Pharmacoepidemiology

III.9

Edeltraut Garbe, Samy Suissa

9.1 Introduction

In the last two decades we have witnessed a tremendous progress in the medical sciences that has led to the development of a great number of new powerful pharmaceuticals. These new medicines enable us to provide much better medical care, but occasionally they will cause harm and give rise to serious adverse reactions that were unexpected from preclinical studies or premarketing clinical trials. Against this background, pharmacoepidemiology has developed as a scientific discipline at the interface between clinical pharmacology and clinical epidemiology (cf. Chap. III.8 of this handbook). Pharmacoepidemiology can be defined as the application of epidemiologic knowledge, methods, and reasoning to the study of the effects and uses of drugs in human populations (Porta-Serra and Hartzema 1997). The application of epidemiological methods – i.e. the use of nonexperimental observational techniques – , the epidemiological perspective with an emphasis on investigations in large unselected populations and long-term studies, the public health approach and the philosophy of epidemiology are all extended to the scope of clinical pharmacology, i.e. the study of the effects of pharmaceuticals in humans.

Pharmacoepidemiology investigates both beneficial and adverse drug effects. Its focus and the one that receives the greatest attention is the assessment of the risk of uncommon, at times latent, and often unexpected adverse reactions that present for the first time after a drug has been marketed. The greatest challenge of pharmacoepidemiology is then to quantify the risk of a drug accurately, relative to one or several alternatives.

The study of adverse drug effects poses a number of methodological difficulties that must be addressed by pharmacoepidemiological research designs: first, drug exposure is not a stable phenomenon. Drug prescription habits may change due to the development of new pharmaceuticals, better knowledge on already available medications or other reasons. Second, drug exposure can be sensitive to a great number of factors that may also be related to the outcome of interest, such as the indication for prescribing, potential contraindications for drug use, the natural course of the disease or disease severity and compliance. Third, the risk of an adverse drug reaction (ADR) is often not constant, but it may change over time which may have important implications for the design and interpretation of pharmacoepidemiology studies.

In this chapter, we will first discuss limitations of premarketing clinical trials; we will then describe the characteristics of spontaneous reporting systems which have been implemented by regulatory agencies and the pharmaceutical industry for postmarketing surveillance. Another issue will be the use of multipurpose cohorts and large administrative healthcare databases for drug effect studies, which have found widespread application in pharmacoepidemiology. We will further discuss several methodological aspects that are unique to pharmacoepidemiological research: as e.g. the phenomenon of "depletion of susceptibles" which is a form of selection bias; or "confounding by indication" which is also referred to as "confounding by disease severity" or "channelling bias". We will present the use of

propensity scores, as a tool to reduce confounding particularly in studies of intended drug effects, and will discuss newer approaches of studying drug effects, such as the case-crossover and case-time-control designs. Characteristics of drug utilization studies and their units of measurement will be discussed.

Limitations of Premarketing Clinical Trials 9.2

Prior to marketing, new drugs are subjected to preclinical animal studies followed by three phases of clinical trials in humans. These phases are divisions of convenience in what is a continuous process of acquiring knowledge on the effects of a new drug in humans. In *phase I studies* humans are exposed to a new drug for the first time. These studies are often conducted in small numbers of healthy volunteers and are intended to explore the tolerability, pharmacokinetic and pharmacodynamic properties of a new drug in humans. In *phase II studies* the optimal dose range of the new drug is investigated and its efficacy and safety are explored in the intended patient population. These studies usually include several hundreds of patients. *Phase III studies* are aimed to prove the efficacy and safety of the new drug under strictly controlled experimental conditions in a larger patient population. They are mostly conducted as randomised controlled clinical trials (RCT) and often include several thousand patients altogether. In spite of the size of phase III studies, these studies have still limited ability to identify rare ADRs, since this would require an even larger number of study participants.

Table 9.1 displays sample size calculations for prospective studies. It shows the number of patients needed to detect a relative risk of a given magnitude in relation to the incidence of the event in the reference group. We can see that for the detection of very rare ADRs with sufficient statistical power, prohibitively large sample sizes in premarketing clinical studies would be needed. It is thus inherent in the drug development process – taking into account the already high cost for development of new pharmaceuticals – that serious rare ADRs will usually only be detected after drug marketing when the drug has been used in large patient populations.

An investigation by the Food and Drug Administration (FDA) into the withdrawal of five pharmaceuticals from the US market between 1997 and 1998 illustrates this point (Friedman et al. 1999). All five drugs were removed from the US market because of the discovery of unexpected serious adverse drug reactions (ADR) in the postmarketing period. The FDA investigated whether this unexpectedly high number of drug removals in only a 12-month-period was related to the expedited drug review and approval process that had been implemented. They calculated the number of patients exposed in the clinical trials before marketing and the approximate number of patients exposed before drug withdrawal (Table 9.2). The figures demonstrate that usually huge numbers of patients need to be exposed before sufficient knowledge on a rare ADR has been accumulated. For example, serious hepatotoxic effects of bromfenac occurred in approximately 1 in 20,000 patients who took the drug for longer than 10 days (Friedman et al. 1999).

Table 9.1. Sample sizes for detection of a given drug risk in a RCT or cohort study (size per study arm)

	Incidence of outcome in the control group			
Relative risk	1/50,000	1/10,000	1/5000	1/1000
2.0	1,177,295	235,430	117,697	23,511
2.5	610,446	122,072	61,025	12,187
3	392,427	78,472	39,228	7832
5	147,157	29,424	14,707	2934
7.5	78,946	15,783	7888	1572
10	53,288	10,652	5323	1059

Calculations are based upon a two-sided significance level α of 0.05, a power of 80% ($\beta = 0.2$), and one control subject per exposed subject

Table 9.2. Drug removals from the US Market between 1997 and 1998. Number of patients exposed to withdrawn drugs in clinical trials compared to actual use after marketing

Removed drug	Number of patients exposed before marketing[1]	Approximate exposure prior to withdrawals
Terfenadine	5000	7,500,000
Fenfluramine	340	6,900,000
Dexfenfluramine	1200	2,300,000
Mibefradil	3400	600,000
Bromfenac	2400	2,500,000

[1] number of patients included in the US premarketing studies

To reliably detect this toxic effect, some 100,000 patients would need to be included in the premarketing clinical studies.

In addition, premarketing clinical studies differ from routine clinical care for a number of other reasons:

1. These studies mostly include a selected study population, defined by strict inclusion and exclusion criteria, which is often not fully representative of subsequent users of the drug. It is well known that premarketing studies tend to under-represent the elderly, patients with comorbid conditions, pregnant women and children. Patients in premarketing trials may even be considered a selected group of patients just because they are willing or able to participate.
2. Premarketing clinical trials are performed at selected sites which are typically better equipped than routine care facilities. They are conducted by specialists in their field and all participating persons have been specially informed and trained. Surveillance of patients is almost by definition more intensive than in routine clinical treatment, if only one considers the frequency and spectrum of laboratory tests or the assessment of therapeutic and unwanted effects.

3. Treatment regimens in premarketing clinical trials are largely fixed and allow almost no individual treatment variations. In contrast, adjustments are constantly made in routine care, depending on the progress of therapy and on the interaction between doctor and patient.
4. Premarketing clinical trials are usually of short duration. This renders it impossible to detect ADRs that only develop after a long induction period or after cumulative drug intake.

For all these reasons, crucial answers to questions of drug safety cannot be provided even by the most valid and complex phase III study.

Characteristics of Spontaneous Reporting Systems 9.3

Description 9.3.1

In the early 1960s, systems evolved in most Western countries that collected spontaneous reports on ADRs from doctors. Establishment of these spontaneous reporting systems was largely a consequence of the "thalidomide disaster", in which children exposed to the hypnotic thalidomide *in utero* were born with phocomelia, a congenital deformity of the limbs resulting from prenatal interference of the drug with the development of the fetal limbs (Wiholm et al. 2000). Worldwide, several thousand cases of limb malformations in newborns observed in the 1950s and 1960s were attributed to the use of thalidomide during pregnancy (Lenz 1987). Based on this experience, spontaneous reporting systems were set up to monitor drug safety in the postmarketing period. In these systems, physicians – in some countries also pharmacists, other health care professionals or patients – report the suspicion of an ADR to the country's drug regulatory agency or to the pharmaceutical company that is marketing the drug. Drug regulatory agencies exchange the ADR reports with the concerned pharmaceutical companies and vice versa. The reports are locally assessed, the reported adverse event terms are coded using a standardized international terminology as e.g. MedDRA (the Medical Dictionary for Regulatory Activities) and are entered in a computerized database. More than 60 countries also forward their ADR reports to the World Health Organization (WHO) Collaborating Centre for International Drug Monitoring in Uppsala (Bate et al. 2002). Through membership in the WHO Programme, one country can know whether similar ADR reports are being made elsewhere.

An ADR report usually contains the patient's demographic information including age and gender; the patient's weight and height; adverse event (AE) information including date and outcome of the event, description of the event, evidence and existing medical history; information of suspect and concomitant medicine(s),

including drug name, dose, application and indication for use; whether the event abated after drug use was stopped ("dechallenge"); and whether the event reappeared after the drug was reintroduced ("rechallenge"); and the reporter's name and address.

Each report is assessed for its causality with drug intake by trained reviewers. Causality assessment is usually based on a set of criteria which includes the time interval between drug administration and the onset of the ADR; the course of the reaction when the drug was stopped; the results of re-administration of the drug; the existence of other causes that could also account for the observed reaction such as patient comorbidity or concomitant drug treatment; the pattern of the adverse effect; and the existence of reliable and specific laboratory test results. Causality assessment is not based on the single case report, but will also take into account other available information on the drug(s). In France, causality assessment criteria have been built into an algorithm (the "official method of causality assessment") that is used throughout the country (Benichou 1994). The French method distinguishes "intrinsic imputability" which takes into account only the single case report information from "extrinsic imputability" which is based on all published data on all drugs. Overall, causality assessment from individual case reports is a complex task and often associated with a high degree of uncertainty, since confounding by concomitant drug therapy or the underlying disease can frequently not be ruled out. Rare exceptions are a positive rechallenge to the drug (which is mostly accidental and involuntary, since this is rarely without risks) or positive results of specific laboratory tests such as the detection of drug-dependent antibodies.

Spontaneous reporting systems have several important advantages. They are relatively inexpensive to operate with respect to staff and basic technical equipment. They have the potential to cover the whole patient population and are not restricted to either hospitalised patients or outpatients. Surveillance starts as soon as a drug is marketed and monitoring of drug safety continues throughout the whole postmarketing life cycle of a drug. The suspicion of an ADR is, in theory, based on the experience of all treating physicians and pharmacists. Spontaneous ADR reporting systems can provide an alert to very rare, but nevertheless potentially important drug toxicity. Spontaneous reporting systems have identified many new, i.e. previously unrecognised drug hazards as e.g. clozapine-induced granulocytopenia, captopril-induced cough and amiodarone-induced hepatotoxicity. Further examples can be found in the article by Rawlins et al. (1989).

9.3.2 Limitations

Many of the successes of spontaneous reporting systems have been in the recognition of ADRs occurring shortly after starting therapy. Spontaneous reporting schemes are much less effective in identifying reactions with a long induction period. An example is the oculomucocutaneous syndrome associated with exposure to practolol which was undetected by the yellow card spontaneous reporting

scheme in the UK (Venning 1983). For the same reason, spontaneous reporting schemes are not well suited to identify drug-induced cancerogenicity.

ADR reports often do not provide sufficient information to confirm that a drug caused an event. For example, the ADR report may not give enough details on comorbidity or other medications to rule out other possible causes for the event in a remote expert assessment. It may be impossible to exclude confounding by indication, i.e. that the cause for the reported adverse event is rather related to the indication for the drug than to the drug itself. As an example, depression has been reported with the anti-acne medication roaccutane (Wysowski et al. 2001). Depression could, however, also be related to psychological disturbances over severe acne in sensitive teenagers rather than to roaccutane itself.

Recognition of an ADR depends on the level of diagnostic suspicion of the treating physician and may be related to the nature of the adverse event. Some ADRs are more likely to be diagnosed and reported than others because of their known association with drug therapy. For example, acute agranulocytosis is attributable to drug treatment in about 60–70% of cases (Kaufman et al. 1996). It may therefore be more likely to be attributed to drug therapy than a disorder as e.g. acute myocardial infarction that is usually not related to drug treatment (Faich 1986).

Spontaneous reporting systems suffer from serious underreporting of ADRs and various biases that affect reporting. Even in the UK, a country with a relatively high reporting rate in relation to its population size, rarely more than 10% of serious ADRs are notified to the regulatory agency (Rawlins et al. 1989). In France, a recent comparison of ADR reports with data about drug-induced hospitalizations in three pharmacoepidemiology field studies indicates that only 5% of ADRs leading to hospitalisations are actually reported in the spontaneous reporting system (Begaud et al. 2002). Lack of knowledge how to report an ADR and misconceptions about the type of ADR that should be reported are important reasons contributing to underreporting (Eland et al. 1999). On the other hand, it has to be taken into account that spontaneous reporting and published case reports may also lead to numerous false alarms.

