies i.e. questionnaires and administrative databases. It concludes with a summary of current knowledge in the field and directions for future research.
4.2.1.1. Assessment of outcomes

A case definition compatible with the observational database 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 (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 not necessarily compatible with the information available in an observational data set. 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 should be provided. Code list generation requires expertise on both the coding system and the disease area. Researchers should also 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):704-07). Coding systems used in some commonly used resources 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. In some circumstances, chart review or 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). In some cases, initial plausibility checks or subsequent medical chart review will be necessary. Where databases have prescription data only, therapy exposure may be used as a marker for an outcome, or linkage of different databases is required.

4.2.1.2. Validation

In healthcare databases, the correct assessment of drug exposure/outcome/covariate will be crucial to the validity of research. The role of automated record linkage in the postmarketing surveillance of drug safety: a critique (Clin Pharmacol Ther 1989;46:371-86) evaluates the validity of research conducted in automated databases according to a standard set of criteria, including validity of exposure, outcome and confounding. It points out that diagnoses obtained from a review of codes of electronic record systems require validation. The validation of electronic information on drug exposure, outcome or covariate definitions should also be included in the technical handbook of each of the databases. Validity of diagnostic coding within the General Practice Research Database: a systematic review (Br J Gen Pract 2010;60:e128-36) and the book Pharmacoepidemiology (B. Strom, S.E. Kimmel, S. Hennessy. 5th Edition, Wiley, 2012) contain examples.

Completeness and validity of all variables should be considered for exposure, outcomes, potential confounders and effect modifiers. Assumptions included in case definitions or other algorithms may need to be confirmed. For resources 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 should however consider the effect of any differences in variables or analyses and subsequent changes to health care, procedures and coding. A full understanding of both the health care system and procedures that generated the data is required. External validation against chart review or physician / patient questionnaire is possible with some resources. While false positives are more easily measured than false negatives (unless the outcome is extremely common), specificity of an outcome is more important than 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). Alternatively, internal logic checks can test for completeness and accuracy of variables. For example, one can investigate wether an outcome was followed by (or proceeded from) appropriate exposure or
4.2.1.3. Drug consumption databases

The Inventory of Drug Consumption Databases in Europe reviews, compiles and updates knowledge about European sources of data on drug utilisation in the out- and inpatient healthcare sector. The inventory describes information available from non-commercial drug consumption data providers in 23 European countries and methods to retrieve this information. It examines the validity of these European national drug consumption databases and explores the availability of inpatient drug consumption data at national level.

4.2.2. Bias and confounding

4.2.2.1. Choice of exposure risk-windows

The choice of exposure risk window can influence risk comparisons. A study of the effects of exposure misclassification due to the time-window design in pharmacoepidemiologic studies (Clin Epidemiol 1994:47(2):183–89) considers the impact of the time-window design on the validity of risk estimates in record linkage studies. In adverse drug reaction studies, an exposure risk-window constitutes the number of exposure days assigned to each prescription. The ideal design situation would occur when each exposure risk-window would only cover the period of potential excess risk. The estimation of the time of drug-related risk is however complex as it depends on the duration of drug use and the onset and persistence of drug toxicity. With longer windows, a substantive attenuation of incidence rates may be observed. The choice of prescription risk windows can, therefore, influence the estimate of exposure risks. Risk windows should be validated or a sensitivity analysis should be conducted accordingly.

4.2.2.2. Time-related bias

Time-related bias in observational studies can produce illusory results in favour of the treatment group and may affect both cohort and case-control studies, mostly database studies. They are most often a form of differential misclassification bias and should be recognised as they can be generally avoided by appropriate accounting of follow-up time and exposure status in the design and analysis of such studies.

4.2.2.2.1. Immortal time bias

Immortal time in epidemiology refers to a period of cohort follow-up time during which death (or an outcome that determines end of follow-up) cannot occur. It is defined in the book Modern Epidemiology (K. Rothman, S. Greenland, T. Lash. 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, e.g., to a drug, during which death 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;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 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;167:492-9) describes various cohort study designs leading to this bias, quantifies its magnitude under different survival distributions and illustrates it by using data from a cohort of lung cancer patients. The author shows that 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. The findings
support the conclusion that observational studies of drug benefit in which computerised databases are used must be designed and analysed properly to avoid immortal time bias.

Survival bias associated with time-to-treatment initiation in drug effectiveness evaluation: a comparison of methods (Am J Epidemiol 2005;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 were excluded from the analysis and the study allowed 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 observation studies can bias the results in favour of the treatment group and how they consider it not difficult to identify and avoid. They recommend that all cohort studies should be assessed for the presence of immortal time bias using appropriate validity criteria. However, Re. 'Immortal time bias on pharmacoepidemiology' (Am J Epidemiol 2009;170: 667-8) argues that sound efforts at minimising the influence of more common biases should not be sacrificed to that of immortal time bias.

4.2.2.2. Other forms of time-related bias

Time-window Bias in Case-control Studies. Statins and Lung Cancer (Epidemiology 2011;22(2):228-31) describes a case-control study which reported a 45% reduction in the rate of lung cancer with any statin use. A differential misclassification bias arose from the methods used to select controls and measure their exposure, which resulted in exposure assessment to statins being based on a shorter time-span for cases than controls and an over-representation of unexposed cases. Properly accounting for time produced a null association.

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 or not, 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 on 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 example of use of inhaled corticosteroids and death in chronic obstructive pulmonary disease patients, no clearly valid approach to data analysis could circumvent this bias. Further research is needed on this issue.

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).

4.2.2.3. Confounding by indication

Confounding by indication refers to an extraneous determinant of the outcome parameter that is present if a perceived high risk or poor prognosis is an indication for intervention. This means that differences in care, for example, between cases and controls may partly originate from differences in indication for medical intervention such as the presence of risk factors for particular health problems. The latter has frequently been reported in studies evaluating the efficacy of pharmaceutical interventions.
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;9:37-41) reviews conceptual issues regarding confounding by indication. It 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, sensitivity analysis for unmeasured confounders, instrumental variables and G-estimation.

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;12:551-8) 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;57:1223-31).

Moreover, claimed advantages of a new drug may channel it to patients with special pre-existing morbidity, with the consequence that disease states can be incorrectly attributed to use of the drug (channelling). How channelling towards high risk gastrointestinal patients occurred in the prescribing of newer NSAIDs is well demonstrated in Channelling bias and the incidence of gastrointestinal haemorrhage in users of meloxicam, coxibs, and older, non-specific NSAIDs (Gut 2003;52:1265–70).

### 4.2.2.4. Protopathic bias

Protopathic bias occurs 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 (Bias: Considerations for research practice. Am J Health Syst Pharm 2008;65:2159-68).

### 4.2.2.5. Unmeasured confounding

Large healthcare utilisation databases are frequently used to analyse unintended effects of prescription drugs and biologics. Confounders that require detailed information on clinical parameters, lifestyle, or over-the-counter medications are often not measured in such datasets, causing residual confounding bias. Using directed acyclic graphs to detect limitations of traditional regression in longitudinal studies (Int J Public Health 2010;55:701-3) reviews confounding and mediation (i.e. intermediate effects) in longitudinal data and introduces causal graphs to understand the relationships between the variables in an epidemiological study.

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 the 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 author 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 spread sheets (for download, e.g. at http://www.drugepi.org/dope-downloads/), sensitivity analyses should be used more frequently, substituting qualitative discussions of residual confounding.
The amount of bias in exposure-effect estimates that can plausibly occur due to residual or unmeasured confounding has been debated. The impact of residual and unmeasured confounding in epidemiologic studies: a simulation study (Am J Epidemiol 2007;166:646–55) 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. With plausible assumptions about residual and unmeasured confounding, effect sizes of the magnitude frequently reported in observational epidemiological studies can be generated. This study also highlights the need to perform sensitivity analyses to assess whether unmeasured and residual confounding are likely problems. 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.

