

# **Some Danish experiences of computer-based civil registration and vital statistics**

## **1. Introduction**

The present paper gives an overview of the computerized civil registration system in Denmark and how this system is used as a source for vital statistics. It draws attention to matters that are important in building and running a system like this.

The paper concentrates on the description of the system of civil registration and vital statistics, but the more technical aspects are not discussed.

## **2. The need for civil registration and vital statistics**

The registration of vital events in the population and the compilation of vital statistics are essential tasks in every country. The registration of vital events like birth, death and marriage is important for the individual. The registration can give the necessary information, for example in form of a certificate, to make it possible for the individual to claim his rights in different areas. It could be the right to get medical care, to get education or to get a passport.

The statistical information is necessary for an efficient administration both on the local and on central level. Most public authorities are depending on vital statistics for development and planning purposes. And for the authorities the information about the population is important when it comes to the collection of taxes. The more advanced the society is, the greater seems the demand for statistics to be.

## **3. History of the vital registration and population statistics in Denmark**

There is a long tradition for registration of vital events and for population statistics in Denmark. Already back in the 16th century events like birth, death and marriage were registered in the church book of every parish and some of these books still exist. It is possible to find regular statistics on births and deaths back to 1743 and in 1769 the first census was conducted. Since 1840 there has been a census every ten year and from 1901 every five year up to 1970.

Since 1924 every municipality in Denmark has kept a local population register, i.e. a file containing information on all persons living in the municipality. These registers contain identifiable information such as name, occupation, date and place of birth, address, family circumstances and nationality. The municipalities are responsible for keeping the files continually updated, with information on births, deaths, etc. obtained from various public authorities. Individual citizens are obliged to report changes in address to the registration office.

Electronic data processing was introduced in 1968, and a Civil Registration System was established. A part of this new system included all the local population registers in one administrative system the Central Population Register (CPR). An essential part of the reform was the introduction of a permanent identification number for each individual person: the person number (PNR). The number was assigned to all residents in the country at that time and later to all the newborn and immigrants. This number was considered a practical necessity for the operation of the Civil Registration System. The main purpose of this new register was administrative, but statistical utilization was also an important element.

A special office under the Ministry of Interior is responsible for the Central Population Register. The office is responsible for the legislation about registration and for checking and controlling of the data in the register. The office is also responsible for creating the person numbers used in the system.

The municipalities are responsible for updating the register with all civil events concerning their resident population.

The responsibility for the general statistical utilization of the information in the Central Population Register rests with Statistics Denmark. Statistics Denmark is an independent institution under the supervision of a Board consisting of the head of Statistics Denmark as chairman and six other members familiar with social and economic conditions. The production of statistics in Denmark is highly centralized and Statistics Denmark is responsible for most of the statistics, not only in the field of population.

During the 1970's the register based computerized vital statistics were established and developed by Statistics Denmark. The basis for these statistics was information on individuals from the CPR, in some cases combined with information from other sources.

The way through the system from the event to the finished statistics and the issues to consider are treated below.

## **4. From event to statistics**

### **4.1 From event to administrative registers**

The quality of the register based computerized statistics can never be better than that of the basic data in the population registers. The greater the number of administrative functions linked to the population registers the more accurate and up to date they are likely to be. The reason why is that, opportunities of updating and correcting become greater with increased use of the register.

Two main points regarding the reliability of the register are completeness and timeliness. Therefore, it is essential to ensure that the relevant events are actually recorded on the register and that it happens within a short period of time. It is sometimes difficult to achieve both objectives, especially at the same time. Some events are more likely to be recorded on the register without involving problems than other. For example, for most people events like births, deaths and marriages are more obvious to register than changes in address or emigration. A linkage of the registration and issuing of required certificates can have a positive effect on the completeness of registration.

Who is responsible for notifying the register is an issue, which is to be considered. In Denmark, for example births are notified by local authorities and by the individual person, whereas for changes in address the individual person is the only one under the obligation to report. Two or more ways for notification of the register can, of course, enhance the possibility of capturing the event, but it should be balanced against the risk of duplicate registration and the risk of over-administration.

As a general rule, the population register keeps - at least for some time - information on persons who have died or left the region. It is therefore important to indicate for each person on the register whether the person is active or not.

