How did the study come about?

The Stockholm Birth Cohort Study (SBC) was created in 2004/2005 by a probability matching of two comprehensive and longitudinal datasets. The first, the Stockholm Metropolitan study 1953–1985, consists of all children born in 1953 and living in the Stockholm metropolitan area in 1963. The second, The Swedish Work and Mortality Database 1980–2002 (WMD), consists of all individuals living in Sweden in 1980 or 1990, and born before 1985. The initiative to create the database was taken by Denny Vägerö at the Centre for Health Equity Studies, CHESS, of Stockholm University/Karolinska Institute and Sten-Ake Stenberg at the Swedish Institute for Social Research, Stockholm University (SOFI). We refer to the resulting database as the SBC, which in effect provides a 50 year long follow-up of the original 1953 birth cohort.

Project Metropolitan began in 1964 on the initiative of the late professor Kaare Svalastoga at the Institute of Sociology, University of Copenhagen. Towards the end of the 1950s he tried to persuade Nordic researchers of the desirability of longitudinal cohort studies. The studies should have a prospective design and cover the period from early school age to early middle life. Svalastoga’s initiative was only successful in Denmark and Sweden. Details of the Danish study are published in an accompanying profile in this issue of the Journal.

At Stockholm University professor Carl-Gunnar Janson shouldered the task and began to collect data for the cohort. Fortunately, in contrast to Svalastoga’s recommendations, professor Janson decided to include not only boys but girls also. Data was de-identified in May 1986, after which no more information could be included. In 2002 all data were moved to the Swedish Institute for Social Research in order to better document the study and to explore more fully its research potential. The Work and Mortality Database (WMD) is a temporary database without any personal identification, created at CHESS in 2003 to allow a study of how work, income, and labour market position combine to influence health, disease, and mortality. Both datasets are anonymous and without personal identification codes. However, we realized that the two datasets included a number of identical variables, presumably derived from the same sources, that could be used to combine the two datasets. This gave rise to the hope that we would be able to create a new 1950s birth cohort, following the examples of the revitalized studies in Copenhagen and Aberdeen (both profiled in this issue of the journal). A very large amount of work went into tracing historical documentation and classification procedures of the original data for those variables that were identical in both registries and, therefore, could serve as the basis for a matching algorithm. Ethical permission from the Stockholm Regional Ethics committee as well as ethical rules imposed by the National Board of Health and Welfare and by Statistics Sweden was a necessary precondition before the probability matching was eventually performed in 2004. Consistency checks and documentation are still taking place. Analysis on this new dataset has barely begun.

The original Metropolitan study was financed by the Bank of Sweden Tercentenary Foundation during the 21 year period, 1966–86, mostly as grants to collect and organize the data. Professor Janson’s research on Metropolitan data was also financed by the Swedish Council for Planning and Coordination of Research (1986–92), the Commission for Social Research (1986–90), the Foundation Wenner-Grenska samfundet (1991), the Bank of Sweden Tercentenary Foundation (1991–96), the Swedish Council for Social Research (1992–95), and the Swedish Council for Humanistic and Social Sciences (1997–2002). Documentation and creation of the SBC has received financial support from the Swedish council for working life and social research and from the Swedish Research Council, including its longitudinal committee.

What does the study cover?

The original study had four main areas of enquiry: (i) social mobility, especially intergenerational mobility, and the role played by education; (ii) memberships of groups and formal associations; (iii) conformity and deviance; and (iv) choice of partner, marital adjustment, and divorce.

The new Stockholm Birth Cohort Study was created to allow the study of intergenerational and life course associations, for instance research questions of how childhood social circumstances influence adult labour market position, unemployment experience or income career as well as adult health, disease, and mortality.

Who is in the sample?

The cohort is defined as all children who were born in 1953 and lived in the Stockholm metropolitan area on November 1, 1963.
How often have they been followed-up and what has been measured?

The first comprehensive data collection was a questionnaire-based school study in 1966, which included practically all public and private schools in the Stockholm area. The questionnaire covered a range of questions, including tests of three components of intelligence, three attitude scales about education and school, five scales measuring leisure time interests, and questions on plans for future education. Questions also covered how attractive the pupil found various occupations, as well as three socio-metric questions, each of which asked for the three closest classmates, given certain specific circumstances.

The Family Survey was finished in 1968, 2 years after the school study. At approximate cohort age 15, a stratified sample of mothers (mostly) of 4021 children were interviewed. Questions were asked about the education and occupation of the interviewee and of the father, about the mother’s opinions on occupational choice for the child, on upbringing, and on society at large, and on her (and the father’s) educational plans for the cohort member. Interactions and relations within the family were covered by specific questions.

The same sub-sample was used for a mail questionnaire study of the cohort members in 1985, at the approximate age of 32. The questions concerned the cohort members’ self-rated health, education, use of, or attitude to, mass media, occupation, employment, family size and composition, as well as education, occupation, and employment of the spouse (if any).

In addition, routine registries were used to collect additional data. Early registry data refer both to family of origin and to the cohort members themselves (1953–72). For specific years the family data cover residential district, family composition, marital status, housing conditions, means-tested social assistance, parents’ employment, income, education, social class, and nationality. Father’s police record, if any, is indicated. There is full obstetric data from cohort members’ birth in 1953; information about decisions by the Child Welfare Committees (CWCs); about offences committed up to the age of 15; school grades and truancy in the sixth and ninth grades of comprehensive school (ages 13 and 16); application to, admission to, and completion of senior secondary school (gymnasium, at age 18); and psychological and physiological data from military conscription (boys only) in 1971–72.

