The Danish Twin Registry:
127 Birth Cohorts of Twins

Axel Skytthe1,2, Kirsten Kyvik1, Niels V. Holm1, James W. Vaupel3, and Kaare Christensen1,2

1The Danish Twin Registry, Epidemiology, Institute of Public Health, University of Southern Denmark, Odense, Denmark
2Danish Centre for Demographic Research, University of Southern Denmark, Odense, Denmark
3Max Planck Institute for Demographic Research, Rostock, Germany

The Danish Twin Registry is the oldest national twin register in the world, initiated in 1954 by ascertainment of twins born from 1870 to 1910. During a number of studies birth cohorts have been added to the register, and by the recent addition of birth cohorts from 1931 to 1996 the Registry now comprizes 127 birth cohorts of twins from 1870 to 1996, with a total of more than 65,000 twin pairs included. In all cohorts the ascertainment has been population-based and independent of the traits studied, although different procedures of ascertainment have been employed. In the oldest cohorts only twin pairs with both twins surviving to age 6 have been included while from 1931 all ascertained twins are included. The completeness of the ascertainment after adjustment for infant mortality is high, with approximately 90% ascertained up to 1968, and complete ascertainment of all liveborn twin pairs since 1968.

The Danish Twin Registry is used as a source for large studies on genetic influence on aging and age-related health problems, normal variation in clinical parameters associated with the metabolic syndrome and cardiovascular diseases, and clinical studies of specific diseases. The combination of survey data with data obtained by linkage to national health related registers enables follow-up studies both of the general twin population and of twins from clinical studies.

The Danish Twin Registry started as the first nationwide twin register in the world, initiated in 1954 by late professor M. Hauge and professor B. Harvald under the supervision of late professor T. Kemp (Harvald & Hauge, 1965; Harvald & Hauge, 1999). By covering birth cohorts since 1870 it is also the twin register with longest follow-up, a feature that has become very useful in the study of aging and age-related conditions (Christensen et al., 1995; Christensen et al., 1999; Christensen et al., 2000; Herskind et al., 1996; Holm, 1983; McGue et al., 1993). During the 1990s twins from the birth cohorts 1953–1982 were included in the Registry as part of a study of type I diabetes mellitus (Kyvik et al., 1995), and these cohorts were the target of a range of studies on migraine (Gervil et al., 1999; Ulrich et al., 1999), eating disorders (Christensen et al., 1995; Christensen et al., 1999; Christensen et al., 2000; Herskind et al., 1996; Holm, 1983; McGue et al., 1993). During the 1990s twins from the birth cohorts 1931–1996 will be described in detail, since the ascertainment of these cohorts has not been presented before.

Material and Methods

Data Sources for the Danish Twin Registry

The ascertainment of a twin pair involves at least two steps: identification of the twins, (i.e., date of birth and names of the twins); and the tracing of the twins, (i.e., their vital status and addresses if alive).

The Danish Civil Registration System (CRS) plays an important role in both the establishment and the maintenance of the Danish Twin Registry. Every person alive at or born after April 1, 1968, has been assigned a unique 10-digit identifier, the Personal Identification Number (PIN). This identifier is the key to individual information in almost all official registers of the Danish population. The CRS covers the total Danish population, since having a PIN is a prerequisite for obtaining any form of social benefit, health care, education, or salary from employer. The CRS also works as a central population register, containing current information about vital status, name, address, place of birth and basic demographic variables such as marital status and children. Also historical information, including historical names, are kept in the CRS.

Initially, the link between family members was made by means of addresses, but the links were erased, when a child reached 18 years, moved away from home, or got married. Since 1979 children has been linked to the mother’s PIN. The CRS has worked its way backwards through earlier birth cohorts in order to re-establish broken links between families, a process which successfully has been accomplished for all children born after the introduction of the CRS in 1968, and to some extent for earlier birth cohorts back to 1953.

Address for correspondence: Axel Skytthe, The Danish Twin Registry, University of Southern Denmark, Sdr. Boulevard 23A, DK-5000 Odense, Denmark. Email: askytthe@health.sdu.dk
Birth Cohorts 1870–1930

Twins from the birth cohorts 1870–1930 have been ascertained by Harvald and Hauge in the 1950s and 1960s by searching the births registers in every parish in the Kingdom of Denmark. The search was carried out by the local clergy who was asked to report twin births, where both twins had survived to the age of 6. For each twin pair, date of birth, names of the twins, names of their parents and their address were requested. Subsequently, all reported twins were searched through available population registers and other public sources, and whenever needed, their closest relatives were sought.

