Existing data sources for clinical epidemiology: Aarhus University Prescription Database

Abstract: Population-based prescription databases in Nordic countries have become a mainstay of epidemiologic research. Denmark has both national and regional population-based prescription databases. Aarhus University Prescription Database collects data on reimbursed medications dispensed at all community pharmacies of the North Denmark Region and the Central Denmark Region. The regions have a combined population of 1.8 million inhabitants, or one-third of the Danish population. Denmark’s primary health care sector, which includes general practitioners, specialists, and dentists, generates about 96% of the prescription sales, most of which are reimbursable and are dispensed by the community pharmacies. The Aarhus University Prescription Database combines the region’s pharmacy records in a single database, maintained and updated for research purposes. Each dispensation record contains patient-, drug-, and prescriber-related data. Dispensation records retain patients’ universal personal identifier, which allows for individual-level linkage to all Danish registries and medical databases. The linked data have many applications in clinical epidemiology, including drug utilization studies, safety monitoring, etiologic research, and validation studies.

Keywords: epidemiology, prescriptions, registry-based research
containing largely overlapping information. The regional databases thus represent an alternative source of population-based prescription data. Recently, Furu et al provided a brief comparative overview of all Nordic prescription databases. The early description of Aarhus University Prescription Database was published in 1996, before Denmark’s municipal reform of 2007. This paper aims to provide a detailed and up-to-date description of the Aarhus University Prescription Database as a resource for epidemiologic research. Other names of this database used in published literature include “Pharmacoepidemiologic Prescription Database of North Jutland” and “Prescription Databases of the North Denmark and Central Denmark Regions”.

Health care and prescribed medicine in Denmark

Denmark is a welfare state, with tax-financed universal access to health services. Except for emergencies, patients’ initial contact with the health care system is through their general practitioners, who provide referrals to hospitals and specialists as necessary. The primary health care sector is represented by about 4100 general practitioners, 1200 specialists, and 4600 dentists, as well as physiotherapists, chiropractors, and home nurses. There are about 300 community pharmacies. Prescriptions originating from the primary health care sector account for approximately 96% of the total volume of medicinal product sales in Denmark.

Most prescription medications in Denmark are eligible for full or partial general reimbursement by the National Health Service (Sygesikringer). Adults (persons aged 18 years or older) become eligible for reimbursement after exceeding 850 Danish kroner (∼114 Euro, ∼151 US dollars) in annual out-of-pocket expenditure for prescription medicine; children (persons under age 18 years) are exempt from the out-of-pocket expense requirement. Drugs ineligible for general reimbursement include sedatives, hypnotics, oral contraceptives, and over-the-counter (OTC) medicine. Conditional reimbursement, however, may be granted for prescription-only and OTC medicines, if they are prescribed for chronic conditions or to certain patient groups (eg, senior citizens). The Danish Medicines Agency (Lægemiddelstyrelsen, www.dkma.dk) regularly posts on its website updated rosters of generally and conditionally reimbursable medicines. Changes in reimbursement policies have been described, since 2005, on a dedicated website (http://www.dkma.dk/1024/visUKLSArtikel.asp?artikelID = 6301).

Population and period covered by the Aarhus University Prescription Database

Aarhus University Prescription Database covers the population of the Central Denmark Region and the North Denmark Region. These are two of the five Danish regions, with a combined population of 1.8 million inhabitants, or about one-third of the total Danish population (Table 1, Figure 1). Following the municipal reform, which went into effect on 1 January 2007, regions replaced counties as administrative units in Denmark. As a result, the Central Denmark Region and the North Denmark Region now encompass four former counties: Aarhus, North Jutland, Ringkøbing, and Viborg, as well as seven municipalities of the former Vejle county (Table 1).

Before the municipal reform, the Aarhus University Prescription Database received and merged prescription data from Aarhus, North Jutland, Ringkøbing, and Viborg counties. Because counties started contributing at different times, they differ with respect to the earliest availability of

<table>
<thead>
<tr>
<th>Region (Danish name)</th>
<th>Population in 2009</th>
<th>Former counties included in the two regions</th>
<th>Coverage since</th>
</tr>
</thead>
<tbody>
<tr>
<td>North Denmark Region (Region Nordjylland) <a href="http://www.nn.dk">www.nn.dk</a></td>
<td>580 515</td>
<td>Aarhus (one municipality)</td>
<td>1996</td>
</tr>
<tr>
<td></td>
<td></td>
<td>North Jutland</td>
<td>1989a</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Viborg (five municipalities)</td>
<td>1998</td>
</tr>
<tr>
<td>Central Denmark Region (Region Midjylland) <a href="http://www.regionmidjylland.dk">www.regionmidjylland.dk</a></td>
<td>1 247 732</td>
<td>Aarhus (all but one municipalities)</td>
<td>1996</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Ringkøbing</td>
<td>1998</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Vejle (seven municipalities)</td>
<td>2007</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Viborg (12 municipalities)</td>
<td>1998</td>
</tr>
</tbody>
</table>

Data sourced from www.regioner.dk

Notes: aComplete coverage since February 1992; aAnnexed to the Central Denmark Region as a result of the 2007 municipal reform and are therefore represented in the regions’ prescription database from 2007 onwards.
the data. Data from the former North Jutland county are available since 1992,\(^6\) from the former Aarhus county since 1996, and from the former Viborg and Ringkjøbing counties since 1998. After the municipal reform, the Aarhus University Prescription Database has been merging prescription data from the community pharmacies of the Central Denmark Region and the North Denmark Region.