Medical or mass media attention can stimulate reporting in a distorted manner and give rise to differential reporting in a dramatic way (Griffin 1986). An example of such "media bias" is that of central nervous side effects following treatment with the benzodiazepine triazolam. After van der Kroef published a case series of these ADR in 1979 (van der Kroef 1979), these side effects received extensive media coverage on Dutch television. As a consequence, the Netherlands received 999 ADR reports related to triazolam in 1979 out of a total of 1912 annual reports in 1979 overall (Griffin 1986). Reporting can also be affected by the market share of the drug; the quality of the manufacturer's surveillance system; reporting regulations and the length of time a drug is on the market (Griffin 1986; Lindquist and Edwards 1993). It has been shown that reporting rates do not remain stable over time, but usually peak during the first or second year after a drug has been introduced into the market and then progressively decline over the following years (Haramburu et al. 1992, 1997).

9.3.3 Statistical Approaches: Reporting Rates and Proportional Reporting Ratio

An ideal early warning system should not only recognize new hazards but also provide an estimate of their incidence. Spontaneous reporting systems may provide alerts of drug hazards, but they cannot be used to calculate incidence rates of adverse events related to a specific drug. Calculation of an incidence rate requires accurate numerator and denominator information, both of which are not available from the spontaneous reporting systems. First, the extent of underreporting for an individual drug is very difficult to assess and may even differ between drugs of the same pharmacological class. Second, the population exposed to the drug (i.e. the population at risk) is unknown and cannot be determined from drug sales data, since the duration of drug use, the dose regimen and compliance in individual patients are unknown.

Instead of the calculation of incidence rates, reporting rates (number of AE reports per market share) based on sales data are sometimes computed as an alternative approach (Pierfitte et al. 2000; Moore et al. 2003). Calculation of reporting rates is based on the assumption that the magnitude of under-reporting is reasonably similar for similar drugs that share the same indication, country and period of marketing (Pierfitte et al. 1999). The comparison of ADR reporting rates should therefore be restricted to drugs of the same category used for the same indication. Factors that may bias the comparison of reporting rates include differences in the length of time the drugs are on the market; differences in exposure populations; secular reporting trends; reporting variations; diagnosis and prescription variation; and the publicity of an ADR. Statistical corrections for year of marketing, secular trends of all-drug-all-adverse event reporting, and drug usage have been proposed (Tsong 1995). These adjustments do not, however, cover all possible sources of bias. In particular do they not erase concerns about differences in the magnitude of under-reporting for different drugs. Interpretation of reporting rates should therefore be conducted with an understanding of their limitations. Differences in reporting rates do not establish differences in incidence rates. They may, however, provide alerts of drug hazards to be investigated by more rigorous pharmacoepidemiological study designs.

The proportional reporting ratio (PRR) has been proposed as another statistical approach for signal generation (Evans et al. 2001). Signals of drug hazards are usually not based on one single ADR report, but on a series of similar suspected reports. The ADR databases maintained by the regulatory authorities and the WHO contain a large number of reports suitable for aggregation, as e.g. 2.5 million reports in the WHO database (Bate et al. 2002), over 2 million reports in the Adverse Event Reporting System (AERS) database maintained by the FDA (Szarfman et al. 2002), and over 350,000 reports in the Adverse Drug Reactions On-line Information Tracking (ADROIT) database of the UK regulatory agency (Evans et al. 2001). The PRR involves calculation of the proportions of specified reactions or groups of

reactions for drugs of interest where the comparator is all other drugs in the database. The PRR is the quotient of $a/(a+c)$ divided by $b/(b+d)$ derived from the reported frequencies of all drug-event pairs in the database arranged in a two-by-two table (Table 9.3).

Table 9.3. Calculation of the proportional reporting ratio (PRR) (a, b, c, d are absolute frequencies of the according combination)

	Drug of interest	All other drugs in the database
Event of interest	a	b
All other events	c	d

The PRR behaves in a similar fashion as the relative risk, i.e. the higher the PRR, the greater the strength of the signal. Statistical association is tested using a chi-squared test on 1 degree of freedom for the null hypothesis of independence. Signals can then be identified based on the PRR, the value of the chi-squared test, and the absolute number of reports. In a proof-of-concept study, Evans et al. (2001) defined a signal as a PRR value of at least 2, a value of chi-squared test of at least 4 and a minimum of 3 cases. Using these criteria on the UK ADROIT data base for 15 newly marketed drugs, they identified 481 potential signals, 339 (70%) of which were recognised ADR, 62 (13%) were considered to be related to the underlying disease and 80 (17%) were signals requiring further evaluation. Statistical approaches such as calculation of PRRs are not a substitute for a detailed ADR review, but they may aid in the decision on which series of cases should be investigated next. Similarly, as already mentioned for reporting rates, the PRR may be affected by differential ADR reporting related to notoriety, surveillance and market size effects. A PRR above 1 may therefore just indicate a higher reporting of a possible reaction under a drug, but not necessarily a differential occurrence (Moore et al. 2003). Recently, some modified statistical approaches to ADR data have been proposed for signal generation (Bate et al. 1998; Szarfman et al. 2002). These more complex statistical approaches are, like the PRR, based on a comparison of observed versus expected frequencies of adverse events under a particular drug, but they differ in the way they relate all drug-event combinations in the database to each other and in the use of Bayesian versus frequentist statistical models.

Sources of Data in Pharmacoepidemiological Research 9.4

A great number of pharmacoepidemiology studies are being conducted as field studies, with data being collected for the specific hypothesis under study. These studies are sometimes conducted in an international setting to increase the number

of cases and to provide more timely results (Spitzer et al. 1996; Abenhaim et al. 1996; Anonymous 1995b, 1986). Increasingly, already existing data sources are being used for pharmacoepidemiological research. Existing data sources include multipurpose cohort studies or large health databases. Studies utilizing such data can be conducted more quickly and are less expensive than field studies, since the data have already been collected.

9.4.1 Multipurpose Cohorts

Multipurpose cohorts are designed to investigate many different research hypotheses. Their study population usually consists of a subset of a defined population that has not been assembled by a specific exposure, but by other factors. For example, in the US Nurses' Health Studies, study participants were defined by age, female gender and profession (Nurses' Health Study I: 121,700 female nurses aged 30 to 55 years at baseline in 1976; Nurses' Health Study II: 116,671 female nurses aged 25 to 42 years at baseline in 1989). If a multipurpose cohort is used to investigate an association between a specific drug exposure and a disease, its cohort members will usually have sufficient variability in their exposure status for the drug to be investigated: they may currently be exposed or non-exposed, they may be exposed to different doses of the drug, they may have been exposed in the past or they may be exposed in the future. If, in addition, disease occurrence and relevant confounder information has been ascertained, the multipurpose cohort data may be used to investigate a specific pharmacoepidemiological hypothesis. The US Nurses' Health Studies have been extensively used for pharmacoepidemiology research questions. Examples include the association between nonsteroidal anti-inflammatory drugs and risk of Parkinson's disease (Chen et al. 2003), use of estrogens and progestins and risk of breast cancer in postmenopausal women (Colditz et al. 1995), postmenopausal estrogen and progestin use and risk of cardiovascular disease (Grodstein et al. 1996), oral contraceptives and the risk of multiple sclerosis (Hernan et al. 2000), aspirin, other nonsteroidal drugs and risk of ovarian cancer (Fairfield et al. 2002), calcium intake and risk of colon cancer (Wu et al. 2002) and many more associations (Grodstein et al. 1998; Hee and Grodstein 2003; Hernandez-Avila et al. 1990; Weintraub et al. 2002). Other multipurpose cohorts that have been less frequently used for pharmacoepidemiological research include the Health Professionals Follow-up Study (Giovannucci et al. 1994; Chen et al. 2003; Wu et al. 2002); the National Health and Nutrition Examination Survey I (NHANES) epidemiologic follow-up study (Lando et al. 1999; Funkhouser and Sharp 1995); the Framingham cohort study (Worzala et al. 2001; Abascal et al. 1998; Kiel et al. 1987; Felson et al. 1991); and the Rotterdam Study (Schoofs et al. 2003; Beiderbeck-Noll et al. 2003; Feenstra et al. 2002).

9.4.2 Record Linkage Studies

Large health databases have emerged as another important data source for pharmacoepidemiology research. In the United States and Canada, administrative health

databases have been set up for the administration of reimbursement payments to health care providers in nationally funded health care systems or managed care organizations. In the United Kingdom, Scotland and some other countries, large health databases consist of data entered by general practitioners (GP) into their practice computers.

Administrative Databases in the US and Canada

These databases usually consist of patient-level information from two or more separate files which can be linked via a unique patient identifier contained in each file. The unique patient identifier often consists of the social security number of the patient which is "scrambled" to ensure patient confidentiality. Information contained in the different files usually consists of demographic patient information; information on drug dispensations from pharmacies; information on hospitalisations; and information on ambulatory physician visits (Fig. 9.1). Through record linkage, person-based longitudinal files can be created for particular research questions. In some databases, record linkage is possible with cancer registries or birth malformation registries to investigate hypotheses of drug carcinogenicity or teratogenicity. Researchers usually have to submit a study protocol for review by an ethics committee and they only receive subsets of the files which are extracted to investigate the particular research hypothesis. Fees are charged for the time needed to extract the necessary data from the entire database. All statistical analyses are done on the anonymized data.

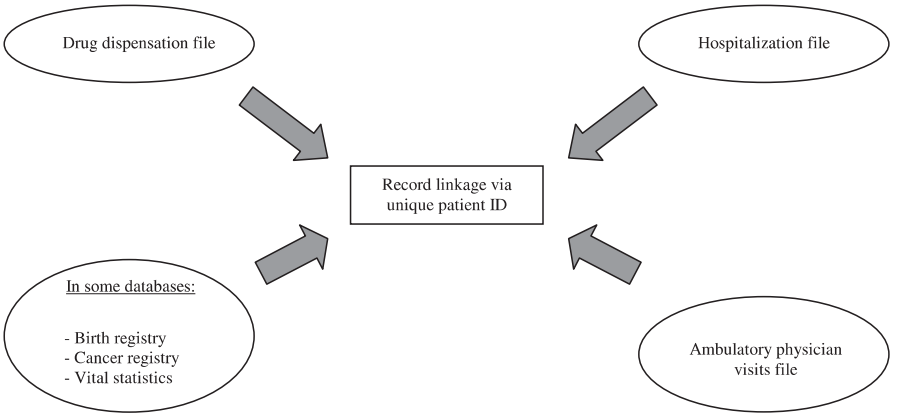

Figure 9.1. Record linkage in administrative health databases

A considerable number of administrative databases in the US and Canada are now available for pharmacoepidemiology research. A brief overview of some of these databases is given in Table 9.4. More detailed information on these databases can be found in the textbook "Pharmacoepidemiology" by Strom (2000).

Saskatchewan's Health Databases have been used extensively for pharmacoepidemiological studies. These databases will be used to illustrate which information may be expected in an administrative healthcare database (Table 9.5, adapted

Table 9.4. Examples of administrative health databases in the US and Canada

Database	Characteristics	Eligible population	Drug dispensations since
Saskatchewan's Health Databases, Saskatchewan, Canada	Provincial health plan	1 million	1975
RAMQ database, Quebec, Canada	Provincial health plan for the elderly	750,000	
Group Health Cooperative, Washington, US	HMO	460,000	1977
Kaiser Permanente, Northern California, US	HMO	2.8 million	1994 (from all pharmacies)
Kaiser Permanente Northwest Division, US	HMO	430,000	1986
Harvard Pilgrim Health Care, New England, US	HMO	1.1 million	
Tennessee Medicaid database, US	Health insurance for recipients of social welfare	1.4 million	1977
New Jersey Medicaid Database	Health insurance for recipients of social welfare	700,000	1980

HMO = Health Maintenance Organizations

from http://www.health.gov.sk.ca/mc_dp_phb_infodoc.pdf). Health Databases in Saskatchewan are based on the universal health insurance programme in this Canadian Province. Differently from the Medicaid program, there is no eligibility distinction based on socio-economic status. Record linkage is possible with the province's cancer registry. Medical records in hospitals are accessible upon approval from individual district health boards and affiliated facilities. Physician records may also be accessed for specific studies. A wide range of conditions has been validated by hospital chart review, including rheumatoid arthritis (Tennis et al. 1993), hip fractures (Ray et al. 1989), gastrointestinal bleeding (Raiford et al. 1996), asthma-related conditions (Spitzer et al. 1992) and others.

