Cohort profile: the DANish LIFE course (DANLIFE) cohort, a prospective register-based cohort of all children born in Denmark since 1980

Jessica Bengtsson, Nadya Dich, Andreas Rieckmann, Naja Hulvej Rod

ABSTRACT

Purpose The DANish LIFE course (DANLIFE) cohort is a prospective register-based study set up to investigate the complex life course mechanisms linking childhood adversities to health and well-being in childhood, adolescence and young adulthood including cumulative and synergistic actions and potentially sensitive periods in relation to health outcomes.

Participants All children born in Denmark in 1980 or thereafter have successfully been included in the cohort totalling more than 2.2 million children. To date, the study population has been followed annually in the nationwide Danish registers for an average of 16.8 years with full data coverage in the entire follow-up period. The information is currently updated until 2015.

Findings to date DANLIFE provides information on a wide range of family-related childhood adversities (eg, parental separation, death of a parent or sibling, economic disadvantage) with important psychosocial implications for health and well-being in childhood, adolescence and young adulthood. Measurement of covariates indicating demographic (eg, age, sex), social (eg, parental education) and health-related factors (eg, birth weight) has also been included from the nationwide registers. In this cohort profile, we provide an overview of the childhood adversities and covariates included in DANLIFE. We also demonstrate that there is a clear social gradient in the exposure to childhood adversities confirming clustering of adverse experiences within individuals.

Future plans DANLIFE provides a valuable platform for research into early life adversity and opens unique possibilities for testing new research ideas on how childhood adversities affect health across the life course.

INTRODUCTION

The fact that many adulthood diseases have roots in earlier life is by now quite well documented and hardly debatable. The life course approach to studying health, which has increasingly been gaining attention in the past two decades, acknowledges the long-term biological, psychosocial and behavioural processes that operate across the life span, linking earlier life exposures to health in later life. These processes are extremely complex and far from being fully elucidated, but a growing amount of evidence suggests that in early life, when the brain and the physiological systems are not fully mature, exposure to psychosocial adversities may be particularly detrimental for future health and developmental trajectories. Advances in neuroscience, molecular biology and epigenetics have put forward an intriguing hypothesis that adverse events experienced early in life, when the human stress response system is still under development, may produce long-lasting alterations in the way the brain and the body respond to stress. These alterations, which can be traced not only on the behavioural but also on the epigenetic level, are, on one hand, an adaptation to a stressful environment; on the other hand, they may have long-lasting costs to future mental health and the normal regulation of the metabolic and cardiovascular systems, leading to physical health consequences.
Given that childhood may constitute a sensitive period for a whole range of exposures, including the exposure to stress, the timing of these exposures is central to life course epidemiology. Furthermore, researchers in child development have long argued that while exposure to a single risk factor hardly ever causes enduring harm, experience of multiple risk factors often has consequences for health. Thus, research of adversity needs to account for the accumulation of stressors from different domains as well and across the life span. The empirical testing of the hypotheses regarding the exact timing of exposures as well as the cumulative and interactive effects of stressors requires large data materials with frequent repeated measures over time. The DANish LIFE cohort is a large-scale repeated-measure dataset established to provide information on how childhood adversities affect health across the life course.

There exist a number of birth cohorts and other longitudinal studies with rich repeated information, which provide a valuable platform for life course research on early life adversities and health. The DANLIFE cohort is different from the existing datasets in several key ways and thus provides new opportunities within this research area. DANLIFE utilises data from nationwide registers covering all children born in Denmark since 1980. Because the data are exclusively register-based, the DANLIFE sample is unselected, thus solving the problem of selection bias that observational studies are prone to. Furthermore, using the national registers ensures a very large sample size. The population covered by DANLIFE totals more than 2.2 million children. Thus, the cohort provides strong statistical power needed to test hypotheses involving complex interactions.

Using register-only information has its limitation in that the psychosocial measures available are not very detailed. However, as we explain below, due to the unique Civil Registration System (CRS) available in Denmark, it is also possible to link DANLIFE to other existing (non-register-based) studies of Danish participants or invite DANLIFE participants for more assessments. All register-based data on major life events, social adversity and health outcomes are routinely collected and continuously updated for all children included in the cohort. The precise recording of timing of the information in DANLIFE enables the assessment of temporality between exposures, outcomes and covariates as well as identification of potentially critical and sensitive periods of exposure to childhood adversities. The cohort was specifically set up to test hypotheses on complex life course mechanisms including:

a. Investigating whether specific childhood adversities act independently, cumulatively or synergistically on clinical health outcomes across life stages.

b. Identifying potentially sensitive periods (eg, infancy, childhood, adolescence, early adulthood) for childhood adversities in relation to health outcomes.