### Data flow

Ambulatory patients redeem prescriptions at community pharmacies, which electronically transmit dispensation data to the regional subdivisions of the National Health Service for reimbursement. Prescription reimbursement records must be de-identified within two months of sale.\(^7\) In order to enable individual-level linkage of prescription data to other medical databases, patient records need to remain identifiable. According to an agreement with Aarhus University, the National Health Service subdivisions of the Central Denmark Region and the North Denmark Region transfer individually identifiable dispensation data collected from the community pharmacies to the Department of Clinical Epidemiology at Aarhus University Hospital. The transfer is conducted using a password-protected encrypted connection, and the data are kept on a secure server. Data protection is further discussed below. The dispensation data are also reported to the Registry of Medicinal Products Statistics.\(^1\)

### Variables

One drug dispensation represents the unit of observation in the Aarhus University Prescription Database. Each line of record contains the following variables:

1. Patient’s identifier (the so-called CPR number), which is a unique 10-digit numeric code assigned at birth or immigration by Denmark’s Central Personal Registry. The first six digits of the CPR number correspond to the day, month, and year of birth, and the last digit encodes sex (odd digits for male; even digits for female).\(^12\) Prescriptions for children younger than 16 years were registered under the CPR numbers of their parents until April 1, 1996 and under the children’s own CPR numbers thereafter;
2. Date of sale, which, in combination with the CPR number, enables calculation of the patient’s age at dispensation;
3. Type of drug, which is coded according to the anatomical therapeutic chemical (ATC) classification;\(^13\)
4. Universal product number (Varenummer), which encodes medication name, strength, pack size, dose units (eg, mg or mL), defined daily dose (DDD), and manufacturer;
5. Provider code, which corresponds to a clinic, a practice, or a single physician. The provider code enables identification of the prescriber’s specialty, name, and address, and in case of single practitioners, their age and sex.

### What is registered in the Aarhus University Prescription Database?

Aarhus University Prescription Database records prescription medications that receive general or conditional reimbursement dispensed at the community pharmacies of the Central Denmark Region and the North Denmark Region. In 2009, 14.9 million dispensations were reported to the database. Table 2 and Figure 2 show the distribution of these dispensations by the major ATC classification group. Dispensations in group C (cardiovascular system) account for more than one-quarter of all dispensations, followed by dispensations in group N (nervous system), which account for more than one-fifth (Figure 2).

Thirty ATC codes account for half of all dispensations. The five most frequently dispensed medications are simvastatin (ATC code C10AA01), acetylsalicylic acid (ASA, B01AC06), paracetamol (N02BE01), metoprolol (C07AB02), and bendroflumethiazide (C03AB01) (Table 3).

The distribution of dispensations reflects the extent of by-prescription use of ASA and paracetamol, which can be sold over the counter. ASA is reimbursed (and therefore recorded) if prescribed for specific indications, including primary and secondary prophylaxis of cardiovascular diseases.\(^10\)
Therefore, prescriptions for ASA recorded in the Aarhus University Prescription Database are expected to originate primarily from persons older than 55 years, or those with increased risk of cardiovascular morbidity. Paracetamol is reimbursable if prescribed for chronic pain. An estimated 11% of ASA and 80% of paracetamol packages prescribed in the primary sector are not reimbursed and therefore not recorded in the Aarhus University Prescription Database. Thus, for conditionally reimbursed medications, the proportion of dispensations recorded in the Aarhus University Prescription Database varies by pharmaceutical. The Internet-based database www.medstat.dk provides on-demand statistics (in Danish) on annual sales of pharmaceuticals with and without reimbursement in all sectors of health care. The aggregate statistics are available for each medication sold in Denmark and can be defined in packages or in daily doses. Using these data, degree of underascertainment of medication use due to lack of reimbursement can be estimated on a case-by-case basis.

What is not registered

The Aarhus University Prescription Database does not contain data on drugs that do not receive general or conditional reimbursement or drugs dispensed outside the community pharmacies of the two regions. The non-reimbursable medications include, for example, oral contraceptives and OTC medications (unless prescribed for chronic conditions, as discussed above). Reimbursed medications dispensed outside community pharmacies include medications dispensed directly to patients in hospitals or outpatient clinics. These include, for example, outpatient anti-neoplastic treatments, antiretroviral treatments for patients with human immunodeficiency virus infection, or substitution therapy for substance abusers. For drug-specific information regarding reimbursement rules and distribution of pharmaceutical sales according to reimbursement status and health care sector, the readers are referred to the relevant publications by the Danish Medicines Agency that are freely available on the Internet, the latter publication also in English.