4.2.3. Methods to handle bias and confounding

4.2.3.1. New-user designs

The practice of most observational studies to include many prevalent users, i.e. patients taking a therapy for some time before study follow-up began, can cause two types of bias. First, prevalent users are "survivors" of the early period of pharmacotherapy, which can introduce substantial bias if risk varies with time. Second, covariates for drug users at study entry are often plausibly affected by the drug itself. New user designs help avoid the mistake of adjusting for factors on the causal pathway which may introduce bias. Evaluating medication effects outside of clinical trials: new-user designs (Am J Epidemiol 2003;158 (9):915–20) reviews such designs, which avoid these biases by restricting the analysis to persons under observation at the start of the current course of treatment. In addition to defining new-user designs the article explains how they can be implemented as case-control studies and describes the logistical and sample size limitations involved.

4.2.3.2. Case-only designs

Case-only designs reduce confounding by using the exposure history of each case as its own control and eliminate between-person confounding by constant characteristics, including chronic diseases.

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). In this study, the risk of depression associated with cardiovascular drugs was estimated by analysing the nonsymmetrical distribution of prescription orders for cardiovascular drugs and antidepressants.

The case-crossover design studies transient exposures with acute effects (The Case-Crossover Design: A Method for Studying Transient Effects on the Risk of Acute Events. Am J Epidemiol 1991;133:144-53). The case-time-control design (Epidemiology 1995;6(3):248-53) is an elaboration of the case-crossover design. It uses exposure history data from a traditional control group to estimate and adjust for the bias from temporal changes in prescribing (Case-crossover and Case-Time-Control Designs as Alternatives in Pharmacoepidemiologic Research. Pharmacoepidemiol Drug Saf 1997; Suppl 3. S51-S59). However, if not well matched, the control group may reintroduce selection bias. In this situation, a ‘case-case-time-control’ method may be helpful as explained in Future cases as present controls to adjust for exposure trend bias in case-only studies (Epidemiology 2011;22:568–74).

The self-controlled case series (SCCS) design was primarily developed to investigate the association between a vaccine and an adverse event but is increasingly used to study drug exposure. In this design, the observation period following each exposure for each case is divided into risk period(s) (e.g. number(s) of days immediately following each exposure) and a control period (e.g. the remaining observation period). Incidence rates within the risk period after exposure are compared with incidence...
rates within the control period. A SCCS analysis adjusting for age effects has the advantage of an implicit control of any known or unknown confounders which are stable over time. The Tutorial in biostatistics: the self-controlled case series method (Stat Med 2006; 25(10):1768-97) and the associated website http://statistics.open.ac.uk/sccs explain how to fit SCCS models using standard statistical packages. Like cohort or case-control studies, the SCCS method remains, however, susceptible to bias if exposure is timed to minimise the risk of an adverse event. Relevant time intervals for the risk and control periods need also to be defined and this may become complex, e.g. with primary vaccination with several doses. 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 assumes that the event itself is not a contra-indication for the exposure. A modification of the method (pseudolikelihood 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). Based on a review of 40 vaccine studies, 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, 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.

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. It demonstrates that the self-controlled case-series and case-cross over designs were more susceptible to bias, but this bias was removed when follow-up was sampled both before and after the outcome, or when a case-time-control design was used. When should case-only designs be used for safety monitoring of medicinal 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 (165 positive controls and 234 negative controls) in 5 observational databases and 6 simulated datasets. Four outcomes and five design choices were assessed. The predictive accuracy of the most optimal design across all 20 outcome-database scenarios yielded an AUC>0.5, AUC>0.75 in 12 scenarios, and AUC>0.85 in 4 scenarios. The coverage probability based on the negative controls in real data and relative rates in simulated data was poor: the highest value was 64% for 1 outcome for negative controls and it decreased with increasing relative rates (0% for relative rates>=1.5). A limitation of the study is that the performance of SCCS is tested on many drug-outcome pairs without taking into account whether critical assumptions of the method are met (outcome influences future exposure or outcome leads to censoring.

More research on the performance of the SCCS design for various types of drug exposure in empirical and simulation studies are needed to assess its validity.

4.2.3.3. Disease risk scores

An approach to controlling for a large number of confounding variables is to construct a multivariable confounder score which summarises potential confounding factors in a single score. Stratification by a multivariate confounder score (Am J Epidemiol 1976;104:609-20) 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;18:67-80) 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. Large variation exists with differences in terminology and methods used for score derivation.

4.2.3.4. Propensity scores

Databases used in pharmacoepidemiologic 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;150:327–33) 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 emerging high dimensional propensity score (hd-PS) model approach. It attempts to empirically identify large numbers of potential confounders in healthcare databases and, by doing, extract more information on confounders and proxies thereof. Covariate selection in high-dimensional propensity score analyses of treatment effects in small samples (Am J Epidemiol 2011;173:1404-13) 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;20:849-57) evaluates the use of hd-PS in a primary care electronic medical record database. In addition, the article Using high-dimensional propensity score 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) tested several methods for 1:n propensity score matching in simulation and empirical studies and recommended 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) developed and tested 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;20:1115-29) and further evaluated in Propensity score balance measures in...
pharmacoepidemiology: a simulation study (Pharmacoepidemol Drug Saf 2014; Epub ahead of print). In most situations, the standardised difference performs best and is easy to calculate (Balance measures for propensity score methods: a clinical example on beta-agonist use and the risk of myocardial infarction, Pharmacoepidemol Drug Saf 2011;20(11):1130-7). Metrics for covariate balance in cohort studies of causal effects (Stat Med 2013;33:1685-99) assessed in a simulation study that the c-statistic and general weighted difference perform as well as the standardized difference and are preferred when an overall summary measure of balance is requested.

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, by using additional covariate measurements observed in a validation study.

Although in most situations propensity score models, with the exception of hd-PS, do not have any advantages over conventional multivariate modelling in terms of adjustment for investigator 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 mechanical 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.

4.2.3.5. Instrumental variables

Instrumental variable (IV) methods were invented over 70 years ago, but were used by epidemiologists only recently. Over the past decade or so, non-parametric versions of IV methods have appeared that connect IV methods to causal and measurement-error models important in epidemiological applications. An introduction to instrumental variables for epidemiologists (Int J Epidemiol 2000;29:722-9) 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. Where IV assumptions are questionable, the corrections can still serve as part of the sensitivity analysis or external adjustment. Where the assumptions are more defensible, as in field trials and in studies that obtain validation or reliability data, IV methods can form an integral part of the analysis.

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). This article also proposes a potential approach to controlling for confounding by indication in non-experimental studies of treatment effects. It illustrates this in a study comparing the effect of exposure to COX-2 inhibitors with non-selective NSAIDs on gastrointestinal complications. Contrary to results from randomised controlled trials showing that COX-2 inhibitors lead to a reduced risk of gastro-intestinal toxicity relative to non-selective NSAIDs, the author’s conventional multivariable analysis found no evidence of a gastro-protective effect attributable to COX-2 inhibitor use. In contrast to the conventional analysis, 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. The authors however point out to the possibility that a physician can influence the outcome in ways other than through the prescribing of NSAID. For example, physicians who frequently prescribe COX-2 inhibitors may also be more likely to co-prescribe proton pump inhibitors for additional gastro-protection. In such a situation, the protective effect due to COX-2 exposure is partly attributable to the use of a proton pump inhibitor.

Instrumental variable methods in comparative safety and effectiveness research (Pharmacoepidemol Drug Saf 2010;19:537–54) is a practical guidance on IV analyses in pharmacoepidemiology.
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; 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.

4.2.3.6. Handling time-dependent confounding in the analysis

4.2.3.6.1. G-estimation


4.2.3.6.2. Marginal Structural Models (MSM)

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

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;11:561–70)).

Effect of highly active antiretroviral therapy on time to acquired immunodeficiency syndrome or death using marginal structural models (Am J Epidemiol 2003;158:687-94) 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 beta(2)-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.

Beyond the approaches proposed above, traditional and efficient approaches to deal with time dependent variables in the design of the study, such as nested case control studies with assessment of time varying exposure windows, should be considered.