Physician-based Databases

The General Practice Research Database (GPRD) is a large physician-based computerized database of anonymized longitudinal patient records from hundreds of general practices in the UK, containing more than 35 million patient years of data. Currently, information is collected on approximately 3 million patients, equivalent to approximately 5% of the UK population. The database was created in June

Table 9.5. Information contained in different files of Health Databases in Saskatchewan

Population registry	Prescription drug database	Hospital separation[1] database	Physician services database
• Name • HSN[2] • Sex • Marital status • Date of birth • Date of death (if applicable) • Mailing address • Location code • If recipient of Saskatchewan welfare plan • Dates of coverage initiation and termination	*Patient information* • HSN[2] • Sex • Year of birth • Designation of special status[3] *Drug information* • Pharmacologic-therapeutic classification • Drug identification number • Active ingredient number of drug • Generic and brand names • Strength and dosage forms • Manufacturer of drug • Date dispensed • Quantity dispensed • "No substitution" indicator, if applicable *Prescriber information* • Prescriber identification number *Dispensing pharmacy information* • Pharmacy identification number *Cost information* • Unit cost of drug materials • Dispensing fee • Markup • Consumer share of total cost • Drug plan share of total cost • Total cost	*Patient information* • HSN[2] • Sex • Month and year of birth *Diagnostic and treatment information* • Up to 3 discharge diagnoses (4-digit ICD-9) • Up to 3 procedures (4-digit CCP[4]) • Accident code (ICD-9 external cause code) • Other • Admission date • Discharge date • Level of care codes • Length of stay • Admission and separation types • Case mix group • Resource intensity weight • Attending physician • Attending surgeon (if applicable) • Hospital identification number	*Patient information* • HSN[2] • Age • Sex • Location of residence • Indicator for registered Indian status *Physician information* • Physician specialty • Referring physician • Clinic • Age • Sex • Place and year of graduation • Practice type[5] • *Diagnostic and service information* • Date of service • Type of service • Primary diagnosis (3-digit ICD-9 code) • Location of service (e.g., office, inpatient, outpatient, home, other) • Billing information (amount paid, date of payment)

[1] separation defined as discharge, transfer, or death of an inpatient
[2] health services number [3] e.g. welfare recipient, palliative care registrant, long-term care home resident
[4] CCP: Canadian Classification of Diagnostic, Therapeutic, and Surgical Procedures
[5] solo, association, rural, urban

1987 as the Value Added Medical Products (VAMP) research databank. VAMP provided practice computers and general practice software to general practitioners (GPs) and, in return, GPs consented to undertake data quality training and to contribute anonymized data to a central database for subsequent use in public health

research. During the 1990s, VAMP research databank underwent several organisational and management changes. The database was renamed General Practice Research Database (GPRD) in 1994 when it was donated to the UK Department of Health. In 1999, management responsibility for the database was transferred to the UK Medicines Control Agency which became part of the newly created Medicines and Healthcare Products Regulatory Agency (MHRA). The database has been used extensively for pharmacoepidemiology and clinical epidemiology research. A bibliography of studies using GPRD data can be found on the webpage of the GPRD under www.gprd.com.

The database includes the following information: Demographics, including age and gender of patient (information on race is not collected); medical diagnosis, including comments; all prescriptions; events leading to withdrawal of a drug or treatment; referrals to hospitals; treatment outcomes, including hospital discharge reports where patients are referred to hospital for treatment; and miscellaneous patient information e.g. smoking status, height, weight, immunisations, and for a growing number of patients also lab results. Validation studies of the GPRD have shown that the recording of medical data into GPs' computers is almost complete (Garcia Rodriguez and Perez 1998).

Besides the GPRD, other physician-based databases are the MediPlus databases from IMS Health. The MediPlus databases are available in different countries and, like the GPRD, contain anonymized longitudinal patient records. Depending on the particularities of the respective health care system, different data are available for research. A description of the German IMS Disease Analyzer-MediPlus database can be found in an article by Dietlein and Schroder-Bernhardi (2002). The IMS databases have not been used extensively for pharmacoepidemiology research. Studies which examine data validity and comprehensiveness are mostly lacking. The German IMS Disease Analyzer-MediPlus database does not contain patient hospitalisation data. It also lacks diagnostic or treatment information from all physician specialists, since it is usually based on one panel of doctors only, e.g. on a panel of GPs and internists, or gynecologists, or urologists etc. The database derived from the panel of gynecologists would therefore not include information on ambulatory physician contacts in GPs' or internists' offices and vice versa. The German IMS Disease Analyzer-Mediplus database has been used for several drug utilization studies, e.g. on the dosing of cava-cava extracts (Dietlein and Schroder-Bernhardi 2003); whether hospitals influence the prescribing behavior of general practitioners (Schroder-Bernhardi and Dietlein 2002); or how doctors treat Helicobacter pylori infections (Perez et al. 2002) etc. In the UK, the MediPlus database contains similar patient information as the GPRD database.

9.4.3 Advantages and Limitations

Large health databases offer several important advantages:

1. They are usually large, with patient numbers ranging from several hundred thousand to well over several millions. This makes it possible to study rare adverse events of pharmaceuticals in large populations.

2. Medication information is usually more accurate than self-recorded exposure information. Drug histories obtained from patients have limited reliability particularly for drugs that are used only intermittently and not on a regular basis (Kelly et al. 1990). Database information probably represents the most accurate information on drug utilization that can be obtained in elderly patients (Tamblyn et al. 1995). In drug dispensation databases, fairly accurate information can be obtained on drug intake that occurred a long time ago. Exposure information is available also for patients who are deceased or too ill to answer questions, without having to rely on proxy information. There is no potential for recall or interviewer bias which is always of concern with primary data collection.
3. Because the data are collected in an ongoing manner as a by-product of health care delivery, epidemiological studies can be undertaken in reasonable time and at relatively low cost. The study variables are already available in computerized form and need not be obtained in time-consuming and expensive processes of data collection.
4. Some databases provide population-based data which cover the entire population of a geographical region and are thus fully representative of the population. Database studies do not require an informed patient consent and are therefore less prone to selection bias which may be a consequence of a low response rate in the study population.

Use of computerized databases for pharmacoepidemiology research is, however, not undisputed (Shapiro 1989) and there are a number of important limitations. A major concern is related to the validity of the diagnostic information contained in the database. In administrative health databases, diseases are primarily coded for billing and not for research purposes. There is no incentive for the health care provider to use specific codes as e.g. "duodenal ulcer with bleeding" instead of "upper gastrointestinal bleeding otherwise not specified". Diseases are often coded according to the International Classification of Diseases (ICD-9 or ICD-10 coding schemes) and many different ICD-codes may be compatible with the same disease process. A combination of several diagnostic codes into a single "broader" code may therefore be necessary (Garbe et al. 1997, 1998a). Strom and Carson (1990) have described this problem by stating that researchers using diagnostic codes in a computerized database must be "lumpers" rather than "splitters".

The validity of diagnostic coding also depends on the ability of a diagnostic code to rather selectively represent the condition in question and therefore varies with the condition. Strom conducted a validation study of ICD-9 coding of Stevens-Johnson Syndrome in the COMPASS Medicaid database in the US (Strom et al. 1991). Records of 3.8 million patients in five US states were searched for ICD9-CM code 695.1 which codes for Stevens-Johnson syndrome, but also for several other, less serious conditions. In an expert medical record review, only 14.8% of patients with ICD9-CM code 695.1 whose medical records could be reviewed were judged to have Stevens-Johnson syndrome. Thus, studies of Stevens-Johnson-Syndrome in

databases with ICD9-coding cannot be conducted without additional validation of the diagnosis. Whenever possible, validity of disease coding should be quantified for each condition studied.

Validation studies usually use the paper medical or pharmacy record as the "gold standard". Access to patient charts for validation purposes is often obtained via the scrambled social security (patient identification) number. All personal-identifying information is removed before copies of the charts are made available. The validation study by Strom also illustrates another problem: Only 51% of the medical records that were sought for in the study could actually be obtained (Strom et al. 1991). The authors state several reasons why they did not obtain access to the medical records: refusal of hospitals (30%); transcription errors (27%); translation of ID-number not possible (17%), no location of medical record possible (22%); other reasons (4%).

Administrative databases usually contain information on large numbers of patients, however, the amount of information per patient is limited:

1. Information about disease severity is mostly lacking and it may not be possible to exclude confounding by disease severity. In some instances, it is possible to construct an index of disease severity based on the patient's pharmacotherapy (Spitzer et al. 1992).
2. Relevant other confounder information for the association under study may not be contained in the database, creating a potential for bias. For example, most administrative databases do not contain information on smoking or alcohol use (Friedman et al. 2000) or age at menopause and reproductive history in women.
3. Administrative databases usually do not contain data on laboratory values or clinical measurements, although, in some databases, linkage with laboratory files has now become possible.
4. If relevant confounder information is missing and additional data collection required, it will make a study considerably more expensive and time-consuming, thereby diminishing some of the advantages connected with database research. It has to be decided on a case-by-case basis whether the information contained in a database is sufficient for the investigation of an association of interest and how much time and cost would be incurred if additional data have to be collected.

Although the medication information in databases is one of their major strengths, it also has some limitations: Information on drugs bought over-the-counter (OTC) and not prescribed by a physician is not available in the database; the patient's compliance with the prescription is unknown; in-hospital medication is usually not contained in the database; the prescribed daily dose is not documented in most databases and the average daily dose has to be calculated instead based on the duration of drug use and the quantity of drug prescribed; the prescription file does not contain the indication for drug prescribing, however, in many instances, this information may be deduced from diagnostic coding in the ambulatory physician file; medication data will be trun-

cated, if the database does not exist long enough or the database only includes elderly subjects. Truncation will limit the study of cumulative drug toxicity. Confounding by previous drug use may be avoided if a prior period of follow-up in the database is defined and the risk is only investigated in new users of the drug (Garbe et al. 1998b); many of the newest and/or most expensive drugs may not be available for study, if they are not included on the drug formulary.

Other important issues include whether the population contained in a database is representative of the source population and stable over time. For example, Saskatchewan Health Databases, which cover the population of the whole province of Saskatchewan, consist of a representative and fairly stable database population (Downey et al. 2003). In contrast, Medicaid Databases in the US are not representative of the US population, since they only include social welfare recipients and thereby over-represent children, females and non-whites in comparison with the total US population (Strom and Carson 1990). The skewedness of Medicaid Databases may not compromise the internal validity of a pharmacoepidemiology study conducted with these databases, but it may be a serious threat to its external validity, particularly when the data are being used for drug utilization research. In Medicaid Databases, turnover of the database population is high due to changing eligibility for Medicaid. Over a five-year time period, only 35% of Michigan and 38% of Tennessee Medicaid enrollees were still in the system, with loss of eligibility being greatest in children and young adults (Ray and Griffin 1989). High patient turnover may also make it difficult to locate patient files for validation studies as has been reported in the study by Strom (Strom et al. 1991). Data from health maintenance databases are also not fully representative of the US population. Members of these organizations tend to be less frequently black or poor and have higher educational achievements. Turnover in membership at HMOs is usually less than in Medicaid databases (Saunders et al. 2000; Friedman et al. 2000).

Methodological Approaches for Pharmacoepidemiology Studies 9.5

The strategies employed to verify hypotheses on drug risks or benefits are similar to those used in other fields of epidemiology. The case-control design is the design of choice for the investigation of rare drug risks, particularly if multiple countries are necessary to attain sufficient power, while the cohort approach is preferably used to assess the risk of more frequent events or if more than one outcome has to be considered simultaneously. The special nature of drug exposure and the availability of existing databases in pharmacoepidemiology have given rise to specific challenges and preferred solutions to estimate risk and benefit.

9.5.1 Case-Control Studies

Field studies that employ a case-control design (cf. Chap. I.6 of this handbook) are not common in the evaluation of the risks and benefits of prescribed drugs. Drugs available over-the-counter without prescription are not recorded systematically in computerized databases can therefore not be studied with linked databases. Thus studies of drugs such as analgesics, vitamin supplements, anorexiants etc, can only be studied by directly obtaining exposure information from subjects. With this design, cases with the outcome under study are identified from a given population, usually in hospitals or specialised clinics since they mostly involve serious outcomes. The approach to select the controls varies across studies. The International Agranulocytosis and Aplastic Anemia Study (IAAAS) that evaluated the effect of different analgesics on the risks of agranulocytosis and aplastic anemia used hospital-based controls (Anonymous 1986). The International Primary Pulmonary Hypertension Study (IPPHS) of the risk of primary pulmonary hypertension associated with anorexiant agents used patients treated by the same physician as the source of controls (Abenhaim et al. 1996). The Yale Hemorrhagic Stroke Study assessing the risk of diet and cough/cold remedies containing phenypropalonamine used population-based controls identified by random-digit dialling (Viscoli et al. 2001) (see also Chap. I.10 of this handbook). Finally, the Transnational study on oral contraceptive risks used both hospital and population-based controls (Spitzer et al. 1996). Such field studies that collect information from patients, physicians or medical charts are the exception because of the resources, expense and time required to complete the study.