In appendix 1 is shown a draft of the stages from event to administrative register, to statistical register and finally to statistics for some vital events in Denmark

### **4.2 From administrative registers to statistical registers**

An administrative register, like the Central Population Register, cannot be directly used for compiling statistics. The information extracted from the administrative register has to be selected and corrected before it is ready to be recorded on a statistical register, which is the basis for the statistics. In this process there are some things to be considered.

As a general rule, the administrative registers contain a vast quantity of information, which is not relevant for the statistics. The first thing to do is therefore to decide which data to select from the comprehensive file. It is important to be careful in the selection of data. All relevant data, which are going to be used here and now must of course be selected, but also data which in all probability will be of interest within the foreseeable future should be considered. On the other hand, data that cannot be used directly or indirectly (for example, for checking) should be avoided. It can be expensive to store and to treat data and data that are not regularly used often causes confusion in the system.

The extracts of data from an administrative register, like the Central Population Register, to a statistical register are mainly of two different types.

The first type, the status extract, is an extract of all relevant data for all persons who are registered as active persons on the register at a specific time. A status extract typically forms the basis for a yearly survey of the total population and its composition.

The second type of extract, the change extract, concerns all relevant changes in the population register, such as births, deaths etc. for a period of time.

A change extract typically forms the basis for statistics of population changes – births, deaths, marriages, migration etc.

These change extracts must be subjected to an initial processing, during which the category of change must be established for each individual case. It must particularly be decided whether a change has actually taken place, or whether it is only a case of correcting wrong data.

Owing to delays in reporting births, deaths and migrations etc., the population register never corresponds as regards contents to the resident population as it is at a specific time in question. The delays can differ widely from event to event. The delays concerning births and deaths are usually not very long, whereas reports concerning emigration are those with the longest delays in Denmark.

Danish experience has shown that within 30 days, more than 99 per cent of the events are reported. When making a status extract these delays have to be taken into account. An extract from the Central Population Register, as it is at a specific date, e.g. 1st January, is adjusted with regard to births, deaths, migration etc., which have taken place before 1st January, and were reported within 30 days of the new year. By means of this period of collection, it is ensured that all reports, which are not subject to extraordinary delays, are included in the statistics.

As for status extracts, attention has to be paid to delays in the reports, when processing change extracts. It is also necessary to decide what to do with changes reported to the register too late. One possible solution is to transfer these late reports to the next period (year or quarter of a year). In this way, all changes are registered in the statistics, but some of them in the wrong period. Another solution is to exclude these very late registrations from the statistics of changes and record them as corrections in a new status.

It is possible to create a new status with a status extract as basis and with regular change extracts. However, it is the experience that it is difficult to keep the quality of the status over the years by using this procedure. Small errors in the first status or in the change extract can over time develop, in such a way that after some years there are increasing differences between the calculated status and the real status. Therefore, processing the status extract regularly, e.g. once a year gives a better quality of data.

Like all data used for statistics, the extracts - status as well as change extracts - have to be checked for errors. Checks of computer-based information are a comparatively quick process, which makes it possible to effect a more extensive checking than in the old-fashioned processes. There are mainly two different types of error-checks to consider, validity checks and combination checks or logical checks.

A validity check ensures that for every single item only valid codes exist. For example, the item of sex will normally have only two acceptable codes, one for men and one for women, all other codes will be understood as errors.

The logical checks are based on comparison between two or more pieces of information concerning the same person. If the event e.g. is a birth, the mother's age is supposed to be in the interval 14-49 years. A marriage should be between a man and a woman, and there can be an age limit for marriages. In the Danish data extracts from the CPR today, less than 1 per cent of the items of information are defective. This can, of course, be attributed to a well-developed check for errors in the CPR.

The data extracts can, after check for errors and corrections, provide inputs to statistical registers. Statistical registers are usually organized after the type of statistic that is the final output. There is no one-to-one relation-

ship between extracts from administrative registers and statistical registers. The data extracts from one administrative register often constitute the input to several statistical registers and one statistical register often receives inputs from more than one administrative register, and sometimes the input is a combination of administrative extracts and data collected in other ways.

Once a year Statistics Denmark receives a *status extract* from the CPR. This extract forms together with 30 days' change extracts the basis of the statistical register, which in turn constitutes the basis of the annual statistics of the population and its composition according to various criteria. The unit in this statistical register is naturally the person, who is unambiguously identified by his/her person number.