4.2.4. Effect modification

An important question that may arise when studying the effects of medicines is whether such effects differ between subgroups of patients (effect modification). To answer this question, one can stratify the study population, e.g. by gender, and compare the effects in these subgroups. In 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;326:219), it is recommended to perform a formal statistical test to assess if there are statistically significant differences between subgroup for these effects. 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;18:805-35).

Effect modification 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 most appropriate. The standard measure for interaction on the additive scale 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). With sufficient sample size, most interaction tests perform similarly with regard to type 1 error rates and power 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). In small samples (<250), the Breslow-Day and Tarone test performed best for interactions on the odds-ratio scale, whereas Likelihood Ratio and RERI-based tests performed better on RD scale. When exposure prevents outcome, in small samples the RERI-based test is relatively underpowered compared to other tests. Possible solutions include choosing an alternative interaction test, or recoding exposure categories taking the category with the lowest risk as reference.

Strengthening the Reporting of Observational Studies in Epidemiology (STROBE): explanation and elaboration (Epidemiology 2007;18:805-35) and Recommendations for presenting analyses of effect modification and interaction (Int J Epidemiol 2012;41:514-20) recommend that effect modification should be reported as follows:

1. separate effects (RRs, Odds Ratios or RDs, 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;
2. effects of the exposure within strata of the potential effect modifier;
3. measures of effect modification on both additive (e.g. RERI) and multiplicative (e.g. S) scales with confidence intervals;
4. confounders for which the association between exposure and outcome was adjusted for.

It should be kept in mind that past drug use should be considered as a potential effect modifier in studies assessing the risk of occurrence of events associated with recent drug use. 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.

### 4.3. Ecological analyses and case-population studies

Ecological analyses should not be considered hypothesis testing studies. As explained in Control without separate controls: evaluation of vaccine safety using case-only methods (Vaccine 2004; 22(15-16):2064-70), they assume that a strong correlation between the trend in an indicator of vaccine coverage (in this
example) and the trend in incidence of a disease that is a presumed effect of the vaccine (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. An example of a multinational case-population study aiming to estimate population rates of a suspected adverse event using national sales data is provided in Transplantation for Acute Liver Failure in Patients Exposed to NSAIDs or Paracetamol (Acetaminophen) (Drug Saf 2013;36:135–44). 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) compares 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 therefore give erroneous results, with the choice of the right denominator depending on the hazard function of the adverse event.

A practical attitude towards case-population studies is recommended: in situations where nation-wide or region-wide electronic health records are available that allow assessing the outcomes and confounders with sufficient validity, a case-population approach is not necessary nor wanted, 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 databases (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, with the limitation that the ecological nature of the reference data restricts the ability to control for confounding to some basic variables, such as sex and age, and precludes an exhaustive control for confounding.

4.4. **Hybrid studies**

The term ‘hybrid studies’ refers to efforts at bridging the pharmacoepidemiological principles and practices of interventional and non-interventional study design, conduct and analysis. One of the primary aims for doing this is to better reflect ‘real life’ populations and circumstances.

4.4.1. **Large simple trials**

RCT are considered the gold standard for demonstrating the efficacy of medicinal products. This design can also be used to obtain unbiased estimates of the risk for adverse outcomes. However, large sample sizes are required when the risk is small or delayed (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 (e.g. risks of stroke and of myocardial infarction) or when many confounding factors need to be balanced between treatment groups. In such circumstances, the cost and complexity of a RCT may outweigh its advantages over observational studies. A study design which, ethical considerations permitting, allowed drug allocation to be randomised in an otherwise normal clinical setting, and which relied upon the routine collection of primary and secondary healthcare records, could overcome the size limitations and atypical settings of conventional clinical trials. It would also avoid the channelling bias that may, in some cases, make it impossible to interpret the results of purely observational studies. A Large Simple Trial (LST) is a study design that keeps the volume and complexity of data collection to a minimum.
Outcomes that are simple and objective can be measured from the routine process of care using epidemiological follow-up methods, for example by using questionnaires or hospital discharge records. LST methodology is discussed in Chapters 36 and 37 of the book *Pharmacoepidemiology* (Strom BL, Kimmel SE, Hennessy S. 5th Edition, Wiley, 2012), which includes a list of conditions appropriate for the conduct of a LST and a list of conditions which make a LST feasible. Examples of published LSTs are *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 ‘LST’ refers to data structure and not data collection. It is used in relation to situations in which a small number of outcomes are measured. The term may not adequately reflect the complexity of the studies undertaken.

### 4.4.2. Randomised database studies

Randomised database studies can be considered a special form of an LST 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 Chapter 17 of *Pharmacoepidemiology and Therapeutic Risk Management* (A.G. Hartzema, H.H. Tilson and K.A. Chan, Editors, 1st Edition, Harvey Whitney Books Company, 2008), 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.

Another use of databases in RCT is the long-term follow-up of patients in observational studies after RCT termination, for example to assess long-term safety and effectiveness at regular intervals using objective outcomes.

### 4.5. Systematic review and meta-analysis

Several studies may be available to answer a research question, and it is important to identify and integrate this evidence. In epidemiology, the focus of this activity is often more to learn from the diversity of designs, results and associated gaps in knowledge than to obtain risk estimates.

A systematic review is a review of the literature aiming to answer a specific and clearly formulated research question. Systematic reviews use systematic and explicit methods to identify, select and critically appraise relevant research, and to analyse the data included in the review. The methods used to minimise bias should be explicit and the findings should be reproducible as stated in the *Cochrane Handbook for Systematic Review of Interventions*.

For example, it has long been recognised that persons using NSAIDs are at a significantly increased risk of gastrointestinal complications, for instance, injury to the intestinal lining that can result in ulcers.
and/or gastrointestinal bleeding. To reduce the morbidity associated with NSAIDs, specific estimates for individual drugs and individual groups of patients with different risk profiles are needed. Therefore, a systematic review of a number of studies is appropriate to determine specific pharmacologic features of NSAID-associated gastro-intestinal toxicity and to explore multi-factorial determinants in the risk of gastro-intestinal bleeding among NSAID users, including clinical background, use of concomitant medications or a possible genetic susceptibility.

Meta-analysis is a statistical technique used to analyse and summarise the findings of a systematic review by quantitative pooling of the data from individual studies addressing the same question. Quantitative synthesis in systematic reviews (Ann Intern Med 1997;27:820-6) shows how meta-analysis can provide more precise estimates of the effects of healthcare than those derived from the individual studies. In addition, a meta-analysis evaluates the consistency of results across studies and facilitates the exploration of their heterogeneity (clinical, methodological and/or statistical). Indeed, as shown in Investigating causes of heterogeneity in systematic reviews (Stat Med 2002;21:1503-11), when very significant heterogeneity exists, the heterogeneity itself may deserve more emphasis than the pooled summary estimates. In addition to direct comparisons, consideration should be given to the appropriateness of indirect comparisons through network meta-analysis of existing trials.

Systematic review and meta-analysis can be conducted with different sources of information (including clinical trials and epidemiological studies) for the assessment of safety and tolerability profiles of therapeutic interventions. An example is provided in Risk of venous thromboembolism from oral contraceptives containing gestodene and desogestrel versus levonorgestrel: a meta-analysis and formal sensitivity analysis (Contraception 2001;64:125-33). Any systematic review and meta-analysis will, however, have the same limitations as the sources of information they use. There are also additional limitations pertaining to the actual statistical combination of data via a meta-analytic approach.

RCTs are considered the gold standard for causality assessment. They frequently have limitations relating to sample size, narrow population characteristics and indications, and short follow-up duration. Therefore RCTs alone and subsequent systematic review or meta-analysis of RCTs will not address issues relating to the incidence of diseases. They have little value in detecting rare events and in the evaluation of outcomes that are far in the future. Systematic review and meta-analysis of observational studies and other epidemiological sources are therefore becoming as common as those of RCTs. Challenges in systematic reviews that assess treatment harms (Ann Intern Med 2005;142:1090-9) explains the different reasons why both provide relevant information and knowledge for pharmacovigilance. An important element to consider is the source of information used in meta-analyses. It explains the different results found for the risk of adverse events associated with Cox-2 inhibitors in Vascular and upper gastrointestinal effects of non-steroidal anti-inflammatory drugs: meta-analyses of individual participant data from randomised trials (Lancet 2013;382:769-79), which used individual-participant data from RCTs, and Myocardial infarction and individual nonsteroidal anti-inflammatory drugs meta-analysis of observational studies (Pharmacoepidemiol Drug Saf 2013;22(6):559-70), which used summary-level data from publications.

