**BMJ Open** National, clinical cohort study of late effects among survivors of acute lymphoblastic leukaemia: the ALL-STAR study protocol

Liv Andrés- Jensen,1 Mette Tiedemann Skipper,2 Kristian Mielke Christensen,1 Pia Hedegaard Johnsen,3 Katrine Aagaard Myhr,4 Martin Kaj Fridh,1 Kathrine Grell,1,5 A. M. L. Pedersen,6 Sune Leisgaard Mørck Rubak,2 Martin Ballegaard,7,8 Arne Hørlyck,9 Rikke Beck Jensen,10 Trine-Lise Lambine,11 Kim Gjerum Nielsen,1,8 Ruta Tuckuviene,12 Peder Skov Wehner,13 Birgitte Klug Albertsen,2,14 Kjeld Schmiegelow,1,8 Thomas Leth Frandsen1

**ABSTRACT**

**Introduction** More than 90% of patients diagnosed with childhood acute lymphoblastic leukaemia (ALL) today will survive. However, half of the survivors are expected to experience therapy-related chronic or late occurring adverse effects, reducing quality of life. Insight into underlying risk trajectories is warranted. This is established a Nordic, national childhood ALL survivor cohort, to be investigated for the total somatic and psychosocial treatment-related burden as well as associated risk factors, allowing subsequent linkage to nation-wide public health registers.

**Methods and analysis** This population-based observational cohort study includes clinical follow-up of a retrospective childhood ALL survivor cohort (n=475), treated according to a common Nordic ALL protocol during 2008–2018 in Denmark. The study includes matched controls. Primary endpoints are the cumulative incidence and cumulative burden of 197 health conditions, assessed through self-report and proxy-report questionnaires, medical chart validation, and clinical examinations. Secondary endpoints include organ-specific outcome, including cardiovascular and pulmonary function, physical performance, neuropathy, metabolic disturbances, hepatic and pancreatic function, bone health, oral and dental health, kidney function, puberty and fertility, fatigue, and psychosocial outcome. Therapy exposure, acute toxicities, and host genome variants are explored as risk factors.

**Ethics and dissemination** The study is approved by the Regional Ethics Committee for the Capital Region in Denmark (H-18035090/H-20006359) and by the Danish Data Protection Agency (VD-2018–519). Results will be published in peer-reviewed journals and are expected to guide interventions that will ameliorate the burden of therapy without compromising the chance of cure.

**INTRODUCTION**

Five-year overall survival rates for childhood acute lymphoblastic leukaemia (ALL) has climbed from less than 10% to above 90% during the last seven decades as a result of chemotherapy intensification, refined risk stratification and improved supportive care.1,3,4 However, cure comes at a cost.4 All patients experience acute, but transient toxicities during therapy (eg, infections, mucositis and peripheral neuropathy), and half of the patients are burdened by at least one severe, potentially life-threatening organ toxicity, such as acute pancreatitis, thromboembolism, or severe neurotoxicity.3–8 Acute organ damage resulting from leukaemia and/or therapy may persist (eg, insulin-dependent
Diabetes or emerge (eg, osteonecrosis) several years into survival as late effects. The 30-year-old survivor will have experienced an average of 5.4 health conditions, and half of survivors face at least one chronic health condition 20 years from ALL diagnosis, both measures significantly exceeding those among siblings and community controls. This disproportion in morbidity rate seems to amplify with increasing age. Certain therapeutic exposures have been reduced over time (eg, reduction in cardiotoxic anthracycline exposure and limited use of cranial irradiation in first-line therapy protocols), which has decreased the risk of early mortality and lowered the prevalence of treatment-related cancers, severe cognitive deficits, hypothalamic-pituitary dysfunction and immunological disease. However, the number of health conditions per individual survivor is still double of that in controls, now dominated by endocrine system disorders, and musculoskeletal, neurological and cardiovascular conditions, maintaining risk of impaired physical, cognitive and emotional functions, socioeconomic achievements and health-related quality of life. Therapy exposure and genetic susceptibility are considered the two major risk factors predicting late morbidity. Risk of specific late effects are associated with cumulative chemotherapy exposures but no safe lower doses have been established. Specific genetic variants have been linked with survivor obesity, reduced bone mineral density (BMD), cardiotoxicity, skeletal muscle dysfunction and neurocognitive dysfunction; however, clinical recommendations await harmonisation of outcomes, larger sample sizes and replication in independent cohorts.

To date, many late effects studies are observational and focus on one or few organ systems, cross-sectional in design, or emerge from single centres with resulting risk of selection bias and limited sample power. A few very large (>10 000 patients), multi-institutional childhood cancer survivor cohort studies exist, such as the US Childhood Cancer Survivor Study, the Nordic Adult Life after Childhood Cancer in Scandinavia study and the British Childhood Cancer Survivor Study. Most studies emerging from these cohorts rely on self-reported and register-based data, while few collect systematic clinical data, none of which include Nordic survivors.

During the 10-year period of 2008–2018, patients diagnosed with Philadelphia chromosome-negative (Ph−) ALL at age 1–45 years in the Nordic and Baltic countries, have been treated according to the Nordic Society of Paediatric Haematology and Oncology (NOPHO) ALL2008 protocol described elsewhere. Two thirds of this cohort were below 18 years of age at diagnosis. The aim of the Acute Lymphoblastic Leukemia Survivor Toxicity And Rehabilitation (ALL-STAR) study is to establish a national, Danish NOPHO ALL2008 survivor cohort, to be assessed for the total somatic and psychosocial treatment-related burden, using both self-report and systematic clinical data that will supplement data from the Danish nationwide public demographic, socioeconomic and health registries.