DANLIFE is placed at a secure server at Statistics Denmark. Access to Danish administrative and health registers is granted by Statistics Denmark and The Danish Health Data Authorities. Data are provided for research purposes in an anonymous and secure form.

COHORT DESCRIPTION

Study population

All children born in Denmark in 1980 or thereafter to mothers who had a Civil Personal Registration (CPR) number at the time of birth have been identified in the Danish CRS and have successively been included in DANLIFE. The CPR number is a unique 10-digit number for all Danish residents who become members of the Danish Labour Market Supplementary Pension Fund or pay taxes in Denmark. The CPR number permits exact linkage on individual level between national administrative, clinical and health research registries in Denmark and thus provides an opportunity for country-wide population-based studies on public health matters by covering information such as drug prescriptions, hospitalisations, income and employment (figure 1). Some register-based information can have missing data, for example, children of immigrated parents will have missing data on parental education if their parents completed their education abroad. Since the data from CRS provided to DANLIFE are updated annually and include the population on the 1st of January each year, children who are live-born but die or emigrate within the same calendar year are not included in the data from CRS. Children who died or emigrated in their first year of life were instead identified in the Danish Medical Birth Registry (MBR) (n=13756). MBR includes all children born in Denmark to mothers with residence in Denmark at the time of birth. Immigrants not born in Denmark are not included in DANLIFE because no information on childhood adversities is available for immigrants prior to immigration. Valid and continuously updated information on many socioeconomic, demographic and health-related factors from the Danish nationwide registers was provided by Statistics Denmark and the Danish Health Data Authorities to DANLIFE.

The study participants have been followed until emigration, death or the end of follow-up. The information is currently updated until the 31st of December 2015, but we will continue with annual updates when the data become available. The current version of DANLIFE includes 2,223,927 individuals followed for an average of 16.8 years corresponding to more than 37 million person-years (figure 2). Individuals emigrating during follow-up (n=164,348; 7.4%) are censored at the date of emigration and are not re-entered into the cohort if ever returning to Denmark, since there would be an information gap in the period the person spent outside of Denmark. Censoring on date of death was possible using information from CRS (n=20,514; 0.9%). Studies based on Danish registers do not require informed consent or involvement by
the population nor is an ethical approval by the Danish National Committee on Health Research Ethics required.

**Patient and public involvement**

Patients or the public were not involved in the design of this register-based cohort.

**Study population: DANLIFE**

**Family linkage**

The CPR numbers of the parents in DANLIFE have been retrieved from the child’s first year of life from CRS supplemented by information from MBR. The parents registered in CRS are the legal parents, and biological parents can therefore not be distinguished from adoptive parents. However, the number of children who are born in Denmark and adopted is approximately 1%. The parents registered in MBR are assumed to be the biological parents. The CPR numbers of the parents were used to identify siblings as persons linked to the same mother and father and to identify twins as persons born to the same mother on the same day (+/-1 day). The linkage to parents and siblings in DANLIFE enables identification of exposure to a range of family-related childhood adversities, such as severe illness in the family. The family linkage also enables identification of family history of specific diseases through linkage to the Danish National Patient Registry. Children with missing information on both parents in both CRS and MBR were not included in the DANLIFE cohort (n=3103), since it would be impossible to identify...
the siblings of these children as well as the exposure to most of the family-related childhood adversities.

Data sources
The registers combined in DANLIFE are updated for research purposes once a year (figure 1). Information comes from nine registers: CRS captures sex, date of birth, date of death, family identifier, emigration, nationality and parental separation;16 MBR captures birth-related information;16 the Integrated Database for Labour Market Research captures parental long-term unemployment;21 the Register of Support for Children and Adolescents captures foster care; the Income Statistics Register captures family poverty (only from 1987 and onwards);22 the Population Education Register captures parental education;23 the Danish National Prescription Registry captures drug prescriptions,24 which we use to identify parental alcohol and drug abuse; the Danish Psychiatric Central Research Register captures admissions due to psychiatric diagnoses,25 which we use to identify parental and sibling psychiatric illness and parental alcohol and drug abuse; and the Danish National Patient Register captures hospital admissions with diagnose codes,18 which we use to identify parental and sibling somatic illness and parental alcohol and drug abuse. To date, DANLIFE includes information from the 1st of January 1980 until the 31st of December 2015 and will continue to be updated as children are born and registered. The continuity of the Danish registries provides full data coverage in the entire follow-up period, except for information on family poverty, which is based on the household equivalised disposable income, which was only available from 1987 onwards. Most factors have been specified with an exact date such as date of birth, emigration, death, highest attained education and all somatic and psychiatric diagnoses registered in the National Patient Registry.18 Information on personal income and transfer payments,21 labour market affiliation21 and parental separation14 are registered across a given calendar year, and the timing of these adversities is therefore set to the year of occurrence (table 1). Since the objective of DANLIFE is to enable assessment of the effects of adversities experienced in childhood on health and well-being, we restricted the exposure period to 0–18 years of age. The exposure period can easily be altered to accommodate research projects with alternative objectives.

