The Danish Civil Registration System

A cohort of eight million persons

Carsten Bøcker Pedersen, MSc[1], *Heine Gøtzsche*[1], *Jørgen Østrup Møller*[2] *and Preben Bo Mortensen, MD, DrMedSc*[1]

1) National Centre for Register-based Research, University of Aarhus, Taasingegade 1, 8000 Aarhus C, Denmark. 2) The Central Office of Civil Registration, Danish Ministry of Interior Affairs and Health, Denmark.

Correspondence: Carsten Bøcker Pedersen, National Centre for Register-based Research, University of Aarhus, Taasingegade 1, 8000 Aarhus C, Denmark,

E-mail: cbp@ncrr.dk

Dan Med Bull 2006;53:441-9

ABSTRACT

Background: The Danish Civil Registration System (CRS) was established in 1968, where all persons alive and living in Denmark were registered. Among many other variables, it includes individual information on personal identification number, gender, date of birth, place of birth, place of residence, citizenship, continuously updated information on vital status, and the identity of parents and spouses.

Methods: To evaluate the quality and completeness of the information recorded on persons in the CRS, we considered all persons registered on November 4, 2005, i.e. all persons who were alive and resident in Denmark at least one day from April 2, 1968 to November 4, 2005, or in Greenland from May 1, 1972 to November 4, 2005.

Results: A total of 8,176,097 persons were registered. On November 4, 2005, 5,427,687 (66.4%) were alive and resident in Denmark, 56,920 (0.7%) were alive and resident in Greenland, 2,141,373 (26.2%) were dead, 21,160 (0.3%) had disappeared, and 528,957 (6.5%) had emigrated.

Among persons born in Denmark 1960 or later the CRS contains complete information on maternal identity. Among persons born in Denmark 1970 or later the CRS contains complete information on paternal identity. Among women born in Denmark April 1935 or later the CRS contains complete information on all their children. Among males born in Denmark April 1945 or later the CRS contains complete information on all their children.

The CRS contains complete information on: a) immigrations and emigrations from 1971 onwards, b) permanent residence in a Danish municipality from 1971 onwards, c) permanent residence in a municipality in Greenland from May 1972 onwards, and d) full address in Denmark from 1977 onwards.

Conclusion: Data from the CRS is an important research tool in epidemiological research, which enables Danish researchers to carry out representative population-based studies on e.g. the potential clustering of disease and death in families and the potential association between residence and disease and death.

The first national census of the Danish population occurred in 1769 estimating that 794,584 persons lived in Denmark at that time. However, national registration of Danish residents was first established in Denmark in 1924, where individual information concerning members of each Danish family was registered manually on index cards. The information on the index cards was updated continuously by the local municipality registration offices. This registration was used until 1968, at which time it was replaced by the currently used Danish Civil Registration System (CRS), employing information recorded electronically (1-3).

In November 2005, the CRS included information on 8.2 million persons, of which 5.4 million were alive and living in Denmark. This information has been made available for research purposes by Danish legislation (4). Among many other variables, it includes individual information on CPR-number, gender, date of birth, place of birth, place of residence, citizenship, continuously updated information on vital status, and CPR-number of parents and spouses. The CPR-number is used as a personal identifier in all Danish national registers enabling accurate linkage between all national registers.

In recent years, the CRS has been used as an important research tool in epidemiological research (5), e.g., as regards cancer and psychiatric epidemiology. Linking data from the CRS with information on disease occurrences in the Danish population, it is possible to investigate the association between disease or death in individuals and e.g.: a) disease or death in family members (6, 7), b) place of birth and residence (8), c) immigration (9) and emigration, e) foreign adoptions (10), f) environmental exposures at place of residence (11, 12), g) the geographical distance to environmental exposures (13), h) fertility (14), and i) sibship characteristics (15, 16) using a population-based sample of the Danish population. Other branches of research utilising data recorded in the CRS include clinical research (17) and research based on questionnaires (18).

Individual information included in the CRS will be described with focus on information available on family members, and permanent residence. The quality and completeness of this information will be evaluated and discussed.

METHODS

PERSONS REGISTERED

The CRS (1-3) was established on April 2, 1968, where all persons alive and living in Denmark were registered for administrative use, especially to collect tax from Danish residents. From that time, all persons who took permanent residence in Denmark were registered in the CRS, including every live born baby and every new inhabitant. Stillborn children were not registered in the CRS, but children alive at birth, even if only for a short time, were. In addition, on May 1, 1972 all persons alive and living in Greenland were included in the CRS. Today, the CRS includes information on all persons who were alive and had a permanent residence in Denmark on April 2, 1968 or later, or in Greenland on May 1, 1972 or later. In November 2005, the CRS included information on 8,176,097 persons, of which 5,427,687 were alive and living in Denmark, and 56,920 were alive and living in Greenland. Note that although the Kingdom of Denmark consists of Denmark, Greenland and the Faroe Islands, only persons living in Denmark and Greenland are included in the CRS.

INFORMATION REGISTERED

For each person registered, the CRS contains information on the CPR-number, name (first, middle, and last), gender, date of birth, place of birth, place of residence, citizenship, continuously updated information on vital status, CPR-number of parents and spouses, along with an additional 150 variables. In 1968-1988 information in the CRS was updated weekly, while from 1989 onwards information was updated on a daily basis. In the following, individual information registered on persons in the CRS relevant to epidemiological register-based research is described.

THE CPR-NUMBER

All persons registered in the CRS are assigned a unique personal identification number, called the CPR-number, which is used in all national registers, enabling accurate linkage between all national registers. CPR is an abbreviation of "Centrale Person Register" which is the Danish name of the CRS. The CPR-number has ten digits, where the first six digits indicate the date of birth (day two digits, month two digits, year two digits), the next three digits indicate a serial number to distinguish between persons born on the same day, and the last digit is a control digit introduced to minimize recording errors and also indicates the sex. Once a person has been assigned a CPR-number, the same CPR-number will not be assigned to other persons and the CPR-number follows the person forever. If errors are encountered in the assigned CPR-number, e.g., incorrect

date of birth or gender, the person will be assigned a new CPR-number and the CRS keeps a record of the historical CPR-number(s). A CPR-number for a male born on July 7, 1961 may look this way: 070761-4285, with the first six digits referring to his date of birth, and the last four digits being the serial number, with the tenth odd digit denoting the male sex (2).

GENDER AND DATE OF BIRTH

Information on gender and date of birth is both contained in the CPR-number and as separate variables. If a person has had more than one CPR-number, the gender and date of birth is based on the current CPR-number only. Persons who change sex will obtain a new CPR-number.

VITAL STATUS

Vital status is updated continuously and records whether the person is alive and resident in Denmark, alive and resident in Greenland, disappeared (persons whose residence is unknown to Danish authorities), emigrated, or dead along with the date of these events. For persons who have disappeared or emigrated, information on death is available only if death occurred in Denmark, or the Danish authorities were informed of the death. Since the CRS was established in 1968, historical information on vital status is available by using the information on place of residence, emigration, immigration and disappearance described later.