In the birth cohorts 1870–1910 twin pairs of same sex and opposite sex were reported, while only same sexed twin pairs were reported for the cohorts 1911–1930. It may be noted that although Nordschleswig became a part of Denmark in 1921, the search for twins was restricted to the same area as before 1920, the Kingdom of Denmark. A more detailed description of the ascertainment procedure is given by Hauge et al. (1968) and Hauge (1981).

Birth Cohorts 1931–1952

Smaller groups of twins from the cohorts 1931–1952 ascertained from different twin panels were known to the Twin Registry before ascertainment of the total cohort took place. The largest of these was a manually kept register comprising data on male twins supplied by the conscript boards, with approximately 9000 twin pairs. Another twin panel established to study febrile convulsions consisted of twins from Copenhagen. They were identified in 1969 from the newly established CRS as persons born on the same date and having the same address (Schioet Christensen, 1970). In that study a total of 969 twin pairs born before 1953 were identified. Finally, a minor twin panel consisted of 361 same-sexed twin pairs born 1931–1940 on the island of Funen.

In 1996 ascertainment of the birth cohorts from 1931 to 1952 was initiated on a nationwide scale. The most complete ascertainment would have been achieved by scrutinising parish registers, but due to changed administrative regulations and procedures it was not possible to ask the local clergy to report twin births. Parish registers are to be filed in one of four regional archives 30 years after the last entry, but unfortunately not all parish registers are actually handed over to the regional archives (e.g., only half of the parish registers from parishes on the island of Funen and covering birth cohorts 1950–1952 are available at the regional archive). Thus the use of birth registers would only ascertain an incomplete sample.

Therefore another strategy was chosen, taking advantage of the information available in the CRS. The ascertainment of twins was based on the fact that almost all twins are born on the same date, born in the same parish, and are given the same surname at birth.

In CRS place of birth is given by a code indicating either the parish or the municipality where a person is born. About 87% of the registrations in CRS have a parish code indicating place of birth, while a municipality code indicates place of birth in 10%. In the remaining 3% the information is missing. The search algorithm was designed to handle the occurrence of both types of birth place. The first registered historical surname in CRS was used as a proxy for the surname given at birth. In order to match twins who have slightly different spelling of their surname in the CRS the match was truncated to the three first letters in the historical surname or current surname, if no historical surname was registered.

A search in CRS extracted all sets of persons, who were born in Denmark in 1931–52, fulfilling the above conditions, a total of 71,400 persons. The search produced twins as well as singletons, (i.e., pairs fulfilling the criteria but born by different mothers). Confirmation of the twin status was obtained by a mailed one-page questionnaire to 49,699 persons. The questionnaire also included questions on birth weight, parents’ age and vital status, and four questions on similarity in order to determine the zygosity of those who were twins. Approximately 14,500 persons were born in the three largest parishes where they corresponded to one fourth of all persons born in these parishes, indicating that the proportion of singletons was extremely high. In order not to cause unnecessary concern about family relations among too many singletons the twin status for these 14,500 persons was determined by checking birth registers.

Twin status for dead and emigrated persons was clarified by searching either the existing twin panels previously mentioned or available birth registers at the regional archives. Finally, twins from the existing twin panels not found in the CRS sample were added to the register.

Using CRS as the main ascertainment source requires both twins to be alive at April 1, 1968, in order to be selected. In order to evaluate the CRS ascertainment method the Register of Twin Conscripts was used in a capture-recapture analysis to estimate the number of intact pairs with at least one male twin on April 1, 1968 (Hook & Regal, 1995). To obtain a measure for same-sexed female twin pairs, birth registers from a number of parishes were searched for twins, and the number of twin pairs found in birth registers was compared to the number of twin pairs identified through the CRS. This also opened the possibility to estimate the magnitude of twin infant mortality.

Birth Cohorts 1953–1982

Twins from the birth cohorts from 1953–1982 have been ascertained by Kyvik et al. (1995) in the beginning of the 1990s using a single centralized source, the CRS. The primary key to the identification was the link between mother and children. Pairs of persons with the same mother and born within 3 consecutive days were assumed to be twins. Due to the method of establishing the links between parents and children it was only possible to perform a reliable search from 1953. The number of twin pairs identified from the earliest cohorts was low, but all twin pairs born in the period from April 1, 1968, to December 31, 1982, were ascertained by this method if both twins were live born. Pairs with one or both stillborn were not ascertained, because stillborn children do not get a PIN. Since the identification source was the CRS, all additional information needed for tracing was immediately at hand.
## Birth Cohorts 1983–1996

The youngest twins from birth cohorts 1983–1996 have been selected from the Danish Medical Birth Register where all births, including stillbirths, are registered since 1973. Multiple births are marked by a special code, and information about the twins and their mother have been extracted, including the personal identification number. Tracing the twins was done by linkage to the CRS.