Methods
Study aims
Primary aim
To quantify the burden of health conditions occurring during the first decade of survivorship among survivors of childhood and young adult ALL.

Secondary aims
To describe, in detail, organ-specific health among survivors as compared with matched controls and normative data; and to investigate host genome variants, therapy exposure and acute toxicities as risk factors for organ-specific late effects and for the overall cumulative burden. Furthermore, to establish a platform for life-long follow-up in this survivor cohort, facilitating longitudinal studies and linkage to nationwide public, socioeconomic and health registries.

Study design
The ALL-STAR study is a national, Danish, population-based, observational cohort study, including prospective clinical follow-up in a retrospective childhood and young adult ALL survivor cohort. This present protocol primarily focuses on the initial, cross-sectional phase of the project.

Participants
Study population
Survivors treated in Denmark according to the NOPHO ALL2008 protocol and minimum 1 year from therapy cessation (3.5 years from ALL diagnosis) are eligible. This corresponds to all patients (n=475) aged 1–45 years at diagnosis with Ph− B-cell precursor or T-cell ALL, diagnosed between July 2008 and October 2018 (the complete NOPHO ALL2008 protocol period). A control group consisting of age- and sex-matched individuals (1:1) is recruited for the study. First-degree relatives to survivors and individuals with previous or ongoing cancer, previous or ongoing chemotherapy, and/or previous or ongoing radiation therapy are excluded as controls. Study recruitment for the initial cross-sectional studies opened in February 2019 and will terminate in September 2022, allowing for the entire survivor cohort to meet the inclusion criteria.

Recruitment
Eligible survivors are identified through the NOPHO ALL2008 registry. Contact information is obtained through medical charts and provided by treating physicians to the research team. Written information is sent to parents if survivor’s age is <15 (addressed to survivor and parents), to parents and the survivor if aged 15–17,9, and to survivors if aged ≥18 years. Contact is obtained through subsequent phone calls. All survivors and/or families are invited to an on-site meeting for further information.
Endpoints
Primary endpoints and hypothesis
The two primary endpoints are cumulative incidence and cumulative burden of 197 health conditions presented in table 1.

We hypothesise that the cumulative incidence and cumulative burden are significantly higher among survivors of childhood cancer and young adult ALL already during the first decade of survival, when compared with age- and sex-matched controls, and normative data.

The measure of cumulative burden considers occurrence of multiple health conditions and recurrent events, thereby encompassing the total disease burden more comprehensively than cumulative incidence, which includes only first occurrence of an event.40 41 The 197 conditions investigated in the ALL-STAR study are defined and graded according to the St Jude Children’s Research Hospital modification of the National Cancer Institute’s Common Terminology Criteria for Adverse Events (CTCAE; V.4.03),42 developed to optimise characterisation of long-term and late-onset health conditions among childhood cancer survivors.43 Ten additional health conditions (generalised muscle weakness, reduced cardiopulmonary fitness, sarcopenic obesity and seven oral/dental conditions) have been added. Definitions and grading criteria for all conditions are provided in online supplemental table 1.

Secondary endpoints
Secondary endpoints include: (1) organ-specific outcome, including metabolic syndrome (MetS) and body composition, hepatic function, pancreatic function, bone health, cardiac function, pulmonary function, peripheral and autonomic neuropathy, physical performance, renal function, pubertal timing, fertility, oral and dental health, and (2) psychosocial outcome, including emotional distress, fatigue and health-related quality of life.

An overview of the ALL-STAR study is provided in figure 1.

Data collection
Data are collected through questionnaires and clinical evaluations. An overview of ALL-STAR source data used for evaluation of the 197 health conditions is provided in online supplemental table 1.

Questionnaire data
An electronic questionnaire is sent to participants prior to clinical examinations. The questionnaire is adjusted according to self-report (age ≥15 years), proxy report (age <18), age, sex, and status as survivor or control. Items include demographics, use of healthcare services, medicine and nutritional supplements, medical history (including age at diagnosis of a confirmed condition), pubertal status and fertility, familial dispositions, health behaviour, mental health, quality of life and socioeconomic status. An overview of questionnaire versions and content is provided in online supplemental table 2.

Clinical data
All clinical examinations are performed at either Copenhagen University Hospital or Aarhus University Hospital, as requested by the participant. Examinations are performed during 10 hours in a single day. The examination programme and description of clinical investigators are provided in online supplemental table 3.

