13 October 1995

The National Birth Cohort in Denmark

'Better health for mother and child'

A short description of the basic cohort

Jørn Olsen, Mads Melbye, Thorkild IA Sørensen, Peter Aaby
The Danish Epidemiology Science Centre, Denmark

Contents

Summary

Objective: We propose to establish a cohort of 100,000 pregnant women in order to obtain exposure data which will facilitate studies of the role of different conditions during fetal life and infancy for health and morbidity in childhood, adolescence and in adult life.

Background: Data from animal experiments and a growing number of epidemiological studies indicate that prenatal exposures (eg infections, medication, eating habits, life style, occupational and environmental exposures) may play an important role for diseases which occur not only immediately after birth but also later in life. Many reproductive failures still have unknown etiology and often these failures have serious health consequences for the child, the family and the society. Since the cause-effect relation may span many years and involve rare exposures and rare outcomes, a study that focuses on these issues must necessarily be large and involve subjects followed for many years.

Conditions for the study: Denmark has unique possibilities for long-term follow-up studies of health outcomes due to the existence of a large number of population-based health registers and a personal identification number (ID) which permits the linkage of these registers. However, the utilisation of this resource has been insufficient because of limited data on potentially relevant exposures.

Design: The study will collect self-reported exposure information from pregnant women and their children and a series of blood samples during pregnancy which will permit the detection of specific infections, drugs, toxic agents, hormones and genetic characteristics which may influence the development of the fetus. Computer assisted telephone interviews regarding exposures during pregnancy will be conducted at the end of the 1st, 2nd and 3rd trimester. Another interview at 18 months of age of the child will collect information on health outcomes and health exposures during infancy. Blood samples will be collected at weeks 8-12 and at week 26 of gestation and at birth.

Context and implications: The main project presented in this summary protocol will provide the infrastructure for a large number of specific studies. It is expected that the study will contribute significantly to improved antenatal and infancy care as well as to the prevention of chronic disease in childhood, adolescence and adulthood.

Objective and background

The epidemiology of reproductive health has been an active discipline in the Nordic countries for many years. Some of the birth registers have been maintained in closed collaboration with research centers and have been used extensively in research. The main content of the birth registers, and health registers in general, can be characterised as outcome variables. These health registers could be used to a much greater extent in research if systematicly collected data on potential etiological exposures were available.

We therefore propose to establish a cohort of 100.000 pregnant women in order to obtain exposure data on important events during fetal life and infancy. The proposed project will facilitate large-scale studies of the role of different conditions in early life for health and morbidity in childhood, adolescence and in adult life.

Small and medium-sized studies in pregnancy with short-term follow-up have provided very useful information and much of our antenatal care is based upon these studies. However, further improvements in health care are warranted and these require larger studies with longer follow-up.

Justification of the project and aim

There are numerous benefits of establishing the proposed exposure register. In the short run, the study will identify and document causes of ill health in pregnancy and early childhood. In the longer term, it will facilitate the study of the impact of exposures during pregnancy and early childhood on serious illnesses in adolescence and adult life.

There are many reasons to believe that exposures in utero may manifest themselves later in life. Exposure of female fetuses to dietylstilboestrol (DES) and vaginal cancer during adolescence is one well-documented example, but hypotheses related to other illnesses exist. Infertility, cancer of the testis, breast cancer and schizophrenia may for instance be related to in utero exposures. The proposed exposure study provides opportunities for studying such hypotheses.