Recently, the Medical Birth Register has also been used to supplement the birth cohorts from 1973 to 1982 with stillborn twins, so that the Danish Twin Registry has a complete registration of all twin births in Denmark since 1973.

### Zygosity Assessment

Throughout the entire period of existence of the Danish Twin Registry zygosity determination among same-sexed twins has been based on questions about degree of similarity between co-twins. All initial questionnaires to the twins (or their relatives, if the twin was dead or had emigrated) included three or four questions about this. Based on the answers to these questions the zygosity has been established as monozygotic (MZ) or dizygotic (DZ). In case of disagreement between the two twins or inconsistent answers the zygosity has been classified as unknown (UZ). If answer was available from one twin only, the zygosity diagnosis has been based on that answer alone.

In studies where biological material has been sampled, the zygosity can be further assessed by means of either serological or genetic markers. However, with the current technological advances in molecular genetics the use of serological markers has become obsolete.

## Data Content of the Danish Twin Registry

Besides the basic data on each twin individual and twin pair as collected through the ascertainment process, the twins are followed by yearly updates of vital status, addresses, and marital status. Furthermore, the register is linked to other national health related registers on a regular basis. The Cancer Registry comprises all diagnoses of malignant tumours since 1943, the Registry of Causes of Death comprises all deaths occurring in Denmark since 1943, the Danish Discharge Registry covers all admissions to hospitals since 1977 and all out-patient contacts since 1996. In addition, information about pregnancy and birth has been obtained for all twins since 1973 by linkage to the Medical Birth Register.

In addition to register linkage, surveys are used to get information which is accessible only from the twins themselves, like occupation and lifestyle habits. Since 1995 household surveys have been performed every second year among twins aged 70 years or older as part of the Longitudinal Study of Aging Danish Twins (LSADT). Surveys include health related interviews as well as objective tests of physical and cognitive function, and sampling of DNA (Christensen et al., 1999; Christensen, 2000; McGue & Christensen, 2001). Using a similar instrument as for LSADT basic setup, 4314 twins from the birth cohorts 1931–1952 participated in a household survey in 1998 (Gaist et al., 2000). In 1994 more than 34,000 twin individuals from the cohorts 1953–1982 participated in a questionnaire survey about various health problems and life style (Brix et al., 2000; Gervil et al., 1999; Kjeldsen et al., 2001; Kortegaard et al., 2001; Skadhauge et al., 1999; Ulrich et al., 1999). A questionnaire survey is currently being planned to take place in spring 2002 among 46,000 twin individuals from the birth cohorts 1931–1982.

The questionnaire surveys have been the basis for clinical studies of subgroups of the total twin population, e.g., diabetes mellitus (Kyvik et al., 2000).

## Results

### Number of Twin Births

Table 1 shows the number of twin births in Denmark from 1870–1996 according to official statistics. The numbers include all deliveries that resulted in the birth of two children, alive or stillborn.

There is some uncertainty about the actual number of twin births in the Kingdom of Denmark in the period from 1921 to 1930 due to change of the Denmark-Germany border. No information about the number of twin births in Nordschleswig is available, and the coverage that can be calculated for this period is therefore not precise.

### Number of Ascertained Twin Pairs

Table 2 shows the number of twin pairs identified by name and date of birth from the different periods as shown in Table 2. In the period 1870–1930, however, only pairs with stillborn twins are included. M/M = male/male, M/F = male/female, F/F = female/female.

### Table 1

<table>
<thead>
<tr>
<th>Cohorts</th>
<th>M/M</th>
<th>M/F</th>
<th>F/F</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>1870–1910</td>
<td>5071</td>
<td>5498</td>
<td>4731</td>
<td>37,914</td>
</tr>
<tr>
<td>1911–1930</td>
<td>7222</td>
<td>8576</td>
<td>6902</td>
<td>22,700</td>
</tr>
<tr>
<td>1931–1952</td>
<td>8050</td>
<td>9037</td>
<td>7487</td>
<td>24,574</td>
</tr>
<tr>
<td>1953–1967</td>
<td>4855</td>
<td>4939</td>
<td>4538</td>
<td>13,102</td>
</tr>
<tr>
<td>1968–1982</td>
<td>3712</td>
<td>2845</td>
<td>3299</td>
<td>9856</td>
</tr>
<tr>
<td>1983–1996</td>
<td>3792</td>
<td>3723</td>
<td>3667</td>
<td>11,182</td>
</tr>
</tbody>
</table>

Note: Separate numbers for each combination of sex are not given 1870–1895.