Evaluation of metabolic syndrome and body composition
Endpoints include MetS, insulin sensitivity, lipid profile, abdominal circumference, body mass index (BMI) for adults and BMI standard deviation (SD) scores for children based on national reference material,44 45 lean body mass, and fat mass. Adult MetS is defined according to the National Cholesterol Education Program Adult Treatment Panel criteria,46 and includes central obesity, dyslipidaemia, hypertension and elevated fasting glucose. Paediatric MetS is defined using the same measures.47 Questionnaire items include metabolic disease and family disposition, prescription medicine, nutritional supplements and lifestyle (diet, level of physical activity, and sedentary behaviour). Clinical investigations include anthropometrics (weight, height and abdominal circumference) and blood pressure (described in cardiovascular section). Fasting blood samples include glucose, glycated haemoglobin, insulin, proinsulin c-peptide, total cholesterol, high-density lipoprotein cholesterol, low-density lipoprotein cholesterol, very low-density lipoprotein cholesterol, triglycerides, glucagon and leptin. Homeostatic Model Assessment for Insulin resistance score is calculated as fasting plasma glucose (mmol/L)×fasting plasma insulin (μU/L)/22.5.38 Body composition (fat and lean body mass, and android/gynoid fat distribution) is assessed with dual-energy X-ray absorptiometry (DXA) (Lunar Prodigy, GE Healthcare, Wisconsin, USA, or Hologic Horizon A, Hologic, Marlborough, Massachusetts, USA), and is adjusted for sex and pubertal stage.

Evaluation of hepatic function and gallbladder
Endpoints include signs of hepatocellular damage, impaired liver synthesis, iron overload, hepatic fibrosis, portal hypertension and gall bladder disease (cholelithiasis). Questionnaire items include symptoms of hepatic disease (fatigue, abdominal pain, nausea, jaundice, discoloured urine and itchy skin), hepatic diagnoses, alcohol consumption and use of prescription medicine. Fasting blood samples include alanine transaminase, aspartate transaminase, alkaline phosphatase, gamma-glutamyl transferase, bilirubin, lactate dehydrogenase, prothrombin time, international normalised ratio, albumin, ferritin, transferrin saturation,
Table 1  Health conditions investigated in the ALL-ST AR study

<table>
<thead>
<tr>
<th>Auditory-hearing</th>
<th>Gastrointestinal</th>
<th>Infections</th>
<th>Neurological</th>
<th>Pulmonary</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cholesteatoma</td>
<td>Bowel perforation</td>
<td>Bronchial/fung infection C/R</td>
<td>Autonomic dysfunction</td>
<td>Asthma</td>
</tr>
<tr>
<td>Hearing loss</td>
<td>Coeliac disease</td>
<td>Endocarditis</td>
<td>Cavernoma</td>
<td>COPD</td>
</tr>
<tr>
<td>Tinnitus</td>
<td>Constipation</td>
<td>Gastrointestinal infection</td>
<td>Cerebellar dysfunction</td>
<td>Epistaxis, C/R</td>
</tr>
<tr>
<td>Vertigo</td>
<td>Dysphagia</td>
<td>Genitourinary infection</td>
<td>Cerebral necrosis</td>
<td>Obstructive sleep apnoea</td>
</tr>
<tr>
<td></td>
<td>Enterocolitis</td>
<td>Hepatitis B, chronic</td>
<td>Cerebrovascular accident</td>
<td>Obstructive ventilatory defect</td>
</tr>
<tr>
<td>Cardiovascular</td>
<td>Oesophageal varices</td>
<td>Hepatitis C, chronic</td>
<td>Cerebrovascular disease</td>
<td>Pleural space disorders</td>
</tr>
<tr>
<td></td>
<td></td>
<td>HIV infection</td>
<td>Cranial nerve disorder</td>
<td>Pneumomits</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Lymphatic infection</td>
<td>Dysarthria</td>
<td>Pulmonary diffusion defect</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Meningoencephalitis</td>
<td>Generalised muscle weakness</td>
<td>Pulmonary embolism</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Osteomyelitis</td>
<td>Headaches, C/R</td>
<td>Respiratory tract haemorrhage</td>
</tr>
<tr>
<td>Cardiovascular</td>
<td>Gastrointestinal fistulas</td>
<td>Otitis media, C/R</td>
<td>Hydrocephalus</td>
<td>Restrictive ventilatory defect</td>
</tr>
<tr>
<td></td>
<td>Gastrointestinal haemorrhage</td>
<td>Pelvic inflammatory disease</td>
<td>Hydroxyresinogomyelia</td>
<td>Tracheal aspiration</td>
</tr>
<tr>
<td></td>
<td>Gastrointestinal necrosis</td>
<td>Pharyngitis/tonsillitis, C/R</td>
<td>Intracranial haemorrhage</td>
<td>Tracheal stenosis</td>
</tr>
<tr>
<td></td>
<td>Gastrointestinal obstruction</td>
<td>Sinusitis, C/R</td>
<td>Movement disorders</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Gastrointestinal strictures</td>
<td>Soft tissue infection</td>
<td>Multiple sclerosis</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Gastrointestinal ulcer</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Heart valve disorder</td>
<td>Gastroaparesis syndrome</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>High total cholesterol</td>
<td>Malabsorption syndrome</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hypertension</td>
<td>Pancreatic insufficiency</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hypertiglycidaemia</td>
<td>Pancreatitis</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LV systolic dysfunction</td>
<td>Proctitis</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pericarditis</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Prolonged (QTc) interval</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pulmonary hypertension</td>
<td>Cholecytitis/cholelithiasis</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Raynaud phenomenon</td>
<td>Fibrosis/cirrhosis</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RV systolic dysfunction</td>
<td>Hepatic failure</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tachycardia, sinus</td>
<td>Hepatitis</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Thromboembolic event</td>
<td>Portal hypertension</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Vascular disease</td>
<td>Veno-occlusive disease</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Endocrine</td>
<td>Haematological</td>
<td>Oral/dental</td>
<td>Neoplasms</td>
<td>Reproductive/genital</td>
</tr>
<tr>
<td>Abnormal glucose metabolism</td>
<td>Anaemia</td>
<td>Dental caries</td>
<td>Benign neoplasms</td>
<td>Abnormal sperm concentration</td>
</tr>
<tr>
<td>Adrenal insufficiency</td>
<td>Coagulopathy</td>
<td>Dental erosion</td>
<td>Malignant neoplasms</td>
<td>Cervical dysplasia</td>
</tr>
<tr>
<td>Adult GH deficiency</td>
<td>Iron overload</td>
<td>Gingivitis</td>
<td></td>
<td>Dysfunctional uterine bleeding</td>
</tr>
<tr>
<td>Childhood GH deficiency</td>
<td>Neutropaenia</td>
<td>Periodontitis</td>
<td></td>
<td>Endometriosis</td>
</tr>
<tr>
<td>Diabetes insipidus</td>
<td>Polycythaemia</td>
<td>Salivary gland dysfunction</td>
<td>TMJ disorder</td>
<td>Erectile dysfunction</td>
</tr>
<tr>
<td>GH excess</td>
<td>Thrombocytopenia</td>
<td>TMJ disorder</td>
<td></td>
<td>Genitourinary adhesions</td>
</tr>
<tr>
<td>Hyperparathyroidism</td>
<td>Thrombocytosis</td>
<td></td>
<td></td>
<td>Hypogonadism, central</td>
</tr>
<tr>
<td>Hyperprolactinaemia</td>
<td></td>
<td></td>
<td></td>
<td>Leydig cell insufficiency</td>
</tr>
<tr>
<td>Hyperthyroidism</td>
<td>Autoimmune disorders</td>
<td></td>
<td></td>
<td>Polycystic ovarian syndrome</td>
</tr>
<tr>
<td>Hypoparathyroidism</td>
<td>Graft-versus-host disease</td>
<td></td>
<td></td>
<td>Precocious puberty</td>
</tr>
<tr>
<td>Hypothyroidism</td>
<td></td>
<td></td>
<td></td>
<td>Primary ovarian failure</td>
</tr>
<tr>
<td>Overweight/obesity</td>
<td>Immunodeficiency</td>
<td></td>
<td></td>
<td>Prostatic hypertrophy, benign</td>
</tr>
<tr>
<td>Sarcopenic obesity</td>
<td></td>
<td></td>
<td></td>
<td>Vaginal fistula</td>
</tr>
<tr>
<td>SIAHD</td>
<td></td>
<td></td>
<td></td>
<td>Vaginal stenosis</td>
</tr>
</tbody>
</table>