PARENTS

The CRS includes information on the parents' CPR-number. This information has been subject to some changes. The forerunner of the CRS was the municipality registration offices established in 1924, registering manually on index cards members of each family residing at the same address. These index cards were kept at the municipality where the family lived. In 1924-1967 the index card contained information on name, gender and date and place of birth of the head of the family (typically the father), the secondary head (typically the mother), and all children younger than 15 years living at home and who themselves were neither married nor had children. Beginning in 1968, the CRS established links to parents by means of families residing at the same address by using the information on the index cards from the municipality registration offices. In 1968-1978 this link was erased when either the child moved away from home, when a parent moved, when the child itself had children, or when the child reached 18 years. After 1978, the links were changed from being based on addresses to being based on legal relationships, and from 1978 onwards links to parents were kept permanently. In 1978 an extensive revision process was initiated to establish links to parents for persons born in Denmark in 1969-1978, where information on parental links for all persons born in 1969-1978 was verified using the information recorded in the parish registers. Today, parental links for persons born in Denmark in 1969 or later are considered to be 100% correct, and parental links for persons born in Denmark in 1960-1968 are considered to be almost correct (19).

The links to parents are based on the legal relationship, i.e., for each person registered in the CRS, the parental links registered indicate the legal parents and are updated whenever changes occur. Danish legislation prohibits the CRS of keeping information on previous legal relationships. Thus, for example, for adopted children, the CRS contains information on the legal parents only. However, information on historical legal parents is available for research purposes at Statistics Denmark, which obtains information on any changes occurring in the CRS.

SPOUSE

When the CRS was established in 1968, information on the spouse's CPR-number was registered along with information on the date of marriage. From that time, information on spouse was updated continuously, without historical data being deleted from the CRS. Consequently, for each person, the CRS contains information on all spouses from the time of first registration in the CRS.

PLACE OF BIRTH

For persons born in 1977 or earlier, the CRS contains information on the physical place of birth (e.g. municipality of hospital at birth). This is registered as the authority who recorded the birth and is a mixture of Danish parishes (77.8% of persons), Danish municipalities (9.0% of persons), religious communities and congregations (0.1% of persons), Greenlandic parishes (0.8% of persons), country of birth if born abroad (10.6% of persons), or unknown (1.7% of persons).

For persons born in 1978 or later, place of birth is registered as the maternal residence at child's birth, and is a mixture of Danish parishes (84.1% of persons), Danish municipalities (4.3% of persons), Greenlandic parishes (1.5% of persons), country of birth if born abroad (9.8% of persons), or unknown (0.3% of persons).

Denmark is currently subdivided into 2152 parishes located in 271 municipalities. From 2007 onwards, Danish municipalities will be grouped into 98 larger municipalities according to the stipulation that each municipality should include at least 20,000 persons.

Whenever Danish residents require certificates of birth, name, marriage or death from the public administration, the information recorded in the CRS concerning name, date and place of birth and identity of spouse and parents is cross-linked with the information recorded in the parish registers, and information in the CRS is corrected if errors are encountered. As a consequence, the completeness and the quality of the information recorded in the CRS increase over time. Recall that for persons born in Denmark in 1969 or later this information has been verified. In addition, the CRS keeps a record stating that this information had been verified.

PLACE OF RESIDENCE, EMIGRATION, IMMIGRATION AND DISAPPEARANCE

For all persons living in Denmark or Greenland, the CRS contains information on the full address (municipality, road, and house number) and the date when they moved to that address. For immigrants, information on country of immigration, and for emigrants, information on the country of emigration is recorded along with dates of immigration and/or emigration. For persons who have disappeared (persons whose residence is unknown to the Danish authorities), information is recorded concerning their date of disappearance. For persons who are retrieved, information is recorded on the date of retrieval. Changes in this information are updated without old data being deleted from the CRS. Combining these data sources, the CRS contains information on the place of residence – including residence in foreign countries – of all persons who were ever registered in the CRS from the time of first registration in the CRS. Information on place of residence before 1986 has been made available for research purposes only with the understanding that the CRS does not guarantee the quality of these data.

According to Danish legislation, all residents in Denmark are obliged to inform the authorities about any change of permanent address within five days. Failure to supply this information will result in inability to receive supplementary benefit (e.g. unemployment, sickness or disablement benefits, and educational aid from public funds), to attend day nursery, nursery school, and primary and lower secondary school, to avail of free national health care, and to obtain a tax deduction card (required to have paid work), etc. It is very unlikely that this mandatory information is not reported.

CITIZENSHIP

When the CRS was established, information on citizenship was registered and thereafter continuously updated. However, historical information on citizenship was not kept until 1991 onwards. For persons with Danish citizenship in addition to other citizenships, only the Danish citizenship was registered; and for persons with

more than one citizenship, none of which was Danish, citizenship was somewhat randomly registered. Citizenship is registered at the country level, with Greenland being a part of Denmark.

SIBLINGS AND TWINS

The CRS does not contain direct information on the CPR-number of siblings and twins, if any. However, such information can be obtained using maternal CPR-number. Siblings can be identified as persons having the same mother, and twins can be identified as siblings who are born on the same day (+/- one day for the identification of twins born on either side of midnight). Alternatively, one may identify siblings as persons having the same mother and father, or as persons having the same father only. Usually, in register-based research, siblings are identified as persons having the same mother (see e.g. (6, 8, 13, 15, 16, 20, 21)). Data from the CRS was used as the primary source of information when establishing the youngest part of the Danish Twin Register (22, 23).

The above procedure can identify *all* siblings, only if all siblings have information on their mother's CPR-number. Recall, that in April 1968 links to parents were established using the information on the index cards from the municipality registration offices, meaning that children younger than 15 years, living at home with their mother and who themselves were neither married nor had children, obtained information on the mother's CPR-number, whereas other children would not obtain this information. Usually, the earliest age at any of these events is the 15th birthday. Therefore, complete information on all siblings using maternal identity will be available if the mother's first born child was born April 1953 or later (April 1968 minus 15 years). Say also, that the earliest age at time of first birth is the mother's 18th birthday, then for women who were born April 1935 or later (April 1953 minus 18 years) we have complete information on all her children using her CPR-number. This is equivalent to having complete information on all siblings using maternal CPR-number. This was first identified in the pioneer thesis by Westergaard (24).

One technicality in the above reasoning is that it pertains only to mothers who were residing in Denmark when the CRS was established. As this criterion is difficult to manage in practice, in register-based research one usually considers having complete information on all siblings, if the mother was born *in Denmark* April 1935 or later. Among children born in Denmark whose mothers were born in Denmark April 2, 1935 or later, 99.97% of those born in 1969 or later had their maternal link verified, 93.5% of those born in 1960-1968 had their maternal link verified, and 57.5% of those born in 1959 or earlier had their maternal link verified.

Throughout this paper, *complete* information indicates that for the vast majority of persons (i.e. > 99%) information is accurate, *almost complete* indicates that for the majority of persons (i.e. > 95%) information is accurate, and *somewhat complete* indicates that for most persons (i.e. > 90%) information is accurate. Therefore these terms are useful for register-based researchers, but are less useful for administrative purposes. Also, *stabilized* indicates a time point for which it is likely that time trends observed in data are due to time trends in the population rather than time trends in the registration of the population. Generally, we expect time trends in the registration of the population to occur faster than time trends in the population.