The study will furthermore provide data to examine Barker's 'programming hypothesis'. According to this hypothesis, interference with fetal or early childhood growth at critical stages may modify organ functions with later health implications, eg cardiovascular disease. The study will show whether exposures influencing fetal growth are linked with later health problems, such as diabetes, hypertension, a cardiovascular risk profile and childhood cancers. Additionally, the study may enable evaluation of whether the basis for the 'programming' is in fact genetic, is due to gene-environment interactions, or whether it is due to avoidable

exposures during pregnancy. Results linking fetal and early childhood growth to cardiovascular and other chronic diseases have been presented in more than 50 papers and two books. These growth-related associations could, however, have many causes and more detailed data are needed to approach information which may be implemented into prevention. A recent study among Swedish twins showed no substantial cardiovascular disease patterns in their adult lives despite the fact that they were growth-retarded at birth. Similar findings are seen among Danish twins. These studies underline the need for a more detailed classification of fetal growth as risk indicators of subsequent diseases.

Fetal life, once a subject of interest to only a minority of researchers, has now the attention of anyone concerned with the prevention of heart diseases, cancer, hypertension, diabetes or psychiatric diseases. Furthermore, the fast progress in genetic research opens up new possibilities for genetic epidemiology which should be explored.

Overall design of the basic project

The basic structure of the study is described in the following. Linked to the main project is a number of smaller studies which usually add new data for specific subgroups.

The study is based upon a cohort of consecutively sampled pregnant women who are asked to join the project early in pregnancy. The aim is to obtain a sample size of 100,000 pregnant women. The total population in Denmark is about 5 million with approximately 60,000 births per year. Women are invited to participate in the study if they speak Danish well enough to take part in telephone interviews and intend to carry their pregnancy to term.

Participation in the project involves the following:

1) A computer-assisted telephone interview on pre-pregnancy and first trimester exposures in weeks 12 to 14 of pregnancy.
2) A computer-assisted telephone interview on exposure in pregnancy in weeks 28 to 30.
3) A blood sample in week 26 of pregnancy and a blood sample from the umbilical cord after delivery.
4) Permission to store left-over blood from routine syphilis screening early in pregnancy as well as the two above-mentioned blood samples in a biological bank.
5) Answering a questionnaire shortly after delivery on exposures in the third trimester and events in connection with delivery.
6) Computer-assisted telephone interview on exposures in early childhood when the child is about 18 months old.
7) Permission to link information from existing health registers to the exposure data.

The pregnant women are recruited at their first pregnancy visit to the general practitioner (GP), which for 90% of all pregnant women takes place within the first 12 weeks of pregnancy. As part of this visit, a blood specimen is submitted for syphilis testing and the woman is referred to further health controls at the Antenatal Care Centres (ANC).

At the GP, she receives an envelope containing a letter, a leaflet describing the background for the project, a summary of what she is asked to provide, the consent form and a short questionnaire on medication around the time of conception. The consent form is returned to the coordinating centre together with information on the last menstrual period (LMP), her telephone number and most convenient time periods for a telephone interview.

These data are entered into the database and the subsequent flow of information is 'controlled' by the same database. Self-reported exposure information is entered directly into the database in the computer-assisted telephone interviews. Software, questionnaires and registration forms have been developed and are being pilot tested at the moment. Data will be stored in such a way that they are linkable with for instance the Medical Birth Registry, the National Hospital Registry and the Danish Cancer Registry.

Women who return the consent form receive a label to be placed in their own pregnancy file which is used at each visit to health professionals during pregnancy and at delivery. Attached to this label is a set of stickers with the woman's project identification number to be used when labelling blood samples or registration forms.

Self-reported exposure data in pregnancy

Data will be collected twice in pregnancy by means of a telephone interview. The timing of the second interview in pregnancy is a compromise between obtaining exposure data for most of the pregnancy period and still reaching the majority of the women who give preterm birth. Shortly after birth, the women are asked to fill in a third questionnaire regarding exposures in the third trimester and events in connection with delivery.

The aim of the interviews is to collect data at relevant time periods to identify exposures which are easily forgotten, such as minor diseases, short-term medication, changing exposures, etc. Data on eg drinking and smoking habits or exposure to stress are vulnerable to information bias and have to be collected during pregnancy. The data collection covers suspected and known risk factors of fetal growth and pregnancy outcome as well as factors which may act as confounders in specific studies.

