

TWIN REGISTERS AND INTERNATIONAL COOPERATION

Neurological Dept., Glostrup Hospital, Denmark

Computer Selection of Twin Pairs ***E. Schiöttz-Christensen**

A new method of cross-sectional twin sampling from a population may be used in countries where authorities utilize electronic data processing of National Register data.

In Denmark, all National Register data are collected on magnetic tape. In January 1969 it was consequently possible, by the help of a 360-COBOL programme, to select all individuals in the Copenhagen area with identical birthday, month and year, and identical address. The selection was limited to the years 1943 to 1968.

Data from every person were tested against data from all other persons, to prove whether the person had one or more partners with identical data, i. e., twin, triplet, or quadruplet partners. Names and current addresses of the twins were written on self-adhesive labels intended for later correspondence. Had it been desirable, the computer might have produced, in the same operation, index cards and questionnaires ready for mailing. In the primary sample, a few "false" twin pairs from great collective housekeepings (as, for example, colleges) were included. These "false" pairs were excluded after comparing the partners surnames.

From the total population of 1 712 000 people, 4595 twin, 25 triplet, and 1 quadruplet sets were found. As many as 3754 twin pairs were born in 1951-1967. Of these, 34.68% were of opposite sex. Of the remaining pairs, 642 were born in 1943-1950, 28.19% being of opposite sex.

Corresponding to four suburban municipalities with a total population of 171 245 and a population of 51 660 in 1951-1967, it was possible to calculate the twin frequencies within the single years, based on the current population. The mean frequency of twin individuals in complete pairs, born in these years was 2.07%. In the same years, the proportion of opposite sexed pairs in this district was 35.93%. In the metropolitan district, with a total population of 835 246 (of which 152 500 were born in 1951-1967), the corresponding mean twin frequency was 1.68% (this

* This investigation was supported by a research grant from the Danish Epilepsy Association.

frequency having been calculated from the estimated number of people, the second decimal is unreliable). In the metropolitan district the proportion of opposite-sexed pairs in 1951-1967 was 33.24%. The combined metropolitan and suburban district showed a mean twin frequency of 1.73%.

This twin compilation method yields a well defined twin sample including every complete cohabiting pair, but no pair separated by death or different residence. Because of the intrapair mortality correlation, the exclusion of pairs broken by death will exclude physically weaker individuals from the sample, and make it healthier

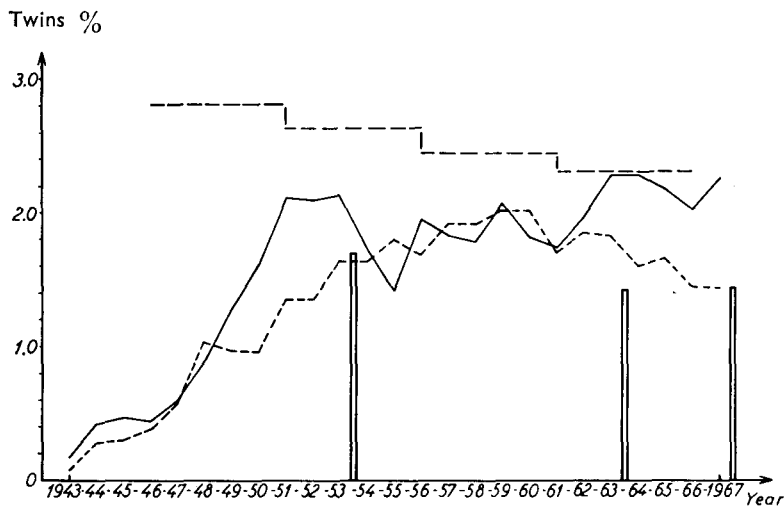


Fig. 1. Frequencies of twin individuals in complete pairs calculated per year
 --- At birth, calculated for the whole of Denmark (Vital Statistics, 1969).
 — In four suburban municipalities of Copenhagen.
 ···· In the metropolitan district of Copenhagen.
 - · - · Expected frequencies according to survival rates observed by Essen-Möller (1941).