SIBSHIP CHARACTERISTICS

Among persons with complete information on siblings, defined as persons whose mother was born in Denmark April 1935 or later, we can calculate the birth order, the birth interval to nearest older and younger sibling, information on multiple births, number of siblings at the 15th birthday and any measure of sibship characteristics using the information recorded on each individual (see e.g. (15, 16, 20)).

ACCESS TO DATA

According to Danish legislation researchers can gain access to information from the CRS (3, 4). Nevertheless, it is a complex and time-demanding task to transform the original data from the CRS into easy accessible research databases. To gain easier, faster and more secure access to research data of high quality, the National Board of Health has established a database for research purposes, which includes information on all persons registered in the CRS linked individually with data from the Medical Birth Register (25) (births from 1973 onwards), the Psychiatric Central Register (26) (psychiatric admissions from 1969 onwards, psychiatric out-patient contacts from 1995 onwards), the National Hospital Register (27) (somatic admissions from 1977 onwards, somatic out-patient contacts from 1995 onwards), the In Vitro Fertilisation Register (28) (induced fertility treatments from 1994 onwards), the Cancer Register (29) (diagnoses with cancer from 1943 onwards), the Register of Congenital Malformations (30) (births from 1983 onwards), and the Register of Causes of Death (31) (deaths from 1970 onwards). For further information please contact the Danish National Board of Health (32). Alternatively, Statistics Denmark offer online access to almost identical data linked individually with information on occupation, gross income, wealth, cohabitant, recipients of social benefits, educational achievement, dwelling, biological parents, prescriptions, etc (14, 33, 34).

STUDY POPULATION

To evaluate the quality and completeness of the information recorded on persons in the CRS and that calculated concerning siblings, we consider all persons registered in the CRS on November 4, 2005, i.e., all persons who have been alive and resident in Denmark at least one day from April 2, 1968 to November 4, 2005, or in Greenland from May 1, 1972 to November 4, 2005.

ETHICAL CONSIDERATION

This study was approved by the Danish Data Protection Agency.

RESULTS

Overall, 8,176,097 persons (4,122,581 males and 4,053,516 females) were alive and resident in Denmark or Greenland at least one day from April 2, 1968 to November 4, 2005. On the latter date 5,427,687 (66.4%) were alive and resident in Denmark, 56,920 (0.7%) were alive and resident in Greenland, 2,141,373 (26.2%) were dead, 21,160 (0.3%) had disappeared, and 528,957 (6.5%) had emigrated.

ANNUAL NUMBER OF BIRTHS

Overall, 7,131,608 (87.2%) persons were born in Denmark, 80,198 (1.0%) were born in Greenland, 851,184 (10.4%) were born abroad, and 113,107 (1.4%) had unknown place of birth.

The annual number of persons born in Denmark varied greatly with year of birth (Figure 1). The annual number of births peaked at

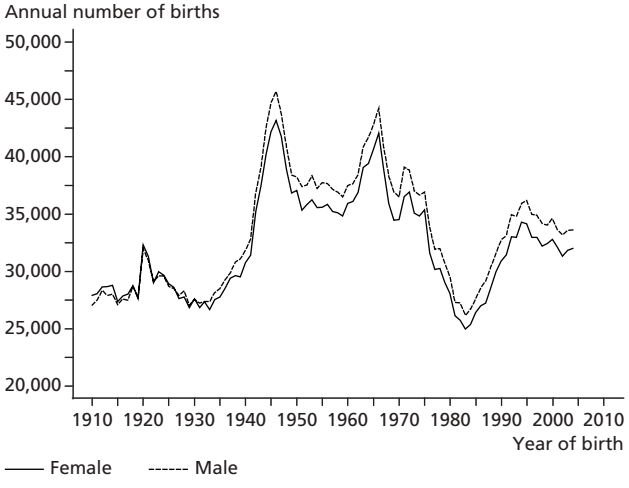

Figure 1. Annual number of persons born in Denmark by gender.

45,664 males and 43,122 females born in 1946. Among persons born in Denmark 1969 or later, 51.36% were male births and 48.64% were female births. Recall that only persons who were alive and resident in Denmark when the CRS was established in 1968 were registered in the CRS.

The annual number of persons born in Greenland (Figure 2) showed somewhat the same tendency of births as that presented for persons born in Denmark (Figure 1). Figure 3 shows the annual number of persons born abroad (and who were registered in the CRS on November 4, 2005). Interestingly, male immigrants were overall born earlier compared to female immigrants.

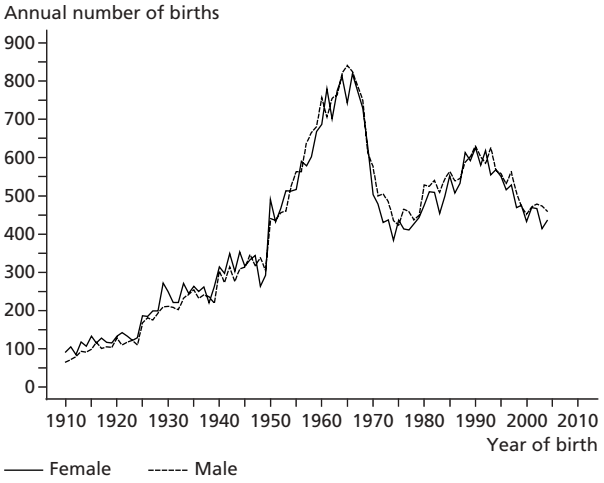

Figure 2. Annual number of persons born in Greenland by gender.

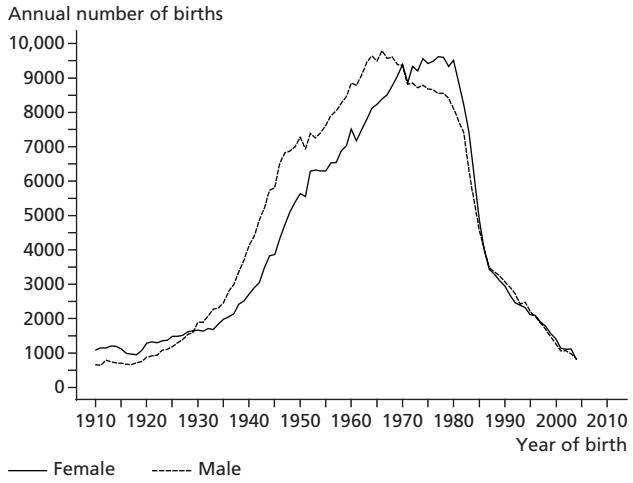

Figure 3. Annual number of persons born abroad (not including Greenland) by gender.

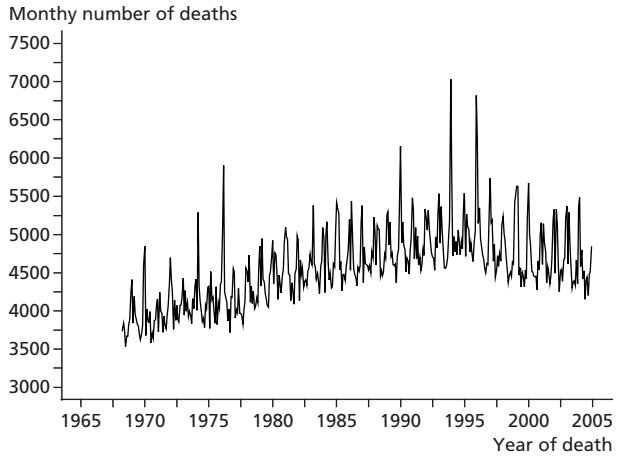

Figure 4. Monthly number of deaths of persons born in Denmark.

