Computerised data bases on prescription drug use and health care in the community of Tierp, Sweden: Experiences and challenges from a study of antidepressant-treated patients

Kerstin Bingefors
Department of Pharmacy, Uppsala university, Box 580, SE-751 23 Uppsala, Sweden, and Department of Neuroscience, Psychiatry, University Hospital, SE-751 85 Uppsala, Sweden

ABSTRACT

Much of our knowledge of drugs originates from clinical trials of drug efficacy performed on stringently selected patient groups, often without multiple concurrent diseases. However, the effectiveness of treatment under conditions of use in ordinary clinical practice may be very different to conditions in the randomised clinical trial. Use of large computerised data bases and record linkage has thus become increasingly common in pharmacoepidemiologic research. The greatest advantages of using routinely collected data are the minimisation of study costs and time required to complete a study, considerations that are particularly relevant for longitudinal studies. The advantages of using data bases also include the possibility of obtaining large sample sizes and to retrospectively study long-term outcomes. The risk for recall bias, a significant problem in interviews and questionnaires, is also reduced. However, computerised data bases also have some potentially serious disadvantages, primarily in the areas of data validity and data availability. The Tierp study, including individually based data bases of prescription drug use, will be used here as an example of research. In this paper an example of a comprehensive data base study concerning health care and drug utilisation in depressed patients is presented. Methodological considerations in data base research are discussed in relation to experiences from the antidepressant study. A well planned and research oriented computerised data base on prescription drugs represents an important tool in the study of the outcome of drug treatment in real world clinical practice.

INTRODUCTION

Much of our knowledge of drugs originates from clinical trials of drug efficacy performed on stringently selected patient groups, often without multiple concurrent diseases. However, the effectiveness of treatment under conditions of use in ordinary clinical practice may be very different to conditions in the randomised clinical trial. Figure 1 describes the stages of drug development pre- and post registration. After registration, drugs are used in a variety of patients and settings not included in the pre-marketing clinical studies. Therefore the study of medication use in actual clinical practice is important.

Observational studies of drug treatment in clinical practice are often based on computerised prescription data bases. However, limitations such as confounding by indication, channelling (a drug is given to a selected group of patients, e.g. with a more severe disease), and lack of important information have to be considered in the interpretation of results (1). Prescription drug data bases have also been used to analyse adverse outcomes of drug therapy, such as side-effects on the central nervous system of beta-blocking agents (2-5).

Use of large computerised data bases and record linkage have thus become increasingly important in pharmacoepidemiologic research. The greatest advantages of using routinely collected data are the minimisation of study costs and time required to complete a study, considerations that are particularly relevant for longitudinal studies. The advantages of using data bases also include the possibility of obtaining large sample sizes at low cost, and the retrospective study of long-term outcomes. The risk for recall bias, a significant problem in interviews and questionnaires, is also reduced. However, computerised data bases also have some potentially serious disadvantages, primarily in the areas of data validity and data availability. The majority of data bases cover only selected segments of the population, e.g. Medicaid beneficiaries or Health Maintenance Organisation enrollees (USA) (6). Some data bases, however, like the Canadian Saskatchewan health data base (7) or the Danish Funen data base (8) cover all inhabitants in a defined area, and are truly population based. In Sweden there are two population-based data bases that include prescription drug use; the Jämtland study (9) covering a sample of the population in the county of Jämtland, and the Tierp study (10,11) from which I will give an example of research.

Medication use has been employed in epidemiological studies of health problems in general populations (12-14) and patient groups (15,16). In these studies the prescription is used as a measure of a decision to treat by a physician; the prescription has therefore been cal-
Figure 1. The drug development process and factors influencing efficacy and effectiveness of drug treatment.

led the final common pathway in therapeutic decision making (17). However, drug prescribing is influenced by several non-medical factors, such as physician prescription habits, type of practice, and patient load (18). According to Avorn (17), forces and disciplines such as anthropology, decision science, health economics, ethics and politics are all involved as well as pharmacology and clinical medicine.

**DEPRESSION AS A PUBLIC HEALTH PROBLEM**

The Global Burden of Disease Study performed by the WHO estimated depression as the fourth leading cause of disease burden in the world, measured as disability adjusted life years (DALYs) (19). The lifetime risk for depression has been estimated to 7-12 percent for men and 20-25 percent for women (20). This gender difference is found in community samples, and are thus not due to differences in help seeking behaviour. As many as 10-15 percent of individuals may require professional treatment for depression during their lifetime (20).