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. The authors emphasise that GRADE addresses questions of alternative management strategies, interventions or policies but has not been developed for questions about risks or prognosis.

Section 3.3 further describes different approaches to integrating studies and pooling data.
4.6. Signal detection methodology and application

A general overview of methods of 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.

Quantitative analysis of spontaneous adverse drug reaction reports is increasingly used in drug safety research. The role of data mining in pharmacovigilance (Expert Opin Drug Saf 2005;4(5):929-48) explains how signal detection algorithms work and address questions regarding their validation, comparative performance, limitations and potential for use and misuse in pharmacovigilance. Quantitative signal detection using spontaneous ADR reporting (Pharmacoepidemiol Drug Saf 2009;18:427-36) 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;16:359–65) provide useful points to consider before incorporating data mining as a routine component of any pharmacovigilance program and review data mining methodologies and their limitations.

Multivariate methods such as logistic regression (that may use propensity score-adjustment) have the theoretical capability to reduce masking and confounding by co-medication and underlying disease. 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 OMOP and concludes that logistic regression performs better than traditional disproportionality analysis. 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.

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.

The Guideline on the use of statistical signal detection methods in the Eudravigilance data analysis system describes quantitative methods of disproportionality implemented in signal detection by the European Medicines Agency (EMA) together with the elements for their interpretation and their potential limitations in the frame of pharmacovigilance. It encompasses the use of quantitative methods in EudraVigilance applied to the evaluation of Individual Case Safety Reports (ICSRs) originating from healthcare professionals and involving authorised medicinal products.
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 allowing a comparison only to drugs for which the adverse event was not considered to be 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 section 4.8 of the Summary of Product Characteristics (SPC) of medicinal products authorised in the 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 ongoing initiatives develop observational data as electronic systems that will complement existing methods of safety surveillance e.g. the PROTECT, EU-ADR, OMOP and Mini-Sentinel projects.

The EU Guideline on good pharmacovigilance practices (GVP) Module IX - 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 pharmacoepidemiologic studies.

5. Statistical and epidemiological analysis plan

5.1. 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). This book also gives useful advice on how to plan complex hypotheses in a way that controls 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 design. 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’. 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 alternative answers to the questions. It will also explain the rationale behind this decision making process and the way that this rationale has influenced the study design. 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 in the same way 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. 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 chance observations in the data. 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.

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.

5.2. Statistical plan

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 source, the intended study population and the study design with discussion of strengths and weaknesses.

2. The effect measures and statistical models used to address each primary and secondary objective.

3. Formal definitions of any outcomes e.g. 'fatal myocardial infarction' 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.

4. Formal definitions for other variables – e.g. thresholds for abnormal levels of blood parameters.

5. 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 a relevant consideration 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. Blinding to exposure variables of evaluators making subjective judgments about the study.

7. Methods of adjusting for confounding, including

    7.1. Which confounders will be considered;
7.2. Criteria for any selection of a subset of confounders;
7.3. Methods for assessing the level of confounding adjustment achieved;
7.4. Sensitivity analyses for residual confounding.

8. Handling of missing data, including
8.1. How missing data will be reported;
8.2. Methods of imputation;
8.3. Sensitivity analyses for handling missing data;
8.4. How censored data will be treated, with rationale.

9. Fit of the model – if considered for a predictive model, including
9.1. Criteria for assessing fit;
9.2. Alternative models in the event of clear lack of fit.

10. Interim analyses – if considered:
10.1. 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.

11. How the achieved patient population will be characterised:
11.1. Description of target population;
11.2. Description of the analysis population if different, e.g. after PS matching or in IV analyses.

12. Treatment of multiplicity issues not elsewhere covered.

5.3. Handling of missing data

Missing data, or missing values, 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.

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 creating 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 (*Indicator and Stratification Methods for Missing Explanatory Variables in Multiple Linear Regression*. J Am Stat Assoc 1996;91(433):222–230).

A concise review of methods to handle missing data is also provided in the section Missing data of the *Encyclopedia of Epidemiologic Methods* (Gail MH, Benichou J, Editors. Wiley 2000). The relevance of defining the pattern of missing data is outlined, since some methods for handling missing data assume a defined pattern of missingness. Biased results may be obtained if it is 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, 1987) 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) and *Recovery of information from multiple imputation: a simulation study* (Emerg Themes Epidemiol 2012;9(1):3 doi: 10.1186/1742-7622-9-3).

6. Quality control and quality assurance

Quality control (QC) and quality assurance (QA) systems together constitute key quality systems that are parts of quality management (QM). QA defines the standards to be followed in order to meet the quality requirements for a product or service, whereas QC ensures that these defined standards are followed at every step. A third element of the quality of research is the introduction of an independent and objective audit of the QA/QC system and its outcomes. The book *Modern Approaches to Quality Control* (A.B. Eldin, Editor. Croatia: InTech Open Access, 2011) presents quality control processes in a variety of domains including, in Chapter 14, medical processes.

Rules, procedures, roles and responsibilities of QA and QC 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, AN George, SP Shearn, eds., Pharmaceutical Press, London, 2010), the *ICH Guideline for Good Clinical Practice E6(R1)*, 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):477-93).

A long standing rule for RCTs, the practice of quality assurance has now been introduced for observational studies. For post-authorisation safety studies, Article 36 of the *Commission Implementing Regulation (EU) No 520/2012* states that the marketing authorisation holder conducting a post-authorisation safety study shall ensure that all study information is handled and stored so as to allow for accurate reporting, interpretation and verification of that information, and that the analytical dataset and statistical programmes used for generating the data included in the final study report are kept in electronic format and are available for auditing and inspection. The *FDA’s Best Practices for Conducting and Reporting Pharmacoepidemiologic Safety Studies Using Electronic Health Care Data Sets* specifies that investigators should ensure that they are aware of the QA and QC procedures used by data holders and how procedures could affect data integrity and the study. It includes a list of topics which investigators should address in the design and conduct of the study. 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. It emphasises that the magnitude of 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:Doc29: 1-14) proposes measures of quality assurance that can be applied at different stages of noninterventional 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. Relevant guidance also includes the *ISPE GPP*, which points out (‘Archiving’ section) that copies of all quality assurance reports and audits should be included within the archived documents, the CIOMS *International Ethical Guidelines for Epidemiological Studies* and the AGENS, DGSMP and DGEpi *Good Practice in Secondary Data Analysis Version 2*. 

Guide on Methodological Standards in Pharmacoepidemiology (Revision 3) EMA/95098/2010 Page 31/74
The following articles are practical examples of quality control implementation in pharmacovigilance and pharmacoepidemiological 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 multisite study. It highlights that use of results of pharmacoepidemiological studies requires at least some consideration to and knowledge of the quality and validity of the data, ideally including some level of validation of the recording and coding for electronic data sets. It recommends that quality assurance should be mentioned in the study protocol and this may lead to proposals of sensitivity analyses.

- **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). It concludes that variability in interviewers' ability to ascertain and code information is a possible source of bias in interview-based case–control studies when blinding cannot be achieved.

- **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.

- **Feasibility study and methodology to create a quality-evaluated database of primary care data** (Inform Prim Care 2004;12(3):171-7) describes a study conducted to build and test a model for collection of computerised retrospective primary care data in the UK, to assess its quality for use in medical and pharmaceutical research. The main quality outcome measures were indicators of the completeness of data recording.

- **Validation and validity of diagnoses in the General Practice Research Database (GPRD): a systematic review** (Br J Clin Pharmacol 2010;69:4-14) assessed 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.


- **Data quality management in pharmacovigilance** (Drug Saf 2004;27(12):857-70) focusses on three first 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.