MONTHLY NUMBER OF DEATHS

The monthly number of deaths of persons born in Denmark was stabilized from the initial operation of the CRS (Figure 4), which was also supported by data on monthly number of deaths from Statistics Denmark (35).

One technicality in Figure 4, is that among persons who died during the initial operation of the CRS, the percentage of persons with unknown place of birth was extremely high; among persons who have died in 1968, 1969, 1970, 1971, 1980, 1990, and 2000, the percentage of persons with unknown place of birth was 82.2%, 39.9%, 3.8%, 2.3%, 1.8%, 0.8%, 0.9%, respectively. Therefore, in Figure 4, we have chosen to include both persons born in Denmark and persons with unknown place of birth. Note that, choosing a cohort of persons known to be born in Denmark will indirectly include only persons who survived until approximately 1971.

The number of deaths of persons born in Greenland stabilized at 30 deaths per month from June 1972 onwards (results not shown), and the number of deaths of persons born abroad stabilized at 125 deaths per month from December 1969 onwards (results not shown).

INFORMATION ON PARENTS' IDENTITY

Overall, 3,803,664 (46.5%) had information on maternal CPR-number and 3,707,334 (45.3%) had information on paternal CPR-number.

Persons born in Denmark

Among persons born in Denmark (N=7,131,608), the percentage of those with a link to a mother increased rapidly from 8.0% in 1950 to 98.7% in 1960 and 99.9% in 1970, remaining at that level (Figure 5). Similarly, the percentage of persons with a link to a father increased rapidly from 7.6% in 1950 to 95.7% in 1960 and 99.2% in 1970, remaining at that level (results not shown).

Persons born in Greenland

Among persons born in Greenland (N=80,198), the percentage of those with a link to a mother increased rapidly from 9.9% in 1950 to 89.9% in 1960 and 98.1% in 1970 increasing slightly thereafter (Figure 6), and the percentage of persons with a link to a father increased rapidly from 7.9% in 1950, 79.6%% in 1960, 87.9% in 1970, to 93.8% in 1985, remaining at that level (results not shown). Recall that persons living in Greenland were included in the CRS in 1972.

Note that, among persons born in Denmark or Greenland in 1950-1960, the percentage of males with a link to a parent was higher than the percentage of females with a link to a parent (Figures 5 and 6). This was likely to be due to the fact that male children

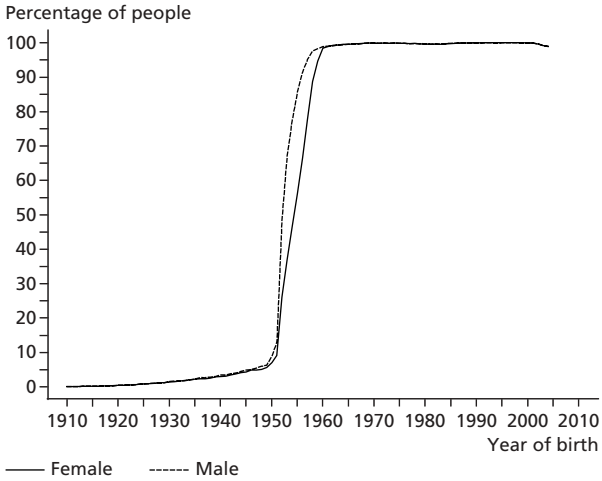

Figure 5. Percentage of persons born in Denmark who have information on maternal identity by gender.

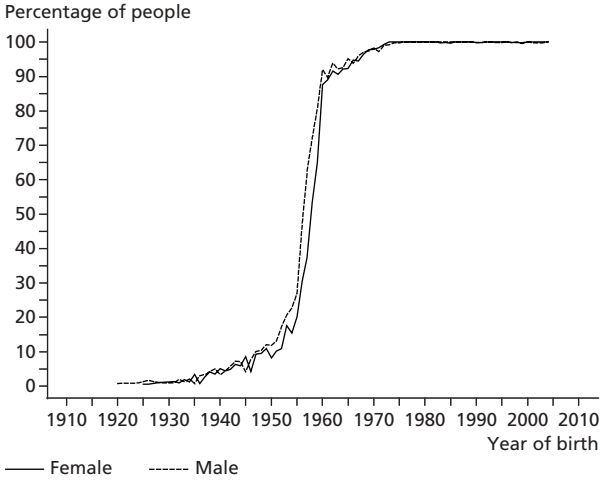

Figure 6. Percentage of persons born in Greenland who have information on maternal identity by gender.

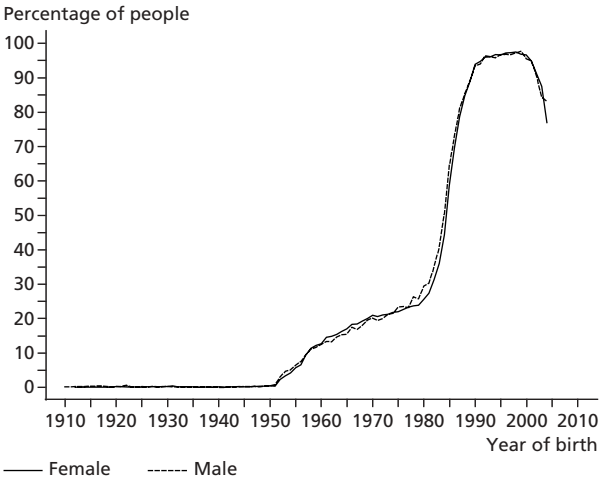

Figure 7. Percentage of persons born abroad (not including Greenland) who have information on maternal identity by gender.

Figure 8. Percentage of persons born abroad (not including Greenland) and resident in Denmark when the CRS was established, who have information on maternal identity by gender.

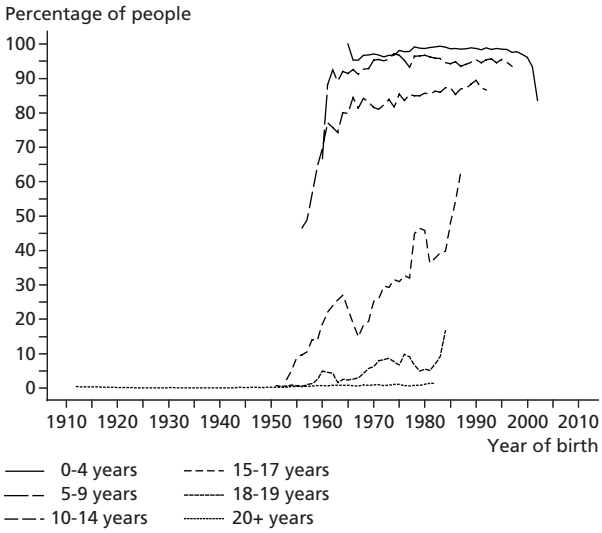

Figure 9. Percentage of persons born abroad (not including Greenland) and not resident in Denmark when the CRS was established who have information on maternal identity by age at immigration to Denmark.

moved away from home later than female children, thus implying that when the CRS was established, more male children obtained a link to a parent as they lived at the same address.