Background characteristics

CPR-number, date of filling in the questionnaire, maternal height, maternal pre-pregnant

weight, paternal height, paternal age, paternal weight, ethnicity of pregnant woman and spouse.

Medical and obstetric history

Last menstrual period (LMP) (date, characteristic of bleeding, cycle length), use of anticonception, present pregnancy planning and waiting time to pregnancy, use of contraceptive methods, specific chronic diseases, diseases developed during pregnancy, fever during pregnancy, uro-genital infections, medication prior to pregnancy (including vitamins), medication during pregnancy, outcome and time of previous pregnancies, infertility treatment, hyperemesis, bleeding episodes.

Socio-economic and work-environmental characteristics

Marital and cohabiting status, education of pregnant woman and spouse, employment status and job title, work-status during pregnancy, number of hours at work, physical exertion, hours of standing and walking, amount of lifting, chemical exposures (list), radiation.

Life style factors and life events

Smoking prior to and during pregnancy (including exposure to environmental tobacco smoke), alcohol consumption, dietary factors, consumption of caffeine, pica, stressful life events, leisure time activities (TV, fitness).

Data collection of exposures and outcomes in early childhood

Studies on programming have shown that not only fetal but also early childhood growth is associated with diseases later in life. Early childhood growth as measured by the weight and height at one year of age is also associated with a number of diseases with later occurrence. It is also well accepted that breast feeding, contact to parents, social factors and exposures to environmental hazards such as infectious agents are important. These data are not readily available in any existing, general database in Denmark. Early childhood development is usually recorded, but the data are not in general available for research.

The following scheme for data collection is therefore planned:

A A computer-assisted telephone interview on exposures in early childhood when the child is about 18 months old. The questionnaire includes data on breastfeeding, growth, diseases, treatment, immunizations, motoric and psychological development and social factors.

B Most children are examined at least once at home by the child health nurse, and often two or three times. All existing records should be copied and stored in a central archive in a way which makes linkage with the cohort possible. The information will only be coded if the children are selected as cases or controls in more targeted projects.

C Most counties offer 'The child's book' to parents. We intend to harmonise the content of this book or supplement it with a diary type note book with critical information on events and developments in early childhood. These notes will serve as an important memory aid when filling in questionnaires or when parents are selected as cases or controls in subsequent studies.

Selection of subcohorts

Data should also be stored in such a way that specific sub-samples may be identified for more intensive follow-up. Such cohorts may be identified according to the exposure (eg a given drug exposure during pregnancy) or according to specific growth characteristics. It would, for example, be of importance to be able to follow a cohort of children born after recently introduced treatment of infertility such as IVF and hormonal induction of ovulations. Such cohorts may be referred to specific clinical or development assessments after birth. Due to the data collection strategy, data will be without differential recall problems and the data are collected close to events under study in order to reduce random misclassification.

Establishment of a biological bank as part of the cohort studies

Most published studies dealing with the consequences of intrauterine exposures for the child's health are questionnaire or register-based and lack biological measurements. Important information regarding external exposures and the physiological characteristics of the pregnant women can only be obtained through analyses of biological materials. A critical element in the National Birth Cohort protocol is therefore the establishment of a biological bank.

With the materials included in the biological bank, it will be possible to detect, for example, DNA, antibodies, hormones (eg thyroid hormone, estrogene products), total IgE, ferritin, heavy metals (eg lead), trace metals, steroids, mannan binding protein, Gm-allotypes, certain vitamines, cholesterol, glycosis, alpha-tocoferol, triglycerides, pharmacological products, pesticides, alcohol dervatives and chemical compounds (environmental toxins). The rapid development in molecular biology will in the coming years expand our possibilities to test biomarkers relatef to eg DNA and protein adducts, disease-associated genes and gene products. An increasing number of exposures can be traced in biologic materials and the amount of material required for such analyses is often very small.