* The numbers given include twin births from Nordschleswig since 1921.

* Pairs with stillborn twins are included. M/M = male/male, M/F = male/female, F/F = female/female.

### Table 2

<table>
<thead>
<tr>
<th>Cohorts</th>
<th>M/M</th>
<th>M/F</th>
<th>F/F</th>
<th>Unknown</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>1870–1910</td>
<td>3989</td>
<td>5142</td>
<td>4197</td>
<td>0</td>
<td>13,237</td>
</tr>
<tr>
<td>1911–1930</td>
<td>2948</td>
<td>3201</td>
<td>0</td>
<td>6149</td>
<td></td>
</tr>
<tr>
<td>1931–1952</td>
<td>5948</td>
<td>6607</td>
<td>4301</td>
<td>163</td>
<td>17,019</td>
</tr>
<tr>
<td>1953–1967</td>
<td>3660</td>
<td>3717</td>
<td>3138</td>
<td>13</td>
<td>10,528</td>
</tr>
<tr>
<td>1968–1982</td>
<td>3630</td>
<td>2814</td>
<td>3234</td>
<td>1</td>
<td>9679</td>
</tr>
<tr>
<td>1983–1996</td>
<td>3788</td>
<td>3728</td>
<td>3667</td>
<td>0</td>
<td>11,183</td>
</tr>
</tbody>
</table>

Note: M/M = male/male, M/F = male/female, F/F = female/female.

* Pairs with stillborn twins are included.
twin pairs where both members have survived to the age of 6 have been included. Approximately 30,000 twin pairs are broken or lost before the age of 6, while the remaining approximately 12,000 twin pairs remain untraced, among them all opposite-sexed pairs from 1911–1930.

The low number of included twin pairs born before 1931 compared to the number of twin pairs born may be attributed to the high infant mortality leading to a low proportion of unbroken twin pairs. In Figure 1 this is illustrated by the proportion of twin pairs included in the register and the estimated proportion of twin pairs not included due to the death of one or both of the twins prior to age 6. The estimate is based on the assumption that in the period 1870–1910 only between 40 and 45% of the twin pairs both twins reached age 6 (Hauge, 1981). Since no opposite-sexed twins are included for the period 1911–1930, the shown completeness is representing the proportion of same-sexed twins — furthermore the missing data on twin births in Nordschleswig makes the calculated completeness for the period 1921–1930 somewhat misleading. Looking at same-sexed twin pairs alone brings the completeness to the same level as for the surrounding years.

From birth cohort 1931 and onwards the numbers include twin pairs with stillbirths and infant deaths to the extent, they have become known to the Registry. A total of 17,019 twin pairs born in Denmark 1931–1952 have been ascertained from the Medical Birth Registry have not been included due to the death of one or both of the twins prior to age 6. The majority of the non-identification was expected to be due to infant death of one or both of the twins. In the capture-recapture analysis a completeness of more than 90% was found for same-sexed male twin pairs where both members were alive on April 1, 1968. For same-sexed female twin pairs a lower completeness, about 80%, was found. The high completeness for male-male twin pairs may be attributed to the use of the Register of Twin Conscripts as an auxiliary source, while the lower completeness for female-female twin pairs may be due to errors in CRS regarding maiden names.

From the comparison with the total number of twins found in selected birth registers it was shown that the twin infant mortality in the period 1931–1952 was 3–5 times the general infant mortality. Within the first year of life one or both twins had died in 20–30% of the pairs (Skytthe, 2000).

A shift in the curve is noted in 1953, when the ascertainment procedure changed. Especially in the birth cohorts 1953–1960 a low completeness is found among same-sexed female twin pairs only. The explanation for this is the underlying method of establishing the relationship between parents and their children in the CRS.

From April 1968 there is a complete ascertainment of twin pairs with both members live born, and as mentioned earlier there is a complete ascertainment of all twins since 1973 due to the use of the Medical Birth Register.