ALL-ST AR, Acute Lymphoblastic Leukemia Survivor Toxicity And Rehabilitation; COPD, chronic obstructive pulmonary disease; C/R, chronic/recurrent; GH, growth hormone; GORD, gastro-oesophageal reflux disease; LV, left ventricle; OD, right eye; OS, left eye; QTc, corrected QT interval; RV, right ventricle; SIAHD, syndrome of inappropriate antidiuretic hormone secretion; TMJ, temporomandibular joint.
alfa-1 antitrypsin, platelets and immunoglobulins. Serum is stored for evaluation of viral antibodies (hepatitis A, B and C, cytomegalovirus and Epstein-Barr virus) in case of elevated transaminases. Hepatic ultrasound is performed using a convex, abdominal probe (C1-5, GE Logiq E9 or E10, GE Healthcare, Chicago, USA). The liver, gallbladder, gall ducts, liver artery and spleen are evaluated for appearance and size. Portal vein flow is measured using Doppler ultrasound. Ultrasound 2D shear wave elastography is performed on the right liver lobe for evaluation of parenchymal stiffness and fibrosis.49

**Evaluation of pancreatic function**

Endpoints include pancreatic disease, endocrine and exocrine function, and morphological appearance. Questionnaire items include pancreatic symptoms and disease, diet and alcohol consumption. Fasting blood samples include pancreatic amylase, lipase, glucose, glycosylated haemoglobin, insulin and proinsulin c-peptide. Faecal elastase-1 is evaluated in survivors only and associated with occurrence of asparaginase-associated pancreatitis during ALL therapy. Pancreatic ultrasound is performed using a convex, abdominal probe (C1-5, GE Logiq E9 or E10, GE Healthcare, Chicago, USA) for anterior–posterior measurement of caput and corpus, and for evaluation of morphological changes (tumour, inflammation, oedema, pseudocysts and ductal ectasia).

**Evaluation of bone mineral density**

Endpoints include bone mineral density (BMD), bone fractures and biomarkers of bone metabolism. Questionnaire items include bone symptoms and joint symptoms, history of fractures and low BMD, age of menarche, hormone supplements, calcium and vitamin D intake, and family history of bone disease. Biomarkers include ionised calcium, phosphate, alkaline phosphatase, parathyroid hormone and 25-OH vitamin D. Bone mineral content, bone area, and BMD is determined by DXA of total body, and lumbar spine (L1–L4) and dual-femur scans. Adult bone mineral values are analysed as age-specific and sex-specific SD and expressed as z-scores and T-scores. Paediatric bone mineral values are analysed as age-specific and sex-specific SD and expressed as z-scores. All paediatric bone mineral values are corrected for height.