A few examples of questions that can be addressed based on the biological samples are: which intrauterine infections cause hearing and vision defect? is there a correlation between the mother's concentration of oestrogene metabolites and male genital malformations? does influenza cause skizofrenia? is there an association between specific IgE responses in the mother and allergic conditions in the child?

The testing of specific hypotheses of a potential causality between an exposure and a disease will be based on a nested case-control design. This approach ensures that little material is used for specific studies, and thus the bank can be a rich source for many studies in the years to come.

Number of blood specimens planned for the birth cohort study

For many adverse outcomes, the most critical period is the first 16-18 weeks during pregnancy. This is particularly the case with respect to exposures such as infectious diseases, teratogenic medication, and x-ray irradiation. However, some of the critical events in the development of the sensory organs and brain appear later in pregnancy and external exposures during this period may cause severe adverse effects as well. Therefore, it is planned to collect biological materials early (weeks 8-12) in pregnancy, in week 26 and again at delivery.

1. An early specimen (pregnancy week 8-12): A blood specimen is routinely taken for syphilis testing on all Danish women at the time of confirmation of their pregnancy. All specimens will be directly forwarded to Statens Seruminstitut where this analysis is performed and where the biological bank will be established. Approximately 2 ml of serum will be available from the syphilis screening for the project. To preserve DNA-material, three blood spots will be placed on filter paper, and the blood clot will be frozen at -80 C (1 aliquot). Serum will be frozen at -20 C in five aliquots.

2. Pregnancy week 26: The second blood specimen will be taken at the control visit to the private physician in week 26. Ten ml of heparinized whole blood will be collected and forwarded to Statens Seruminstitut for processing. Plasma will be frozen in five vials, buffy coat in 2 aliquot, and red blood cells in one vial. Finally, cell-rich material will be placed on filter paper which eases the accessibility and increases the number of laboratories that can make use of the material. Thus, an increasing number of tests (DNA, hormone, antibody, enzyme-analyses) can be performed on dried blood spots placed on filter paper.

3. Time of delivery: Cord blood specimens will be taken and processed as described for the

specimen taken in week 26.

4. Day 5: As part of a national screening programme for PKU, dried blood spots on filter paper are forwarded to Statens Seruminstitut on all newborns (day 5). This specimen is not directly part of the biological bank but routines will be established to ensure easy access to this resource. All specimens are registered by the mother's CPR-number which allows for identification.

The costs associated with the future management of the biological bank will be covered by the specific projects and Statens Seruminstitut.

Health registers to be used in the project

The existence of many population-based health registers in Denmark is critical to the success of the proposed study. In the following, the most important of these are mentioned:

The Medical Birth Registry

Started in 1973 and covers all live and stillborn in Denmark. It is based upon specific forms which are filled in by the midwives and sent to the National Board of Health. The register covers hospital births (more than 95%) and home births.

Place of birth, parents' age, residency and occupation as well as birthweight, gestational age, diseases in pregnancy, complications at delivery, congenital malformations, etc. are routinely recorded in the Medical Birth Registry or the National Hospital Registry.

Work by Barker and others has shown that the placental weight (especially in combination with birthweight) is a powerful predictor of subsequent diseases. The same may be true for liver size at birth, which is estimated by measuring abdominal circumference. Placental weight, abdominal as well as head circumference will be recorded routinely in the Medical Birth Registry according to our agreements with the National Board of Health as long as needed for the cohort.

The National Hospital Registry

Started in 1977 and has computerised data on all in and outpatients (from 1995) indicating diagnoses, surgery, hospital codes, etc. The register includes data on spontaneous and induced abortions.

Private hospitals are few and small. They will report to the registry from 1996. The register covers somatic and psychiatric patients.

The Danish Cancer Registry
Started in 1943 and includes all diagnosed cancer patients in Denmark, also deceased only diagnosed at autopsy.