Case-control studies using existing health databases are much more common. Besides the usual concerns with case-control studies in general, some are specific to studies conducted from databases. For example, the General Practice Research Database (GPRD) was used to evaluate the impact of inhaled corticosteroids on the risk of hip fracture (Hubbard et al. 2002). The entire GPRD was used to identify 16,341 cases of hip fracture and a random sample of 29,889 subjects selected as controls. This design is attractive because of its efficiency in using only a sample of subjects to estimate an effect for an entire population. Such an approach, however, can be deceiving for specific diseases such as asthma because of the illusion of large sample sizes. Indeed, all cases of hip fracture selected within a population such as the GPRD suggests a very large sample size for the study, along with the very large number of controls. However, in assessing the effect of inhaled corticosteroids, a drug pertinent exclusively to the population of asthma or chronic obstructive pulmonary disease (COPD) patients, a large proportion of the cases and the controls are in fact irrelevant to the question at hand. Thus, for example, for the GPRD case-control study of hip fracture risk, only 878 of the 16,341cases and 1335 of the 29,889 controls were subjects with asthma or COPD (Hubbard et al. 2002). With its 16,341 cases and 29,889 controls, the study appears at first more powerful than the Quebec study, based on 3326 cases of hip fracture and 66,237 controls selected from the asthma/COPD population (Suissa et al. 2004). In fact, the inference on the effect of inhaled corticosteroids is actually based on

many fewer subjects than believed. This point is particularly relevant for studies that find no significantly increased risk since, despite appearances, conclusions are based on fewer numbers of cases and controls with respiratory disease and thus lower power than expected. Furthermore, the estimate of effect may be biased if the outcome of interest, in this instance fractures, is also associated with the disease itself (asthma or COPD) and not only with the medications used to treat these conditions. In this case, the bias due to the association with the disease can only be eliminated by restricting the analyses to the population of patients who have the disease (Suissa et al. 2004).

Another issue in such database case-control studies is the manner by which controls are selected and particularly the index date for controls. In the GPRD study of the impact of inhaled corticosteroids on the risk of hip fracture, controls were matched to cases on age, sex, general practice and date of entry into the database (Hubbard et al. 2002). While the index date from which exposure was assessed was the fracture date for the cases, the index date for the controls was the same date as the matched case. To allocate such a date, one must be assured that the control is at risk on that date. Indeed, there could be control subjects with the same age, sex, general practice and date of entry, but who are dead or not in the practice at the time of the matched case's fracture. These will be necessarily currently "unexposed", which could bias upward the rate ratio.

Cohort Studies 9.5.2

Observational database studies that use a cohort design (cf. Chap. I.5 of this handbook) differ primarily with respect to their definition of cohort entry or time zero. The Saskatchewan asthma cohorts have defined asthma as well as its onset, by the dispensing of medications used to treat the condition, without the use of diagnostic codes from physicians (Blais et al. 1998b; Suissa et al. 2002). Patients were considered to have asthma as of the first time they received three prescriptions for an asthma medication, including bronchodilators, inhaled steroids and other asthma drugs, on at least two different dates within a one-year period. The date of the third prescription defined the onset and diagnosis of asthma and patients were then followed from that point on for the occurrence of asthma outcomes. Such a definition is not entirely accurate for two reasons: subjects with asthma may be hospitalized at their initial presentation and medications for asthma are used for other conditions such as COPD. In an attempt to exclude patients with COPD, age criteria were used, including only patients to the age of 44, and also excluding oral corticosteroids as one of the defining drugs for asthma.

Alternatively, cohort entry may be defined by calendar time. For example, a cohort formed from a health maintenance organization in eastern Massachusetts defined cohort entry as October 1, 1991 (Donahue et al. 1997; Adams et al. 2002). This cohort of 16,941 asthma patients was followed from this date or registration in the insurance plan to September 30th 1994. Such calendar time based definitions of cohort entry will inherently define cohorts with patients who have varying

durations of disease at time zero (cohort entry). Such a "prevalent" cohort, to be distinguished from an "incident" cohort defined by patients with new onset of asthma, can be subject to serious biases when evaluating the association between drug use and asthma outcomes. Indeed, if the risk of the asthma outcome and of being dispensed the drug under study are both associated with the duration of asthma, such prevalent cohorts will produce biased estimates of this association unless the duration of the condition can somehow be adjusted for. A source of selection bias for such prevalent cohorts is that the treatment itself may change because of prior events that are not included in the period of observation. For example, if a patient was hospitalized for asthma in the past and, as a result, was prescribed inhaled corticosteroids, such a patient may be at increased risk of a further hospitalization and of being dispensed inhaled corticosteroids subsequently. For such studies to be valid, information on the history of asthma prior to cohort entry, which includes the duration of the disease and prior outcomes such as asthma hospitalizations as well as prior drug exposures, are required for purpose of adjustment or for testing for effect-modification. A frequent problem with computerized database studies is that these historical data on the duration and history of the disease before cohort entry are rarely available.

The third type of cohort defines cohort entry by a specific clinical event, such as hospitalization, emergency room visit or a physician visit. Here again, these cohorts can be incident or prevalent if these cohort defining events are either the first one ever or rather the first to occur after a certain date. An example of this approach is a study from the Saskatchewan databases of asthma patients, with entry defined by the first time they received three prescriptions for an asthma medication within a one-year period, after a two-year span with no asthma medications. The study cohort consisted of all subjects hospitalized for asthma for the first time after cohort entry and followed until readmission. The use of inhaled corticosteroids subsequent to the first hospitalization was evaluated with respect to the rate of readmission (Blais et al. 1998a). A similar cohort definition was used with the Ontario database, although this cohort was based on elderly COPD patients and the COPD hospitalisation defining cohort entry was not necessarily the first one to occur in their disease (Sin and Tu 2001a).

9.5.3 Nested Case-Control Studies

The complexity in data analysis is greater in the field of database studies because of the technical challenges presented by their large size. Indeed, the asthma cohort formed from the health maintenance organization in eastern Massachusetts included 742 asthma hospitalisations occurring during the three year follow-up period (Donahue et al. 1997). With over 16,000 patients in the cohort, an analysis based on the Cox proportional hazards model with time-dependent exposure (cf. Chap. II.4 of this handbook) would require 742 risk sets (all patients in the cohort on the day of hospitalization) each containing approximately 16,000 observations with information on exposure and confounding factors measured at the point in time when the case occurred. Such an analysis would therefore require to generate

close to 12 million observations (742 × 16,000), each with dozens of variables. Another example is the cohort study that included over 22,000 elderly patients hospitalised for COPD in Ontario, Canada, of whom around 8000 either died or were re-admitted for COPD (Sin and Tu 2001a). A proper time-dependent analysis could include up to 140 million observations, creating a serious technical challenge in statistical computing. As a result of this complexity, the temptation to analyse these cohorts with exposures that are assumed not to change over time is attractive but, as described below, can cause severe immortal time bias (see Sect. 9.6.1).

Rather than analyzing such cohort studies with proper but complex time-dependent techniques, methods based on sampling can produce practically the same results at greater efficiency. The nested case-control design (cf. Chap. I.7 of this handbook), nested within the cohort, is precisely such an approach (Suissa 2000; Essebag et al. 2003). It is based on using data on all the cases with the study outcome that occur during cohort follow-up. These represent the case-series. A random sample of person-moments, namely time points during the subjects' follow-up, is then selected from all person-moments in the cohort to provide the control group for the nested case-control approach. For the cases, the index date, on which the timing of the exposure to the drug of interest is based, is simply the time at which the outcome occurred. For controls, the index date is the random person-moment(s) selected for that subject during follow-up, or the same point in time of the corresponding case (Suissa 2000). Because of the highly variable nature of drug exposure over time, it is important that person-moments are selected properly from all person-moments of follow-up for all members of the cohort. Thus, a subject may be selected more than once at different moments of their follow-up, and particularly person-moments preceding the index date of a case are valid control person-moments. For practical reasons and to conform to the Cox proportional hazards model, person-moments are usually selected from the risk set of each case. This approach involves identifying, for each case, all subjects who are at risk of the event at the time that the case occurred (the risk set) and controls are selected from this risk set (incidence density sampling, cf. Chaps. I.6 and I.7 of this handbook). Part of the simplicity of this approach is that all subjects in a risk set are allocated the same index date as the set-defining case.

The advantage of this approach is the direct relationship between the Cox proportional hazard model with time-dependent exposure and the conditional logistic regression analysis (cf. Chap. II.3 of this handbook) that is used to analyse such nested case-control data. Thus, instead of using exposure data on all members of the risk set, as the Cox model would require, data on only a few subjects (usually 4 or 10 controls per case) are sufficient to provide a very efficient estimator of the rate ratio. Such ease of data analysis with the large size databases that are used in pharmacoepidemiology is crucial. As an example, with the asthma cohort study of Donahue, if 10 controls per case were used for each of the 742 cases, the analysis would be based on 7420 observations instead of the 12 million observations necessary with the Cox model analysis.

In studying the effectiveness of drug treatment, one of the major problems is confounding by indication. The nested case-control approach becomes more useful, as it allows cases and controls to be matched on several measures of disease severity. Thus, the effect of a drug can be isolated, independently of the effects of the severity markers. Such a matched nested case-control study was used to evaluate the effectiveness of inhaled corticosteroids on asthma death (Suissa et al. 2000a). In that study, cases were identified from a cohort of over 30,000 asthma patients, from which 66 died of asthma. The Cox analysis would have required almost 2 million observations to be processed. For each case, however, the only members of the risk sets that were identified as controls were those with the same disease severity characteristics as the case, namely: prior hospitalization for asthma, oral corticosteroid use, number of canisters of beta-agonists, use of theophylline and nebulized beta-agonists. Thus, cases and controls were similar on all these severity markers, expect with respect to inhaled corticosteroids. As a result, the effect of inhaled corticosteroids could be assessed independently of these potential confounding factors.

9.5.4 Case-Crossover Design

Pharmacoepidemiology is frequently faced with the assessment of the risk of rare acute adverse events resulting from transient drug effects. Although the case-control approach can be used, the acuteness of the adverse event and the length of the drug's effect, as well as difficulties in determining the timing of drug exposure, induce uncertainty about the proper selection of controls. Moreover, confounding by indication may often be a problematic issue in such a design. In this situation, within-subject approaches have been proposed, including the case-crossover design and its extension the case-time-control design which was devised to counter time trend biases. The principle is that, when studying transient drug effects and acute outcome events, the best representatives of the source population that produced the cases are the cases themselves (cf. Chap. I.7 of this handbook).

To carry out a case-crossover study, three critical points must be considered. First, the study must necessarily be dealing with an acute adverse event which is alleged to be the result of a transient drug effect. Thus, drugs with regular patterns of use which vary only minimally between and within individuals are not easily amenable to this design. Nor are latent adverse events which only occur long after exposure. Second, since a transient effect is under study, the effect period (or time window of effect) must be precisely determined. An incorrect specification of this time window can have important repercussions on the risk estimate. Third, one must obtain reliable data on the usual pattern of drug exposure for each case, over a sufficiently long period of time.

The case-crossover study is simply a crossover study in the cases only. The subjects alternate at varying frequencies between exposure and non-exposure to the drug of interest, until the adverse event occurs, which happens for all subjects in the study, since all are cases by definition. With respect to the timing of the adverse

event, each case is investigated to determine whether exposure occurred within the predetermined effect period. In the VACCIMUS study of hepatitis B vaccination and the risk of a multiple sclerosis relapse, spontaneous reports indicated that such an effect could occur within two months of the vaccination (Confavreux et al. 2001). Thus, the case-crossover design used as the risk-period the 2-month period prior to the onset of the relapse and any vaccination in this period to determine exposure status. To obtain control exposure, data on the average drug use pattern are necessary to determine the typical probability of exposure to the time window of effect. This is done by obtaining data for a sufficiently stable period of time prior to time of the event occurrence and its exposure period. For the VACCIMUS study, there were four control periods consisting of the four 2-month periods prior to the 2-month risk period. The estimation of the odds ratio is based on any appropriate technique for matched data (4 controls per case), such as conditional logistic regression.

This design has been used in pharmacoepidemiology (Fagot et al. 2001; Neutel et al. 2002; Etienney et al. 2003; Ki et al. 2003; Confavreux et al. 2001; Barbone et al. 1998; Sturkenboom et al. 1995).

Case-Time-Control Designs 9.5.5

One of the limitations of the case-crossover design, particularly in the context of drug exposures, is that the exposure pattern may have changed over time, and particularly between the control and risk periods. For example, a rapid increase in vaccination rates over time during the span of the case ascertainment for the VACCIMUS study, particularly if this span had been short, would have biased the estimate of the odds ratio. Indeed, this estimate would also include the effect of the natural time trend in exposure. If control subjects are available, the case-time-control design can be used to separate the time effect from the drug effect (Suissa 1995). In simple terms, the time effect is estimated from the case-crossover odds ratio of exposure among the control subjects. The net effect of exposure on event occurrence is then computed by dividing the combined time and drug effect estimated from the case-crossover odds ratio of exposure among the case subjects by the time effect (cf. Chap. I.7 of this handbook).