(WHO Report, 1966). Furthermore, the fact that broken pairs will not be ascertained may introduce difficulties in twin studies concerning severe diseases, which often lead one or both cotwins to death or residence in an institution. These deficiencies in a computer selected series must, however, be compared with the sometimes considerable loss of cases in twin compilation from birth records, military rosters, etc., which introduces an unknown bias in such series.

Agreement between a twin sample and the parent population, with respect to the proportion of opposite-sexed pairs, is perhaps the most useful evidence of unbiased sampling (Allen, 1955). At birth, this proportion in 1959-1967 was 33.36%, calculated for the whole of Denmark. If correction is made in regard to the different stillbirth rate and neonatal mortality of same-sexed and opposite-sexed pairs (Yerushalmy and Sheerar, 1940), the proportion will increase to 34.35% at the age of one month, which agrees with the here observed proportion of 34.68% in 1951-

1967. The present proportion of opposite-sexed pairs is also close to the proportion of 36.64, which may be calculated from the twin series of Harvald and Hauge (1965) based on birth records from the years 1870-1910. The smaller number of opposite-sexed pairs in 1943-1950 suggests that the sample from these years is biased.

The twin frequencies within the single years are pictured in Fig. 1. From about 1951 to 1943, these frequencies drop steeply, as the pairs are probably separated by marriage and education. In 1950-1953 and 1962-1967 the frequencies in the suburban district are significantly higher than in the metropolitan district, but in 1954-1961 they agree closely. This may be caused by the fact that large families, since 1967, get rent subsidies when moving from a small, old flat to a new and larger one. The agreement of twin frequencies in 1954-1961 may reflect a lesser inclination to move when families have children in the compulsory school age.

The expected frequencies of twins in unbroken pairs in the years 1953, 1957, and 1967 have been calculated according to survival rates observed by Essen-Möller (1941). This is an approximate estimate, as it is uncertain whether these survival rates may be applied to the present series. However, the calculated frequencies closely agree with the observed ones. A mean twin frequency of 1.73% in the present series may also be compared with the frequency of 1.6% found by Husén (1953), who examined Swedish men liable for military service. In his series — from a generation with a supposed twin birth frequency of 1.46% — individuals from pairs broken by death were excluded. The here found frequency of 1.73%, for twins in unbroken pairs, may finally be compared with the USA frequency of 1.90% calculated by Allen (1955). In his calculation, also singletons are included. This comparison between Denmark and the USA is justified by the rather close agreement with respect to twin birth frequency, stillbirth rate, and general neonatal mortality.

It may be concluded that computer selection of twin pairs yields all cohabiting twin pairs in a given district, but no broken pairs. The twin series will be representative in the younger generations only. It must be stressed that the method is very quick, relatively cheap, and that it delivers the information in a practical form.

References

- ALLEN G. (1955). Comments on the analysis on twin samples. *Acta Genet. Med. Gemellol.*, **4**: 143-159.
- ESSEN-MÖLLER E. (1941). Empirische Ähnlichkeitsdiagnose bei Zwillingen. *Hereditas*, **27**: 7-16.
- HARVALD B., HAUGE M. (1965). Hereditary Factors Elucidated by Twin Studies. *Gen. and the Epid. of Chron. Dis. Public Health Service Publication N. 1163*.
- HUSÉN T. (1953). *Tvillingstudier*. Almqvist & Wiksell.
- WORLD HEALTH ORGANIZATION (1966). The use of twins in epidemiological studies. *Acta Genet. Med. Gemellol.*, **15**: 111-128.
- YERUSHALMY J., SHEERAR S. E. (1940). Studies on twins. II. On the early mortality of like-sexed and unlike-sexed twins. *Hum. Biol.*, **12**: 247-263.

E. SCHIØTTZ-CHRISTENSEN, M. D., Skyttehusene 62, 2620 Albertslund, Denmark.