Childhood adversities and covariates
The uniqueness of DANLIFE lies within the definition and construction (eg, death of siblings needs to be constructed using a family identifier) of measures of selected childhood adversities with important psychosocial implications for health and well-being in childhood, adolescence and young adulthood. In childhood, the family environment plays a crucial role for development and well-being. At the same time, a straining family environment, and the social circumstances in which it takes place, may be major sources of stress in children.8 26–29 Two aspects of family environment that have been shown to be important sources of stress in children are unstable family dynamics with lack of responsive caregiving and loss or the threat of loss within the family.22 23 Material deprivation in the family is also a source of stress in children as it may affect parenting skills and family climate, as well as the availability of social and material resources, such as high-quality housing, proximity of good schools and recreation areas.34–36 The linkage between child, parents and siblings in DANLIFE enables measurement of a range of childhood adversities that are likely to affect the family dynamics and quality of caregiving (ie, parental separation, being placed in foster care, parental psychiatric illness, sibling psychiatric illness and parental alcohol or drug abuse); indicate loss or the threat of loss, within the family (ie, parental somatic illness, sibling somatic illness and death of a parent or a sibling) and material deprivation (ie, family poverty and parental long-term unemployment). Table 1 provides an overview of the adversities included in DANLIFE.

An important and growing field within the life course framework is the assessment of mediation and interaction.1 The following covariates for confounding, mediation and interaction assessment have been incorporated in DANLIFE so far: sex (male, female), age at any given time (based on date of birth), birth order, birth weight (in grams), parental age at any given time (based on date of birth), ethnicity (Danish, non-Danish if at least one parent has a nationality other than Danish) and parental highest attained education at the time of birth (low<9 years, ie, mandatory education in Denmark; middle 10–12 years, ie, youth education and vocational education and high ≥12 years). In Denmark, ≤9 years of schooling is equivalent to basic education mandatory by law, 10–12 years of schooling is equivalent to high school in the USA and ≥12 years of schooling is any additional schooling (eg, university). The definitions of the covariates can be altered to accommodate specific research objectives. In principle, the CPR number enables the inclusion of information on exposures, outcomes and covariates from any Danish register or any other study on Danish participants, which has recorded CPR numbers, into the DANLIFE cohort given granted permission from the authorities responsible for data security. An overview of the nationwide registers that have been linked in DANLIFE and which specific adversities and covariates each register has provided information on is shown in figure 1. Information from several registers has been combined to define the adversities representing parental alcohol and drug abuse (table 1 and figure 1).

FINDINGS TO DATE
DANLIFE is a newly established dataset, and no results based on its data have therefore been published to date but several studies investigating the effect of
childhood adversities on health outcomes such as cardiovascular disease, type 1 diabetes and all-cause mortality are ongoing. Here, we present findings regarding the prevalence of adversities among Danes and the social gradient in the exposure to adversities. These findings are summarised in Table 2. Table 2 gives an overview of the exposure to the specific childhood adversities in the total study population of more than 2.2 million children. Parental separation was the most frequently experienced adversity during follow-up (29%) followed by parental long-term unemployment (25%). In this large study population, several thousand had experienced even the rarest adversities (ie, sibling death and sibling psychiatric illness). Table 2 further shows a clear social gradient in the exposure to childhood adversities in DANLIFE. The proportions