By way of illustration, the data contents in the status extract and change extract from the CPR are shown in appendix 2.

Every week Statistics Denmark receives an *extract from the CPR of the changes*, which are of interest. The information from the weekly change extracts is compiled, after error checking, for statistical registers with quarterly information, which at a later stage is compiled for annual registers. Like the status register the unit is here the person. The statistical registers of changes are used as basis for the vital statistics.

### **4.3 Further processing of the statistical information on the registers**

The basis of the population statistics in Denmark is, as stated above, the statistical registers, which are established on the basis of the extracts from the CPR – the status register and change register.

The information contained on the statistical register (see appendix 2) is subject to a further processing in order to provide the final information, which is to be included in the statistics. For example, information is constructed on the basis of the person number concerning date of birth, age and the last digit of the number is converted into the person's sex. Information on address is converted by means of the ancillary registers for various administrative distributions such as parishes and counties.

As already stated, the unit in the status register as well as the change register is the person, but information on the level of persons can be used for constructing other units. This is done in two cases on information from the status register.

The address, which is so exact that every apartment or dwelling has a unique address, is used for constructing households. All persons with the same address are referred to the same household.

The information on person number for mother, father and spouse is used together with the information on address for constructing families. Persons living at the same address and who have joint reference numbers are referred to the same family. This is the principal rule, which is subject to exceptions such as age of the child, etc. (they must be under the age of 18 years in order to be considered as children, and if they are married or if they have children of their own they constitute their own family).

For each of the new units a number of new data can be constructed. For households, data concerning, e.g. size of household and number of children in the household can be constructed. For families, a family type can be indicated, e.g. whether the family consists of a single person, a single person with children or a married couple with or without children. The size of the family is another piece of information, which can now be added to the information already recorded on the register.

For all units, a range of variables such as groupings of age, place of birth or citizenship can be constructed. It can be practical to have groupings of this type available as ready-made data.

The construction of new enumeration units on the basis of the existing material and the construction of new variables and variable groupings are processes, which call for major screenings and comparisons of variable values for several individuals, tasks which, in practice, can only be solved if the data material is computer-based.

In many cases, it is practical to divide the vast data material contained on one register into several sub-registers.

It will be natural to divide the data in accordance with the unit to which the information relates, i.e. information that concerns families separately and data concerning households separately. Data for the same unit – such as data on individuals – can of course also be divided according to any criteria, which are on the register.

In certain cases, the information extracted from the administrative registers does not cover all the statistical fields for which information is required. In such cases, it is necessary to replenish with information from other sources. In Denmark this concerns particularly one area, i.e. medical information in connection with births and deaths. Demographic information from the CPR is here supplemented with medical information from the birth and death certificates. The birth and death certificates are collected by Sundhedsstyrelsen (the National Board of Health), which is responsible for providing the information with code numbers, and thereafter submits the data to Statistics Denmark. The certificates contain the person number and information from the certificates can, therefore, be linked to information from the CPR.

There are several reasons why the medical information is not recorded on the CPR. First, the CPR is, as already mentioned, an administrative register and medical information does not belong to an administrative register. Medical information is considered by many people to be highly confidential, and should therefore be recorded in as few places as possible, and only in places where there is a qualified need for the data. The medical data often require a more qualified evaluation during the coding process than the general demographic data, and this is one of the reasons why that treatment of the medical data is the responsibility of Sundhedsstyrelsen.

By linking information from birth and death certificates with information from the CPR, it is possible to check the extent to which the register based information and information from the certificates correspond. Our experience shows that in about 99 per cent of all cases, it is possible to link the person number by computer. The remaining part, which cannot be linked, is usually evenly distributed between errors on the register and missing or defective certificates.

#### **4.4 The statistics in Denmark**

As already mentioned, once a year statistics of the population are compiled in Denmark as at 1st January. These status statistics contain information on the distribution of the population by sex, one-year-age groups, by marital status and by citizenship and country of origin. Furthermore, statistics of families by type and size and households by size are compiled.

All the statistics mentioned are compiled for all Denmark, but also for each municipality (Denmark is divided into 275 municipalities). Address information, which is available on the register, makes it possible to compile the same statistics for a range of other fields. However, as a standard only summary statistics of the population in parishes are compiled (2000 parishes).

