BMJ Open

Cohort-based association study of germline genetic variants with acute and chronic health complications of childhood cancer and its treatment: Genetic Risks for Childhood Cancer Complications Switzerland (GECCOS) study protocol

Nicolas Waespe ,[1,2,3,4] Sven Strebel,[1,2,5] Tiago Nava,[1,6] Chakradhara Rao S Uppugunduri,[1] Denis Marino,[1] Veneranda Mattiello,[1,6] Maria Otth,[2,3,7] Fabienne Gumy-Pause,[1,6] André O Von Bueren,[1,6] Frederic Baleydier,[1,6] Luzius Mader,[2] Adrian Spoerri,[8] Claudia E Kuehni,[2,4] Marc Ansari[1,6]

To cite: Waespe N, Strebel S, Nava T, *et al*. Cohort-based association study of germline genetic variants with acute and chronic health complications of childhood cancer and its treatment: Genetic Risks for Childhood Cancer Complications Switzerland (GECCOS) study protocol. *BMJ Open* 2022;**12**:e052131. doi:10.1136/bmjopen-2021-052131

► Prepublication history and additional supplemental material for this paper are available online. To view these files, please visit the journal online (http://dx.doi.org/10.1136/bmjopen-2021-052131).

Received 06 April 2021
Accepted 22 December 2021

For numbered affiliations see end of article.

Correspondence to
Professor Marc Ansari;
marc.ansari@hcuge.ch

ABSTRACT

Introduction Childhood cancer and its treatment may lead to various health complications. Related impairment in quality of life, excess in deaths and accumulated healthcare costs are relevant. Genetic variations are suggested to contribute to the wide inter-individual variability of complications but have been used only rarely to risk-stratify treatment and follow-up care. This study aims to identify germline genetic variants associated with acute and late complications of childhood cancer.

Methods and analysis The Genetic Risks for Childhood Cancer Complications Switzerland (GECCOS) study is a nationwide cohort study. Eligible are patients and survivors who were diagnosed with childhood cancers or Langerhans cell histiocytosis before age 21 years, were registered in the Swiss Childhood Cancer Registry (SCCR) since 1976 and have consented to the Paediatric Biobank for Research in Haematology and Oncology, Geneva, host of the national Germline DNA Biobank Switzerland for Childhood Cancer and Blood Disorders (BISKIDS).

GECCOS uses demographic and clinical data from the SCCR and the associated Swiss Childhood Cancer Survivor Study. Clinical outcome data consists of organ function testing, health conditions diagnosed by physicians, second primary neoplasms and self-reported information from participants. Germline genetic samples and sequencing data are collected in BISKIDS. We will perform association analyses using primarily whole-exome or whole-genome sequencing to identify genetic variants associated with specified health conditions. We will use clustering and machine-learning techniques and assess multiple health conditions in different models.

Discussion GECCOS will improve knowledge of germline genetic variants associated with childhood cancer-associated health conditions and help to further individualise cancer treatment and follow-up care, potentially resulting in improved efficacy and reduced side effects.

Strengths and limitations of this study

- The strength of the Genetic Risks for Childhood Cancer Complications Switzerland study is the recruitment of childhood cancer patients and survivors from the national population-based Swiss Childhood Cancer Registry (SCCR) with data from 8163 participants.
- The SCCR contains an extensive dataset including demographic, treatment, outcome, follow-up and survival information which is then used for genotype–phenotype association analyses.
- The germline DNA collection within the Germline DNA Biobank Switzerland for Childhood Cancer and Blood Disorders allows storage of samples and sequencing data creating an increasing collection of genetic material and data for future use.
- While the cohort for patient recruitment is large, the population with a specific outcome of interest might be small for specific populations. This limitation will be counteracted by actively seeking international collaborations with pooling of available data.
- We will explore novel association models to account for the complex interactions of treatment exposure, genetic predisposition and environmental factors.

Ethics and dissemination The Geneva Cantonal Commission for Research Ethics has approved the GECCOS study.
Research findings will be disseminated through national and international conferences, publications in peer-reviewed journals and in lay language online.
Trial registration number NCT04702321.