**Evaluation of cardiovascular function**

Endpoints include left and right ventricular systolic and diastolic dysfunction, ventricular volumes and left ventricular mass, myocardial fibrosis, myocardial iron deposits, atrial thrombosis, conduction abnormalities and hypertension. Questionnaire items include cardiovascular symptoms, diagnoses, prescription medicine and familial disposition to cardiovascular disease. Automated oscillometric blood pressure is measured in all, and manual auscultatory systolic blood pressure is measured in children (<18 years). Cardiac biomarkers include N-terminal pro b-type natriuretic peptide and troponin-T. Cardiac electrical activity is assessed with a 12-lead ECG. Imaging includes speckle tracking echocardiography and cardiac MRI. Echocardiography is performed using a Vivid E95 ultrasound scanner (GE Vingmed Ultrasound AS, Horten, Norway) to obtain measurements of left ventricular output and volumes, as well as global longitudinal systolic strain. An intravenous contrast agent (SonoVue) is used to improve imaging.50 Cardiac MRI is performed using a whole body 1.5T scanner for assessment of left ventricular systolic function, volumes, and...
mass. Multiparametric characterisation, including T1, T2, T2* and mapping is performed for visualisation of possible myocardial oedema and fibrosis, and to evaluate myocardial iron content. Intravenous gadolinium-containing contrast (0.15 mmol/kg Gadovist) is used for assessment of extracellular volume fraction in survivors (not controls).

**Evaluation of pulmonary function**

Endpoints include function of conductive and acinar airways, resistance of airways, alveolar volumes, alveolar membrane diffusion, lung clearance index, capillary volume, total diffusion capacity, spirometric volumes, and flows and reversibility. Questionnaire items include respiratory symptoms, pulmonary diagnoses, smoking and use of prescription medicine. Prebronchodilator and postbronchodilator spirometry is performed using a Jaeger MasterScreen Bodybox and Jaeger Vyntus Spiro (CareFusion, Hochberg; Vyaire Medical, Bayern) for measurement of forced expiratory volume during the first second, forced vital capacity (FVC), and forced expiratory flow between 25% and 75% of FVC. Impulse oscillometry, including reversibility test, is performed at University Hospital of Aarhus using a Jaeger MasterScreen Bodybox and Jaeger Vyntus IOS system (CareFusion, Hochberg; Vyaire Medical, Bayern) to evaluate lung volumes, resistance and elasticity of the lung. Single breath multiple breath washout is performed using an Exhalizer D, N2 and SF6 option (Eco Medics AG, Dürnten, Switzerland) to assess Lung Clearance Index, including indexes of acinar and conductive airways ventilation heterogeneity. Single breath diffusion capacity for carbon monoxide and nitrogen oxide, corrected for haemoglobin, is performed using a Jaeger Vyntus Body (CareFusion, Hochberg) for real-time single breath diffusion to evaluate diffusion membrane capacity and pulmonary capillary volume. All pulmonary measures are compared with normative data and with data from the matched controls.

**Evaluation of peripheral and autonomic neuropathy**

Endpoints include sensory nerve, motor nerve and autonomic nerve dysfunction. Peripheral neuropathy is assessed according to recommendations from the American Academy of Neurology, the American Association of Electrodagnostic Medicine and the American Academy of Physical Medicine and Rehabilitation, and using the Total Neuropathy Score, which includes self-report of sensory, motor and autonomic symptoms, pin sensibility, vibration sensibility (quantitative threshold using Vibrameter type IV (Somedic Electronics, Solna, Sweden) and Biothesiometer model PVD-LP (Biomedical Instruments, Ohio, USA), muscle strength (Medical Research Council grade 5 score), deep tendon reflexes, and amplitude of sural and peroneal nerves. Abnormal values corresponding to quantitative vibration threshold are calculated as percentages relative to the upper limit of normal values based on data from matched ALL-STAR controls and previously published normative data. Electroneurography is performed using standard surface recording techniques (Neuroline 715, Ambu). A G3 Keypoint platform (Dantec, Natus, USA) is used for amplification, filtration and storing of signals (motor: 2 Hz to 10 kHz; sensory: 20 Hz to 10kHz). Electroneurography includes unilateral evaluation of sensory nerve action potential from the sural nerve, compound motor action potential from the peroneal nerve, the distal motor latency in the peroneal nerve and both sensory and motor nerve conduction velocities. Recorded values are compared with ALL-STAR control values and with unpublished national (Danish age-specific and gender-specific reference values), and reported as SD from the expected mean (z-scores). Autonomic neuropathy is evaluated with the validated 31-item composite autonomic symptom questionnaire, COMPASS 31, which assesses orthostatic intolerance, vasomotor, secretomotor, gastrointestinal, bladder and pupillomotor function. Objective evaluation of autonomic dysfunction is performed by analysis of heart rate variability (HRV) obtained from a 21 min Holter recording (15 min reclined resting period followed by shift to standing position with continued ECG recording for 8 min). The RR-interval data set derived from the ECG segments is analysed for heart rate and HRV. Analysis of HRV is performed in both the time and frequency domain and changes elicited by the standing position is calculated.