The majority of depressed patients are seen in primary care and not in the specialised mental health sector. It has been estimated that most depressed patients are not properly recognised. The Depression Research in European Society (DEPRES) study found that only 57 percent of individuals with depression consulted health care, only 31 percent received drug treatment; out of these 31 percent only one quarter received antidepressant drug treatment (21).

A number of studies have shown that individuals suffering from depression or depressive symptoms are high utilisers of health care, and that they are as functionally impaired as patients with severe chronic diseases (22,23).

The low detection rate of depression and uncertain diagnostic practices in primary care presents a problem in the case-finding phase of data base studies if one wants to study depression with a population based approach. The "gold standard" in case finding would be to screen the target population or a representative sample with one of the validated instruments available, and then follow up detected cases by the registries (24-26). However, this procedure requires very substantial input of time and money in order to assemble a cohort of patients and controls large enough for a conclusive study. In the studies presented here, we chose an alternative approach to a resource-consuming screening procedure, namely to concentrate efforts on the study of an easily defined register cohort; patients who have received antidepressant drug treatment. Antidepressant-treated patients and their concomitant somatic health problems are of interest in themselves, considering the problems of treating multiple illnesses in this patient group. Consequently, the results presented here deal with antidepressant-treated patients in ambulatory care, regardless of the reasons for treatment.

**THE TIERP STUDY: STUDY POPULATION AND METHODS**

The study area covered by the data base is the community of Tierp in mid eastern Sweden with a population of around 20,000, which has remained relatively stable over time. The studies shown here originated in the 80's for purposes of follow-up. The population in Tierp was then somewhat older than in Sweden as a whole; 22 percent was 65 years or older as opposed to 16 percent in the country. The community includes a few small industrial townships and a relatively large rural area. Tierp is a well defined geographical area.

The Uppsala University Centre for Primary Care Research has collected data on individual health care utilisation for the Tierp population since 1972. Visits
in outpatient care, prescriptions filled, and hospital care are all recorded in a database using the unique civic registration number as identifier. The database is linked to information on population changes, such as births, migration and deaths.

All prescription drug purchases made by the community residents from local Tierp pharmacies are recorded. Data on type of drug, date of prescription, date filled, and prescribing doctor are transferred from prescriptions to a computerised research register. Prescription drugs purchased outside the community are not included, but the magnitude of this attrition has been determined to be only about 5 percent of the total prescriptions (10). Prescriptions are registered according to the pharmacological classification used in Sweden, since 1988 this is the Anatomic Therapeutic Chemical classification (ATC) system (27).

All visits to the health care centre made by community residents are registered. In the earlier years this was done through forms filled in by the attending physician and sent to the Research Centre. Nowadays a record linkage system gather diagnoses directly from the computerised medical records. The diagnoses are coded according to the International Classification of Diseases (ICD) system adapted for primary care (28).

The registration of health centre visits is almost complete, but there are missing data on visits to private physicians outside the community (11). In addition, visits made by community residents to outpatient departments in Uppsala hospitals are registered. Information on the number of inpatient stays at Uppsala hospitals, including discharge diagnoses, is also collected for each individual. The hospital data are based on administrative routines at the Uppsala county office.

In the studies discussed in this paper, we used two approaches. One was cross-sectional, where medication and health care use among individuals who had received at least one prescription for antidepressants in a calendar year (users) were compared with the rest of the population (non-users) (29). The second approach was longitudinal, first time users of antidepressants were defined as individuals who filled a prescription and had no prescriptions for antidepressants for five years before the date for this prescription (index date). Data was retrospectively collected for 5 years prior to the index date and participants were then followed for 6 years after the index date. They were compared with an age and sex matched referent group (29).

**Statistical methods**

In the cross-sectional study, mean health care utilisation was analysed using the Wilcoxon rank-sum test, a non-parametric test for differences between means. The proportions of users of different drugs were compared using Mantel-Haenzel stratified analysis with Chi-square tests and Fisher exact test when this was applicable (30).

In the longitudinal studies, analyses of repeated measures were performed using generalised estimating equations (GEE) as proposed by Zeger et al. (31). We analysed our data using the SAS macro developed by the Johns Hopkins group (32).

**THE USE OF PHARMACEUTICALS AMONG PATIENTS TREATED WITH ANTIDEPRESSANTS**

The cross-sectional study showed that patients treated with antidepressant drugs in ambulatory care were heavy concurrent users of health services and of prescription drugs of virtually all pharmacological classes (29). In the age group 45-64 years, the relative risks of receiving drugs from five or more ATC classes were 3.7 (men) and 3.1 (women). Other studies of depressed patients have confirmed this higher concurrent use of health services for somatic complaints (33,34).