INTRODUCTION

Childhood cancers have become curable in ≥85% of patients in developed countries.[1] Current treatment protocols are multimodal with varying combinations of surgery, chemotherapy, radiation, haematopoietic stem cell transplantation (HSCT) and immunotherapies. The price of increased survival is a wide range of acute and chronic health conditions. Cancer treatments are associated with acute complications such as transient nausea and vomiting, mucositis and fatigue but also pneumonitis, cardiomyopathy, encephalitis and life-threatening infections.[2] While many of these conditions are potentially reversible, some are not and become chronic or develop over time like cardiac, pulmonary, auditory, endocrine, reproductive and neurocognitive health conditions, and second primary neoplasms (SPNs).[3] A recent publication found a cumulative incidence of severe chronic health conditions of 96% in childhood cancer survivors aged 50 years. The number of severe chronic health conditions was two times higher in survivors compared with matched community controls.[4] In more recent decades, chronic health conditions were reduced following treatment adaptations to reduce adverse events.[5] Mortality is significantly increased in survivors compared with the general population[6] and varies depending on the treatment exposure over time.[7] Recurrence and SPNs are the leading causes of death in the first two decades after cancer diagnosis followed by diseases of the cardiovascular and respiratory systems thereafter.[6 8] Because of their young age at diagnosis, survivors have decades of life time ahead. The burden of chronic conditions, related impairment in quality of life, excess in deaths and accumulated healthcare costs is therefore of great relevance for them.[9]

Only few genetic variants modifying the risk of acute and chronic toxicities in children with cancer have so far led to personalised treatment protocols or follow-up care. An example is 6-mercaptopurine, where dosing is routinely adapted in contemporary treatment protocols for patients with thiopurine methyltransferase variants which increase the risk of acute haematological toxicity in the treatment of childhood acute lymphoblastic leukaemia (ALL).[10] Other genetic markers have been associated with various outcomes but so far not implemented in treatment protocols, such as nudix hydrolase 15 with haematological toxicity after purine analogue treatment,[11 12] dihydrofolate reductase with overall survival in some subtypes of ALL[13 14] and glutathione S transferase genes with various outcomes after HSCT.[15] Also, the effect of genetic variation on late toxicities, chronic health conditions arising after the end of childhood cancer treatment, has been investigated.[16 17] Different clinical outcomes such as hearing loss,[18–21] cardiomyopathy,[22 23] metabolic syndrome[24 25] and gonadal impairment have been studied.[26 27] Exome-wide (EWAS) and genome-wide association studies (GWAS) have led to the identification of genetic variants associated with drug toxicities (eg, osteonecrosis in children with ALL *(BMP7* and *PROX1-AS1)*, asparaginase hypersensitivity *(GRIA1)*, or vincristine-associated peripheral neuropathy in children with ALL *(CEP72))*,[28–31] metabolic syndrome,[24] sinusoidal obstruction syndrome after HSCT[32] and hearing loss.[18] Acylphosphatase 2 (*ACYP2*) was associated with hearing loss after platinum treatment exposure in several independent datasets but failed to be replicated in a recent large candidate gene analysis while a variant in solute carrier family 22, member 2 (*SLC22A2*) was found to be associated with mild hearing loss.[20 33] Genetic modifiers have been implicated for the development of SPNs including breast cancer,[34] CNS tumours[35] and leukaemia.[36] Replication of findings is necessary before allowing translation of these findings into clinical practise.[37] For many SPNs (such as thyroid cancer), no data is available.[38] To address the contribution of genetic risk variants in the development of late toxicities, large cancer survivor studies such as the Childhood Cancer Survivor Study[36] and SJLIFE cohort[39] in the US and the French Childhood Cancer Survivor Study for Leukaemia (LEA Cohort)[40] are collecting DNA systematically to conduct genotype–phenotype analyses.

Most studies on genetic risk variants used a candidate-gene approach, had a small sample size of less than 200 participants, heterogenous cohorts with various treatment exposures and inconsistent outcome assessments.[17] Many health conditions have not been investigated, like renal insufficiency, pulmonary and ocular complications. It is likely that several distinct pathways and their corresponding gene variants are involved in the development of complex phenotypes like pulmonary dysfunction. Therefore, many candidate gene variants and treatment exposures need to be considered. This is possible with hypothesis-free EWAS or GWAS. For specific outcomes, case–control designs using samples from different cohorts have also been successfully used.[41] New analytical approaches which combine clinical, pharmacological and genetic data into integrative models have been developed and are showing promising results.[42 43] Network analyses, machine learning and clustering methods might help to understand the impact of genetic variations on complex phenotypes with biological pathways and their corresponding genes. The Genetic Risks for Childhood Cancer Complications Switzerland (GECCOS) study will enable large genotype–phenotype association studies in childhood cancer patients and survivors in Switzerland and international collaborations. Selected organ dysfunctions and SPNs will be studied in subprojects, and in silico and in vitro methods will be used to further explore mechanisms associated with genetic variation and outcomes.