### Table 1 Definition and timing of the childhood adversities included in the DANish LIFE course (DANLIFE) cohort

<table>
<thead>
<tr>
<th>Adversity</th>
<th>Definition</th>
<th>Date</th>
</tr>
</thead>
<tbody>
<tr>
<td>Foster care</td>
<td>Being placed in out-of-home care</td>
<td>Date of placement</td>
</tr>
<tr>
<td>Parental psychiatric illness</td>
<td>A parent's admission for at least 1 day with an International Classification of Diseases (ICD) version 8 (ICD-8) or version 10 (ICD-10) diagnostic code related to psychiatric illness. ICD-8 codes: psychoses (290–299 except 291); neuroses, personality disorders and other nonpsychotic mental disorders (300–309 except 303–304); and mental retardation (310–315); ICD-10 codes: mental, behavioural and neurodevelopmental disorders (F00–F99)</td>
<td>Date of diagnosis</td>
</tr>
<tr>
<td>Parental somatic illness</td>
<td>A parent diagnosed with one of the ICD-8 codes included in the Charlson comorbidity index in the period 1980–1993 or the ICD-10 codes included in the updated version of the Charlson comorbidity index in the period 1994–2015</td>
<td>Date of diagnosis</td>
</tr>
<tr>
<td>Parental death</td>
<td>Death of a parent</td>
<td>Date of death</td>
</tr>
<tr>
<td>Sibling psychiatric illness</td>
<td>A sibling’s admission for at least 1 day with an ICD-8 or ICD-10 code related to psychiatric illness. ICD-8 codes: psychoses (290–299); neuroses, personality disorders and other nonpsychotic mental disorders (300–309); and mental retardation (310–315); ICD-10 codes: mental, behavioural and neurodevelopmental disorders (F00–F99)</td>
<td>Date of diagnosis</td>
</tr>
<tr>
<td>Sibling somatic illness</td>
<td>A sibling diagnosed with one of the following somatic illnesses related to mortality in children aged 0–18 years (ICD-8/ICD-10 codes): malignant neoplasm (140–199/C00–C96); congenital anomalies of the heart and circulatory system (746–747/Q20–Q28); congenital anomalies of the nervous system (743/Q00–Q07); cerebral palsy (343–344/G80–G83); epilepsy (345/G40–G41); cardiomyopathy (425/I42–I43); congenital disorders of lipid metabolism (272/E75)</td>
<td>Date of diagnosis</td>
</tr>
<tr>
<td>Sibling death</td>
<td>Death of a sibling</td>
<td>Date of death</td>
</tr>
<tr>
<td>Parental alcohol abuse</td>
<td>A parent diagnosed with an illness related to alcohol abuse or receiving a prescription of a drug used in treatment of alcohol addiction. ICD-8 codes: alcoholic psychosis (291); alcoholism (303); alcoholic cirrhosis of the liver (571.09); alcoholic steatosis of the liver (571.10); ICD-10 codes: alcohol psychosis and abuse syndrome (F10); alcoholic polyneuropathy (G62.1); alcoholic cardiomyopathy (I42.6); alcoholic-induced acute (K85.2) and chronic (K86.0) pancreatitis; alcoholic liver disease (K70); alcoholic gastritis (K29.2); Anatomical Therapeutic Chemical (ATC) Classification System codes: drugs used in alcohol dependence (N07BB)</td>
<td>Date of diagnosis/prescription</td>
</tr>
<tr>
<td>Parental drug abuse</td>
<td>A parent diagnosed with an illness related to drug abuse or receiving a prescription of a drug used in treatment of drug addiction. ICD-8 codes: drug dependence (304); ICD-10 codes: opioids (F11); cannabinoids (F12); sedatives/hypnotics (F13); cocaine (F14); other stimulants (F15); hallucinogens (F16); other and multiple drugs (F18–F19); ATC codes: drugs used in opioid dependence (N07BC)</td>
<td>Date of diagnosis/prescription</td>
</tr>
<tr>
<td>Parental separation</td>
<td>Separation of the parents</td>
<td>Year of separation</td>
</tr>
<tr>
<td>Parental long-term unemployment</td>
<td>A parent being unemployed for at least 12 months within two consecutive years</td>
<td>First year of unemployment</td>
</tr>
<tr>
<td>Poverty</td>
<td>Family income below 50% of the median national family income in a given year, in three consecutive years</td>
<td>Third consecutive year of poverty</td>
</tr>
</tbody>
</table>
of children exposed to the specific childhood adversities are consistently increasing with decreasing levels of maternal education measured at the time of birth. The gradient is strong for the adversities reflecting family poverty, parental unemployment, and parental alcohol and drug abuse. The social gradient is also pronounced for parental separation where as many as 40% of the children born to mothers with a low level of education have experienced separation compared with only 19% of the children born to mothers with a high level of education. The social gradient is evident for all childhood adversities included in the cohort.