---

**Figure 2.** Average use of non-psychotropic prescription drugs five years before the index date antidepressant prescription (periods I-5 to I-1) and six years after the index date antidepressant prescription. Antidepressant-treated (●) and referents (■).
The longitudinal studies showed that these patients had been heavy users of prescription drugs for a long time before the index date, Fig 2. Furthermore, they had considerably more visits to primary care physicians for somatic complaints for many years prior to receiving antidepressive drug therapy. This heavy use of physician services and medications lasted for several years after the initial antidepressant prescription (11,35). Furthermore, only forty percent of the patients initially received an adequate drug treatment in terms of dose and duration (11).

Persons with chronic disease commonly use more drugs also for other indications, which partly may be because they more often encounter a prescriber (15). However, somatic comorbidity is presumably the most important explanation of the very frequent use of non-psychotropic medications found among antidepressant-treated patients in the community. Some of the concurrent medications could possibly also have been prescribed to counteract possible side effects of tricyclic antidepressants (e.g. dry mouth, urinary retention, and constipation). Several commonly used drugs have also been suspected to cause depressive or other psychiatric symptoms, but the evidence for these associations rests mainly on case reports and inadequately controlled studies (20).

The high rate of concomitant medication use and polypharmacy in antidepressant-treated patients may be a serious case for concern because polypharmacy is linked to increased problems with adverse drug reactions, interactions, and non-compliance, and increased health care costs.

METHODOLOGICAL CONSIDERATIONS, EXPERIENCES AND CHALLENGES

A computerised data base is always subject to certain limitations. Stergachis (36) listed a number of factors that determine the utility of automated data bases. These factors will be discussed here in relation to research in the Tierp data base.

Data completeness

In a population-based data base, all events experienced by the population in the recorded types of health services use should appear. This is, however, rarely the case. In the studies discussed here, patient identification was based on the filling of antidepressant prescriptions. Hence, an important issue is whether a larger proportion of the prescriptions for these drugs were filled outside Tierp. There may be two reasons for getting a prescription filled outside the community; one is due to convenience, e.g. when visiting a clinic in a neighbouring town. Another plausible reason may be concerns regarding anonymity. Some patients may feel that antidepressant use is such a sensitive issue that they prefer to fill their prescriptions in a larger town where they are unknown. In our studies we had no way of determining attrition specifically for antidepressant-treated patients, but in Tierp only five percent of the total number of prescriptions are filled outside the community, according to an earlier evaluation (10).

Patient identification

One advantage of conducting research in the Nordic countries is that patient identification rarely poses a problem. The unique identifier required in Sweden is a 10-digit civic registration number for each inhabitant and this number is routinely used in all health care contacts. A full civic registration number is recorded in the Tierp data base.

Follow-up

Missing values caused by losses to follow-up represents a particular problem in longitudinal studies (37-39). In our studies, the major reason for loss to follow-up was mortality. We included those lost to follow-up in the analyses for as long as data were available for them because the aim for our investigation was to analyse the uses of prescription drugs and health care from an epidemiological perspective. An exclusion of those who died during the study, would therefore represent a selection bias. In order to test the effects of missing values, statistical analyses (GEE) were also performed with exclusion of those who were lost to follow-up. These results were not significantly different, which suggests that the missing values, although not missing at random, did not represent a significant error in the final studies.

Losses to follow-up through migration was relatively low, about 7 percent in nine years, and did not differ between groups. Individuals who left the community were relatively young.

Internal validity

Internal validity refers to the degree that a measurement records what it purports to measure, i.e. the percentage of individuals with a given characteristic in the data base who truly has that attribute. Data validity involving drug exposures is influenced by treatment compliance; a filled prescription does not automatically mean that the patient has actually taken the drug. Because the prescription was not primarily used to ascertain drug exposure in our studies, but to measure the physician’s intention to treat, compliance issues were not so important here.