The approach is illustrated with data from the Saskatchewan Asthma Epidemiologic Project, a study conducted to investigate the risks associated with the use of inhaled β-agonists in the treatment of asthma. Using databases from Saskatchewan, Canada, a cohort of 12,301 asthmatics was followed during 1980–87. All 129 cases of fatal or near-fatal asthma and 655 controls were selected. The amount of β-agonist used in the year prior to the index date, namely high (more than 12 canisters per year) compared with low (12 or less canisters), was found to be associated to the adverse event. Of the 129 cases, 93 (72%) were high users of β-agonists, compared with 241 (37%) of the 655 controls. The resulting crude odds ratio for high β-agonist use is 4.4 (95% confidence interval (CI): 2.9–6.7). Adjustment for all available markers of severity, such as oral corticosteroids and prior asthma hospitalizations as confounding factors, lowers the odds-ratio to 3.1 (95% CI: 1.8–5.4),

the "best" estimate one can derive from these case-control data using conventional tools.

The use of inhaled β-agonists, however, is known to increase with asthma severity which also increases the risk of fatal or near-fatal asthma. It is therefore not possible to separate the effects of the drug to the risk from that of disease severity, so that a within-subject design may be preferable. To apply the case-time-control design, exposure to β-agonists was obtained for the one-year current-period and the one-year reference-period. Among the 129 cases, 29 were currently high users of β-agonists and were low users in the reference period, while 9 cases were currently low users of β-agonists and were high users previously. The case-crossover estimate of the odds ratio is thus 29/9 (OR 3.2; 95% CI:1.5–6.8). However, the high use of β-agonists may have increased naturally over time, so that the control subjects were used to estimate this effect. Among the 655 controls, 65 were currently high users of β-agonists and were low users in the reference period, while 25 were currently low users of β-agonists and were high users previously, for an odds ratio of the time trend of 65/25 (OR 2.6; 95% CI:1.6–4.1) . The case-time-control odds ratio, using these discordant pair frequencies for a paired-matched analysis, is given by (29/9)/(65/25) = 1.2 (95% CI: 0.5–3.0). This estimate, which excludes the effect of unmeasured confounding by disease severity, indicates a minimal risk for these drugs.

The case-time-control approach provides a useful complement to the case-crossover design when the probability of drug exposure is not stable over time, particularly between the control and risk periods (Donnan and Wang 2001; Hernandez-Diaz et al. 2003). However, its validity is subject to several assumptions, including the homogeneity of the odds-ratio across subjects (Greenland 1996; Suissa 1998).

9.6 Some Methodological Challenges

9.6.1 Immortal Time Bias in Cohort Studies

A challenge of cohort studies is in their data analysis. Since drug therapy, the exposure of interest, often changes over time, data analysis must take this variability into account. However, such variability in exposure over time is not simple to incorporate in the analysis. Due to the complexity of such analyses, several of the studies mentioned above employed a time-fixed definition of exposure, by invoking the principle of intention-to-treat analysis. This principle, borrowed from randomised controlled trials, is based on the premise that subjects are exposed to the drug under study immediately at the start of follow-up. This information is unknown in database studies.

To emulate randomized controlled studies in the context of cohort studies, some authors have looked forward after cohort entry for the first prescription of the drug under study. In this way, a subject who was dispensed a prescription for such drug

was considered exposed and a subject who did not was considered unexposed. Different time periods of exposure assessment were used. For instance, in the context of COPD, a prescription for inhaled corticosteroids during the period of 90 days after cohort entry was used to define exposure (Sin and Tu 2001b). In other studies, periods of one year and three years were used to consider subjects exposed to inhaled corticosteroids in assessing their impact on mortality (Sin and Tu 2001a; Sin and Man 2002). This approach, however, leads to immortal time bias, a major source of distortion in the rate ratio estimate (Suissa 2003).

Immortal time bias arises from the introduction of immortal time in defining exposure by looking forward after cohort entry. Indeed, if exposed subjects were classified as such because they were observed to have been dispensed their first prescription for an inhaled corticosteroid 80 days after cohort entry, they necessarily had to be alive on day 80. Therefore, this 80-day period is immortal. While some exposed subjects will have very short immortal time periods (a day or two), others can have very long immortal periods. On the other hand, unexposed subjects do not have any immortal time, and in particular the subjects who die soon after cohort entry, with too little time to receive the drug under study. Therefore, the exposed subjects will have a major survival advantage over their unexposed counterparts because they are guaranteed to survive at least until their drug was dispensed.

This generation of immortal time in exposed subjects, but not in the unexposed subjects, causes an underestimation of the rate of the outcome among the exposed subjects. This underestimation results from the fact that the outcome rate in the exposed is actually composed of two rates. The first is the true rate, based on the person-time cumulated after the date of drug dispensing that defines exposure (post-Rx), while the second is that based on the person-time cumulated from cohort entry until the date of drug dispensing that defines exposure (pre-Rx). The first rate will therefore be computed by dividing all outcome events in that group by the first rate person-time, while the second rate will by definition divide zero events by the second rate person-time. For example, the rate in the exposed

$$\text{rate} = \text{deaths/total person-years}$$

consists in fact of two rates:

$$\text{rate pre-Rx} = 0/\text{person-years pre-Rx}$$

and

$$\text{rate post-RX} = \text{deaths/person-years post-Rx}\ .$$

The zero component of the rate will necessarily bring down the exposed rate. Since there is no such phenomenon in the unexposed group, the computation of the rate ratio will systematically produce a value lower than the true value because of the underestimation of the exposed rate. In particular, if the drug under study is altogether unrelated to the outcome, so that the true rate ratio is 1, this approach

will produce rate ratios lower than 1, thus creating an appearance of effectiveness for the drug.

The immortal time in exposed subjects also causes an overestimation of the rate of the outcome among the unexposed subjects. This is because the zero component of the rate in the exposed group should in fact be classified in the unexposed group. Indeed, subjects are in fact unexposed to the drug under study between cohort entry until the date of drug dispensing that defines exposure. They only start to be exposed after the drug is dispensed. Thus, the zero rate should in fact be combined with the unexposed rate.

Immortal time bias is thus the result of simplistic yet improper exposure definitions and analyses that cause serious misclassification of exposure and outcome events. This situation is created by using an emulation of the randomised controlled trial to simplify the analysis of complex time-varying drug exposure data. However, such studies do not lend themselves to such simple paradigms. Instead, time-dependent methods for analysing risks, such as the Cox proportional hazard models with time-dependent exposures or nested case-control designs, must be used to account for complex changes in drug exposure and confounders over time (Suissa 2003, 2004; Samet 2003).

9.6.2 Confounding by Indication

The indication for which a medication is given may act as a confounder in observational studies, particularly when assessing the effectiveness of a drug (Slone et al. 1979; Horwitz and Feinstein 1981; Strom et al. 1983). Such confounding by indication will be present if the indication for the prescription of the medication under study is also a determinant of the outcome of interest. Generally, a drug is more likely to be prescribed to a patient with more severe disease who, in turn, is more likely to incur an adverse outcome of the disease. Thus, patients prescribed the drug under study will have higher rates of outcome than the subjects not prescribed the drug. Such an appearance of lack of effectiveness could simply be a reflection of the effect of indication, in this case disease severity.

Confounding by indication is often difficult to control, primarily because the precise reason for prescribing is rarely measured. This may preclude the study of drug effectiveness with observational designs (Miettinen 1983). Yet, a clinical trial to answer this question would require the follow-up of thousands of patients over a long time, which may simply be unfeasible. Observational studies become the tool of choice as long as validity of the study is not compromised by intractable confounding by indication (Miettinen 1983). If such an observational study produces lower rates of outcome for the drug under study, one may conclude that these medications are effective. On the other hand, if users of the drug are found to be at equal or increased risk of the outcome relative to nonusers, it would not be possible to conclude on the absence of a protective effect of these medications.

This problem of confounding by indication is compounded with the use of computerized databases, because of their lack of information on important con-

founders (Shapiro 1989). The absence of information on drug indication precludes the control of confounding by adjustment in the analysis. Thus, control for confounding by indication must be tackled at the design level. One approach is to restrict the study to a group of patients homogeneous with respect to disease severity. For example, in a study of the effectiveness of inhaled corticosteroids in asthma, Blais et al. (1998a) identified a point in time at which users and nonusers of inhaled corticosteroids would have a similar level of asthma severity. The study was thus restricted to patients who had just been hospitalized for asthma, with the discharge date taken as time zero, which would greatly reduce heterogeneity in disease severity. The rate of a readmission for asthma was then assessed according to the use of inhaled corticosteroids after this initial hospitalization.

Another approach is to compare two medications prescribed for the same indication (Strom et al. 1983, 1984). In this case, relative effectiveness as opposed to absolute effectiveness will be evaluated. An example of this approach was also used in a study of the effectiveness of early use of inhaled corticosteroids in asthma. Blais et al. (1998b) identified a cohort of newly treated asthma patients and compared regular users of inhaled corticosteroids with regular users of either anti-allergic agents or theophylline and matched for the duration of asthma at the initiation of therapy.

Depletion of Susceptibles 9.6.3

In general, the risk of an ADR associated with drug use does not remain constant over time, and can change in different ways from the start of its use. The risk may increase with cumulative drug exposure (e.g. the risk of cardiomyopathy associated with cumulative anthracycline exposure or the risk of cataract associated with continued glucocorticoid use), but it may also decrease after an initial period of sharp increased risk. Therefore, in using a case-control study that evaluates the effect of current use of a drug, past history of use of a drug or a class of drugs must be accounted for as it may modify the risk of an ADR associated with current use of the drug.

A decreasing risk after an initial period of increased risk is probably more important than cumulative drug toxicity and may, at the population level, lead to a phenomenon which has been described as "depletion of susceptibles": patients who remain on the drug are those who can tolerate it while those who are susceptible to adverse drug reactions will stop the drug and thereby select themselves out of the exposed cohort (Moride and Abenhaim 1994). Such a pattern has been demonstrated for the gastrointestinal toxicity of nonsteroidal antiinflammatory drugs (NSAIDS). It has been shown that the risk of upper gastrointestinal bleeding (UGIB) was highest after the third NSAID prescription and thereafter decreasing (Carson et al. 1987). Moride and Abenhaim (1994) empirically showed a depletion of susceptibles effect in a hospital-based case-control study of NSAIDS and the risk of UGIB. They investigated the risk of UGIB associated with recent NSAID use stratified by past or no past NSAID use. The risk of UGIB was significantly greater

for those patients who used NSAIDS for the first time in 3 years (OR = 22.7) than for those who had used these drugs before (OR = 3.0) (Yola and Lucien 1994).

The enormous importance of accounting for changes in drug risk over time in the design and/or analysis of pharmacoepidemiology studies was highlighted in the debate about the risk of venous thromboembolism (VTE) associated with second or third generation oral contraceptives (OC). Several pharmacoepidemiology studies published in 1995/1996 reported an increased risk of VTE among users of newer OC preparations compared with those of older OC preparations (Bloemenkamp et al. 1995; Anonymous 1995a; Spitzer et al. 1996; Jick et al. 1995). Additional analyses suggested that the magnitude of the risk estimates for individual OC were closely linked with the time of market introduction of the respective OC, with increasing risk for the newer preparations (Lewis et al. 1996). Since a larger proportion of users of older OC preparations were long-term users compared with those using newer OC, a depletion of susceptibles effect was postulated to be active within these studies. It was hypothesized that individuals with good tolerance were preferentially long term users of older OC preparations, whereas groups with shorter duration of use might be more frequently using the newer OC preparations and thereby constitute a different subpopulation.

The phenomenon of depletion of susceptibles can lead, if not accounted for properly, to a comparison of OC medications with different years of entry into the market and result in an overestimation of the risk associated with the most recently introduced medications. To properly account for depletion of susceptibles, the approach to statistical analysis must take account of the duration and patterns of OC use, and of course have the available data to do so. In this example, the pattern and duration of OC use are not confounders, but effect modifiers of the risk of VTE associated with recent OC use. OC pattern and duration can therefore not be simply "adjusted for" in the statistical analysis, but a stratified analysis has to be conducted which compares the risk of VTE for the different OC preparations for the different durations and patterns of use. When the analysis was restricted to the same pattern of OC use, distinguishing between first time users, repeaters and switchers, the risk of VTE as a function of the duration of oral contraceptive use was essentially the same for second and third generation pills relative to never users (Suissa et al. 1997, 2000b) .

9.6.4 Use of Propensity Scores in Pharmacoepidemiology

In a randomized trial, randomization of study subjects to different treatment regimens aims to assure the absence of systematic differences between the patients in terms of measured and unmeasured confounders. In observational studies, direct comparisons of the outcomes of treated and untreated patients may be misleading because of systematic differences between those patients who have and have not received treatment. The propensity score has been proposed as a method of adjusting for covariate imbalances in an observational study and has recently been proposed also for pharmacoepidemiology research (Perkins et al. 2000; Wang and Donnan 2001; Wang et al. 2001). The propensity score

$\pi(X) = Prob(\text{exposed}|X)$ is defined as the conditional probability of receiving a particular treatment (i.e. being exposed) given the set of observed covariates X. The propensity score thus represents a summary of the covariates X that are associated with treatment allocation in the form of a single variable.