The Registry of Causes of Death
Started as a computerised register in 1970. Covers all deaths including stillbirths. Includes also data from autopsies.

Cerebral Palsy Registry
Started in 1925 and covers all such patients in the country.

More than a dozen other disease specific registers exist.

Exposure data from routine registers
Most existing registers to be used in the study are of outcome type, but some of the population registers could provide useful information. Residency is recorded, together with the type and size of the house or flat. Job title, supply of drinking water, unemployment, etc. are also recorded in computerised registers.

Ethical considerations and data protection

The study will serve only research interests and will pose no direct risk to the participants. The study is based upon written informed consent and all possible safeguards against unwanted disclosure will be taken. It will be possible, at all times, to withdraw data from the research file for those who may subsequently regret their consent. Data will be safely stored with no direct possibility for personal identification for people without access to the link file which will be kept in a separate, locked room.

All data are collected and analysed with the sole aim of providing information on the causes of diseases which may be related to exposures in pregnancy or early childhood. Data are stored in ways ensuring maximum security and reducing all possible risks of accidental leak of individual information. Rules stated by Good Epidemiological Practice will be followed and the regulations for storing and the use of data will be given by the Data Protection Agency.

The project has been approved by the Ethical Committees in Denmark and is supported by the National Board of Health and a number of other health care organisations.

All research protocols using data from the study will be submitted to the scientific

ethical committees for approval and the committee may, at any time, ask for a renewal of the informed consent.

Data registration

All data units (questionnaires, blood samples, analyses, etc.) are given a unique ID-number. Data are stored in a database with good facilities for adding longitudinal data. At regular time intervals, the cohort will be linked with the National Hospital Registry, the Danish Cancer Registry and the Registry of Causes of Death.

Power considerations

It is estimated that a study of 100,000 pregnancies will lead to around 10,000 hospital registered spontaneous abortions, 470 stillbirths, 750 infant deaths (including 100 sudden infant death syndromes). Approximately 2,000 children will be born with congenital malformations. Over a 30-year period, 500 will develop schizophrenia, 900 cancer (including 200 leukemias/lymphomas), and 300 will be expected to develop cardiovascular diseases.

As illustrated by the enclosed list of proposed studies linked to the core project, the new exposure research registry is providing data to many subprojects. The following table illustrates that studies of very rare outcomes and/or caused by infrequent exposures have rather low statistical power, especially if the statistical association is low. Most other studies have good power and many studies need not be based upon the entire cohort. Explorative analyses will be restricted to no more than half of the cohort and statistical associations emerging from this exercise may then be closer scrutinised in the remaining data.

Power values according to the exposed fraction in the cohort, the estimated outcome frequencies (test level 0.05) and the underlying relative risk (RR)

Exposure fraction	RR	Outcome frequency 0.5%	1.0%	1.5%
1%	1.5	12%	20%	28%
	3.0	41%	69%	85%
	4.5	56%	84%	96%
5%	1.5	42%	69%	86%
	3.0	97%	100%	100%
	4.5	99%	100%	100%
10%	1.5	68%	93%	99%
	3.0	100%	100%	100%
	4.5	100%	100%	100%

Budget

The Danish Epidemiology Science Centre covers expenses to the scientific staff and some of the technical staff (secretaries and programmers). The Danish Epidemiology Science Centre furthermore provides facilities for the data collection and scientific supervision of the subprojects associated with the study.

Local health authorities have kindly promised to cover the expenses associated with sampling blood to the biological bank. Costs are indicated in units of DKK 1,000 (approx. USD 180).

