

Håkan Sjöberg, Andreas Blomquist

Swedish registers - a gold mine for medical research

Sweden and other Nordic countries are unique in the sense that we hold detailed comprehensive registers with information on all the registered population. The records are of great value for Swedish research institutions and often yield competitive advantages towards researchers outside of the Nordic countries, unable to access similar data.

An independent international group of six renowned scientists, including two from the U.S. and one each from Great Britain, Holland, Denmark and Finland, have evaluated the Swedish public health research. Their positive review was summarized by the following statement:

"Sweden is one of the world leaders in PHR (public health research) – in several areas of PHR it is at the fore-front of research exemplified by publications in top international scientific journals. Among these areas are epidemiology and register based research, ... This is due to a unique combination of an excellent data infrastructure, an enlightened public sector and a productive public health research workforce." (Kamper-Jørgensen, 2005)

The prerequisite for the positive rating is a tradition of having comprehensive national data registers and in a system where personal identification number makes it possible to track individuals over long periods of time in longitudinal studies.

Some of the Swedish registers at Statistics Sweden

Total Population Register (TPR)

The stock registers of the TPR are annual registers containing data on the registered population as of 31 December each year, from 1968 onwards. TPR is mainly used as background information in medical research and examples of some of the variables are sex, age, municipality and country of birth.

Also, the TPR contain annual events registers with events that took place during the year. TPR contains the following event registers:

- Births, live births
- Births, stillborn
- Deaths
- Immigrations
- Emigrations
- Migrations within Sweden

- Changes in civil status
- Changes in citizenship

Multi-Generation Register (MGR)

The MGR was created focusing on the child-parent (biological and adoptive) relationship. With a target population defined as all individuals born in 1932 or later and registered in Sweden at any time since 1961, MGR enable the study of biological inter- and intra-generational relationships, of great interest for medical research. Currently the target population is some 10 700 000 individuals and the whole register contains a total of 15 500 000 individuals, with linkage to parents outside of the target population. Due to birth and immigration the register increases each year. Thus, Statistics Sweden can deduct children, grandchildren, siblings, parents, grandparents, cousins, etc. within a defined population.

The Swedish Occupation Register

The Swedish Occupation Register covers all registered persons 16 years and older in Sweden as of 31 of December each year. The register is primarily aimed at providing annual information on the working population's occupational distribution and to highlight occupational development in various industries and sectors. Within the research field of occupational medicine, for example, the register can be used to establish relationships between occupational exposure and health.

LISA Database

The LISA database (by Swedish acronym) is a longitudinal integration database for health insurance and labour market studies. LISA is made up of annual registers since 1990 and includes all individuals 16 years of age and older, registered in Sweden as of December 31 each year. The database integrates existing data from registers on the labour market, education, social sectors using background information from population registers. Each year a new annual integrated version is added. The primary objects in LISA are individuals, with connections to family, employers and location of workplaces, available through special retrieval tables and descriptive data.

Other registers

Some of the other registers at Statistics Sweden used in medical research are:

- Educational attainment of the population
- Register based labour market statistics
- Income and Tax register
- Geographical information database

The use of registers in medical research

Data linkage

It is possible to obtain databases with linkage of data from more than one authority holding registers. Statistics Sweden often collaborate with authorities such as The National Board of Health and Welfare, The Swedish National Council for Crime Prevention and The National Service Administration. This is possible using key files of personal identification numbers and randomized serial numbers uniquely created for each project database. Using the serial number, the researcher can merge data from the separate deliveries made by different authorities within a project. Examples of useful registers not located at Statistics Sweden are:

- Cancer Register
- Cause-of-death Register
- Medical Birth Register
- Health Data Register

Cohort studies

At Statistics Sweden a study population may be defined as a cross section of the population, a randomized sample or be derived using specific characteristics such as age, sex, residential area, etc. using population registers. Study populations derived from diagnosed disease or special risk factors yields collaboration with other authorities.

Case control studies

In case control studies two existing groups, differing in outcome or exposure, are identified and compared on the basis of some supposed causal attribute. Case-control studies are often used to identify factors that may contribute to a medical condition by comparing subjects with a condition/disease (the cases) with individuals without the condition/disease but otherwise similar (the controls). A major issue in case and control studies is the high risk of bias between cases and controls. If, for example, the controls are younger than the cases and the study outcome is correlated with age the results of the study may be incorrect. One way to avoid this is to try to get controls that resemble the cases in factors that could affect the outcome. Due to the Total Population Register at Statistics Sweden it is possible to match cases and control individuals by age, sex and residential area at the time of outcome. Since the total population is available for control selection, it is possible to match more than one control individual per case, thus increasing the power of the study.

GIS and epidemiology

Data on the spatial patterns of disease and exposure can provide useful tools for analysis. The Geographical information database at Statistics Sweden makes it possible to link all residents and most workplaces in Sweden to coordinates and enable the use of Geographic information systems (GIS).

This technology is increasingly used to analyze geographical distribution of diseases as well as relationships between pathogenic factors and their geographic environments. Basic and analytical applications of GIS in epidemiology can help in visualizing and analyzing geographic distribution of diseases through time, thus revealing spatio-temporal trends, patterns, and relationships that would be more difficult or obscure to discover in tabular or other formats. GIS can provide a means to meet the demands of outbreak investigation and response, where understanding the spatial spread and dynamics of an outbreak is central to the design of prevention and control strategies.