The propensity score approach is a two-stage approach. At the first stage, the propensity score is estimated for each study subject based on the values of the observed covariates. The most commonly used approach is to obtain the propensity score estimates in a logistic regression model, where treatment allocation is used as the dependent (response) variable and observed potential confounders X are used as explanatory variables: $\log(\pi(X)/(1 - \pi(X))) = X\beta$, where the regression coefficients β are fitted by maximum likelihood.

Having obtained the estimated propensity score, it can be used by a number of approaches at the second stage: study subjects may be matched or stratified based on their propensity scores or the propensity scores may be adjusted for in a regression model (Wang and Donnan 2001).

Many published applications use stratification by the propensity score. A common approach is to stratify by quintiles of the distribution of the estimated propensity scores and to test the balance of each confounder between the treatment groups in each stratum (Wang and Donnan 2001). Having patients with similar propensity scores in each stratum, it may be assumed that the covariate distributions in the two treatment groups are equally similar within each stratum, so that the treatment assignment within the strata can be functionally regarded as random. If unbalanced confounders are still found, the propensity score model may be re-estimated with modifications until balance is achieved.

Stratification by the propensity score cannot control confounder effects within a single stratum. The somewhat arbitrary choice of five strata can be viewed as a compromise between reduction of bias and robustness of the results: an increasing number of strata will reduce the bias in the stratification estimate, but it will at the same time decrease the robustness of the results when the sample size of the smaller arm in a stratum becomes too small. Control of bias through use of propensity scores is based on the following assumptions:

- All subjects must have some non-zero probability of receiving each treatment (referred to as the "strongly ignorable assumption"). This ensures independence of treatment assignment and response variable within propensity score strata.
- Treatment assignment depends solely on the observed covariates, i.e. *all* confounders are included in the propensity score model.

If *all* confounders were not ascertained or confounder measurement was associated with bias, the use of propensity scores will not eliminate bias. In fact, the study will be subject to the same bias as an observational study that did not measure all confounders and could only incompletely adjust for known confounders. The use of propensity scores may, however, help to detect incomparability between treatment groups (i.e. lack of overlap in covariate values) that may remain undetected in a standard regression model. It also provides an additional tool to assess

the performance of the traditional regression model (Wang and Donnan 2001). Propensity scores have more often been used in cohort studies of drug effectiveness (Seeger et al. 2003; Mojtabai and Zivin 2003; Schroder et al. 2003; Young-Xu et al. 2003), but they have also been applied to studies of drug safety (MacDonald et al. 2003). Nevertheless, the use of propensity scores may be a particular challenge in the common situation in pharmacoepidemiology of time-dependent exposures and covariates. Moreover, with the very large sizes of databases used in pharmacoepidemiology, the need to reduce the number of covariates to a single score is not crucial and thus, the advantage of propensity scores compared to including the confounders directly in the model of data analysis becomes less evident.

9.7 Drug Utilization Studies

Drug utilization studies are an important tool in improving rational drug use and providing data for cost/benefit considerations. Drug utilization has been defined as the "prescribing, dispensing, administering, and ingesting of drugs." (Serradell et al. 1991). This definition implies that several steps are involved in drug utilization and that, consequently, in each of these steps problems in drug use can arise. The World Health Organization defines drug utilization in a broader sense as the "marketing, distribution, prescription and use of drugs in a society, with special emphasis on the resulting medical, social and economic consequences" (World Health Organization 1977), thereby including also the effects of drug use on the population. Apart from examining drug use, goals of drug utilization studies include the identification of problems of drug utilization with respect to their importance, causes and consequences; the establishment of a scientific basis for decisions on problem solving and the assessment of the effects of actions taken. Some examples for studies that illustrate these goals are the following: What is the prevalence, pattern and risk factors of use for benzodiazepines in Italy? What is the quality of NSAID prescribing in Croatia and Sweden (Vlahovic-Palcevski et al. 2002)? Are labelled contraindications to the use of cisapride adhered to (Weatherby et al. 2001)? What is the impact of safety alerts on the prescribing of a drug (de la Porte et al. 2002; Weatherby et al. 2001)? After a drug has been withdrawn from the market, in which way does drug utilization of related drugs change (Glessner and Heller 2002)? What are characteristics of physicians and practices that make early use of new prescription drugs (Tamblyn et al. 2003)?

Drug utilization studies can be qualitative and quantitative. *Quantitative* drug utilization studies are conducted for a number of purposes: to ascertain the quantities of drugs consumed in a specific period and in a specific geographical area; to investigate the development of drug utilization over time; to compare and contrast the use of a drug between different geographical areas; to identify possible over- or underutilization of drugs; to determine trends in drug use according to population demographics; to estimate the prevalence of illness based on the consumption

of drugs utilised in its treatment and to compare the prevalence of an illness in different areas.

The main aim of *qualitative* drug utilization studies is to determine the appropriateness of drug prescribing. They require the *a priori* establishment of quality indicators against which drug utilization is compared. National or international expert panels are sometimes used to help defining quality indicators in a consensus process (McLeod et al. 1997). Quality indicators may be based on the following parameters: the medical necessity for drug treatment; adherence to labelling with respect to labelled indications, contraindications or interactions; duration and dose of treatment; use of fixed drug combinations when only one of its components would be justified; availability of treatment alternatives which are more effective or less hazardous; availability of an equivalent less costly drug on the market; etc. In North America, these studies are known as drug utilization review (DUR) studies. DUR studies are aimed at detecting and quantifying problems of drug prescribing. They should be distinguished from DUR programs which are interventions in the form of an authorized, structured and ongoing system to improve the quality of drug prescribing (Lee and Bergman 2000). In contrast to DUR studies which provide only minimal feedback to the involved prescribers and are not interventional by their design, DUR programs include efforts to correct inappropriate patterns of drug use, and include a mechanism for measuring the effectiveness of corrective actions taken to normalize undesirable patterns of drug use (Hennessy and Strom 2000).

For quantitative and qualitative studies, it would be ideal to have a count of the number of patients who either ingest a drug of interest during a certain time frame or who use a drug inappropriately in relation to all patients who received the drug during a given time frame. The available data are often only approximations of the number of patients and may be based on cost or unit cost, weight, number of prescriptions written or dispensed and number of tablets, capsules, doses etc sold (Lee and Bergman 2000). Drug cost data have a number of limitations, since the price of a drug is not the same within and across countries. Drug pricing may be affected by different drug distribution channels, the quantities of drugs purchased, exchange rate fluctuations, different import duties and regulatory policies that affect pricing (Serradell et al. 1991). Studies based on the overall weight of a drug sold are similarly limited, since tablet sizes vary which makes it difficult to translate weight even into the number of tablets sold (Lee and Bergman 2000). The number of prescriptions written or dispensed for a particular drug product is a measure that is frequently used in drug utilization studies. However, the number of prescriptions for different patients in a given time interval varies and also the supply of drugs prescribed. To estimate the number of patients, one must divide by the average number of prescriptions per patient. The number of tablets, capsules etc. sold is often used in conjunction with the defined daily dose (DDD) measurement unit for drug use.

The DDD is the assumed average maintenance dose per day for a drug used for its main indication in adults. Use of the DDD underlies two basic assumptions: that patients are compliant and that the doses used for the major indication are

the average maintenance doses (Serradell et al. 1991). The DDD dosing levels are assigned per ATC 5th level by the WHO Collaborating Centre for Drug Statistics Methodology in Norway based on recommendations in the medical literature. The DDD provides a fixed unit of measurement independent of the price and formulation of a drug. It can be used to examine changes in drug consumption over time and permits international comparisons. The DDD is a technical unit of measurement and does not necessarily reflect the recommended or prescribed daily dose (PDD). Doses for individual patients and patient groups will often differ from the DDD based on individual patient characteristics such as age, weight, and pharmacokinetic considerations. DDDs may be used to obtain crude estimates of the number of persons exposed to a particular drug or class of drugs and are sometimes used as denominator data for crude estimation of ADR rates (Kromann-Andersen and Pedersen 1988; Leone et al. 2003). This use of the DDD methodology is rather limited in drugs with more than one indication, particularly when the drug dose differs for each indication. It is also limited, when the duration of drug treatment varies greatly between patients. The DDD does not take into account pediatric use of a drug. DDDs are not established for topical preparations, sera, vaccines, antineoplastic agents, allergen extracts, general and local anesthetics and contrast media.

The Defined Daily Dose (DDD) is usually used in conjunction with the Anatomical Therapeutic Chemical (ATC) coding system. Coding for the ATC system at the WHO Collaborating Centre for Drug Statistics Methodology in Norway is based on requests from users including manufacturers, regulatory agencies and researchers. In the ATC system, drugs are classified in groups at five hierarchical levels. The drugs are divided into fourteen main groups (1st level), with one pharmacological/therapeutic subgroup (2nd level). The 3rd and 4th levels are pharmacological/therapeutic and chemical subgroups and the 5th level is the chemical substance. Table 9.6 illustrates the structure of the ATC coding system using metformin as an example. In the ATC system all plain metformin preparations are thus given the code A10B A02.

Table 9.6. The structure of the ATC coding system for metformin

A	Alimentary tract and metabolism (1st level, anatomical main group)
A10	Drugs used in diabetes (2nd level, therapeutic subgroup)
A10B	Oral blood glucose lowering drugs (3rd level, pharmacological subgroup)
A10B A	Biguanides (4th level, chemical subgroup)
A10B A02	Metformin (5th level, chemical substance)

Coverage of the ATC system is not comprehensive: complementary and traditional medicinal products are generally not included in the ATC system; ATC codes for fixed combination drugs are assigned only to a limited extent; some drugs may not be included in the system since no request for coding has been received by the WHO Collaborating Centre; a medicinal product that is used for two or more

equally important indications, will usually be given only one code based on its main indication which is decided from the available literature. On the other hand, a medicinal product can have more than one ATC code if it is available in two or more strengths or formulations with clearly different therapeutic uses.

Use of computerized databases has greatly facilitated drug utilization research. Databases can be distinguished into those which include both drug and diagnostic data (examples have been given in Sect. 9.4) and into those which include only drug data (e.g. Denmark's Odense Pharmacoepidemiologic Database, Denmark's Pharmacoepidemiologic Prescription Database of the County of North Jutland, Spain's Drug Data Bank, Sweden's County of Jämtland Project, etc.). Some more databases used in drug utilization research and a description of these databases can be found in references (Lee and Bergman 2000; Serradell et al. 1991).

Interpretation of drug utilization data needs appropriate care. Observed geographic or time differences in drug utilization may be caused by many factors different from prescribing behaviour as e.g. differences in the age and sex distribution, different patterns of morbidity, change in diagnostic criteria, differences in the access to healthcare etc. If used appropriately, drug utilization research provides a powerful scientific tool to identify factors that influence drug prescribing and to develop strategies to modify prescribing behaviour. Further research is needed to determine which characteristics of inappropriate prescribing are susceptible to modification and what are the most efficient intervention strategies.

9.8 Conclusions

Pharmacoepidemiology is still a relatively young scientific discipline. Over the last 20–30 years there has been enormous progress in the improvement of its methods and development of new approaches to studies of drug safety and effectiveness. Pharmacoepidemiology has taken advantage of the rapidly expanding methods in epidemiology and has developed sophisticated methods to cope with problems that are specific to the field. New statistical approaches have been developed for signal generation based on data from the spontaneous reporting systems. Large computerized health databases are now widely used for research into beneficial and harmful drug effects, their use being facilitated by the development of more and more powerful computer technologies. With the experience gained through the use of these data and a careful understanding of the underlying health care system in which the data were generated, computerized databases provide a highly useful data source for pharmacoepidemiology studies. There has been a progressive refinement of case-control and cohort studies and efficient sampling strategies within a cohort are now often employed. The case-crossover and case-time-control designs are being used for the study of acute transient drug effects to eliminate control selection bias and confounding by indication or other factors. Propensity scores are increasingly used as a method to minimize confounding in the study of intended drug effects. Other new methodologies are likely to become of more

importance in pharmacoepidemiology over the coming years as e.g. neural networks, sensitivity analysis, etc. Drug utilization review programs are now required in all US hospitals and have been implemented voluntarily in many other health care programs which will lead to further refinement of drug utilization research. A great challenge ahead is linkage of pharmacoepidemiology studies with the latest techniques of genetics, biochemistry, immunology and molecular biology. It is of particular interest to understand why individuals respond differently to drug therapy, both in terms of beneficial and adverse effects. Investigation of the genetic make-up of study patients on a population level will be greatly facilitated through the enormous progress in pharmacogenomics and molecular biology. New study designs may emerge as a consequence of these developments. It remains to be explored to which extent database studies may be used to include moleculargenetic or immunologic investigations.