**Evaluation of physical performance**

Endpoints include cardiorespiratory fitness, level of physical activity, physical function and muscle strength. Questionnaire items include levels of sedentary behaviour and of physical activity, which is objectified with a 7-day continuous accelerometer measurement (ActiGraph model GT3X+, ActiGraph, Pensacola, Florida, USA). Cardiorespiratory fitness is evaluated using an electronically braked cycle ergometer (Lode Corival Pediatric or Monark Ergomedic 839 E) following a modified Godfrey protocol. Ventilation and gas exchange data are determined breath-by-breath (INNOCOR ergospirometry-system, INNO00010, Innovision, DK-5260 Odense, Denmark or Jaeger Master Screen Vyntus CPX and JLAB software package). Peak oxygen uptake (VO2 peak) is defined as the highest mean over 60 s. Results are reported as SD from the expected mean (z-scores), derived from the age-matched and sex-matched ALL-STAR controls. Muscle strength is evaluated as isometric muscle strength (knee extensor, upper body (Gym 2000) with a strain gauge (US2A100 kg, Hottinger, Germany) and custom-made amplifier), and hand grip strength (Saehan hand dynamometer, Glanford Electronics, Scunthorpe). Dynamic muscle strength is evaluated by the counter movement jump test (FP4, HUR Labs Oy, Tampere, Finland) and physical functional tests (sit-to-stand 30 s and 60 s and timed up and go).

**Evaluation of kidney function**

Endpoints include kidney function and kidney size. Questionnaire items include urinary tract symptoms and disease, and use of prescription medicine. Fasting blood samples, include creatinine, urea, cystatin C, sodium, potassium,
phosphate, albumin, ionised calcium and parathyroid hormone. Estimated glomerular filtration rate is evaluated using creatinine in adults, and using creatinine, urea and cystatin in children. Proteinuria is determined by the albumin–creatinine ratio in a morning urine specimen. A spot urine dip stick test is performed to evaluate haematuria and urine pH. Kidney size is evaluated by ultrasonography using a convex abdominal probe (C1-5, GE Logiq E9 or E10, GE Healthcare, Chicago, USA) as longitudinal anterior–posterior length, width and cross-sectional anterior–posterior length, and abnormal appearance (eg, hydronephrosis, cysts or tumours) is registered. Kidney size is related to body weight, height and lean mass.

**Evaluation of pubertal status and gonadal function**

Pubertal stage is evaluated in all participants aged >8 years at time of evaluation. Endpoints include age at menarche/spermarche, pubertal stage and levels of reproductive hormones. Questionnaire items include age at menarche and use of contraception for females, age at spermarche for males and hormonal substitution. Pubertal stage is evaluated according to Tanner in combination with testicular volume for males, estimated by comparative palpation using a Prader orchidometer. Results are compared with Danish puberty normograms. Measurements of serum follicle stimulating hormone, luteinising hormone, estradiol, testosterone, insulin-like peptide 3, inhibin B, anti-Müllerian hormone and sex hormone binding globulin are performed, and levels are compared with sex-matched and age-matched national reference material as well as with ALL-STAR controls.

**Evaluation of fertility**

Fertility is evaluated in all participants aged ≥18 years. Endpoints include impaired fertility (defined as previously diagnosed impaired fertility or infertility and/or use of assisted reproductive technology), number and health of offspring. Questionnaire data include items regarding health behaviour, previous relationship/marital status, attempts to conceive, menopause, results from fertility investigations, natural conceptions, use of assisted reproductive technology, pregnancies, miscarriages, voluntary terminations, stillbirths, live births and health of offspring. Reproductive hormones and testicular volume are assessed as described in the previous section.

**Evaluation of oral health and salivary secretion**

Oral health assessments are performed at the Copenhagen University Hospital. Endpoints include oral symptoms, impaired oral and dental health (mucosal changes, caries, erosions, gingivitis, periodontitis, malocclusions, temporomandibular dysfunction and salivary gland dysfunction) and changes in salivary microbiome and proteome. Questionnaire items include oral symptoms, current medication, dietary supplements, tobacco and alcohol consumption and oral hygiene habits. Xerostomia is assessed by means of the bother index, xerostomia inventory and visual analogue scale (0–10). Additional evaluation of oral health includes a clinical examination with palpation of the temporomandibular joint, salivary glands and regional lymph nodes; assessments of the morphological and functional occlusion, dental status (decayed missing filled teeth scores, dental erosion scores, dental malformations and delayed eruptions), periodontal status (gingival inflammation and dental plaque indices and assessment of periodontal probing depth) and mucosal status (mucosal changes and signs of dryness using clinical oral dryness score). Unstimulated and chewing-stimulated whole saliva flow rates are measured according to the drooling method. DNA is extracted from the saliva samples and the microbiome characterised by Human Oral Microbe Identification (using next-generation sequencing) with bacterial species-level identification based on 16S rDNA comparisons. Salivary proteome analyses are performed by means of 2D electrophoresis and matrix-assisted laser desorption/ionisation time-of-flight mass spectrometry. Western blot analysis and ELISA are used for validation of the 2DE results.

**Evaluation of mental health and quality of life**

Endpoints include psychiatric disease, social relations, emotional distress (anxiety, depression and fatigue), executive function and health-related quality of life. Questionnaire items include psychiatric diagnoses and psychopharmaceuticals. Additional psychosocial data are collected via instruments from Paediatric Quality of Life Inventory, Patient Reported Outcomes Measurement Information System, and the Behavior Rating Inventory of executive Function system. Specific instruments and versions are provided in online supplemental table 4.