STUDY OBJECTIVES

The main objectives of the GECCOS study are

1. To identify genetic variants associated with health conditions after childhood cancer and its treatment using genotype–phenotype association methods.
2. To evaluate the biological function of genetic variants associated with health conditions through in silico and in vitro methods.

3. To identify genetic variants associated with multiple health conditions after childhood cancer using models integrating multiple outcomes.

The secondary objectives are

1. To create a common framework for sub-studies using genotype–phenotype associations with germline genetic material and data of childhood cancer patients and survivors.
2. To feed germline genetic data generated in substudies into a biobank database for future research and create a growing repository of genetic sequencing data.
3. To facilitate research using a common structure that can be used for collaborations.

METHODS

Study design

The GECCOS study is a nationwide cohort study in collaboration of the Paediatric Biobank for Research in Haematology and Oncology (BaHOP), Geneva University Hospital, Switzerland, and the Institute for Social and Preventive Medicine at the University of Bern. The Geneva Cantonal Commission for Research Ethics has approved the GECCOS study (approval 2020-01723), and the BaHOP biobank (approval PB_2017-00533). Recruitment of patients into the biobank started in September 2019 and inclusion in the GECCOS study in December 2020. The GECCOS study is serving as backbone for several subprojects on various outcomes with the end date of the study set to December 2037. Within subprojects assessing specific outcomes (such as hearing loss or pulmonary complications), we will sample patients and survivors according to risk exposure (for a cohort design), or according to the outcome of interest (for a case–control or case–cohort design).

Data sources

The GECCOS study uses genetic data and material from BaHOP. Clinical information is collected from (1) the Swiss Childhood Cancer Registry (SCCR) and (2) the Swiss Childhood Cancer Survivor Study (SCCSS; figure 1, online supplemental table 1). The nationwide Germline DNA Biobank Switzerland for Childhood Cancer and Blood Disorders (BISKIDS) was established in May 2019 within the BaHOP with support of all nine paediatric oncology centres caring for childhood cancer patients in Switzerland. Return of results and relevant incidental outcomes to the patient is defined in the BaHOP regulations with oversight of a genetic advisory board. BISKIDS collects germline DNA samples, extracts and stores genomic DNA and genetic data of childhood cancer patients as well as survivors in Switzerland.

The SCCR collects information on childhood cancer patients diagnosed in Switzerland.[44] Children and adolescents aged <21 years with a primary cancer diagnosis according to the International Classification of Childhood Cancer, third edition (ICCC-3), and Langerhans cell histiocytosis (LCH) were registered since 1976 in the

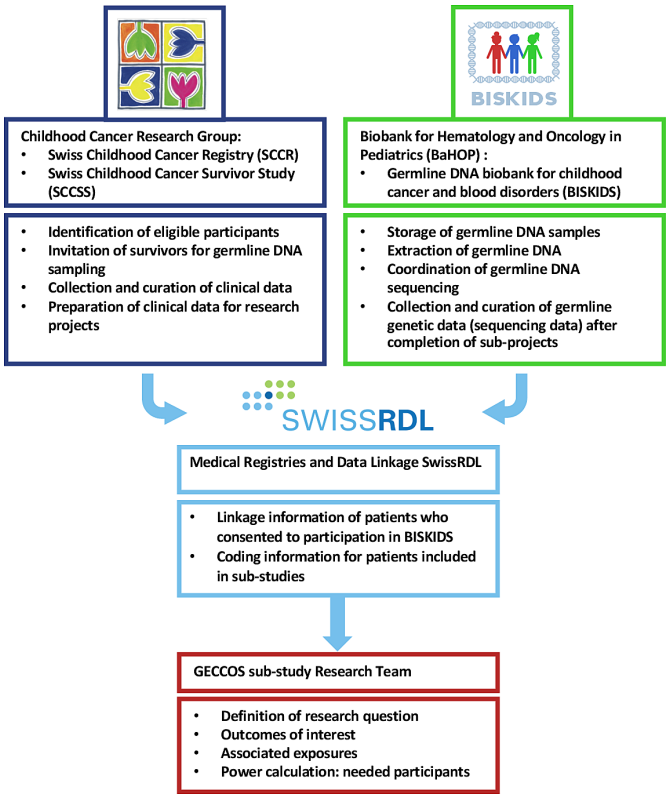

Figure 1 Responsible teams in the Genetic Risks for Childhood Cancer Complications Switzerland (GECCOS) study for germline genetic associations with health conditions in childhood cancer patients and survivors.