Prenatal data collection	
15 full-time interviewers in 2.5 years	6,375
Telephone costs	1,800
Equipment (telephones, PC's, etc.)	100
Printing, mailing related to enrolment	1,400
Total	9,675
Postnatal data collection	
Questionnaire at birth:	
Printing 100,000 x 3.35 DKK each	335
Mailing	4
Other expenses	33
Optical reading	813
Total	1,185
18-month telephone interviewing	
Salary	2,550
Phone costs	600
Other expenses	10
Mailing, printing	450
Total	3,610
Storing of biological material	
Salary	2,741
Rent of laboratory space	507
Major equipment	1,790
Other expenses	4,261
Total	9,299
Grand total	23,769

Time schedule

Done	development of protocols
	recruitment of staff
	constructing questionnaires and registration forms
	consult experts and decision makers
	application for permissions
August 1995 - May 96	pilot testing of logistic procedures
	response rate
	questionnaire reliability
May 1996	adjusting the protocol and data collection forms
July 1996-	establishing the cohort
96 - 99	data collection

List of studies within the National Birth Cohort

The study aims at establishing the necessary data for research in the field of reproductive epidemiology for future generations of epidemiologists. We benefit from far-sighted colleagues who established disease registers such as the Danish Cancer Registry decades ago. In like manner, we should take responsibility for establishing the necessary infrastructure for research in the future. The value of the data will increase with time and it is the nature of research that we cannot foresee how the data will be best used in the future. But unlike most of the 'cancer cohorts', several important studies can be done as soon as the children are born and in their early childhood.

Protocols which are written (in Danish or English)

- Dietary marine n-3 fatty acids during pregnancy and fetal brain development: A substudy of the 100,000 Danish National Birth Cohort project
- The association between smoking during pregnancy, childhood development and reproductive failures
- Alcohol intake in pregnancy / impact of binge drinking and different types of alcoholic beverages
- Preventable birth defects
- Pharmacovigilance in pregnancy
- Work environment and reproductive health. A study based upon the National Birth Cohort
- Gene-environment interaction in facial clefts
- Fever in pregnancy and the risk of second trimester miscarriage and extremely preterm delivery
- Dietary marine n-3 fatty acids during pregnancy in relation to preterm delivery, fetal growth retardation, preeclampsia, and associated complications.
- Incidence and risk factors for 2nd trimester abortion and extreme, premature delivery.
- TGFA interaction in the etiology of facial clefts.
- Screening for short and long-term reproductive side-effects of ovulation-induced treatment of infertility in women.
- Psychosocial stress in pregnancy and rare adverse pregnancy outcome.
- Maternal diet, course of pregnancy, and offspring health.
- Parvovirus B19-infection and pregnancy outcome.

Protocols in progress

- Prenatal exposures associated with postterm delivery.
- The association between exposures in pregnancy to heavy metals and pesticides and early childhood development.
- Prenatal exposures and the occurrence of asthma and allergy.
- Prenatal exposures and colic in early childhood.
- Low levels of mannan binding protein (MBP) as a risk factor of serious infections in childhood.
- Critical studies of 'the programming hypothesis'.
- Intrauterine viral exposures and acute lymphoblastic leukaemia.
- Intrauterine exposures and juvenile type/diabetes.
- Intrauterine viral infections and subsequent hearing and vision defects.
- How is BMI and fat distribution related to fecundity and pregnancy outcome?
- Congenital abnormalities of the reproductive system including hypospadias and chryptorchibism and their relationship to maternal exposure to hormones.
- The role of motherhood in Denmark before the millennial change. An interdisciplinary project related to *Better health for mother and child.*
- Preventive research concerning schizophrenia and cognitive disorders.
- Postnatal nutrition and subsequent growth of body compartments.
- Non-specific effects of immunization in a population based study.
- Relations between pre and postnatal exposure factors and neuropsychological outcome profiles.
- Standardization of a new screening instrument (Bayley Infant Neurodevelopmental Screener (BINS)) in a Danish cohort and evaluation of its predictive power for later neuropsychological dysfunction.
- Molecular, pathophysiological and behavioral investigation of children with attention deficit and hyperactivity disorder (ADHD).
- The importance of family network for development in early childhood.
- Evaluation of The Health Nurse Service during infancy.