The first final results from the annual status statistics, concerning the total population in all Denmark and in each municipality, are published in a newsletter in the middle of February. Simultaneously with the publication of the first available information, detailed information on the level of municipalities is made available for the users who subscribe to Statistics Denmark's data banks. The data banks constitute a system of databases to which the users can obtain on-line access. Today a fee is charged for access to the data banks but from the beginning of 2001 the access is free of charge.

The more detailed information is published, at a later stage, subject to the delays caused by the editing and technical production of major publications.

As mentioned, the weekly change extracts are compiled for quarterly vital statistics. For each quarter the number of births and deaths (preliminary figures), number of external and internal migrations, marriages and divorces are published. The statistics are analysed by sex. Once a year, statistics of naturalizations and adoptions are compiled in addition to the statistics already mentioned.

The quarterly statistics are compiled for all Denmark and for each municipality.

The quarterly statistics, like the status statistics, are also published with a delay, i.e. within 6 weeks after the end of the quarter.

The final statistics of births and deaths, i.e. the information that is constructed by linking information from the CPR and information from the certificates, are only compiled annually. They are normally available with a delay of 6 to 9 months.

## **5. Vital statistics as part of a register system**

The statistics described here constitute today part of a register-based statistical system in Denmark. Through the 1970's a range of administrative EDP registers were established, which by means of various keys can be linked in order to construct statistical information on persons, dwellings and business enterprises.

The information contained in the tax system concerning incomes and economic activity status can be linked to the demographic information in the CPR by means of the person number, which is the identification number in both systems. The registers in the tax system also contain for each person in employment an identification number for the employer. The employer's number is recorded as identification number in the register system for business statistics. This register system contains information on industry, size and ownership of the business enterprises.

The Central Register of Buildings and Dwellings, which contains a range data concerning quality and size for each dwelling, has the same exact address as identification as the one, which is used in the CPR. Consequently, it is possible to link information on the dwelling to individual persons.

The register of educational attainment contains information on persons receiving education or persons who have completed an education. The identification is here the person number.

Linking information from all these administrative register systems made it possible for Denmark in 1981 to conduct a population an housing census based entirely on extracts from registers and without applying directly to the population.

## **6. Centralized versus decentralized data processing**

The data input concerning the basic demographic information has been decentralized in Denmark, i.e. the local population register in each municipality is responsible for recording an event. However, the same event is also recorded on the Central Population Register. This combination consisting of decentralized and centralized registration yields a range of advantages. The local knowledge, which is a great strength of the decentralized data collection, is maintained, while at the same time the centralized registration ensures that, for example, duplicate registration is avoided.

The Danish system enhances the possibilities of a decentralized as well as a centralized production of population statistics. The information on the population stored by the municipalities is similar to the information stored by the central statistical authority Statistics Denmark. The Danish production of official statistics is, however, greatly centralized. Statistics Denmark is responsible for producing most of the statistics in the demographic, social and economic fields. Some of the reasons for a centralized production of the statistics appear from chapter 4.

The numerous processes, which are conducted before the final production of the statistics, can be difficult to handle and a homogeneous treatment of the data is very important for the final results.

The centralized statistical production ensures a high degree of consistency, e.g. the population figures for small areas add up to larger areas, which, in turn add up to figures for all Denmark. It is also possible to ensure consistency between the various statistical fields, e.g. by using the same definitions and classifications, when this is desirable. Furthermore, a centralized production also ensures certain advantages of a large-scale operation. This may involve that the data material from one field to another is re-used.

From a user's point of view, there are often advantages of a centralized production, which make it easier to get an overview of the statistics and to describe the relationship between the various statistics.

The decentralized production can, to a greater degree, be adapted to the local requirements as regards the time of compilations and the level of detail.