SCCR. The dataset is managed by the Childhood Cancer Research Group at the Institute of Social and Preventive Medicine, University of Bern. It has a completeness of coverage of childhood cancer patients in Switzerland aged up to 15 years of >85% since 1985 and >95% since 1995.[45] Mandated by a national cancer registration law in Switzerland enacted on 1 January 2020, registration of new patients with cancer is performed by the federal government from that date onwards.[46] The Institute of Social and Preventive Medicine of the University of Bern was commissioned to perform patient registrations and data collection, and the Childhood Cancer Research Group will continue research activities on these datasets. The GECCOS project will seek to include patients diagnosed with neoplasms during childhood after 1 January 2020 and continue collaboration with the Childhood Cancer Research Group at the University of Bern. The SCCSS (clinicaltrials.gov: NCT03297034) is a population-based, long-term cohort study of all childhood cancer patients who were registered in the SCCR, resident in Switzerland at diagnosis, and survived ≥5 years after initial cancer diagnosis.[47] The SCCSS collects questionnaire-based information from survivors on self-reported health outcomes, sociodemographic information, and environmental exposures (such as smoking). Clinical data on chronic health conditions after childhood cancer from survivorship clinics and hospital records (eg, audiograms and lung function tests) will be extracted for GECCOS.

The datasets used for the association analyses will contain collected follow-up information from medical records, self-reported outcomes and functional outcome data from specific projects on long-term complications.

Study population

Eligible for the participation in the GECCOS study are persons who:

1. Are registered in the SCCR.
2. Were diagnosed with a neoplasm according to the ICCC-3, or LCH before age 21 years.
3. Have consented themselves or through their legal representative to the BaHOP (host biobank for BISKIDS).

Inclusion criteria for subprojects focusing on specific health conditions will be defined (online supplemental tables 2 and 3 for established subprojects). We will identify eligible participants from the SCCR and SCCSS. LCH has been included in the SCCR due to its clonal proliferation of immature cells with somatic activating gene mutations and the need for antineoplastic treatment in an important subset of patients.[48]

Selection of participants

We will identify participants eligible for specific sub-studies with defined in- and exclusion criteria. As of December 2019, 13 029 patients were registered in the SCCR, of which 9306 (69%) were still alive and 8163 (63%) were Swiss residents and potentially eligible for participation in GECCOS. We will use information in the SCCR and SCCSS and assess availability of corresponding germline DNA samples or sequencing data from previously sequenced participants in BISKIDS (figure 2). If clinical data is available for a sufficient number of participants but further genetic samples are needed, we will invite potential participants to contribute germline DNA samples to BISKIDS for research. For collection of biological material within BISKIDS, we will use two pathways: (1) invitations to participate are sent out by the Childhood Cancer Research group at the University of Bern, consisting of germline DNA collection kits (predominantly using saliva samples or buccal swabs) with information on the biobanking project and associated research, and informed consents to the participant's home; (2) participants are invited by healthcare staff in hospitals caring for childhood cancer patients and survivors. These potential participants and their legal representatives are informed of the project and written consent and germline DNA are collected during a medical visit already planned for their treatment or follow-up. All participants consent to have their germline DNA stored in the BISKIDS section of the BaHOP biobank and their health-related data to be used for genotype-phenotype association studies. All specific GECCOS substudies will be reviewed and approved by a national scientific committee and submitted to the responsible ethics committee as amendment to the main GECCOS protocol, insofar as the applicable law requires authorisation.