Persons born abroad

Among persons born abroad (not including Greenland) (N= 851,184), the percentage of those with a link to a mother increased from 0.5% in 1950 to 29.8% in 1980 and 94.2% in 1990 (Figure 7), and the percentage of persons with a link to a father increased from 0.3% in 1950 to 25.0 in 1980 and 76.9% in 1990, remaining at that level (results not shown). Note that information on parental identity required that the parent had a Danish CPR-number, i.e., had been residing in Denmark in 1968 or later. Completeness of information on maternal identity for persons born abroad was subject to complex heterogeneity. First, among persons born abroad who lived in Denmark when the CRS was established (N=147,201), the percentage of those with a link to a mother increased rapidly from 1.1% for those born in 1950 to 93.2% for those born in 1960 and remained at that level (Figure 8), i.e. almost similar to that of persons born in Denmark (Figure 5). Second, among persons born abroad who did not live in Denmark when the CRS was established (N=703,983), the completeness of information on maternal identity depended on age at immigration to Denmark (Figure 9). Among those who immigrated at age 0-4 years, 97.5% of persons had information on maternal identity, while among those who immigrated at age 20 years or above, only 0.6% had information on maternal identity. In addition, among persons who immigrated to Denmark before age 20 years, the younger the age of immigration, the greater the percentage of persons with information on maternal identity (Figure 9). Obviously, young children who immigrated to Denmark along with their mother obtained a link to the mother, while older persons who immigrated to Denmark without their mother did not.

INFORMATION ON SIBLINGS AND CHILDREN

The former information on parental information concerned whether individuals had a link recorded to their parents in the CRS. In this section, we describe whether a parent had information on all children. This is equivalent to individuals having information on all siblings using maternal (and/or paternal) identity. Recall that the CRS contains information on the parents of each person and not vice versa.

Parents born in Denmark

For women born in Denmark in 1910-1960 (N=1,624,634), Figure 10 shows the percentage of women with no children, 1-2 children, and three or more children. The number of children was calculated as the number of live born children of each women on November 4, 2005 as registered in the CRS, i.e., irrespective of the previously defined completeness of all children through maternal identity (see Methods section). Here, the purpose was to evaluate the previously described logical reasoning on the completeness of information on children and siblings using maternal identity. Note, that Figure 10 does not indicate the number of children of each woman, but the

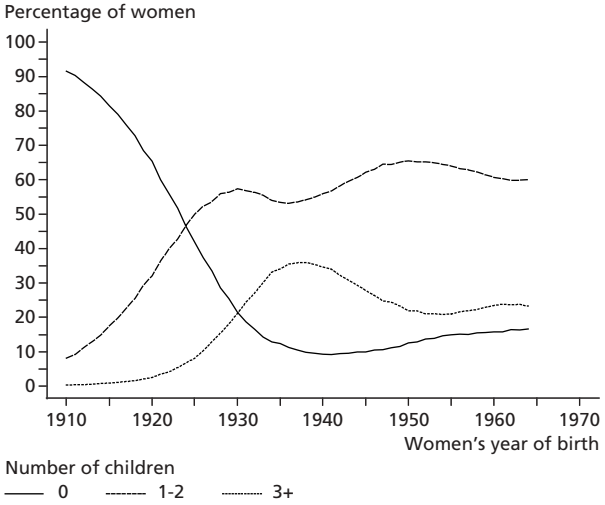

Figure 10. Percentage of women born in Denmark according to their number of children registered in the CRS.

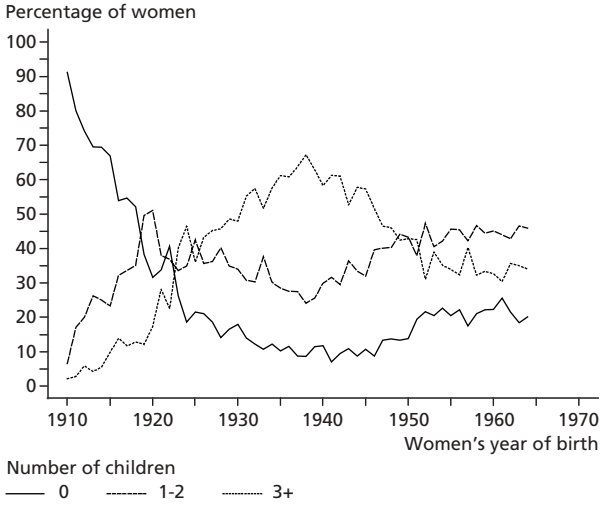

Figure 11. Percentage of women born in Greenland according to their number of children registered in the CRS.

number of children who had a maternal link to each woman in the CRS. Women born 1966 or later were not included in this figure as some of these had not had all their children yet, and would therefore increase the complexity of this figure. The percentage of women born in Denmark registered with no children decreased from 91.5% for women born in 1910 to 11.9% for women born in 1935 and stabilized at that level. The percentage of women registered with 1-2 children increased from 8.2% for women born in 1910 to 53.8% for women born in 1935 and stabilized at that level. The percentage of women registered with three or more children increased from 0.3% for women born in 1910 to 34.3% for women born in 1935 and stabilized at that level. Combining this finding with the finding that among children born in Denmark 1960 or later we had complete information on maternal identity, we found support for the notion that among women born in Denmark April 1935 or later, the CRS contains complete information on all their children.

Similarly, the percentage of males born in Denmark registered with a) no children decreased from 83.2% for males born in 1910 to 16.0% for males born in 1935, b) 1-2 children increased from 15.5% for males born in 1910 to 49.9% for males born in 1935, and c) three or more children increased from 1.3% for males born in 1910 to 34.1% for males born in 1935 (results not shown). This finding combined with the finding that among children born in Denmark 1960 or later we had *almost* complete information on paternal identity, while among children born in Denmark 1970 or later we had complete information on paternal identity, may suggest that among males born in Denmark April 1935 or later the CRS contains *almost* complete information on all their children, while among males born in Denmark April 1945 or later the CRS contains complete information on all their children.

Parents born in Greenland

For women born in Greenland in 1910-1960 (N=14,499), Figure 11 shows the percentage of women with no children, 1-2 children, and three or more children. Following a similar argument to that above, this figure supports the notion that among women born in Greenland 1935 or later the CRS contains *almost* complete information on all their children, while complete information on all their children is obtained for women born in Greenland 1945 or later.