The accumulation of childhood adversities in DANLIFE is presented in table 3. Almost half of the study population did not experience any childhood adversities during follow-up and almost 10% had experienced three or more adversities. The well-known clustering of social disadvantage within individuals can also be seen in table 3. More than 60% of the children born to mothers with a high level of education at the time of birth have not experienced any adversities during follow-up compared with only 25% of the children born to mothers with a low level of education. Accordingly, only 3.5% of the children born to highly educated mothers have experienced three or more adversities.
adversities compared with 20% of the children born to mothers with a low educational level. This is also reflected in the average number of adversities experienced during follow-up among children born to mothers with a high (0.5), middle (0.9) and low (1.5) level of education.

STRENGTHS AND LIMITATIONS
The main strengths of DANLIFE include the unselected sample and the large sample size. The unselected sample ensures unbiased estimates of the effects of adversity on clinical outcomes, and the large sample size and linkage to medical registers give the possibility to study very rare health outcomes, including possibilities for interaction analyses, which tend to require a lot of statistical power. Yearly updated information provides higher time resolution than most available birth cohorts and allows for estimation of exposure trajectories (eg, moving in and out of poverty) and rather precise timing of exposures, which is crucial for investigation of critical and sensitive periods. Finally, information about exposures and outcomes that come from registers is more objective than questionnaire-based data, as they are not subject to recall bias and other types of biases that are associated with self-reporting (eg, participants’ personality affecting their over-reporting of symptoms).37

The DANLIFE sample is still very young, with a maximum age of 36 years, and therefore quite healthy. Even though many diseases will not be manifested until a later age, we still observe more than 20 000 deaths and a considerable number of clinical outcomes such as cardiovascular disease, type 1 diabetes and subtypes of cancer in the DANLIFE sample. Thus, the large sample size provides a unique opportunity to evaluate effects of childhood adversities on clinical health outcomes in early adult life, which is a severely understudied area due to lack of statistical power and few cases in previous life course studies.

On the other hand, defining exposures and outcomes based on registers has important trade-offs. With respect to outcomes, only clinical diagnoses are available for investigation. This for example means that only severe mental health problems resulting in hospital admission will be captured in the data. Similarly, many relevant exposures related to family dynamics and sensitive care cannot be measured using registers. Furthermore, those measures that can be defined via registers are crude and may be vulnerable to misclassification. To illustrate, the effects of parental separation on children’s well-being may be modified by a large number of circumstances, such as the extent to which parents are able to collaborate postdivorce and the shared custody arrangement. It may also be important whether parents enter new relationships once their marriage is dissolved. This information can only to a very limited degree be obtained via registers. As an example of possible misclassification, parental psychiatric illness, which may be severe enough to affect the family functioning, but not severe enough to result in hospitalisation, will not be captured in registers. However, the limitations concerning crude or incomplete measures can be overcome if DANLIFE data are combined with more precise questionnaire-based measures: because the CPR number is available for all participants, it is possible to link DANLIFE to other, non-register-based cohort studies of Danish participants or to invite a subsample of those who are in the DANLIFE cohort for a more detailed assessment. Thus, DANLIFE provides a valuable platform for research into early life adversity and opens a lot of unique possibilities for testing new research ideas.

Acknowledgements We acknowledge the researchers in the Psychosocial Epidemiology Research Group at the Department of Public Health, University of Copenhagen, who have been involved in discussing the theoretical framework of the DANLIFE cohort. We thank professor Ulrik Becker at the National Institute of Public Health, University of Southern Denmark, for valuable advice on the use of Danish registers regarding alcohol and drug misuse.

Collaborators Access to the Danish register data must be granted by Statistics Denmark and The Danish Health Data Authorities and is therefore not open access. Access to the established DANLIFE cohort is available to other investigators through collaborative agreements and a secured access. Please contact Professor Naja Hulvej Rod (nahu@sun.ku.dk) for further information.

Contributors JB, ND and NHR were involved in the conception and design of the study. JB and AR were responsible for data management and analyses. All authors contributed to the writing of the manuscript and have critically revised and approved the final version.

Funding This work was supported by The Innovation Fund Denmark (grant number 5189-00038) and Helsefonden (grant number 17-B-0102).

Competing interests None declared.

Patient consent for publication Not required.

Provenance and peer review Not commissioned; externally peer reviewed.

Data sharing statement Access to DANLIFE is available through collaborative agreements and granted access to the Danish registers by Statistics Denmark and The Danish Health Data Authorities. Please contact Professor Naja Hulvej Rod (nahu@sun.ku.dk) for further information.

Open access This is an open access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited, appropriate credit is given, any changes made indicated, and the use is non-commercial. See: http://creativecommons.org/licenses/by-nc/4.0/.

REFERENCES