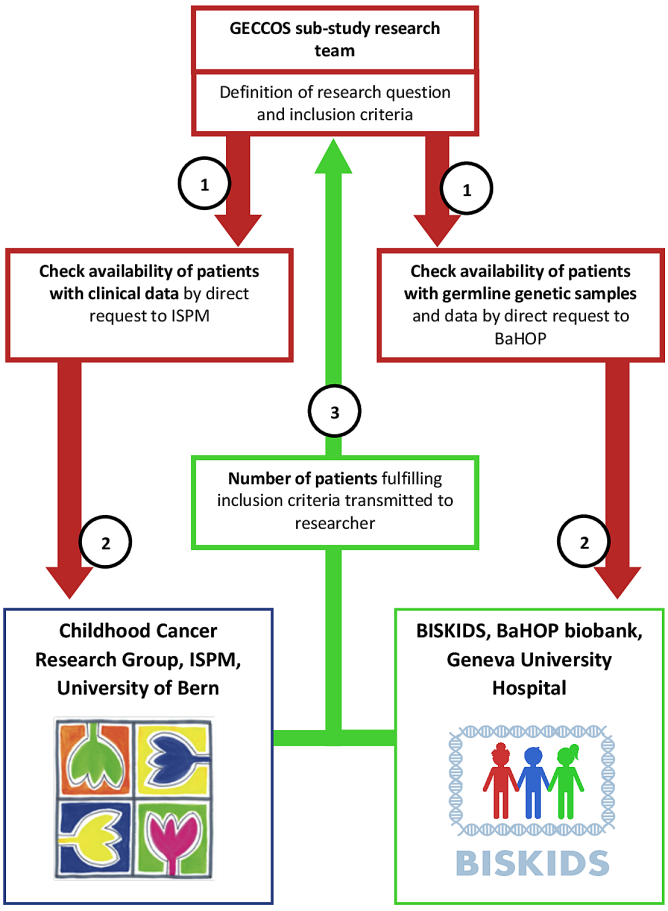

Figure 2 Flow diagram of identification of eligible participants for a specific subproject: (1) the eligibility criteria, as defined by the researchers, will be transmitted to the childhood cancer research group at the Institute of Social and Preventive Medicine (ISPM), University of Bern, and the Germline DNA Biobank Switzerland for Childhood Cancer and Blood Disorders (BISKIDS) at the Geneva University Hospital; (2) the number of eligible participants is compiled by secured data exchange from the BISKIDS project and the childhood cancer research group; (3) the number of eligible participants will be transmitted to the researcher to assess feasibility of a subproject. BaHOP, Paediatric Biobank for Research in Haematology and Oncology; GECCOS, Genetic Risks for Childhood Cancer Complications Switzerland.

Outcomes: health conditions and SPNs

We will assess health conditions in childhood cancer patients and survivors by collecting data on organ function such as pulmonary functions tests for lung conditions, or audiograms for hearing loss. For outcomes that cannot be adequately measured (eg, tinnitus), we will use information from self-assessment questionnaires. We will use the International Agency for Research on Cancer (IARC) criteria for coding SPNs.[49] In brief, we will include neoplasms according to the ICCC-3 classification which originated in different tissues or had a different morphology than the first primary neoplasm. We will not classify progression, transformation, metastasis and relapse of first primary neoplasm as SPN.

Covariates

For specific substudies and analyses, we will extract data on relevant covariates that might influence the outcomes of interest from the SCCR and SCCSS. We will extract demographic, socioeconomic, first primary neoplasm, treatment and follow-up information. For treatment information, we will estimate cumulative doses of chemotherapies using individual treatment protocols, or calculate effective treatment doses from medical records, if available. We will estimate exposure to radiotherapy using radiotherapy field descriptions,[50–52] or calculate organ-specific exposure from effective administered radiotherapy documentation (table 1).

Data linkage

Coded unique identifiers allow linkage of genotype data from the BISKIDS collection to phenotype data from the SCCR and SCCSS. The identifiers are securely stored in a separate trust centre database managed by a third party (Swiss Medical Registries and Data Linkage, Institute of Social and Preventive Medicine, University of Bern). This procedure allows to separate identifying information from clinical and genetic information. We use a web-based secured and personalised access point (currently REDCap V.10.6, Vanderbilt University, Nashville, Tennessee, USA). The key located at the trust centre is used to merge the clinical dataset with the germline genetic information without releasing identifying information to the GECCOS research team or to the BaHOP biobank. The research dataset will only contain a unique study-specific patient-identifier, without any identifying information. Study-specific identifiers will be securely stored in the trust centre database to assure traceability of datasets (figure 3).

Data and sample handling

Clinical data will be transferred in standard data formats. Analysis datasets will not include identifying information. Management of germline DNA samples (eg, saliva or buccal swabs), DNA extraction, aliquoting and storage procedures are clearly defined in the BaHOP biobank regulation. In a first recruitment effort, 928 childhood cancer survivors from the SCCR were asked to participate in germline DNA collection by home saliva collection and 463 (50%) participated.[53] For germline DNA sequencing in the GECCOS study, one of the DNA aliquots will be sent to Campus Biotech, Geneva, a sequencing facility. For genotype–phenotype analyses, we will collaborate with the Swiss Institute of Bioinformatics, Switzerland. Sequencing data and relevant clinical outcome data will be shared in a secured and encrypted way between the sequencing facility, the Swiss Institute of Bioinformatics for analysis, and the research platform for paediatric oncology and haematology in Geneva. Data will be stored and made available through a harmonised nationwide network to support computational biomedical research and clinical bioinformatics.[54]

Table 1 Summary information on main covariates and exposures of interest used in the genotype–phenotype association analyses