Similarly, the percentage of males born in Greenland registered with a) no children decreased from 71.4% for males born in 1910 to 29.0% for males born in 1935, b) 1-2 children increased from 16.1% for males born in 1910 to 18.7% for males born in 1935, and c) three or more children increased from 12.5% for males born in 1910 to 52.4% for males born in 1935 (results not shown). Previously we showed that among persons born in Greenland 1960-1985 roughly 85% had a link to a father while among persons born in Greenland

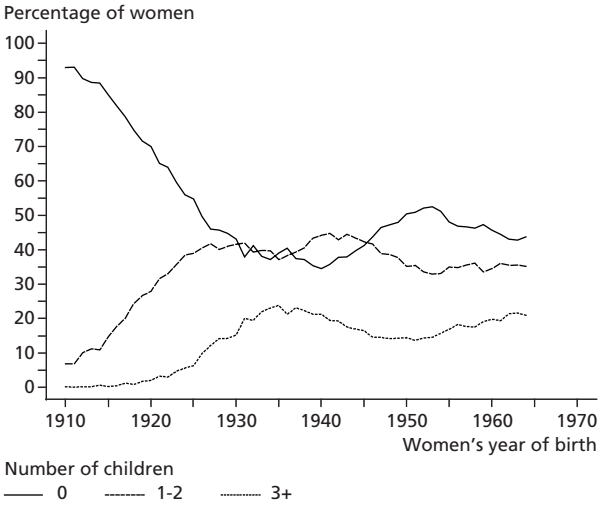

Figure 12. Percentage of women born abroad (not including Greenland) according to their number of children registered in the CRS.

1985 or later roughly 95% had a link to a father. Therefore, these results suggest that among males born in Greenland 1935 or later, the CRS contains *somewhat* complete information on all their children, while among males born in Greenland 1965 or later, the CRS contains *almost* complete information on all their children.

Parents born abroad

For women born abroad 1910-1960 (N=155,248), Figure 12 shows the percentage of women with no children, 1-2 children, and three or more children. Overall, these birth numbers stabilized for women born abroad 1935 or later. However, as the completeness of information on maternal identity for persons born abroad depended on whether the child was living in Denmark at the time the CRS was established and on the age at immigration to Denmark, simple guidelines on how to obtain complete information on all children to a mother could not be given.

SIBLINGS

Among children whose mother was born in Denmark April 1935 or later (N=2,745,113), we calculated the birth order of each child (1 for the first born child, 2 for the second born child, etc), and the birth interval to nearest older sibling. Overall, 47.3% of Danish children were the first-born (Table 1). Figure 13 shows the distribution of the birth interval to nearest older sibling by birth order. Among children with a birth order of 2 the most common birth interval to nearest older sibling was 2.5 years, and the greater the birth order, the shorter the most common birth interval to nearest older sibling.

Table 1. Distribution of birth order among children whose mother was born in Denmark April 1935 or later.

Birth order	No. of children	Percentage of children
1	1,299,215	47.33
2	996,623	36.31
3	345,226	12.58
4	80,008	2.91
5	17,350	0.63
6	4,425	0.16
7	1,381	0.05
8	497	0.02
9	188	0.01
10	85	0.00
11	53	0.00
12-18	62	0.00
Total	2,745,113	100.0

TWINS

The percentage of twin births increased from 1.8% in 1970 to 4.5% in 2004 (Figure 14). The increased use of fertility treatment, which is known to induce twinning pregnancies (36), might explain this finding. Twin births were counted as two children. Overall, among children whose mother was born in Denmark 1935 or later, 97.49% (N=2,676,282) were singletons, 2.45% (N=67,156) were twins, 0.06% (N=1603) were triplets, and 0.00% (N=72) were quadruplets. Information on zygosity of twins is not included in the CRS, but is included in the Danish Twin Register (22, 23). Note, since the CRS (and the Danish Twin Register) includes live-born children only, the number of twin births may be biased towards nil.

PLACE OF RESIDENCE, EMIGRATION, AND IMMIGRATION

First, information on residence is described at the level of municipality and country for those residing abroad: From January 1971 onwards the number of persons with a newly registered permanent residence in a Danish municipality was stabilized (Figure 15), thus suggesting that from 1971 onwards the CRS includes complete information on the municipality of residence for persons with a permanent address in Denmark. Note, the number of newly registered residences in a Danish municipality in 1970-1979 may be artificially high; persons with unknown information on house number in the initial operation of the CRS were treated as persons with newly registered place of residence when the CRS obtained information on the house number.

Similarly, from May 1972 onwards, the number of newly registered permanent residences in Greenlandic municipalities was stabilized (results not shown), from January 1971 onwards, the number of newly registered emigrants from Denmark to a foreign country was stabilized (Figure 16), and from January 1971 onwards, the number of newly registered immigrants to Denmark was stabilized (Figure 16).

The completeness of information on residence in Denmark is as follows (Figure 17): Among persons born in Denmark, a) the completeness of information of "municipality" and of "municipality and road" was almost identical (approximately 99.8% from 1968 onwards), b) the completeness of information on full address (municipality, road, and house number) increased from 84.7% in 1971 to 99.1% in 1977 increasing slightly thereafter, and c) the completeness of geocodeable information increased from 76.0% in 1968 to 98.1% in 1982, increasing slightly thereafter. For information to be geocodeable the information recorded concerning road must correspond to a physical road existing today and the information recorded concerning house number must correspond to a specific house existing on each road today. No historical recording of roads in Denmark exists. Note, that to identify persons living together, information on the full address is sufficient, while to identify the geographical distance to some exposure for each individual the geocoded address information is necessary (See e.g. (13). Also note that the completeness and the quality of geocodeable information may depend on whether a) addresses were geocoded using the Danish Address and Road Database (37) or using the recent register on official standard addresses and coordinates (38), b) how historical information on residence on an abandoned or renamed road is treated, and c) how invalid and missing information on house number is treated. Also note that among addresses in Denmark in 1968-1976 the more urban the municipality of residence the more complete the information on house number in the CRS (results not shown). Persons resident abroad (overall 2.0% of residences) and persons with un-

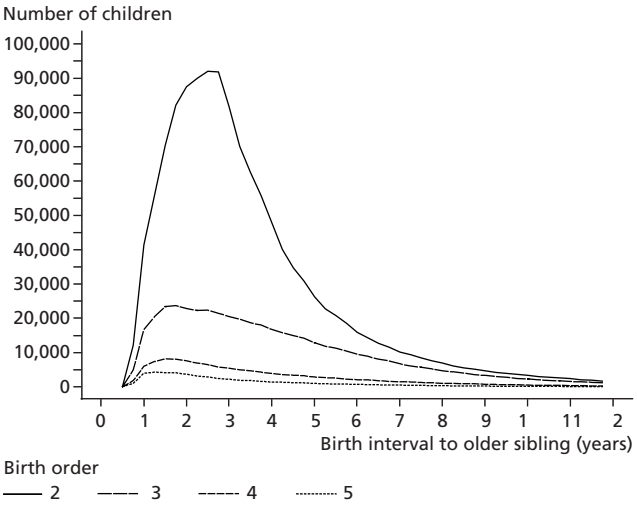

Figure 13. Distribution of birth interval to nearest older sibling according to birth order among children whose mother was born in Denmark April 1935 or later.

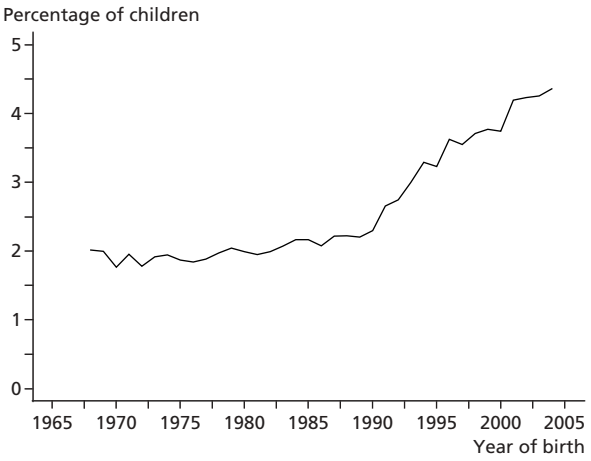

Figure 14. Percentage of multiple births among children whose mother was born in Denmark April 1935 or later. Twins were calculated as two births by the same women occurring on the same day.