- **A systematic literature review: Prescribing quality indicators for type 2 diabetes mellitus and cardiovascular risk management** (Pharmacoepidemiol Drug Saf 2010;19(4):319-34) provides an example of the assessment of validity of existing prescribing indicators.

### 7. Communication

#### 7.1. Principles of communication

Aspects of research communication and reporting include, but are not limited to, reports to health authorities, sponsors, presentations in scientific fora, scientific publications, patient focused communications and websites.
The Declaration of Helsinki includes a section on Research Registration and Dissemination and Communication of Results stating that:

- Every research study involving human subjects must be registered in a publicly accessible database before recruitment of the first subject.
- Researchers, authors, sponsors, editors and publishers all have ethical obligations with regard to the publication and dissemination of the results of research. Researchers have a duty to make publicly available the results of their research on human subjects and are accountable for the completeness and accuracy of their reports.
- All parties should adhere to accepted guidelines for ethical reporting. Negative and inconclusive as well as positive results must be published or otherwise made publicly available. Sources of funding, institutional affiliations and conflicts of interest must be declared in the publication.
- Reports of research not in accordance with the principles of this Declaration should not be accepted for publication.

One of the objectives of the new EU pharmacovigilance legislation is to increase transparency as regards drug-safety issues. Regulation (EU) No. 1235/2010 (Art. 26) obliges the European Medicines Agency (EMA) to publish online protocols and public abstracts of post-authorisation safety studies (PASS) concerning centrally-authorised medicinal products and imposed as an obligation to the marketing authorisation. Directive 2010/84/EC (Art 102) specifies that Member States shall ensure that the public is given important information on pharmacovigilance concerns relating to the use of a medicinal product. Such information may include protocols and results of PASS. The Guideline on good pharmacovigilance practices (GVP) Module VIII - Post-authorisation safety studies also recommends, for all PASS, registration of study information (including the protocol, amendments to the protocol, progress reports and final study report) in the public register of PASS maintained by the EMA (EU PAS Register - currently the ENCePP E-Register). The purposes of the register are to increase transparency, reduce publication bias, facilitate collaborations within the scientific community and facilitate optimal use of pharmacoepidemiology and pharmacovigilance expertise in Europe by preventing unnecessary duplication of research. Registration of studies in the register is mandatory for studies seeking an 'ENCePP Study Seal'.

**7.2. Guidelines on communication**

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 Guidance on 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 all 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 Guidelines for Submitting Adverse Event Reports for Publication (Pharmacoepidemiol Drug Saf 2007;16(5): 581–7) introduce readers to the key elements that have to be included when someone wishes to report and publish results about adverse drug events. The information is clearly and coherently presented and the data are divided based on three levels of requests: ‘required’, ‘highly desirable’ and ‘if relevant’.

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. The
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 prescribe the reporting of observational research in a rigid format, but to address what should be the critical information that a publication on an observational study should contain. In this regard, the guidance provided is complete, with practical examples that facilitate interpretation and understanding of the recommendations, though it is of limited usefulness for the design and conduct of epidemiological research projects. The recommendations are limited to cohort, case-control, and cross-sectional studies, though other types of epidemiological studies might benefit from most of the recommendations at the time of drafting the manuscript. No recommendation on ethical considerations, ownership of data and criteria for establishing the authorship are given. This is considered a major limitation as these aspects are highly relevant for the reporting and publishing of studies.

The Meta-analysis of Observational Studies in Epidemiology (MOOSE) group has developed a consensus statement and recommendations for reporting meta-analyses of observational studies in epidemiology. It is equivalent to the STROBE Statement Guidelines for reporting observational studies and the Consolidated Standards of Reporting Trials Consolidated Standards for Reporting Trials (CONSORT) 2010 Statement for RCTs, in that they have communication as their primary objective and take the form of a list of minimum requirements for adequate reporting. The MOOSE article is quite similar to the others in its structure, scope, length and depth of detail and is useful for the declared audience of researchers, readers, reviewers and editors. The structure of the article is slightly confusing though, as the formal ‘Results’ includes subheadings such as ‘background’, ‘search strategy’, ‘results’ and ‘discussion’. The authors recommend a broad inclusion of studies and to conduct post-hoc sensitivity 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.

The Preferred Reporting Items for Systematic Reviews and Meta-analyses (PRISMA) Statement is an evidence-based minimum set of items for reporting in systematic reviews and meta-analyses consisting of a 27-item checklist and a flow diagram. 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 a quality assessment instrument to gauge the quality of a systematic review. PRISMA is a successor to the Quality of Reporting of Meta-analyses (QUORUM) Statement and the associated QUORUM flow chart.

Additional guidance is provided in the ENCePP Checklist for Study Protocols and Code of Conduct and the IEA GEP guideline that have been reviewed elsewhere in this Guide.

Some of the points that are emphasised by the cited guidelines are:

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

• Authorship should conform to the guidelines established by the International Committee of Medical Journal Editors (ICJME) ‘Uniform Requirements for Manuscripts Submitted to Biomedical Journals’.

For a case report (or series) on suspected adverse drug reactions, minimum requirements include an account of the patient’s medical history and disposition, a detailed account of the dispensed product (substances, brand, route of administration) and a detailed account of the adverse event (nature, timing, severity, outcome).

8. Legal context

8.1. Ethical conduct, patient and data protection

In Europe, EU and national laws are the keys to what may and may not be done with regard to data access, data linkage and consent issues, including such domains as human rights and duty of confidentiality. Therefore, while data custodians may have differing requirements related to what approvals are needed before data can be released, the requirements will fit within the overall need to meet all applicable EU and national legislation and guidelines for the actual study. This includes situations where multi-country studies are being conducted and there may be transfer of data or information.

The Declaration of Helsinki and the provisions of the legislation on the protection of individuals with regard to the processing of personal data and on free movement of such data, as laid down in Directive 95/46/EC and Regulation (EC) No 45/2001 of the European Parliament and of the Council, need to be followed in the EU in terms of the ethical conduct of studies. For interventional research, the Regulation (EU) No 536/2014 and the Guidelines for Good Clinical Practice (Commission Directive 2005/28/EC) apply. Marketing authorisation holders and investigators must also follow relevant national legislation and guidance of those Member States where the study is being conducted (Directive 2001/83/EC and Directive 95/46/EC).

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

Consideration of ethical issues, data ownership and privacy is an important part of the International Society for Pharmacoepidemiology (ISPE) guideline for Good Pharmacoepidemiology Practices (GPP), section IV. It includes a sub-section (IV.A) on protection of human subjects and a reference to the ISPE guidelines on Data Privacy, Medical Record Confidentiality, and Research in the Interest of Public Health. The ISPE GPP also recommends a stand-alone section within a study protocol that contains a description of plans for protecting human subjects. Such a section should include consideration of the need for submitting the protocol to an Institutional Review Board/Independent Ethics Committee and the requirement of informed consent. The main scope of the International Epidemiological Association (IEA) Good Epidemiological Practice (GEP) guideline is on the ethical principles of pharmacoepidemiological field studies, which could also apply to interventional studies, such as the role of ethics committees, patients’ informed consent, use and storage of personal data and publication of results.

The Council for International Organisations of Medical Sciences (CIOMS) 2002 International Ethical Guidelines for Biomedical Research Involving Human Subjects prepared in collaboration with the World Health Organisation (WHO) consist of a statement of general ethical principles, a preamble and guidelines indicating how the ethical principles that should govern the conduct of biomedical research involving human subjects could be effectively applied. The CIOMS 2009 International Ethical Guidelines for
Epidemiological Studies set forth ethical guidance on how investigators - as well as those who sponsor, review, or participate in the studies they conduct - should identify and respond to the ethical issues that are raised by such research.

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 useful although focused on US law.

The Uniform Requirements for Manuscripts Submitted to Biomedical 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.

It should be kept in mind that the applicability of ethical requirements varies based on the nature of the inquiry and the studies to be conducted. Certain human subject protections applicable to clinical studies (e.g. full informed consent) would not apply to other kinds of research (e.g. review of data from de-identified medical records).