Distribution of zygosity as evaluated by the questionnaire method is shown in Table 3. Since the youngest twins ascertained from the Medical Birth Registry have not been contacted, no zygosity information is available for these cohorts yet. The relatively high number of pairs with unknown zygosity among the twins from 1931–1952 is due to the inclusion of pairs, where one or both of the twins had died in infancy.

### Discussion

The establishment of population-based twin registers has primarily been based on computerized registers covering the majority or whole population — for example, social security files (Goldberg et al., 1997), driver’s license (Afari et al., 2001), fiscal purposes (Salvetti et al., 1997) — involving some kind of algorithm to identify possible twins, involving same birth date, same name, and same place of birth. The success of the procedure relies on the specificity of these data, (i.e., how common are names), what is the tradition regarding change of names, and the size of region considered as same place of birth. In any case it is necessary to verify that a suspected twin actually is a twin, and for same-sexed twins also to assess the zygosity, often by a direct contact to the twin. The experience from the Danish Twin Registry for the birth cohorts 1931–1952 shows that only half of the selected
persons were actually twins. The Finnish Twin Registry has been compiled using the same methodology, but probably due to the greater diversity of Finnish surnames, 83% of the selected persons were twins (Kaprio et al., 1979).

One of the advantages of population-based twin registries is that they are established independently of the traits under study. However, other limitations may be imposed on a population based twin register. The way the register is established might limit the traits that can be studied using the register. For example, the use of driver license application files as ascertainment source implies that only twins being “healthy” at the moment of applying for a driver license are included in the register; thus twins being impaired physically or mentally are not included in the register (Afari et al., 2001).

In the Danish Twin Registry some of the birth cohorts are left truncated, meaning that it is not possible to study traits or events occurring in the early part of life among all birth cohorts, (e.g., twin infant mortality). Likewise a study of long-term effect of the loss of a co-twin in childhood is only possible for the cohorts since 1968; the inclusion of broken twin pairs from earlier cohorts depends on the surviving twin to report his twin status.

The completeness of a twin register can be viewed in several ways. The most stringent way is to calculate the proportion of twin pairs ascertained by the twin register out of the total number of twin births. Another more pragmatic approach is when the denominator consists of all twins reaching a certain age, depending on the phenotype under study. However, when some restrictions have been put on the ascertainment criteria, the completeness, given the restrictions, can be difficult to assess. Whether the completeness is an issue to consider depends on the actual research question. Studying infant death in twins requires a very high completeness in order to be reliable, while the search for genes in disease-discordant twin pairs is much less dependent on the completeness, if not independent.

There is no doubt that the twin infant mortality is the primary cause of the seemingly low completeness found before 1953. The estimates of infant mortality among twins have been found to be 3–5 times higher than in the general population, which means that up to half of all twin individuals born in the 19th century may have died in infancy. The general decline in infant mortality observed in the general population through the first half of the 20th century also affects the twin population (Statistiske Undersøgelser, 1966). This has implications for the number of twin pairs that are available for identification through the CRS.

Danish twins have shown a high willingness to participate in surveys, with interview participation rates between 72 and 83% for the various age groups, and more than 90% of the participants willing to give a biological sample (Christensen, 2000). Questionnaire response rates have been as high as 92% for a one page questionnaire and 86% for a seven page questionnaire administered to the younger twins (Brix et al., 2001; Kyvik et al., 1995).

Despite the various limitations in the completeness and truncations due to the different ascertainment methods the Danish Twin Registry constitutes a valuable research resource, spanning a period of more than 125 years. This opens possibilities for the study of diseases and health related behaviour in all ages. Besides the classical twin study of the relative importance of genetic and environmental factors and twin-control of environmental exposures, it also makes it possible to further study the influence of age in relation to genes, environment and the interaction of these. Furthermore the use of information about family relations obtained from linkage to other registers has broadened the application of twin data to new areas, for example, fertility patterns (Rodgers et al., 2001) and bereavement effects (Tomasini et al., 2002).

Acknowledgments

We want to express our thanks to Gudrun Hauge for her enthusiastic work at the Danish Twin Registry through more than 45 years.

We want to thank the following for their considerable financial contribution to the foundation and the continued support of the Danish Twin Registry: the Danish Research Council, University of Southern Denmark, the Danish National Research Foundation, Helsfonden, the Danish Cancer Society, the Danish Diabetic Association, the Danish Heart Foundation, the Novo Nordisk Foundation, United States National Institute of Aging (P01-AG08761), and National Cancer Institute.

References