Covariate type	Covariates	Unit
Demographic information	Sex	Male/female/other
	Birthdate	Month/year
	Country of origin	Country name
Socioeconomic status	Highest education patient and parents	10-unit scale
	Income patient and/or parents	Monthly net income (Swiss francs)
First primary neoplasm information	Age at diagnosis	Years
	Date of diagnosis	Month/year
	Type of diagnosis	ICCC3 code; ICDO3 morphology, topography, behaviour code
	Laterality	Left/right/ bilateral/ medial/ not applicable
	Relapse date	Month/year
	Relapse type	Local/distant/ systemic/other
	Relapse location	Organ and morphology
Treatment information	Treatment protocol	Name of protocol, arm, randomisation group
Chemotherapy	Separately per antineoplastic agent: Cumulative dose estimated using treatment protocols or if available calculated using extracted data from medical records	mg/m^2, or appropriate metric; cycles (n); dose per cycle (mg/m^2)
	Start date	Month/year
Radiotherapy	Radiation type	Photon, proton, brachytherapy, stereotactic radiation, other
	Radiation field	Description of radiation field (eg, mantle field, whole lung irradiation, total body irradiation)
	Laterality	Left/ right/ bilateral/ medial/ not applicable
	Start date	Month/year
	Separately per radiation field: Cumulative dose estimated on an intention to treat basis using radiotherapy field descriptions or calculated from radiotherapy protocols	
	Concomitant chemotherapy	Antineoplastic agent and dose (mg/m^2)

Continued

Table 1 Continued

Covariate type	Covariates	Unit
Surgery	Location	Organ, site and description of intervention
	Laterality	Left/right/ bilateral/ medial/ not applicable
	Date	Month/year
Haematopoietic stem cell transplantation	Type	Allogeneic, autologous, other
	Donor type	Matched sibling, matched other relative, haploidentical relative, matched unrelated, mismatched unrelated, other
	Donor graft source	Bone marrow, peripheral stem cells (apheresis), umbilical cord blood, other
	Matching degree	No of HLA loci matched of total assessed HLA loci
	Conditioning regimen	Name and treatment details
	Total body irradiation	Yes (including dose in Gy), no
	Date of stem cell transfer	Month/ year
	Acute complications	Sinusoidal obstruction syndrome, infection, acute GvHD (with grading), others
	Chronic complications	Chronic GvHD (with grading), others
Follow-up information	Last vital status	Alive, dead, unknown
	Date of last vital status	Month/year
	Last follow-up information from clinical site	Month/year
	Acute and chronic health complications, environmental and lifestyle exposures	Extracted data from medical records and self-reported information from questionnaires

GvHD, graft-versus-host disease; HLA, human leucocyte antigen; ICCC3, International Classification of Childhood Cancer, third edition; ICDO3, International Classification of Diseases for Oncology, third edition.

Sequencing analyses

We will use genetic information (table 2) from different genomic sequencing methods: (1) gene panel sequencing adapted to specific research questions (using TruSeq DNA PCR free library preparation kit, Illumina, San Diego, USA); (2) whole-exome sequencing (using Illumina HiSeq4000 or NovaSeq 6000 platform with a mean read depth of at least 70x) and (III) whole-genome sequencing (with a mean read depth of at least 30x), depending on the research question. We will use workflows for genotyping implemented in the Genome Analysis Toolkit[55] and adapted to the aim of the study for (1) sequence generation; (2) sequence alignment; (3) variant calling; (4) variant filtering and (5) variant annotation.[56] We will also include data on read quality control. We will perform analyses with any of the following: (1) a candidate gene approach with filtered variants in preselected genes based on scientific hypotheses, (2) hypothesis-free exome-wide or (3) genome-wide association analysis and (4) multivariate approaches such as clustering methods or machine learning to identify associations of genetic variants with outcomes of interest. We will perform meta-analyses using combined cohorts of discovery and replication datasets and previous studies reporting on the same genetic variants, where possible.

We will mainly use the software packages Stata version 17 (StataCorp), R (R Foundation for Statistical Computing, Vienna, Austria) and PLINK[57] for analyses. Depending on the substudy and data availability, we will use different pipelines for quality control, filtering and annotation.[58] Statistical significance tests will be two-sided and appropriate significance levels will be applied, adjusting for multiple testing where appropriate taking into account clinical covariates, where possible (Bonferroni method, False Discovery Rate by Benjamini and Hochberg, or similar). Statistical uncertainty of estimates will be expressed as 95% CIs.