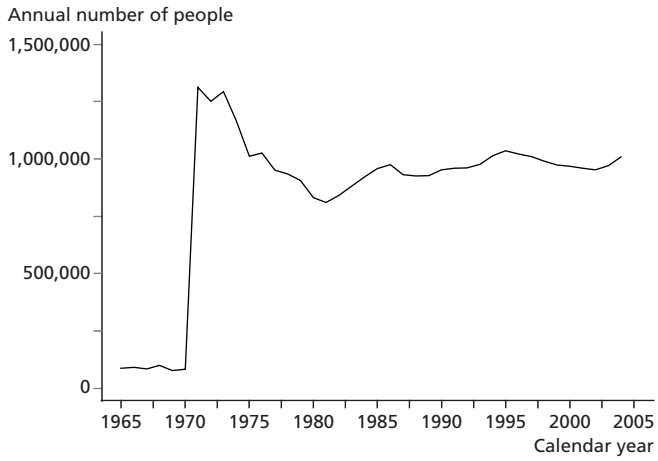

Figure 15. Number of persons with a newly registered permanent residence in a Danish municipality.

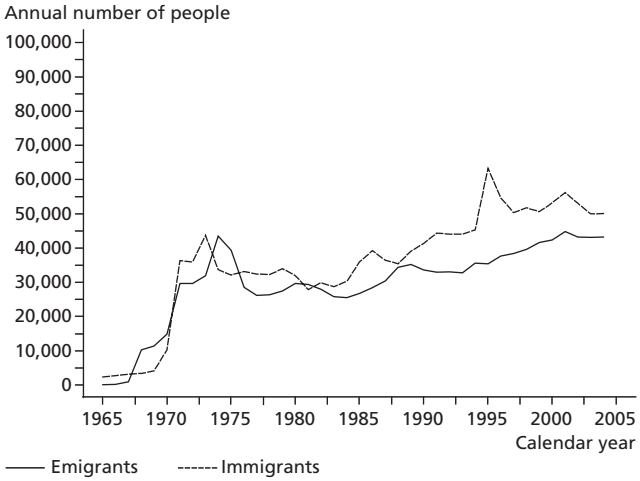

Figure 16. Number of newly registered immigrants to Denmark, and number of newly registered emigrants from Denmark.

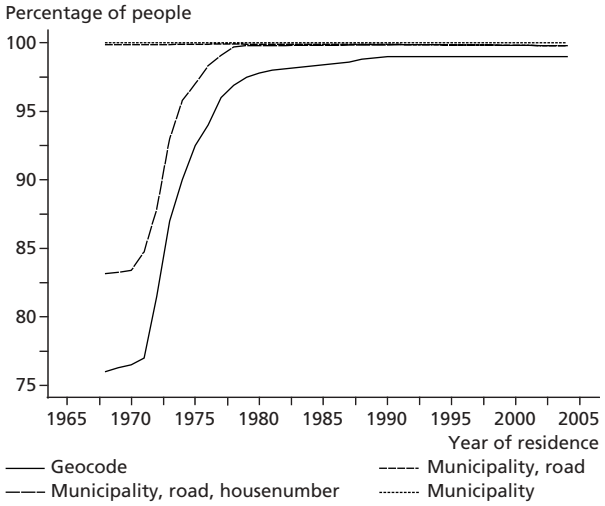

Figure 17. Completeness of information of permanent residence in Denmark according to level of detail of information. Municipality indicate that we have information on the municipality, municipality, road indicate we have information on the municipality and the road, municipality, road, house number indicate we have information on these three variables, and geocode indicate that information on municipality, road and house number is geocodeable, meaning that it is possible to determine an exact geographical position for the residence.

known residence (overall 0.1% of residences) were not included in Figure 17.

Obviously, the level of detail required on place of residence would determine the completeness of the information accessible in the CRS.

CITIZENSHIP

Overall, among persons alive and resident in Denmark on November 4, 2005, 5.154.187 (95.0%) had a Danish citizenship, 273,498 (5.0%) had a foreign citizenship, and 2 (0.0%) had unknown citizenship. In addition, among those born in Denmark, 4,944,687 (99.2%) had a Danish citizenship, among those born in Greenland, 13151 (99.9%) had a Danish citizenship, whereas among those born abroad, 180,107 (44.3%) had a Danish citizenship.

INFORMATION ON SPOUSES

Among persons alive and resident in Denmark on November 4, 2005, 2,516,656 (47.4%) persons have never been married, 2,159,592 (39.8%) were married, 418,097 (7.7%) divorced, 325,275 (6.0%) were widowed, 6,308 (0.1%) had entered civil partnership with a same sex partner, 1,397 (0.01%) were divorced from a same sex partner, and 362 (0.01%) were widowed of a same sex partner. The number of married persons has been almost constant since the CRS was established, while the number of divorced persons has increased 4-fold from 1968 to 2004.

DISCUSSION

The CRS was established in 1968 and has since then recorded current and historical information on all persons living in Denmark. The extensive recording of Danish residents is generally well accepted. This may be due to the fact that national registration of all residents has been required by law since 1924, meaning that persons in Denmark are accustomed to being registered, that Danes in general have confidence in authorities (39), and that there is no history of misuse of information recorded in the CRS.

The CRS was established for administrative purposes independently of health and social factors. Although no studies exist exploring the quality of the information recorded in the CRS, it is generally accepted that the information recorded is of very high quality. First, information in the CRS is used continuously by the administrative system in Denmark, which corrects errors whenever they are encountered. Second, high quality is ensured by the ongoing validation of information recorded. Third, when the CRS was established, all residents obtained a civil registration number certificate including their own personal information so that potential errors could be corrected. Fourth, registration in the CRS is required by law. Fifth, there is a positive public attitude towards registration in the CRS. Sixth, the CPR-number includes easily accessible information on date of birth and gender of each individual thus ensuring a good quality of the registration of gender and date of birth.

Register-based research is often based on genetic hypotheses concerning the potential clustering of diseases in families. However, it should be noted that the parental link included in the CRS represents the link to the legal parents. No studies investigated the potential discrepancy between the legal parents and the biological parents in the Danish population. A recent meta-analysis based on foreign data showed that the paternal discrepancy was 3.7% (40).

Data from the CRS is an important research tool in epidemiological research, enabling Danish researchers to carry out representative population-based studies on e.g. the potential clustering of disease and death in families and the potential association between residence and disease and death. These data constitute an important and rare asset. To our knowledge such data are only currently available in Sweden, Finland and Taiwan, with varying differences in coverage periods, the extent of other available data that could be linked to the population data etc. The emerging possibilities to link these data to collections of biological material only adds to the research potential offered by these data. A large number of scientific papers have been published using data from the CRS and we believe that this system in connection with other registers and biobanks will continue to provide the basis for significant knowledge relevant to the etiological understanding and possible prevention of human diseases.