Generally speaking, data validity in prescription records is far better than in interviews and questionnaires (40). Some misclassifications in the data base were discovered when we selected our antidepressant-treated cohort. All original prescription forms were checked, and the few misclassified patients were deleted from the cohort.
In selecting the study cohort we used antidepressant prescriptions as a proxy marker for depression. It is important to note that users of antidepressants are not equivalent to patients with the diagnosis of depression. A large body of literature discusses the low detection- and treatment rate of depression in primary care, and in general medical settings (20,24,41). The reasons for prescribing antidepressants include a number of different indications (42). However, as discussed previously, prescription drug use can be used as an indicator of an intention to treat by a physician. A chart review of the cohort of first time users showed that a majority (80%) of the index date prescriptions in the longitudinal studies (11,35) were issued with the intention to treat depression or depressive symptoms. The introduction of new treatment entities and/or new licensed indications for a drug can affect the validity of prescription drug use as a marker for disease. However, the studies presented here were all performed before the new generation of antidepressant drugs (e.g. SSRIs) was launched on the market. It has been suggested that the introduction of the newer antidepressants has contributed to more adequate drug therapy for depression in ambulatory care (43,44).

The low detection- and treatment rate of depression in the population most likely means that the referent group also included persons that should have been treated. This may contribute to an underestimation of effects caused by depressive problems on the total prescription drug utilisation, health care use, and mortality. On the other hand, the antidepressant group probably includes those with the more severe symptoms which consequently lead to a higher risk for excess health care use.

**Ethical considerations**

The individual’s right to privacy and confidentiality is always important in the use of research data bases on health care utilisation. In Sweden, the use of the civic registration number, not only in contacts with health care, but also in virtually all contacts with the authorities, poses a particular problem. The individual’s right to privacy has to be balanced against the benefits to the patients and to society, and for research.

Individually recorded data are essential for any valid discussion on causality. A unique identifier is not only essential in order to follow subjects within a certain register, but also to establish record linkage with other data bases. In epidemiological studies, the study of uncommon events require large sample sizes, often only to be found in large population based data bases.

Although the information in the Tierp registers is recorded for each individual, the researcher always works with anonymous data files. The compilation of research data files is done by data base managers, and is restricted to one specific location. Information on the project to patients is given at points of contact with health care in the community. Anyone who does not wish to participate has the right to abstain. As yet, there has been virtually no opposition against the registration. The Tierp study has rather been regarded as an asset to the community (46).

The Tierp project has been approved by the research ethics committee at the Uppsala University Faculty of Medicine and by the Swedish Data Inspection Board.

**CONCLUDING REMARKS**

For many patient groups there are continuing needs for studies concerning the adequacy of health care, including continuity of care. The example presented in this paper concerned antidepressant-treated patients. They had frequent contacts with the health care system during many years. A question may be whether they have been subjected to unnecessary medical interventions. Studies evaluating specific hypotheses concerning adverse reactions to antidepressant treatment, depression as a side effect of somatic treatment, and iatrogenic disease due to drug-drug and drug-disease interactions in these patients would be of importance. Among these patients, the lifetime consequences in the terms of personal suffering, the burden on health care resources, and the economic burden are all factors which need to be further investigated. Our studies demonstrate the usefulness of studying drug treatment in actual clinical practice settings. Antidepressant-treated patients are often elderly, two thirds are women, and they often suffer from long-standing concurrent somatic diseases. In short, the prevailing conditions in this population are altogether different from those in randomised clinical trials.
The analyses presented here share methodological reservations with most other studies based on information from computerised registers. However, it is again important to emphasise the main advantage, namely that the studies are population-based in a naturalistic setting and not based on selected patients from specialised clinics, or selected from health insurance populations. Results, even when interpreted with caution, should represent the prevailing condition in the population studied.

Prescription drug data bases are of considerable value in the study of medication use in real world clinical practice, however it is important to develop adequate methods for the analyses of these data. The studies reported here show that they also can be used to evaluate outcome of drug treatment in a population as well as the burden of disease in a health care system.

It is important to emphasise that a population based computer register of prescription drugs has to be intended for research. Access should only be granted to researchers, and confidentiality is an extremely important issue. In the European Union as well as in the US, there have been intensive discussions on patient confidentiality for many years. Patient safety and trust should always be considered in establishing a data base. It is also important to ensure the collaboration of prescribing doctors. The discussion in Sweden on the use of a computerised register for control purposes has been detrimental to the establishment of a nation-wide prescription data base.

Reimbursement authorities are now faced with an accelerating drug bill which will affect the amount possible to spend on other forms of health care. A more adequate use of drugs and more research into the effectiveness of particularly new and expensive drugs in the population, would probably benefit society in the long run. A well planned and research oriented computerised data base on prescription drugs represents an important tool in the study of the outcome of drug treatment in real world clinical practice.

REFERENCES