Power calculation

For each substudy, we will calculate the power of the planned association analysis using appropriate tools, for example, the Genetic Association Study Power Calculator (https://csg.sph.umich.edu/abecasis/gas_power_calculator).[59] We will perform power calculations that are appropriate for the intended analyses such as GWAS, EWAS or candidate–gene association studies. We will use for the different approaches the expected number of variants after filtering and define the adjusted cut-off p value appropriate for multiple testing. We will estimate the sample size taking into account the outcome of interest incidence, the expected relative risk for possible risk variants and minor allele frequency cut-off values using different models (dominant, additive, recessive, where appropriate). For the subproject on hearing loss, we have collected germline DNA from 426 survivors. Data collection and cleaning for the outcome measure (audiograms) is ongoing. We calculated sufficient power to detect a variant with a relative risk of 2.5 in an exome-wide association analysis using a dominant model.

In silico and in vitro analyses

We will use computational (in silico) tools to estimate the effect of specific gene variants on gene regulation, splicing and expression of proteins (eg, PolyPhen, SIFT, Human Splicing Finder, Matinspector).[60–63] We will

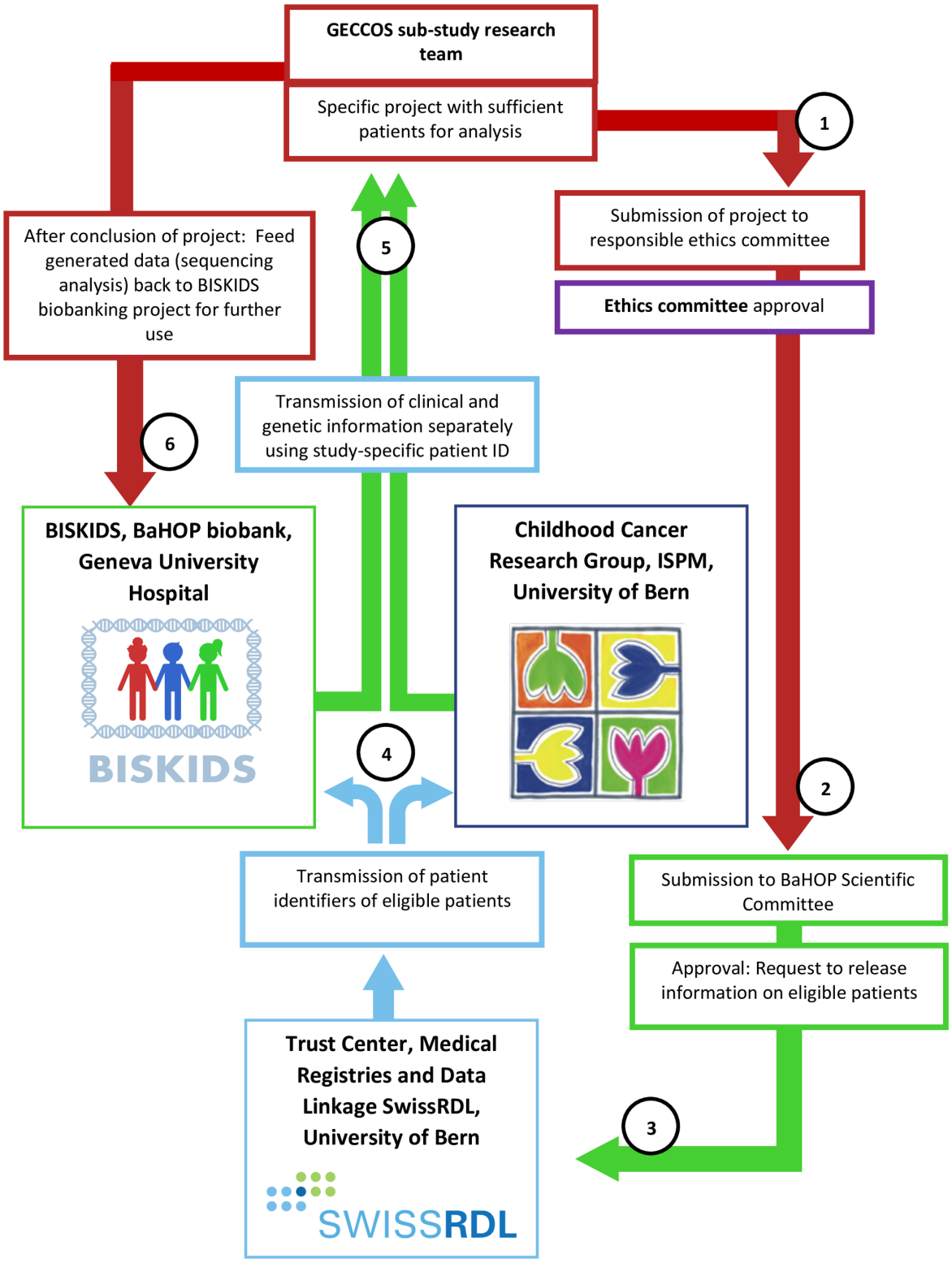

Figure 3 Flow diagram of release of information from the different resources. (1) Submission of subproject to the responsible ethics committee, either as amendment to the main protocol of Genetic Risks for Childhood Cancer Complications Switzerland (GECCOS) or as a separate project; (2) submission of subproject to the scientific committee of the Biobank for Research in Haematology and Oncology (BaHOP), host biobank for the Germline DNA Biobank Switzerland for Childhood Cancer and Blood Disorders (BISKIDS) section at the Geneva University Hospital; (3) release of linking information from Swiss Medical Registries and Data Linkage (SwissRDL); (4) transmission of BISKIDS and Swiss Childhood Cancer Registry (SCCR) identifiers for included patients with a newly generated study-specific patient-identifier to BISKIDS and the Institute of Social and Preventive Medicine (ISPM) respectively, to release variables used for the study; (5) transfer of data for included participants with a data transfer agreement to the researcher with study-specific identifier; (6) after conclusion of the project: storage of acquired germline genetic data in the BISKIDS biobanking database at the Geneva University Hospital for future research projects.

Table 2 Summary information on the genetic information used in the genotype–phenotype association analyses

Information	Data	Format
Germline genetic sequencing data	Raw sequencing data	FASTQ or BAM files
	Identified genetic variants	VCF files
	Documentation of sequencing procedure and analyses performed to allow trackability of downstream analyses	CSV or TXT files to allow long-term readability
	Quality measures	Read depth, score to evaluate correct variant calling, etc.
Underlying genetic condition	Heritable underlying condition	Description of underlying condition
	Gene name of affected gene	HGVS name
	Identified causal variant	HGVS notation: genomic location, coding DNA and expected effect on protein; rs number if available