8.2. Pharmacovigilance legislation

New pharmacovigilance legislation has been implemented in the European Union (EU) since July 2012 (Regulation (EU) No. 1235/2010 and Directive 2010/84/EC). This legislation includes the possibility for regulatory authorities to impose on marketing authorisation holders the conduct of post-authorisation safety studies (PASS) as a condition of the marketing authorisation, a PASS being defined as “any study relating to an authorised medicinal product conducted with the aim of identifying, characterising or quantifying a safety hazard, confirming the safety profile of the medicinal product, or of measuring the effectiveness of risk management measures. Annex III of the Commission Implementing Regulation (EU) No 520/2012 provides the format of protocols, abstracts and final study reports for such PASS imposed as an obligation by a competent authority. Based on these formats, the European Medicines Agency (EMA) published detailed templates for the format and content of the protocol and final study report which it recommends to be used for all PASS.

The Guideline of good pharmacovigilance practices (GVP) Module VIII - Post-authorisation safety studies describes practical aspects for the implementation of the new legislation and the operation of the EU medicines regulatory network. It provides a general guidance on the development, conduct and reporting of PASS conducted by marketing authorisation holders, voluntarily or pursuant to an obligation. Of note, the legislation 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 timelines for the commencement of the study and the submission of the final study report to the competent authorities.

8.3. Reporting of adverse events/reactions

The European Union (EU) obligations to companies sponsoring a post-authorisation study are specified in Module VI of the Guideline on good pharmacovigilance practice (GVP) - Management and reporting of adverse reactions to medicinal products (text under revision at the time of finalisation of Revision 3 of this Guide). For a non-interventional post-authorisation study which is not sponsored by a marketing authorisation holder, there are no legal reporting obligations at the European level. Investigators should however enquire whether national obligations exist. Obligations or recommendations may also be
specified by an ethical committee or a data safety monitoring board. In all circumstances, the adverse events/adverse reactions reported during the study should be summarised in the study report.

If the study qualifies as an interventional trial, the reporting criteria laid down in Regulation (EU) No 536/2014 and Volume 10 of the Rules Governing Medicinal Products in the European Union should be followed.

9. Specific topics

9.1. Comparative effectiveness research

9.1.1. 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 comparisons 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 Agency for Health Care Research (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.

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).

9.1.2. 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 study design and conduct. 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 draft Methodological Report of the Patient Centered Outcomes Research Institute (PCORI) provides standards for patient-centered outcome research that aims to improve the way research questions are selected, formulated and addressed, and findings reported. 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.

9.1.3. Prominent issues in CER

9.1.3.1. Randomised clinical trials vs. observational studies

While randomised clinical trials (RCT) 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 ten distinguishing characteristics between pragmatic and explanatory trials to assist researchers in achieving the primary purpose of a trial. 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 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.

9.1.3.2. **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 (or from a subsample). *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 measurement methods and pharmacy claims data to data from electronic healthcare records participating in DARTNet. *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 randomized 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.

9.1.3.3. **Bias and confounding in observational CER**

Methodological issues and principles of Section 4 (Study design and methods) of the ENCePP Guide apply to CER as well as to research on safety 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* (AHRQ, Report No. 33, May 2012). More recently, high-dimensional propensity score (hdPS) 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, hdPS improved confounding control compared to conventional PS adjustment when benchmarked against results from randomized controlled trials. See section 4.2.3.3 of the Guide for an in-depth discussion of propensity scores.

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* (AHRQ, Report No. 32, May 2012) reviews published CER case.
studies in which investigators had used the incident user design, discusses its strength (reduced bias) and weakness (reduced precision of comparative effectiveness estimates) and provides recommendations to investigators considering to use this design.

9.2. Vaccine safety and effectiveness

9.2.1. Vaccine safety

9.2.1.1. General aspects

Specific aspects of vaccines to be considered in pharmacovigilance and pharmacoepidemiology have been highlighted in several documents. The report of the CIOMS/WHO Working Group on Definition and Application of Terms for Vaccine Pharmacovigilance emphasises that characteristics of the vaccine and the vaccinated population, settings and circumstances of vaccine administration and data analysis issues are worthy of special attention in vaccine safety monitoring. It also provides definitions and explanatory notes for the terms 'vaccine pharmacovigilance', 'vaccination failure' and 'adverse event following immunisation (AEFI)'. 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).

Methods for vaccine pharmacovigilance have been developed by the Brighton Collaboration, which provides resources to facilitate and harmonise collection, analysis and presentation of vaccine safety data, including case definitions, 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 describes pharmacovigilance principles, causality assessment procedures, surveillance systems and factors influencing the risk-benefit balance of vaccines.

9.2.1.2. Signal detection

The GVP Module P.I: Vaccines for prophylaxis against infectious diseases describes issues to be considered when applying methods of disproportionate reporting for signal detection 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. Vaccine-Based Subgroup Analysis in VigiBase: Effect on Sensitivity in Paediatric Signal Detection (Drug Saf 2012;35(4):335-346) further examines the effects of subgroup analyses based on the relative distribution of vaccine/non-vaccine reports in paediatric ADR data.

Comparing data mining methods on the VAERS database (Pharmacoepidemiol Drug Saf 2005; 14(9):601-9) compares four techniques: empirical Bayes geometric mean (EBGM), lower-bound of the EBGM's 90% confidence interval (EB05), proportional reporting ratio (PRR), and screened PRR (SPRR) and concludes the value of each method varies according to the situation.

9.2.1.3. Signal refinement

When prompt decision-making about a safety concern is required and there is insufficient time to review individual cases, the GVP Module P.I: Vaccines for prophylaxis against infectious diseases suggests to conduct observed vs. expected (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.

**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) and **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) 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).

Simple "snapshot" O/E analyses are easy to perform but 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 also affected by uncertainties regarding the numbers of vaccinated individuals and age-specific background incidence rates.

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 VSD 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 (large amount of underreporting and variation of reporting with time since vaccination). The method should be seen as a signal strengthening one for exploring a spontaneous reporting signal prior to a pharmacoepidemiologic study, if any. It 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).

### 9.2.1.4. Hypothesis testing studies

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 are 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 as they only involve cases:

- The case-coverage method 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 and permitting 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 an 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).

- The case-crossover method 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 is described in section 4.2.3 of the Guide.

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.
In situations where primary data collection is needed (e.g. a pandemic), the SCCS may not be timely 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 Safety 2012), 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 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).

Non-specific effects of vaccines, such as a decrease of mortality, have been claimed in observational studies but generally result from bias and confounding. Data collection in observational studies (Trop Med Int Health 2009;14(9):969-76.) and 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.

Ecological analyses should not be considered hypothesis testing studies. See section 4.4. of this Guide.

### 9.2.1.5. Meta-analyses

A systematic review evaluating the potential for bias and the methodological quality of meta-analyses in vaccinology (Vaccine 2007; 25(52):8794-806) provide 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.

### 9.2.1.6. Studies on vaccine safety in special populations

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. Observational studies on vaccine adverse effects during pregnancy (especially on pregnancy loss), which often use pregnancy registries or healthcare databases, are faced with three challenges: embryonic and early foetal loss are often not recognised or recorded, data on the gestational age at which these events occur are often missing, and 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):844-850) 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.

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 etiology of immunocompromise and possible limitations due to residual confounding, differences within and between etiological groups and small sample size in some etiological groups. Further research is needed on this topic.
9.2.2. Vaccine effectiveness

9.2.2.1. Definitions

Vaccine effects and impact of vaccination programmes in post-licensure studies (Vaccine 2013;31(48):5634-42) review and delineate, among the various evaluations of vaccine intervention, what applies to the effectiveness of vaccine and to the impact of vaccination programmes, propose epidemiological measures of public health impact, describe relevant methods to measure these effects and discuss the assumptions and potential biases involved.

9.2.2.2. Traditional cohort and case-control studies

Generic protocols for retrospective case-control studies and retrospective cohort studies to assess the effectiveness of rotavirus vaccination in EU Member States based on computerised databases were published by the European Centre for Disease Prevention and Control (ECDC). 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. 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.