As cohort members left home, data exist (1968–83) on their residential area; their own (and any cohabitant’s) housing conditions, employment, occupation, social class, income and marital status, any university degrees. For certain years the number of days and periods of sick-leave and means tested social assistance, narcotics experience, and criminality are covered. Also included are the number of childbirths to female cohort members and number of children living with cohort member as well as main and secondary discharge diagnoses following inpatient care up to 1983.

As the original cohort was matched to the Swedish Work and Mortality database, inpatient data could be included up to 2002. Income data, together with data on unemployment and social assistance, from the so-called Louise database (administered by Statistics Sweden), from each year 1991–2002 are covered this way. Cohort members are linked to their parents, for whom the same data exist. Any deaths up to the end of 2002 are covered, with data on cause-of-death, resulting in a 50 year unbroken follow-up of mortality for the original cohort.

What is attrition like?

The original 1953 birth cohort consists of 15 117 individuals (7398 girls and 7719 boys) of whom 167 had died before the end of 1980. Probability matching of those individuals in the original cohort still alive at the end of 1980, to the Swedish Work and Mortality database, was undertaken. Thus 14 950 individuals were eligible for matching. The matching procedure was rather difficult and performed in several steps. Details can be found in a previous paper.1 We, finally, ended up with 14 294 matched observations corresponding to 96% of the individuals still alive in 1980. Thus we lost 4% of the cohort in this probability matching. We believe most of the loss is due to emigration.

Early data

Of the original cohort, defined in 1963 (see above), 137 girls and 151 boys were not included in the School study (1966). Five had died and the others had moved out of the Metropolitan area. Of the remaining cohort members, 613 girls and 740 boys, i.e. 8.4 and 9.7%, did not take part in the school questionnaire study. In all, data from the school study are thus available for 6648 girls and 6828 boys, together 89% of the original 15 117 individuals.

The total number of children sampled for the family study in 1968, was 4021 (1972 girls and 2049 boys), whose mothers (mainly) were interviewed. Ninety-one percentage (3651 individuals) of the sub-sample was covered by the study. The same sub-sample was used for the mail-questionnaire in 1985. This time 76% of the sample responded.

Most early data were, however, collected from local registers in the metropolitan area: this is true for deliveries (birth data), parental and juvenile social-register records, (data from social registers by municipality), and comprehensive school records (data by school). Also, records from the 1960 Census of Population and Housing covered only those living in the metropolitan area at the time. However, the local datasets cover the great majority of the cohort members: 82% in 1953, 93% in 1960, 100% in 1963, and ~90% in 1972. A further three
series of data were regionally restricted. The conscript data covered those mustering in the Stockholm area (90% of the male cohort); the narcotics injection scar data were restricted to Stockholm police district. Hospital admissions covered only those taking place within the Stockholm metropolitan area. In 1975, 89% of the cohort lived in the Stockholm county, in 1980–82% did so.

**Later data**

Later data series, derived from the WMD, covered the whole of Sweden. Thus data on inpatient visits from 1981, cause of death data from 1981, data from the 1980 and 1990 Census, data on unemployment, social assistance, income, and education are believed to be virtually complete.

**What has it found? Key findings and publications**

The original study has resulted in more than 100 publications, book chapters, and reports, in Swedish or in English. Some of these are difficult to trace. A full list of all these publications has been compiled by the authors and will be published on our home pages shortly (at www.sofi.su.se or www.chess.su.se). The most common research themes in studies of the original cohort have concerned educational achievement and social mobility, youth delinquency and criminality, or intergenerational transfer of social disadvantage. Lagerström found that babies born small for gestational age, unlike low birth weight babies, in general, performed poorly at school. Walldeén found that sibling position, independent of family size, predicted school achievement.

The new data covered by the Stockholm Birth Cohort Study are still being prepared. However, there is some work under way. Brännström has completed a first study, exploring any area effects on the variation in income and social welfare receipt in greater Stockholm and concluding that these were very small. Preliminary studies of cognitive performance at age 13 and subsequent mortality experience suggest a strong link. Most of the effect seems mediated by educational achievement. Primary and secondary socialization effects on achieved education are being addressed. Studies of childhood experience as predictors for unemployment experience, income careers, and later well-being and health are also under way.

**What are the main strengths and weaknesses?**

A total of 13,859 persons of the SBC were still alive by the end of 2002 and are thus covered by 50 year of follow-up data, from birth to age 50. There are 602 deaths in the study, allowing mortality analysis by a number of specific causes. Unlike the Copenhagen study our study includes women.

Can I get hold of the data? Where can I find out more?

The linked data files are kept at CHESS, the original cohort data are kept at SOFI. We are now working on describing the entire dataset in the form of code books, later on these will form the basis for a web-site. A steering committee consisting of the authors, plus database manager Reidar Österman will take decisions of when to extract data for your purpose. Potential collaborators should discuss ideas informally with the steering group before submitting a proposal. We will need information on what the purpose is, what data you need, and who (if anyone) is your partner at CHESS or SOFI. If several people want to do the same thing, priority access is given to researchers at CHESS and/or SOFI, but in most (or all) cases we hope to be able to collaborate with researchers outside these two host departments.

**References**