References

Abascal VM, Larson MG, Evans JC, Blohm AT, Poli K, Levy D (1998) Calcium antagonists and mortality risk in men and women with hypertension in the Framingham Heart Study. Arch Intern Med 158:1882–1886

Abenhaim L, Moride Y, Brenot F, Rich S, Benichou J, Kurz X, Higenbottam T, Oakley C, Wouters E, Aubier M, Simonneau G, Begaud B (1996) Appetite-suppressant drugs and the risk of primary pulmonary hypertension. International Primary Pulmonary Hypertension Study Group. N Engl J Med 335:609–616

Adams RJ, Fuhlbrigge AL, Finkelstein JA, Weiss ST (2002) Intranasal steroids and the risk of emergency department visits for asthma. J Allergy Clin Immunol 109:636–642

Anonymous (1986) Risks of agranulocytosis and aplastic anemia. A first report of their relation to drug use with special reference to analgesics. The International Agranulocytosis and Aplastic Anemia Study. JAMA 256:1749–1757

Anonymous (1995a) Effect of different progestagens in low oestrogen oral contraceptives on venous thromboembolic disease. World Health Organization Collaborative Study of Cardiovascular Disease and Steroid Hormone Contraception. Lancet 346:1582–1588

Anonymous (1995b) Venous thromboembolic disease and combined oral contraceptives: results of international multicentre case-control study. World Health Organization Collaborative Study of Cardiovascular Disease and Steroid Hormone Contraception. Lancet 346:1575–1582

Barbone F, McMahon AD, Davey PG, Morris AD, Reid IC, McDevitt DG, MacDonald TM (1998) Association of road-traffic accidents with benzodiazepine use. The Lancet 352:1331–1336

Bate A, Lindquist M, Edwards IR, Olsson S, Orre R, Lansner A, De Freitas RM (1998) A Bayesian neural network method for adverse drug reaction signal generation. Eur J Clin Pharmacol 54:315–321

Bate A, Lindquist M, Orre R, Edwards IR, Meyboom RH (2002) Data-mining analyses of pharmacovigilance signals in relation to relevant comparison drugs. Eur J Clin Pharmacol 58:483–490

Begaud B, Martin K, Haramburu F, Moore N (2002) Rates of spontaneous reporting of adverse drug reactions in France. JAMA 288:1588

Beiderbeck-Noll AB, Sturkenboom MC, van der Linden PD, Herings RM, Hofman A, Coebergh JW, Leufkens HG, Stricker BH (2003) Verapamil is associated with an increased risk of cancer in the elderly: the Rotterdam study. Eur J Cancer 39:98–105

Benichou C (1994) Imputability of unexpected or toxic drug reactions. The official French method of causality assessment. In: Benichou C (ed) Adverse drug reactions. A practical guide to diagnosis and management. Wiley, Chichester, pp 271–275

Blais L, Ernst P, Boivin JF, Suissa S (1998a) Inhaled corticosteroids and the prevention of readmission to hospital for asthma. Am J Respir Crit Care Med 158:126–132

Blais L, Suissa S, Boivin JF, Ernst P (1998b) First treatment with inhaled corticosteroids and the prevention of admissions to hospital for asthma [see comments]. Thorax 53:1025–1029

Bloemenkamp KW, Rosendaal FR, Helmerhorst FM, Buller HR, Vandenbroucke JP (1995) Enhancement by factor V Leiden mutation of risk of deep-vein thrombosis associated with oral contraceptives containing a third-generation progestagen. Lancet 346:1593–1596

Carson JL, Strom BL, Soper KA, West SL, Morse ML (1987) The association of nonsteroidal anti-inflammatory drugs with upper gastrointestinal tract bleeding. Arch Intern Med 147:85–88

Chen H, Zhang SM, Hernan MA, Schwarzschild MA, Willett WC, Colditz GA, Speizer FE, Ascherio A (2003) Nonsteroidal anti-inflammatory drugs and the risk of Parkinson disease. Arch Neurol 60:1059–1064

Colditz GA, Hankinson SE, Hunter DJ, Willett WC, Manson JE, Stampfer MJ, Hennekens C, Rosner B, Speizer FE (1995) The use of estrogens and progestins and the risk of breast cancer in postmenopausal women. N Engl J Med 332:1589–1593

Confavreux C, Suissa S, Saddier P, Bourdes V, Vukusic S (2001) Vaccinations and the risk of relapse in multiple sclerosis. Vaccines in Multiple Sclerosis Study Group. N Engl J Med 344:319–326

de la Porte M, Reith D, Tilyard M (2002) Impact of safety alerts upon prescribing of cisapride to children in New Zealand. N Z Med J 115:U24

Dietlein G, Schroder-Bernhardi D (2002) Use of the mediplus patient database in healthcare research. Int J Clin Pharmacol Ther 40:130–133

Dietlein G, Schroder-Bernhardi D (2003) Doctors' prescription behaviour regarding dosage recommendations for preparations of kava extracts. Pharmacoepidemiol Drug Saf 12:417–421

Donahue JG, Weiss ST, Livingston JM, Goetsch MA, Greineder DK, Platt R (1997) Inhaled steroids and the risk of hospitalization for asthma. JAMA 277:887–891

Donnan PT, Wang J (2001) The case-crossover and case-time-control designs in pharmacoepidemiology. Pharmacoepidemiol Drug Saf 10:259–262

Downey W, Beck P, McNutt M, Stang M, Osei W, Nichol J (2003) Health databases in Saskatchewan. In: Strom BL (ed) 3rd edition, Pharmacoepidemiology. John Wiley & Sons, Chichester, pp 325–345

Eland IA, Belton KJ, van Grootheest AC, Meiners AP, Rawlins MD, Stricker BH (1999) Attitudinal survey of voluntary reporting of adverse drug reactions. Br J Clin Pharmacol 48:623–627

Essebag V, Genest J, Jr., Suissa S, Pilote L (2003) The nested case-control study in cardiology. Am Heart J 146:581–590

Etienney I, Beaugerie L, Viboud C, Flahault A (2003) Non-steroidal anti-inflammatory drugs as a risk factor for acute diarrhoea: a case crossover study. Gut 52:260–263

Evans SJ, Waller PC, Davis S (2001) Use of proportional reporting ratios (PRRs) for signal generation from spontaneous adverse drug reaction reports. Pharmacoepidemiol Drug Saf 10:483–486

Fagot JP, Mockenhaupt M, Bouwes-Bavinck JN, Naldi L, Viboud C, Roujeau JC (2001) Nevirapine and the risk of Stevens-Johnson syndrome or toxic epidermal necrolysis. AIDS 15:1843–1848

Faich GA (1986) Adverse-drug-reaction monitoring. N Engl J Med 314:1589–1592

Fairfield KM, Hunter DJ, Fuchs CS, Colditz GA, Hankinson SE (2002) Aspirin, other NSAIDs, and ovarian cancer risk (United States). Cancer Causes Control 13:535–542

Feenstra J, Heerdink ER, Grobbee DE, Stricker BH (2002) Association of non-steroidal anti-inflammatory drugs with first occurrence of heart failure and with relapsing heart failure: the Rotterdam Study. Arch Intern Med 162:265–270

Felson DT, Sloutskis D, Anderson JJ, Anthony JM, Kiel DP (1991) Thiazide diuretics and the risk of hip fracture. Results from the Framingham Study. JAMA 265:370–373

Friedman MA, Woodcock J, Lumpkin MM, Shuren JE, Hass AE, Thompson LJ (1999) The safety of newly approved medicines: do recent market removals mean there is a problem? JAMA 281:1728–1734

Friedman DE, Habel LA, Boles M, McFarland BH (2000) Kaiser Permanente Medical Care Program: Division of Research, Northern California, and Center for Health Research, Northwest Division. In: Strom BL (ed) Pharmacoepidemiology. Wiley & Sons, Chichester, pp 263–283

Funkhouser EM, Sharp GB (1995) Aspirin and reduced risk of esophageal carcinoma. Cancer 76:1116–1119

Garbe E, LeLorier J, Boivin JF, Suissa S (1997) Inhaled and nasal glucocorticoids and the risks of ocular hypertension or open-angle glaucoma. JAMA 277:722–727

Garbe E, Boivin JF, LeLorier J, Suissa S (1998a) Selection of controls in database case-control studies: glucocorticoids and the risk of glaucoma. J Clin Epidemiol 51:129–135

Garbe E, Suissa S, LeLorier J (1998b) Association of inhaled corticosteroid use with cataract extraction in elderly patients. JAMA 280:539–543

Garcia Rodriguez LA, Perez GS (1998) Use of the UK General Practice Research Database for pharmacoepidemiology. Br J Clin Pharmacol 45:419–425

Giovannucci E, Rimm EB, Stampfer MJ, Colditz GA, Ascherio A, Willett WC (1994) Aspirin use and the risk for colorectal cancer and adenoma in male health professionals. Ann Intern Med 121:241–246

Glessner MR, Heller DA (2002) Changes in related drug class utilization after market withdrawal of cisapride. Am J Manag Care 8:243–250

Greenland S (1996) Confounding and exposure trends in case-crossover and case-time-control design. Epidemiology 7:231–239

Griffin JP (1986) Survey of the spontaneous adverse drug reaction reporting schemes in fifteen countries. Br J Clin Pharmacol 22:83S–100S

Grodstein F, Stampfer MJ, Manson JE, Colditz GA, Willett WC, Rosner B, Speizer FE, Hennekens CH (1996) Postmenopausal estrogen and progestin use and the risk of cardiovascular disease. N Engl J Med 335:453–461

Grodstein F, Martinez ME, Platz EA, Giovannucci E, Colditz GA, Kautzky M, Fuchs C, Stampfer MJ (1998) Postmenopausal hormone use and risk for colorectal cancer and adenoma. Ann Intern Med 128:705–712

Haramburu F, Tubert-Bitter P, Begaud B (1992) Trends in spontaneous reporting. Post Marketing Surveillance 6:129–134

Haramburu F, Begaud B, Moride Y (1997) Temporal trends in spontaneous reporting of unlabelled adverse drug reactions. Br J Clin Pharmacol 44:299–301

Hee KJ, Grodstein F (2003) Regular use of nonsteroidal anti-inflammatory drugs and cognitive function in aging women. Neurology 60:1591–1597

Hennessy S, Strom BL (2000) Nonsedating antihistamines should be preferred over sedating antihistamines in patients who drive. Ann Intern Med 132:405–407

Hernan MA, Hohol MJ, Olek MJ, Spiegelman D, Ascherio A (2000) Oral contraceptives and the incidence of multiple sclerosis. Neurology 55:848–854

Hernandez-Avila M, Liang MH, Willett WC, Stampfer MJ, Colditz GA, Rosner B, Chang RW, Hennekens CH, Speizer FE (1990) Exogenous sex hormones and the risk of rheumatoid arthritis. Arthritis Rheum 33:947–953

Hernandez-Diaz S, Hernan MA, Meyer K, Werler MM, Mitchell AA (2003) Case-Crossover and Case-Time-Control Designs in Birth Defects Epidemiology. Am J Epidemiol 158:385–391

Horwitz RI, Feinstein AR (1981) Improved observational method for studying therapeutic efficacy. Suggestive evidence that lidocaine prophylaxis prevents death in acute myocardial infarction. JAMA 246:2455–2459

Hubbard RB, Smith CJ, Smeeth L, Harrison TW, Tattersfield AE (2002) Inhaled corticosteroids and hip fracture: a population-based case-control study. Am J Respir Crit Care Med 166:1563–1566

Jick H, Jick SS, Gurewich V, Myers MW, Vasilakis C (1995) Risk of idiopathic cardiovascular death and nonfatal venous thromboembolism in women using oral contraceptives with differing progestagen components. Lancet 346:1589–1593

Kaufman DW, Kelly JP, Jurgelon JM, Anderson T, Issaragrisil S, Wiholm BE, Young NS, Leaverton P, Levy M, Shapiro S (1996) Drugs in the aetiology of agranulocytosis and aplastic anaemia. Eur J Haematol Suppl 60:23–30

Kelly JP, Rosenberg L, Kaufman DW, Shapiro S (1990) Reliability of personal interview data in a hospital-based case-control study. Am J Epidemiol 131:79–90

Ki M, Park T, Yi SG, Oh JK, Choi B (2003) Risk analysis of aseptic meningitis after measles-mumps-rubella vaccination in Korean children by using a case-crossover design. Am J Epidemiol 157:158–165

Kiel DP, Felson DT, Anderson JJ, Wilson PW, Moskowitz MA (1987) Hip fracture and the use of estrogens in postmenopausal women. The Framingham Study. N Engl J Med 317:1169–1174

Kromann-Andersen H, Pedersen A (1988) Reported adverse reactions to and consumption of nonsteroidal anti-inflammatory drugs in Denmark over a 17-year period. Dan Med Bull 35:187–192

Lando JF, Heck KE, Brett KM (1999) Hormone replacement therapy and breast cancer risk in a nationality representative cohort. Am J Prev Med 17:176–180

Lee D, Bergman U (2000) Studies of drug utilization. In: Strom BL (ed) Pharmacoepidemiology. John Wiley & Sons, Chichester, pp 463–481

Lenz W (1987) The Thalidomide hypothesis: How it was found and tested. In: Kewitz H, Roots I, Voigt K (eds) Epidemiological Concepts in Clinical Pharmacology. Springer Verlag, Heidelberg, pp 3–10