## **7. Data protection**

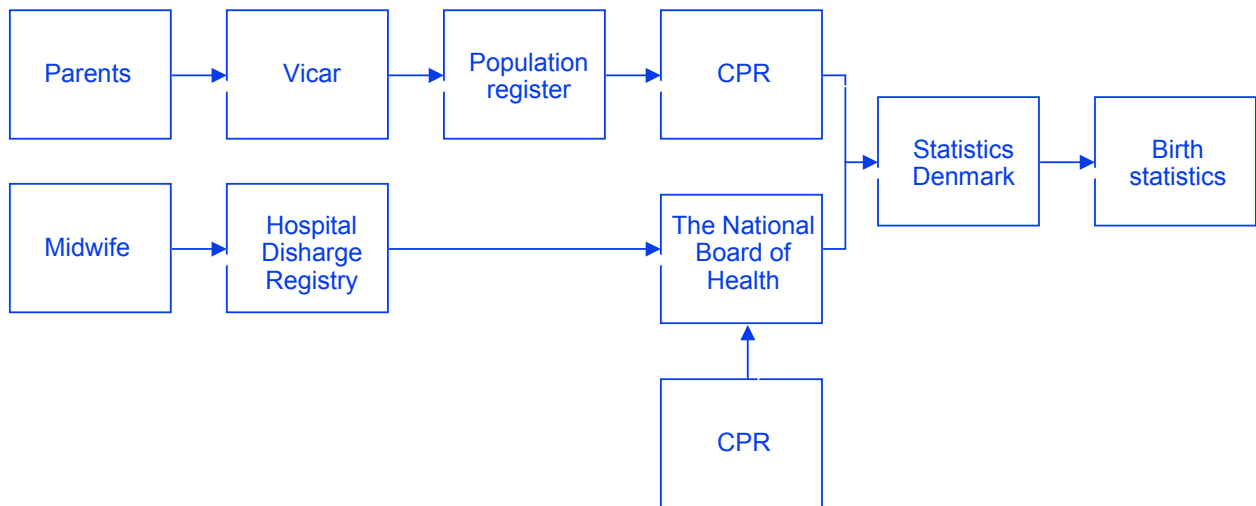
In order to be able to produce and maintain a statistical system, which presupposes that the producer of statistics has access to identifiable information on persons, it is necessary to attach importance to a data policy, which ensures the public against any abuse of the information on individuals.

In Denmark, like most other countries, legislation concerning registers has been laid down. The legislation provides general rules and regulations governing the establishing and operation of public authorities' registers, which contain information on individuals. In some cases, the legislation provides special rules and regulations governing registers that are only used for the production of statistics. The ordinary citizen's right to obtain access to data registered on him/her does not apply to statistical registers. The reason why is that the statistics cannot threaten the privacy of a person, because information on the individual citizens cannot be identified in the tables, and because information on individuals stored on the statistical registers may not be used for administrative purposes. In this context, it must be noted that the legislation provides that *no data referable to an identifiable individual must be passed on from the statistical registers*.

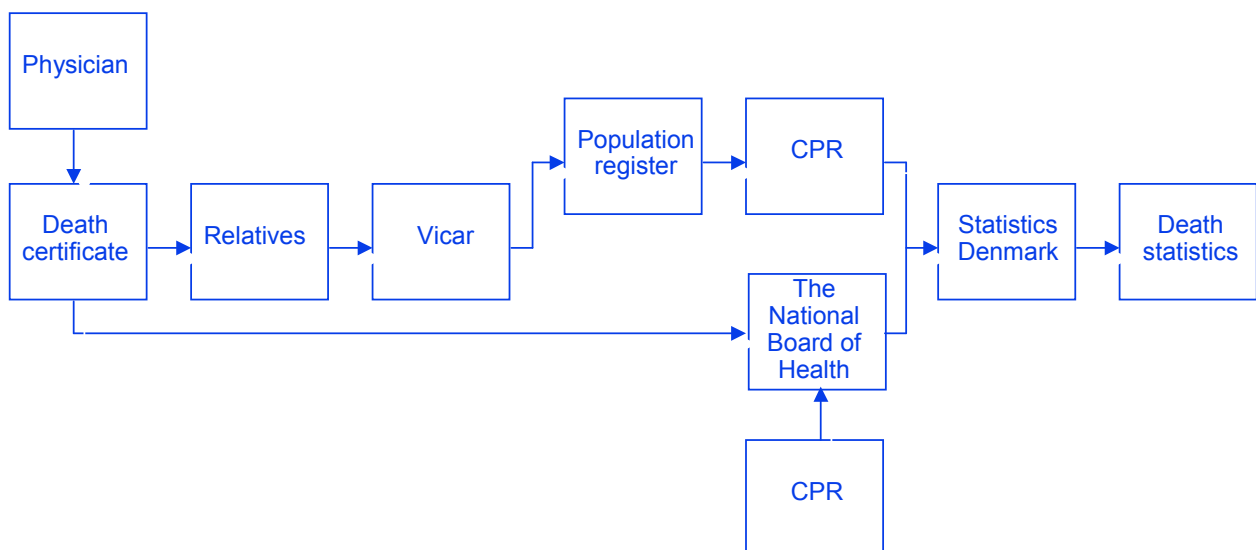
The legislation generally benefits the producer of statistics. Exact rules and regulations have been laid down in respect of a range of issues, and it is an advantage that the producer of statistics is not the only one who guarantees that the information is not being abused.