9.2.2.3. 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.

9.2.2.4. 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.

9.2.2.5. Test negative design

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.

Since 2008/9, the Influenza Monitoring Vaccine Effectiveness (I-MOVE) network has estimated the
effectiveness of seasonal influenza vaccine to prevent medically attended influenza-like illness (ILI) using
laboratory confirmation as influenza in a test-negative case–control study based on a number of
European sentinel surveillance networks. Estimates of Pandemic Influenza Vaccine Effectiveness in
Europe, 2009–2010: Results of Influenza Monitoring Vaccine Effectiveness in Europe (I-MOVE)
Multicentre Case-Control Study (PLoS Med 8(1): e1000388. doi:10.1371/journal.pmed.1000388)
describes in detail the methodologies involved and how the test-negative design is a hybrid design
approaching a density case-control study but differs as former influenza cases in the pandemic are not
excluded from potential controls (ILI testing negative). Early estimates of seasonal influenza vaccine
effectiveness in Europe: results from the I-MOVE multicentre case–control study, 2012/13 (EuroSurveill
2013;18(7):pii=20400) puts the importance of providing early vaccine effectiveness estimates into
context including in terms of defining recommendations for subsequent seasonal vaccine composition.
2011-12 seasonal influenza vaccines effectiveness against confirmed A(H3N2) influenza hospitalisation:
showed that an international multicentre hospital based study is feasible to measure influenza vaccine
effectiveness and allows harmonised practices and collection of good quality data.

9.2.2.6. Case coverage design

This design is described in section 9.2.1.4

9.2.2.7. Impact assessment

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 the introduction of the vaccine to the
incidence/proportion of rotavirus cases in the period following the introduction of the vaccine. These
generic protocols need to be adapted to each country/regions and specific situation.

The impact of the 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. 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.
9.2.2.8. 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.

9.3. Design and analysis of pharmacogenetic studies

9.3.1. Introduction

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 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 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 (see Drug discovery: a historical perspective. Science 2000;287(5460):1960-4).

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

9.3.2. Identification of genetic variants

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

The first approach 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, when information of the genes from previous studies is missing or incomplete, it 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 generalisability of findings to populations of patients, to
classify 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.

9.3.3. Study designs

Several options are available for the design of pharmacogenetic studies. Firstly, randomised controlled
clinical trials (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

Observational studies are the alternative and can be family-based (using twins or siblings) or population-
based (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

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 ENCePP Guide on methodological standards in pharmacoepidemiology 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 recognizable 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. 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.

9.3.4. Data collection

The same principles and approaches to data collection as for other pharmacoepidemiological studies can be followed (primary or secondary data collection or combination of both; see section 3. 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 (Clim 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).

9.3.5. Data analysis

The focus of data analysis should be on the measure of effect modification (see section 4.2.4 of this Guide on Effect Modification). Attention should be given to whether the mode of inheritance is defined a priori (e.g. dominant, recessive or additive) 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 for the analysis 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).

9.3.6. Reporting


9.3.7. 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 (http://www.pharmgkb.org/page/cpic). Furthermore, several clinical practice recommendations have been published, for instances 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...

9.3.8. Resources

An important pharmacogenomics knowledge resource is available through PharmGKB (http://www.pharmgkb.org) 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

All hyperlinks in the document were last accessed on-line on 4 July 2014.

References – Chapter 1

Agency for Healthcare Research and Quality (AHRQ).
http://www.ahrq.gov/

http://www.effectivehealthcare.ahrq.gov/index.cfm/search-for-guides-reviews-and-reports/?pageaction=displayproduct&productid=1166

ENCePP Checklist for Study Protocols


Guidance on the format and content of the Protocol of non-interventional post-authorisation safety studies

ISPE Guidelines for Good Pharmacoepidemiology Practices
http://www.pharmacoepi.org/resources/guidelines_08027.cfm

References – Chapter 2

http://www.mrgibbs.com/tu/research/articles/literature_reviews_researched.pdf

ENCePP Checklist for Study Protocols

Guidance on the format and content of the Protocol of non-interventional post-authorisation safety studies

ISPE Guidelines for Good Pharmacoepidemiology Practices
http://www.pharmacoepi.org/resources/guidelines_08027.cfm

References – Chapter 3


http://www.effectivehealthcare.ahrq.gov/search-for-guides-reviews-and-reports/?pageaction=displayproduct&productID=1897


ARITMO
http://www.aritmo-project.org/


CIOMS. Working Group VIII Practical Aspects of Signal Detection in Pharmacovigilance
IEA Good Epidemiological Practice Guideline
http://ieaweb.org/good-epidemiological-practice-gep/

International Classification for Primary Care (ICPC-2)
http://www.who.int/classifications/icd/adaptations/icpc2/en/

International Classification of Diseases (ICD)
http://www.who.int/classifications/icd/en/

International Society for Pharmacoeconomics and Outcomes Research (ISPOR)
http://www.ispor.org/

ISPE Guidelines for Good Pharmacoepidemiology Practices
http://www.pharmacoepi.org/resources/guidelines_08027.cfm

http://www.ispor.org/TaskForces/documents/RDPartIII.pdf

http://annals.org/article.aspx?articleid=706784


http://www.bmj.com/content/344/bmj.d8012


http://dii.sagepub.com/content/42/5/409.short

Mini-Sentinel
http://www.mini-sentinel.org/

Note for Guidance on Good Clinical Practice

PRISM (Post-licensure Rapid Immunization Safety Monitoring)
http://mini-sentinel.org/work_products/PRISM/PRISM_Summary.pdf

PROTECT
http://www.imi-protect.eu/

http://www.ncbi.nlm.nih.gov/pubmed/PMC2914827/?tool=pubmed


SAFEGUARD
http://www.safeguard-diabetes.org/


http://link.springer.com/article/10.1007%2FBF00607096
References – Chapter 4

http://ncbi.nlm.nih.gov/pubmed/24478163

http://ncbi.nlm.nih.gov/pubmed/12543843

http://smm.sagepub.com/content/18/1/67.short


http://ncbi.nlm.nih.gov/pubmed/21805529
http://aje.oxfordjournals.org/content/150/4/327.abstract

http://onlinelibrary.wiley.com/doi/10.1002/(SICI)1099-1557(200001/02)9:1%3c37::AID-PDS471%3e3.0.CO;2-U/abstract


http://www.nature.com/clpt/journal/v91/n5/full/clpt201226a.html

http://aje.oxfordjournals.org/content/170/5/667.full


http://www.bmj.com/content/340/bmj.b5087

http://www.ncbi.nlm.nih.gov/pmc/articles/PMC1773795/


http://aje.oxfordjournals.org/content/133/2/144.abstract


Medical Dictionary for Regulatory Activities (MedDRA)  
http://www.meddra.org/

http://aje.oxfordjournals.org/content/104/6/609.short

Mini-Sentinel  
http://www.mini-sentinel.org/


http://www.clinmed.rcpjournal.org/content/9/5/486.full

OMOP Observational Medical Outcomes Partnership  
http://omop.org/

PROTECT  
http://www.imi-protect.eu/

PROTECT Database of adverse drug reactions (EU SPC ADR database)  

http://aje.oxfordjournals.org/content/173/12/1404.short


http://aje.oxfordjournals.org/content/158/9/915.long


References – Chapter 6
AGENS, DGSMP and DGEPi Good Practice in Secondary Data Analysis Version 2

http://www.effectivehealthcare.ahrq.gov/search-for-guides-reviews-and-reports/?pageaction=displayproduct&productID=1897


CIOMS. International Ethical Guidelines for Epidemiological Studies


DURQUIM Indicators of prescribing quality in drug utilisation research
http://www.pharmacoepi.org/pub/1c093fe0-2354-d714-51c0-6095720ead95

http://www.intechopen.com/books/modern-approaches-to-quality-control

EuroDURG Quality Indicator Meeting (DURQUIM)
http://www.pharmacoepi.org/eurodurg/durquim.cfm

European Forum for Good Clinical Practice (EFCGP)
http://www.efgcp.be

European Forum for Good Clinical Practice (EFCGP) Guidelines
http://www.efgcp.be/Publications.asp?Tйтpe=EFGCP%20publications&L1=13&L2=1


Imperial College Academic Health Science Centre (AHSC). Quality Control and Quality Assurance SOP JRCO/SOP/025. [https://workspace.imperial.ac.uk/clinicalresearchgovernanceoffice/Public/JRCO_SOP_025_QCQA%20Final%202012.pdf]


References – Chapter 7

CONSORT statement [http://www.consort-statement.org/]


ENCePP E-Register of Studies [http://www.encepp.eu/encepp/studiesDatabase.jsp]

ENCePP Study Seal
References – Chapter 8

Agency for Healthcare Research and Quality (AHRQ).
http://www.ahrq.gov/

http://www.effectivehealthcare.ahrq.gov/search-for-guides-reviews-and-reports/?pageaction=displayproduct&productID=1897

CIOMS. International Ethical Guidelines for Biomedical Research Involving Human Subjects.