Leone R, Venegoni M, Motola D, Moretti U, Piazzetta V, Cocci A, Resi D, Mozzo F, Velo G, Burzilleri L, Montanaro N, Conforti A (2003) Adverse drug reactions related to the use of fluoroquinolone antimicrobials: an analysis of spontaneous reports and fluoroquinolone consumption data from three italian regions. Drug Saf 26:109–120

Lewis MA, Heinemann LA, MacRae KD, Bruppacher R, Spitzer WO (1996) The increased risk of venous thromboembolism and the use of third generation progestagens: role of bias in observational research. The Transnational Research Group on Oral Contraceptives and the Health of Young Women. Contraception 54:5–13

Lindquist M, Edwards IR (1993) Adverse drug reaction reporting in Europe: some problems of comparisons. International Journal of Risk & Safety in Medicine 4:35–46

MacDonald TM, Morant SV, Goldstein JL, Burke TA, Pettitt D (2003) Channelling bias and the incidence of gastrointestinal haemorrhage in users of meloxicam, coxibs, and older, non-specific non-steroidal anti-inflammatory drugs. Gut 52:1265–1270

McLeod PJ, Huang AR, Tamblyn RM, Gayton DC (1997) Defining inappropriate practices in prescribing for elderly people: a national consensus panel. CMAJ 156:385–391

Miettinen OS (1983) The need for randomization in the study of intended effects. Statist Med 2:267–271

Mojtabai R, Zivin JG (2003) Effectiveness and cost-effectiveness of four treatment modalities for substance disorders: a propensity score analysis. Health Serv Res 38:233–259

Moore N, Hall G, Sturkenboom M, Mann R, Lagnaoui R, Begaud B (2003) Biases affecting the proportional reporting ratio (PPR) in spontaneous reports pharmacovigilance databases: the example of sertindole. Pharmacoepidemiol Drug Saf 12:271–281

Moride Y, Abenhaim L (1994) Evidence of the depletion of susceptibles effect in non-experimental pharmacoepidemiologic research. J Clin Epidemiol 47:731–737

Neutel CI, Perry S, Maxwell C (2002) Medication use and risk of falls. Pharmacoepidemiol Drug Saf 11:97–104

Perez E, Schroder-Bernhardi D, Dietlein G (2002) Treatment behavior of doctors regarding Helicobacter pylori infections. Int J Clin Pharmacol Ther 40:126–129

Perkins SM, Wanzhu T, Underhill MG, Zhou XH, Murray MD (2000) The use of propensity scores in pharmacoepidemiologic research. Pharmacoepidemiol Drug Saf 9:93–101

Pierfitte C, Begaud B, Lagnaoui R, Moore ND (1999) Is reporting rate a good predictor of risks associated with drugs? Br J Clin Pharmacol 47:329–331

Pierfitte C, Royer RJ, Moore N, Begaud B (2000) The link between sunshine and phototoxicity of sparfloxacin. Br J Clin Pharmacol 49:609–612

Porta-Serra M, Hartzema AG (1997) The contribution of epidemiology to the study of drug effects. In: Hartzema AG, Porta-Serra M, Tilson HH (eds) Pharmacoepidemiology. The fundamentals. Harvey Whitney, Cincinnati

Raiford DS, Perez GS, Garcia Rodriguez LA (1996) Positive predictive value of ICD-9 codes in the identification of cases of complicated peptic ulcer disease in the Saskatchewan hospital automated database. Epidemiology 7:101–104

Rawlins MD, Breckenridge AM, Wood SM (1989) National adverse drug reaction reporting – a silver jubilee. Adverse Drug Reaction Bulletin 138:516–519

Ray WA, Griffin MR (1989) Use of Medicaid data for pharmacoepidemiology. Am J Epidemiol 129:837–849

Ray WA, Griffin MR, Downey W, Melton LJ, III (1989) Long-term use of thiazide diuretics and risk of hip fracture. Lancet 1:687–690

Samet JM (2003) Measuring the effectiveness of inhaled corticosteroids for COPD is not easy! Am J Respir Crit Care Med 168:1–2

Saskatchewan Health (2004) (http://www.health.gov.sk.ca/mc_dp_phb_infodoc.pdf) Accessed June 03, 2004

Saunders KW, Davis RL, Stergachis A (2000) Group Health Cooperative of Puget Sound. In: Strom BL (ed) Pharmacoepidemiology. John Wiley& Sons, Chichester, pp 247–262

Schoofs MW, van der KM, Hofman A, de Laet CE, Herings RM, Stijnen T, Pols HA, Stricker BH (2003) Thiazide diuretics and the risk for hip fracture. Ann Intern Med 139:476–482

Schroder D, Weiser M, Klein P (2003) Efficacy of a homeopathic Crataegus preparation compared with usual therapy for mild (NYHA II) cardiac insufficiency: results of an observational cohort study. Eur J Heart Fail 5:319–326

Schroder-Bernhardi D, Dietlein G (2002) Lipid-lowering therapy: do hospitals influence the prescribing behavior of general practitioners? Int J Clin Pharmacol Ther 40:317–321

Seeger JD, Walker AM, Williams PL, Saperia GM, Sacks FM (2003) A propensity score-matched cohort study of the effect of statins, mainly fluvastatin, on the occurrence of acute myocardial infarction. Am J Cardiol 92:1447–1451

Serradell J, Bjornson DC, Hartzema AG (1991) Drug utilization studies: Sources and methods. In: Hartzema AG, Porta MS, Tilson HM (eds) Pharmacoepidemiology: An Introduction. Harvey Whitney Books, Cincinnati Ohio, pp 101–119

Shapiro S (1989) The role of automated record linkage in the postmarketing surveillance of drug safety: a critique. Clin Pharmacol Ther 46:371–386

Sin DD, Man SF (2002) Low-dose inhaled corticosteroid therapy and risk of emergency department visits for asthma. Arch Intern Med 162:1591–1595

Sin DD, Tu JV (2001a) Inhaled corticosteroid therapy reduces the risk of rehospitalization and all-cause mortality in elderly asthmatics. European Respiratory Journal 17:380–385

Sin DD, Tu JV (2001b) Inhaled corticosteroids and the risk of mortality and readmission in elderly patients with chronic obstructive pulmonary disease. Am J Respir Crit Care Med 164:580–584

Slone D, Shapiro S, Miettinen OS, Finkle WD, Stolley PD (1979) Drug evaluation after marketing. Ann Intern Med 90:257–261

Spitzer WO, Suissa S, Ernst P, Horwitz RI, Habbick B, Cockcroft D, Boivin JF, McNutt M, Buist AS, Rebuck AS (1992) The use of beta-agonists and the risk of death and near death from asthma. N Engl J Med 326:501–506

Spitzer WO, Lewis MA, Heinemann LA, Thorogood M, MacRae KD (1996) Third generation oral contraceptives and risk of venous thromboembolic disorders: an international case-control study. Transnational Research Group on Oral Contraceptives and the Health of Young Women. BMJ 312:83–88

Strom BL (ed) (2000) Pharmacoepidemiology. John Wiley & Sons, Chichester

Strom BL, Carson JL (1990) Use of automated databases for pharmacoepidemiology research. Epidemiol Rev 12:87–107

Strom BL, Miettinen OS, Melmon KL (1983) Postmarketing studies of drug efficacy: when must they be randomized? Clinical Pharmacology & Therapeutics 34:1–7

Strom BL, Miettinen OS, Melmon KL (1984) Post-marketing studies of drug efficacy: how? Am J Med 77:703–708

Strom BL, Carson JL, Halpern AC, Schinnar R, Snyder ES, Stolley PD, Shaw M, Tilson HH, Joseph M, Dai WS (1991) Using a claims database to investigate drug-induced Stevens-Johnson syndrome. Stat Med 10:565–576

Sturkenboom MC, Middelbeek A, de Jong van den Berg LT, van den Berg PB, Stricker BH, Wesseling H (1995) Vulvo-vaginal candidiasis associated with acitretin. J Clin Epidemiol 48:991–997

Suissa S (1995) The case-time-control design. Epidemiology 6:248–253

Suissa S (1998) The case-time-control design: further assumptions and conditions [comment]. Epidemiology 9:441–445

Suissa S (2000) Novel approaches to pharmacoepidemiology study design and statistical analysis. In: Strom BL (ed) Pharmacoepidemiology. John Wiley & Sons, Chichester, pp 785–805

Suissa S (2003) Effectiveness of inhaled corticosteroids in COPD: immortal time bias in observational studies. Am J Respir Crit Care Med 168:49–53

Suissa S (2004) Inhaled steroids and mortality in COPD: bias from unaccounted immortal time. European Respiratory Journal 23:391–395

Suissa S, Blais L, Spitzer WO, Cusson J, Lewis M, Heinemann L (1997) First-time use of newer oral contraceptives and the risk of venous thromboembolism. Contraception 56:141–146

Suissa S, Ernst P, Benayoun B, Baltzan M, Cai B (2000a) Low-dose inhaled corticosteroids and the prevention of death from asthma. N Engl J Med 343:332–336

Suissa S, Spitzer WO, Rainville B, Cusson J, Lewis M, Heinemann L (2000b) Recurrent use of newer oral contraceptives and the risk of venous thromboembolism. Hum Reprod 15:817–821

Suissa S, Ernst P, Kezouh A (2002) Regular use of inhaled corticosteroids and the long term prevention of hospitalisation for asthma. Thorax 57:880–884

Suissa S, Baltzan M, Kremer R, Ernst P (2004) Inhaled and nasal corticosteroid use and the risk of fracture. Am J Respir Crit Care Med 169:83–88

Szarfman A, Machado SG, O'Neill RT (2002) Use of screening algorithms and computer systems to efficiently signal higher-than-expected combinations of drugs and events in the US FDA's spontaneous reports database. Drug Saf 25:381–392

Tamblyn R, Lavoie G, Petrella L, Monette J (1995) The use of prescription claims databases in pharmacoepidemiological research: the accuracy and comprehensiveness of the prescription claims database in Quebec. J Clin Epidemiol 48:999–1009

Tamblyn R, McLeod P, Hanley JA, Girard N, Hurley J (2003) Physician and practice characteristics associated with the early utilization of new prescription drugs. Med Care 41:895–908

Tennis P, Bombardier C, Malcolm E, Downey W (1993) Validity of rheumatoid arthritis diagnoses listed in the Saskatchewan Hospital Separations Database. J Clin Epidemiol 46:675–683

Tsong Y (1995) Comparing reporting rates of adverse events between drugs with adjustment for year of marketing and secular trends in total reporting. J Biopharm Stat 5:95–114

van der Kroef KC (1979) Reactions to triazolam. Lancet 2:526

Venning GR (1983) Identification of adverse reactions to new drugs. III: Alerting processes and early warning systems. BMJ 286:458–460

Viscoli CM, Brass LM, Kernan WN, Sarrel PM, Suissa S, Horwitz RI (2001) A clinical trial of estrogen-replacement therapy after ischemic stroke. N Engl J Med 345:1243–1249

Vlahovic-Palcevski V, Wettermark B, Bergman U (2002) Quality of non-steroidal anti-inflammatory drug prescribing in Croatia (Rijeka) and Sweden (Stockholm). Eur J Clin Pharmacol 58:209–214

Wang J, Donnan PT (2001) Propensity score methods in drug safety studies: practice, strengths and limitations. Pharmacoepidemiol Drug Saf 10:341–344

Wang J, Donnan PT, Steinke D, MacDonald TM (2001) The multiple propensity score for analysis of dose-response relationships in drug safety studies. Pharmacoepidemiol Drug Saf 10:105–111

Weatherby LB, Walker AM, Fife D, Vervaet P, Klausner MA (2001) Contraindicated medications dispensed with cisapride: temporal trends in relation to the sending of 'Dear Doctor' letters. Pharmacoepidemiol Drug Saf 10:211–218

Weintraub JM, Taylor A, Jacques P, Willett WC, Rosner B, Colditz GA, Chylack LT, Hankinson SE (2002) Postmenopausal hormone use and lens opacities. Ophthalmic Epidemiol 9:179–190

Wiholm B, Olsson S, Moore N, Waller P (2000) Spontaneous reporting systems outside the US. In: Strom BL (ed) Pharmacoepidemiology. John Wiley & Sons, Chichester, pp 175–192

World Health Organization (1977) The selection of essential drugs. Report of a WHO expert committee. Ser. no. 615. World Health Organization, Geneva

Worzala K, Hiller R, Sperduto RD, Mutalik K, Murabito JM, Moskowitz M, D'Agostino RB, Wilson PW (2001) Postmenopausal estrogen use, type of menopause, and lens opacities: the Framingham studies. Arch Intern Med 161:1448–1454

Wu K, Willett WC, Fuchs CS, Colditz GA, Giovannucci EL (2002) Calcium intake and risk of colon cancer in women and men. J Natl Cancer Inst 94:437–446

Wysowski DK, Pitts M, Beitz J (2001) An analysis of reports of depression and suicide in patients treated with isotretinoin. J Am Acad Dermatol 45:515–519

Young-Xu Y, Chan KA, Liao JK, Ravid S, Blatt CM (2003) Long-term statin use and psychological well-being. J Am Coll Cardiol 42:690–697