Quality of data

Pharmacies use bar code scanning to enter drug information into their database, making drug-related data entry errors highly unlikely. Potential dispensation of a wrong drug to a patient cannot be detected by examining the electronic data, but the frequency of such occasions is assumed to be negligible to meaningfully affect most study results. For the population of the two regions, dispensation records of reimbursed medications from community pharmacies, which are reported to the Aarhus University Prescription Database, are expected to...
be the same as records reported to the Registry of Medicinal Products Statistics. A cross-tabulation, at our department, of records with available data from both sources confirmed this expectation (unpublished data).

An estimated completeness of the Aarhus University Prescription Database is 96%, based on cross-tabulation of insulin prescriptions with hospitalization records of diabetes mellitus in the former North Jutland county. However, completeness may vary by drug type. Furthermore, although dispensation record represents patients with primary adherence (those who purchase medications), it reflects neither the exact timing of the actual medication intake.

Record linkage and use of data in clinical research

Thanks to patient-identifiable data in the Aarhus University Prescription Database, it is technically easy to link it to all other Danish registries, both nationwide and regional, using the CPR number. Because Denmark’s registries are numerous and far-reaching even by the standards of Nordic countries, such linkage opens many possibilities for epidemiologic and interdisciplinary research. Linkage can be done, for example, to migration and vital status data in the Civil Registration System, hospitalization records in the Danish National Patient Registry, cancer diagnoses in the Danish Cancer Registry,
or birth records in the Danish Medical Birth Registry.\textsuperscript{16} Linkage to the birth records is a valuable resource for studies of medication safety in pregnancy. Birth records contain both maternal and paternal CPR numbers of live- or still-born babies, allowing linkage of prescription and hospital records of parents and children. Examples of studies involving linkage to different data sources are given below. Even unlinked dispensation data can provide a wealth of information. Because each dispensation is identified by CPR number, it is possible to study patterns of drug utilization by patient age and sex and to reconstruct individual dispensation trajectories over time. The availability of prescription data since the mid-1990s enables the conduct of intergenerational studies. Besides Denmark, Finland is the only other Nordic country in which this is possible, because of the availability of prescription data since the 1990s, eg, parental use of medication before conception or birth can be examined for today’s young adults.\textsuperscript{1}

**Data access**

The Danish Act on Processing of Personal Data (\textit{Persondataloven}) provides the legal basis for the ability of public institutions, including universities, to retain person-identifiable health data for research purposes.\textsuperscript{3,17} In order to access the data from the Aarhus University Prescription Database, researchers may apply to the Department of Clinical Epidemiology, at Aarhus University Hospital (www.kea.au.dk). Furthermore, use of any health data requires project-specific permission from the Data Protection Agency (Datatilsynet, www.datatilsynet.dk), and often additional permission for linkage from the National Board of Health (Sundhedsstyrelsen, www.sst.dk).

**Examples of studies conducted with data from the Aarhus University Prescription Database**

Reports of studies based on data from the Aarhus University Prescription Database have been published in several major international peer-reviewed journals and include both safety investigations and epidemiologic studies that use prescription record as a proxy measure of disease. A non-exhaustive list of examples includes investigations among pregnant women based on linkage to birth registry, including safety of non-steroidal anti-inflammatory drugs (NSAIDs),\textsuperscript{18} loratadine,\textsuperscript{19} fluconazole,\textsuperscript{20} selective serotonin reuptake inhibitors (SSRIs),\textsuperscript{21} and clomifene.\textsuperscript{22} Nørgaard et al used linkage with hospitalization records to investigate risk of miscarriage among pregnant users of pivmecillinam, while using data on the women’s dispensation of anti-epileptic and anti-diabetic medications to control potential confounding by epilepsy and diabetes.\textsuperscript{23} Likewise, by linkage to hospitalization records, Johnsen et al examined the risk of myocardial infarction among users of several NSAIDs.\textsuperscript{24} Fryzek et al studied cancer risk among women using antihypertensive medications.\textsuperscript{25} Nørgaard et al used prescriptions for SSRIs as a marker of depression in a study of acute pancreatitis,\textsuperscript{26} and Lash et al and Cronin-Fenton et al studied the role of tamoxifen and SSRIs in preventing the recurrence of breast cancer.\textsuperscript{27–29} Finally, several investigations demonstrated socioeconomic differences in drug utilization by linking prescription dispensation records from Aarhus University Prescription Database to individual-level socioeconomic indicators.\textsuperscript{30–33}

**Conclusion**

Routinely recorded person-identifiable prescription data have many uses, including post-approval research or etiologic studies. The Aarhus University Prescription Database contains, with few exceptions, dispensation data on reimbursed medicine for the entire population of two Danish geographic regions representing one-third of the country’s population. The data are accessible for research and are linkable to all other Danish databases and registries. The main drawbacks of the Aarhus University Prescription Database are a lack of nationwide coverage and the absence of data of certain medication types. At the same time, a vast majority of prescription dispensations are routinely recorded, with more than 15 years’ worth of follow-up data accumulated to date. Thus, the Aarhus University Prescription Database represents a valuable research tool for many epidemiologic applications.

**Disclosure**

The authors report no conflict of interest for this work.

**References**