**Use of previously collected data for risk factor analysis**

All patient characteristics (date of diagnosis, tumour burden, immunophenotype, cytogenetics, minimal residual disease levels, final risk stratification, participation in protocol randomisations and occurrence of 19 well-defined acute toxicities) are obtained from the NOPHO ALL2008 registry. Medical chart data are used for validation of registered acute toxicities and for validation of self-reported health conditions (including dates). Genotyping for genome wide association studies has been performed previously on germline DNA from samples obtained after clinical remission using the exome-enriched Illumina Infinium Omni2.5Exome-8 BeadChip arrays and is available for approximately 70% of the survivor population. Whole blood samples are collected.
from participants (including controls) not previously genotyped and samples are analysed using the same platform.

Data management
Study data were collected and managed using Research Electronic Data Capture (REDCap) tools hosted at Copenhagen University Hospital. REDCap is a secure, web-based software platform designed to support data capture for research studies, providing: (1) an intuitive interface for validated data capture; (2) audit trails for tracking data manipulation and export procedures; (3) automated export procedures for seamless data downloads to common statistical packages; and (4) procedures for data integration and interoperability with external sources.

Statistical considerations
Cumulative incidence of the 197 health-conditions is calculated (individually and grouped according to organ system and severity grade) for survivors and controls and presented with 95% confidence intervals (CI). Gray’s test is used for comparison between the two groups. Cumulative burden of the 197 health conditions is calculated (individually and grouped according to organ system and severity grades) in survivors and controls by the method of mean cumulative count, estimating the mean number of recurring or multiple health events in the presence of competing risk, and presented with 95% CIs. For this purpose, health conditions are categorised pre-analysis into three subtypes (chronic, non-recurrent; single, recurrent; and chronic, recurrent) based on clinical definitions of recurrence and chronicity, as described by others. Based on a projected sample size of 250 survivors and 250 controls (survival rate of 90% in the Danish NOPHO ALL2008 patient cohort and a conservative recruitment rate of 60%), a 16% prevalence of chronic disease among Danish children, and significance level alpha=0.05, we will have a power of 87% to detect a HR of 2.0 in a Cox regression model, which is the lowest HR previously reported in a study comparing leukaemia survivors with sibling controls. Secondary outcomes (ie, description of organ-specific functions) are compared between survivors and controls using relevant regression models (eg, linear regression models, logistic regression models or Cox regression models). For survivors, associations of primary and secondary outcomes with therapy exposure and occurrence of acute toxicities are explored using multiple regression models. Association with single nucleotide polymorphisms (SNPs) is explored with genome-wide association analysis and using a candidate gene approach where relevant SNPs have been identified.

Patient and public involvement
In preparation of this study, we performed a qualitative interview study among adolescent and young adult NOPHO ALL2008 survivors to identify the challenges they perceive as most important. Results from this study have been incorporated into the ALL-STAR examination programme, securing focus on the issues most important to the survivors.

Longitudinal follow-up
The cohort is invited for follow-up examinations corresponding to 5-year intervals from diagnosis, thereby aligning survival time at data capture. Results from the initial cross-sectional studies will inform which examinations to prioritise and additional endpoints can be added to address new hypotheses. The ALL-STAR database will facilitate repeated invitations and longitudinal data capture. Participants will provide new informed consent before repeated examinations.

ETHICS AND DISSEMINATION
The study has been approved by the Regional Ethics Committee for the Capital Region in Denmark (no. H-18035090 and H-20006359), the Danish Data Protection Agency (no. VD-2018–519) and is conducted in accordance with the Helsinki Declaration. Written informed consent is obtained from all participants prior to enrolment. All participants can choose to participate in only parts of the study, can choose not to be informed of results, and can withdraw consent at any time. Study results will be published in international peer-reviewed scientific journals and presented at relevant conferences.

DISCUSSION
The ALL-STAR cohort is the first Nordic, population-based childhood and young adult ALL survivor cohort treated according to the same ALL protocol, to be uniformly evaluated for treatment-related morbidity across several organ systems, using both self-report data and objective, clinical investigations.

The evolution of long-term, adverse consequences of antileukaemic therapy is complex and dependent not only on therapy exposure and genetic susceptibility, but also on age at diagnosis, pre-existing morbidity, individual resources, resources among family and peers, and on surrounding healthcare structures and societal structures. The comparison of a Nordic survivor cohort with survivor cohorts from other Western countries will provide insight into the consequences of differing systems, as well as differences in treatment regimens used. Scandinavian countries, and particularly Denmark, have access to extensive register data including health information and socioeconomic information, linked to the individual person and surrounding network through personal identification numbers. Coupling of systematic survivor data obtained in the ALL-STAR study to national register data will enable contextualising association studies novel to survivorship research.
Strengths of the ALL-ST AR study include a population-based design, a homogenous and well characterised cohort, the combination of both subjective and objective assessments of outcome, and inclusion of matched controls. The ALL-ST AR cohort has already been characterised by prospective registration of tumour burden, immunophenotype, cytogenetics, treatment response and occurrence of 19 severe toxicities, performed by treating physicians from time of diagnosis until end of therapy. These data enable high precision evaluation of risk factors and confounders associated with our main outcome of interest. Similarly, high quality survivor data is secured through self-report data, medical chart validation and systematic clinical examinations across most organ systems and using both well-established and novel, advanced techniques. This combination of data enables inclusion of survivor perspective while minimising risk of recall bias and inaccurate information. Furthermore, the systematic screening of organ functions facilitates capture of low-grade conditions, which are likely to be underestimated if using retrospective or self-report data only. Since the investigated health conditions are not pathognomonic, the inclusion of matched controls, often missing in survivorship studies, will support conclusions made regarding the risk patterns, which we seek to identify.