## Appendix 1

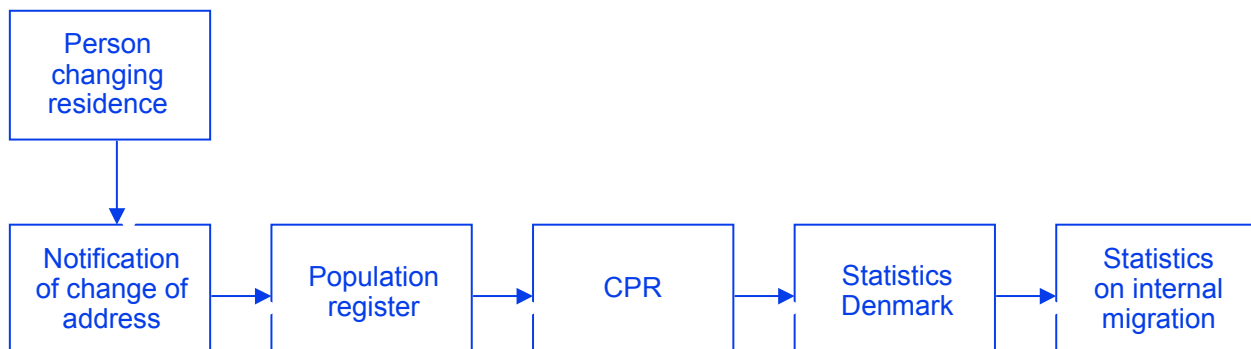
# Birth statistics



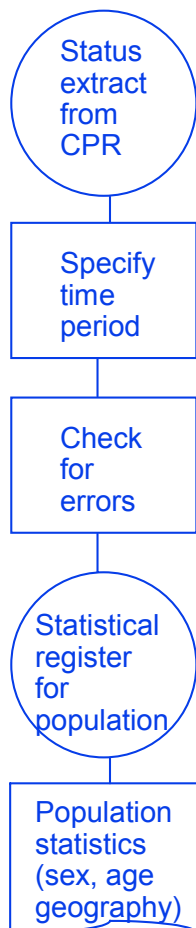
# Death statistics



# Change of address



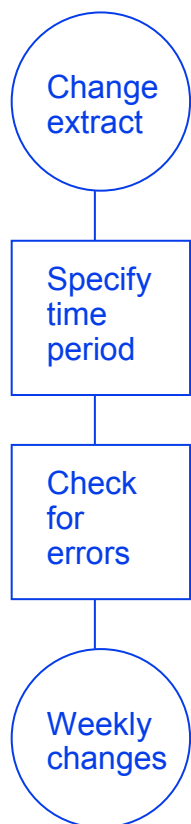
# From extract to statistical register



## Status extract from CPR: (yearly)

**Personnumber**  
**Status (active or not)**  
**Date for change of address (latest)**  
**Code for address (municipality, street, streetnumber, floor, apartment)**  
**Member of church**  
**Marital status**  
**Date for change in marital status**  
**Place of birth**  
**Name**  
**Personnumber of spouse**  
**Personnumber of mother**  
**Personnumber of father**  
**Date for immigration**  
**Citizenship**

# Change extracts, weekly



## CONTENT:

- Personnumber
- Type of change
- Date of change
- Date of registration
- Personnumber of spouse
- Personnumber of mother
- Personnumber of father
- Marital status
- Address
- Citizenship
- Place of birth
- Country of immigration/emigration

**Accumulated to quarters and years**