CIOMS. International Ethical Guidelines for Epidemiological Studies


Council for International Organisations of Medical Sciences (CIOMS)
http://www.cioms.ch/

Declaration of Helsinki


Guidance on the format and content of the final study report of non-interventional post-authorisation safety studies

Guidance on the format and content of the Protocol of non-interventional post-authorisation safety studies

Guideline on good pharmacovigilance practices (GVP) Module VI – Management and reporting of adverse reactions to medicinal products.

Guideline on good pharmacovigilance practices (GVP) Module VIII - Post-authorisation safety studies.


ICMJE Uniform Requirements for Manuscripts Submitted to Biomedical Journals
http://www.icmje.org/urm_main.html

IEA Good Epidemiological Practice Guideline
http://ieaweb.org/good-epidemiological-practice-gep/

International Committee of Medical Journal Editors (ICJME)
http://www.icmje.org/

International Epidemiological Association (IEA)
http://www.ieaweb.org/

International Society for Pharmacoepidemiology (ISPE)
http://www.pharmacoepi.org/

ISPE Guidelines for Good Pharmacoepidemiology Practices
http://www.pharmacoepi.org/resources/guidelines_08027.cfm

ISPE Guidelines on Data Privacy, Medical Record Confidentiality, and Research in the Interest of Public Health.
http://www.pharmacoepi.org/resources/privacy.cfm


Volume 10 Clinical trials guidelines
http://ec.europa.eu/health/documents/eudralex/vol-10/

World Health Organisation (WHO)
http://www.who.int/en/

References – Chapter 9

http://www.effectivehealthcare.ahrq.gov/index.cfm/search-for-guides-reviews-and-reports/?pageaction=displayproduct&productid=1166


AHRQ. What Is Comparative Effectiveness Research.
http://effectivehealthcare.ahrq.gov/index.cfm/what-is-comparative-effectiveness-research1/


http://www.plosone.org/article/info%3Adoi%2F10.1371%2Fjournal.pone.0028435


Brighton Collaboration https://brightoncollaboration.org/public


DARTNet Institute
http://www.dartnet.info


http://www.bmi.com/content/343/bmj.d3908?view=long&pmid=21750072


ECDC Impact of rotavirus vaccination - Generic Study Protocol.


European Centre for Disease Prevention and Control (ECDC)

European Network for Health Technology Assessment (EUnetHTA)
http://www.eunethta.eu/


GRACE (Good ReseArch for Comparative Effectiveness)  
http://www.graceprinciples.org

GRADE working group. 6-part 2008 BMJ series.  

GRADE working group. Journal of Clinical Epidemiology series of articles.  
http://www.gradeworkinggroup.org/publications/JCE_series.htm

Grading of Recommendations Assessment, Development, and Evaluation (GRADE) working group.  
http://www.gradeworkinggroup.org/index.htm


http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2677258/

http://www.cceb.upenn.edu/pages/hennessy/SA14-14-2_Case_Control_Study.pdf


I-MOVE in Europe  
https://sites.google.com/site/epiflu/

Influenza - Monitoring Vaccine Effectiveness (I-MOVE) in Europe project.  
https://sites.google.com/site/epiflu/Home
http://www.pcori.org/assets/PCORI-Methodology-Standards.pdf


http://aje.oxfordjournals.org/content/158/9/915.long


http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2905665/


http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2905665/

http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2905665/


Vaccine Adverse Event Reporting System (VAERS)  
http://vaers.hhs.gov/index

Vaccine Safety Basics e-learning course (World Health Organization)  
http://www.vaccine-safety-training.org/home.html
Vaccine Safety Datalink (VSD)
http://www.cdc.gov/vaccinesafety/Activities/VSD.html


http://www.eurosurveillance.org/ViewArticle.aspx?ArticleId=20400


http://www.bmj.com/content/344/bmj.e55


http://apps.who.int/iris/bitstream/10665/80670/1/9789241505338_eng.pdf


ENCePP Guide, Revision 3 - Authorship

Editor
Xavier Kurz and the ENCePP WG1

Design and layout
Dagmar Vogl

Main authors
1. General aspects of study protocol: Nicholas Moore, Hubert Leufkens, Annie Fourrier-Réglat
2. Research question: Hubert Leufkens, Nicholas Moore, Yola Moride
3.1. Primary data collection: Susana Perez-Gutthann, Hubert Leufkens
3.2. Secondary use of data: Susana Perez-Gutthann, John Parkinson
3.3. Research networks: Miriam Sturkenboom
3.4. Spontaneous report databases: Niklas Norèn, Jim Slattery
4.1. General considerations: Yola Moride, Kevin Blake, Hubert Leufkens
4.2.1. Definition and validation of drug exposure, outcomes and covariates: Alejandro Arana, Gillian Hall, Jesper Hallas, Rachael Boggon
4.2.2. Bias and confounding: Olaf Klungel, Jesper Hallas, Mirko Di Martino, Ursula Kirchmayer, Alejandro Arana
4.2.3. Methods to handle bias and confounding: Rolf Groenwold, Jesper Hallas, Olaf Klungel
4.2.4. Effect modification: Olaf Klungel, Frederica Pisa
4.3. Ecological analyses and case-population studies: Ursula Kirchmayer
4.4. Hybrid studies: John Parkinson, Tjeerd van Staa
4.5. Systematic review and meta-analysis: Victor Kiri, Nawab Qizilbash
4.6. Signal detection methodology and application: Niklas Norèn, Jim Slattery
5. Statistical and epidemiological analysis plan: Jim Slattery, Alejandro Arana
6. Quality control and quality assurance: Milena Jadrijević-Mladar Takač
7. Communication: Kevin Blake, Susana Perez-Gutthann
8. Legal context: Xavier Kurz, Kevin Blake
9.1. Comparative effectiveness research: Annie Fourrier-Réglat, Olaf Klungel, Frederica Pisa
9.2.2. Vaccine effectiveness: Kevin Blake, Alain Moren
9.2.3. Design and analysis of pharmacogenetic studies: Olaf Klungel, Bruce Carleton (in collaboration with the ISPE SIG “Molecular Epidemiology, Biomarkers and Pharmacogenetics”)
**Acknowledgments**

Peer review of one or several chapters was received from:
Ana Marta Anes
Peter Arlett
Gillian Bartlet
Ulf Bergman
Jan Bonhoeffer
Nada Božina
Gonzalo Calvo
Ann Daly
Martin Daumer
Marina Davoli
Mark de Groot
Frank DeStefano
Nancy Dreyer
Paddy Farrington
Xavier Fournie
Amalia Issa
Hervé Le Louet
Geoff Liu
David Martin
Sarah McFarlane
Consuelo Pedros
Munir Pirmohamed
Elisabetta Poluzzi
Stefanie Prilla
Kenneth J. Rothman
Annalisa Rubino
Sebastian Schneeweiss
Ivana Silva
Gunilla Sjolin-Forsberg
Camilla Stephens
Brian Strom
Samy Suissa
Michael Theodorakis
Lina Titievski
Robert Vander Stichele
Wei Zhou
Patrick Zuber