ACKNOWLEDGEMENT

This study was supported by the Stanley Medical Research Institute. Psychiatric epidemiological research at the National Centre for Register-based Research is in part funded through a collaborative agreement with Centre for Basic Psychiatric Research, Psychiatric Hospital in Aarhus. We thank Dr. Elizabeth Cantor-Graae for helpful comments on a draft of this paper.

REFERENCES

1. Willumsen H. Folkeregistreringen 75 år i 1999. The Danish Civil Registration Office, 1999. www.cpr.dk/publikationer/folkeregistrering1999/forside.htm
2. Malig C. The civil registration system in Denmark. Technical Papers IIVRS 1996;66:1-6.
3. The Danish Civil Registration Office. The Civil Registration System in Denmark. The Danish Civil Registration Office, 2005. www.cpr.dk/Index/dokumenter.asp?o=7&n=0&d=141&s=5
4. Danish Act on the Civil Registration System. The Danish Civil Registra-

tion Office, 2000. www.cpr.dk/index.asp?d=140&o=7&n=0&h=7&s=5
5. Frank L. Epidemiology. When an entire country is a cohort. Science 2000;287:2398-99.
6. Mortensen PB, Pedersen CB, Westergaard T, Wohlfahrt J, Ewald H, Mors O et al. Effects of family history and place and season of birth on the risk of schizophrenia. N Engl J Med 1999;340:603-8.
7. Li J, Laursen TM, Precht DH, Olsen J, Mortensen PB. Hospitalization for mental illness among parents after the death of a child. N Engl J Med 2005;352:1190-6.
8. Pedersen CB, Mortensen PB. Evidence of a dose-response relationship between urbanicity during upbringing and schizophrenia risk. Arch Gen Psychiatry 2001;58:1039-46.
9. Cantor-Graae E, Pedersen CB, McNeil TF, Mortensen PB. Migration as a risk factor for schizophrenia: a Danish cohort study. Br J Psychiatry 2003;182:117-22.
10. Teilmann G, Pedersen CB, Jensen TK, Skakkebæk NE. Increased risk of precocious puberty in internationally adopted children in Denmark. Pediatrics 2006;118:e391-99.
11. Pedersen CB, Raaschou-Nielsen O, Hertel O, Mortensen PB. Air pollution from traffic and schizophrenia risk. Schizophr Res 2004;66:83-5.
12. Raaschou-Nielsen O, Hertel O, Thomsen BL, Olsen JH. Air pollution from traffic at the residence of children with cancer. Am J Epidemiol 2001;153:433-43.
13. Pedersen CB, Mortensen PB. Urbanization and traffic related exposures as risk factors for schizophrenia. BMC Psychiatry 2006;6:2.
14. Knudsen LB. The Danish Fertility Database. Dan Med Bull 1998;45: 221-5.
15. Pedersen CB, Mortensen PB. Sibship characteristics during upbringing and schizophrenia risk. Am J Epidemiol 2004;160:652-60.
16. Westergaard T, Rostgaard K, Wohlfahrt J, Andersen PK, Aaby P, Melbye M. Sibship characteristics and risk of allergic rhinitis and asthma. Am J Epidemiol 2005;162:125-32.
17. Jensen LS, Puho E, Pedersen L, Mortensen FV, Sorensen HT. Long-term survival after colorectal surgery associated with buffy-coat-poor and leucocyte-depleted blood transfusion: a follow-up study. Lancet 2005;365: 681-2.
18. Kessing LV, Hansen HV, Bech P. General health and well-being in outpatients with depressive and bipolar disorders. Nord J Psychiatry 2006;60: 150-6.
19. Kirkebogen. Kirkeministeriet, 1998. www.dnk.dk/docs/introduk.htm
20. Westergaard T, Mortensen PB, Pedersen CB, Wohlfahrt J, Melbye M. Exposure to prenatal and childhood infections and the risk of schizophrenia. Suggestions from a study of sibship characteristics and influenza prevalence. Arch Gen Psychiatry 1999;56:993-8.
21. Pedersen CB, Mortensen PB. Family history, place and season of birth as risk factors for schizophrenia in Denmark: a replication and reanalysis. Br J Psychiatry 2001;179:46-52.
22. Kyvik KO, Green A, Beck-Nielsen H. The new Danish Twin Register: establishment and analysis of twinning rates. Int J Epidemiol 1995;24: 589-96.
23. Kyvik KO, Christensen K, Skytthe A, Harvald B, Holm NV. The Danish Twin Register. Dan Med Bull 1996;43:467-70.
24. Westergaard, T. Birth order and cancer risk. Copenhagen, Denmark: Statens Serum Institut, 1999.
25. Knudsen LB, Olsen J. The Danish Medical Birth Registry. Dan Med Bull 1998;45:320-3.
26. Munk-Jorgensen P, Mortensen PB. The Danish Psychiatric Central Register. Dan Med Bull 1997;44:82-4.
27. Andersen TF, Madsen M, Jorgensen J, Mellemkjoer L, Olsen JH. The Danish National Hospital Register. A valuable source of data for modern health sciences. Dan Med Bull 1999;46:263-8.
28. Andersen AN, Westergaard HB, Olsen J. The Danish in vitro fertilisation (IVF) register. Dan Med Bull 1999;46:357-60.
29. Storm HH, Michelsen EV, Clemmensen IH, Pihl J. The Danish Cancer Registry – history, content, quality and use. Dan Med Bull 1997;44: 535-9.
30. Christensen K, Knudsen LB. Registration of congenital malformations in Denmark. Dan Med Bull 1998;45:91-4.
31. Juel K, Helweg-Larsen K. The Danish registers of causes of death. Dan Med Bull 1999;46:354-7.
32. Familiedatabasen. The National Board of Health, 2005. www.sst.dk/Informatik_og_sundhedsdata/Registre_og_sundhedsstatistik/Beskrivelse_af_registre.aspx?lang=da
33. Roed AS, Juhl C, Kamper-Jorgensen F. The Danish Prevention Register. A comprehensive health and socio-economic, individual based register. Dan Med Bull 1999;46:269-72.
34. Statistics Denmark. IDA – en integreret database for arbejdsmarkedsforskning [The integrated database for longitudinal labour market research]. Copenhagen: Statistics Denmark Press, 1991.
35. Statistics Denmark. Befolkningens bevægelser. Statistiske Meddelelser, 1970.
36. Campbell DM. Aetiology of twinning. In: MacGillivray I, Campbell DM, Thompson B, eds. Twinning and twins. Chicester: John Wiley & Sons, 1988:27-36.
37. KampSax. Dansk Adresse og Vej database [Danish address and road database]. Odense, Denmark: KampSax, 1998.
38. Officielle standard adresser og kordinater [Official standard adresses and coordinates]. OIS: Offentlig Informations Server, 2006. www.ois.dk
39. Paldam M, Svendsen GT. Trust, social capital and economic growth: an international comparison. Cheltenham, UK: Edward Elgar, 2006.
40. Bellis MA, Hughes K, Hughes S, Ashton JR. Measuring paternal discrepancy and its public health consequences. J Epidemiol Community Health 2005;59:749-54.