Despite vast amounts of late effects research being performed, no international consensus on how to characterise therapy-related late effects currently exists. Lack of precise definitions and uniform grading result in substantial variability across studies, hindering meaningful pooling and comparisons of results. Cancer and therapy-related acute toxicities are traditionally defined and graded according to the CTCAE system and systematic modifications proposed by St Jude researchers has improved its utility for also addressing the long-term effects among childhood cancer survivors. We have aligned our outcome definitions accordingly and by introducing 10 additional health conditions of relevance for cancer survivorship research, we wish to further increase the relevance of this classification system.

The magnitude of late effects is often described with cumulative incidence; however, we encourage the use of the cumulative burden, which includes recurrent events as opposed to only first events, thereby reflecting the burden more comprehensively. This is illustrated by a study investigating therapy-related cardiovascular conditions among survivors of childhood and young adult Hodgkin’s lymphoma. By the age of 50 years, cumulative incidence of all-grade cardiovascular health conditions did not differ between survivors and controls, while the cumulative burden of the same conditions was almost double in survivors compared with controls. Since none of the measures, however, reflect survivor perspective, we have included the perceived impact of the individual health conditions, psychosocial function and health-related quality of life as reported by the participants. Coupling of objective findings with subjective survivor experiences is essential for interpretation of results and for guiding healthcare workers and researchers when choosing which survivorship issues to prioritise.

There are several limitations to our study. A survival time of 3.5–14 years at evaluation means that the full spectrum of therapy-related conditions will not be captured. Establishment of the ALL-ST AR cohort within the first decade of survival, however, forms the platform for longitudinal follow-up, which is within the overall scope of the project. Interpretation of initial data will be challenged by a heterogeneous distribution in both age and survival time, but this will be adjusted for in the statistical analysis. Several of the health conditions evaluated in the ALL-ST AR study rely on questionnaire data and subsequent medical chart validation. Low-grade conditions, which are not screened for are, therefore, likely to be missed. Finally, the ALL-ST AR study outcomes are not exhaustive in describing the treatment-related burden. Important outcomes, such as semen quality, ovarian reserve and neurocognitive function were omitted to ensure feasibility of the initial ALL-ST AR initiative. Future studies of these important outcomes in the ALL-ST AR cohort will benefit from the infrastructure already created.

In conclusion, the ALL-ST AR study is establishing a Nordic childhood and young adult ALL survivor cohort to be systematically assessed for the cumulative incidence and burden of treatment-related morbidity. Results will supplement existing evidence regarding risk of treatment-related morbidity and the underlying possible biological, physical and psychosocial mechanisms, while providing novel insight into differences in survivorship related to specific therapy regimens, health cultures and healthcare systems. Findings may guide future toxicity risk grouping and preventive interventions during treatment and follow-up, ultimately ameliorating the burden of therapy without compromising chance of cure.

Author affiliations
1Department of Pediatrics and Adolescent Medicine, Copenhagen University Hospital, Copenhagen, Denmark
2Department of Pediatrics and Adolescent Medicine, Aarhus University Hospital, Aarhus, Denmark
3Department of Cardiology, Aarhus University Hospital, Aarhus, Denmark
4Department of Cardiology, Aarhus University Hospital, Copenhagen, Denmark
5Department of Radiology, Aarhus University Hospital, Aarhus, Denmark
6Department of Radiology, Aarhus University Hospital, Copenhagen, Denmark
7Section of Biostatistics, University of Copenhagen, Copenhagen, Denmark
8Institute of Clinical Medicine, University of Copenhagen Faculty of Health and Medical Sciences, Copenhagen, Denmark
9Department of Radiology, University of Copenhagen, Copenhagen, Denmark
10Department of Oncology, University of Copenhagen, Copenhagen, Denmark
11Department of Oncology, University of Copenhagen, Copenhagen, Denmark
12Department of Pediatric Hematology and Oncology, Odense University Hospital, Odense, Denmark
13Department of Pediatric Hematology and Oncology, Odense University Hospital, Odense, Denmark
14Department of Clinical Medicine, Aarhus University Hospital, Aarhus, Denmark
15Department of Clinical Medicine, Aarhus University Hospital, Aarhus, Denmark

Acknowledgements The authors wish to acknowledge valuable contributions made by Dr Margrethe Bastholm Bille; Associate professor Niels Holtum Birkebak; Dr Alfred Peter Born; Dr Vilteke Brix; Sisse Hellund Carlsen; Perrille Fleggaard; Dr Christian Alcaraz Frederiksen; Dr Peter Orm Hansen; Professor Won Yong Kim; Dr
REFERENCES


39 Frank L. Epidemiology. when an entire country is a cohort. Science 2000;287:2398.