BAM, compressed binary file used to represent aligned sequences; CSV, comma separated values; DNA, deoxyribonucleic acid; FASTQ, text file containing the sequence data from the clusters that pass filter on a flow cell; HGVS, Human Genome Variation Society; rs, reference single nucleotide polymorphism identifier; TXT, standard unformatted text document; VCF, variant call format.

choose in silico methods to identify genetic variants associated with complex disease mechanisms. Examples of such models are clustering methods including similarity network fusion[42] and PEGASUS[64] and deep learning methods to explore interactions between genes and outcomes of interest.[65] We will use multiple outcomes in combination with multiple treatment exposures in suitable models to test their association with genetic variant data and identify genetic variants associated with multiple complications.[66 67]

We will perform in vitro analyses using cell culture models relevant to the outcomes of interest. We will treat cell lines with antineoplastic agents or irradiate them and then perform transcriptomic analyses to identify genes that are differentially expressed after exposure to specific treatment modalities. We will use adapted approaches depending on the outcomes of interest to prioritise genes for further use in genotype–phenotype association studies. Through differential analysis of change in gene expression after treatment exposure, we will seek to identify candidate genes for further use in association studies. In hypothesis-free analysis methods (WGS or WES) we will perform in vitro studies to clarify the biological function of identified genes and genetic variants.[60 61]

Validation and replication

We will seek to validate variants identified by next-generation sequencing that were associated with the respective outcome of interest using a different method (eg, Sanger sequencing or real-time PCR). After successful validation, we will seek to proceed to replication of identified variants. We will reach out to independent cohorts of childhood cancer patients and survivors containing similar outcome information as analysed in the primary dataset such as the SJLIFE cohort or the French LEA cohort.[39 40] We will assess the power to identify an association in the replication cohort (using minor allele frequency of the identified variant in the respective population, the suspected effect size and sample size of the cohort). If patient numbers are deemed insufficient, we will consider pooling of data from multiple cohorts.

DISCUSSION

The GECCOS study will enable genotype–phenotype association studies focusing on various health conditions in childhood cancer patients and survivors. The large interindividual variability in response to antineoplastic treatments and occurrence of early and late complications is currently addressed mainly in a trial-and-error approach, that is, by delaying and adjusting treatment after occurrence of complications. Follow-up care is stratified by treatment exposure but not by genetic predictors.[68] The advantage of germline genetic risk variants is that this knowledge can be assessed when the workup of the neoplasm is made and then used early in the course of the treatment as they do not change over time. Such knowledge would allow to personalise treatment and follow-up care for individual patients before clinical signs of complications are present. Knowledge of genetic variants associated with treatment response will help maximising treatment effect while reducing the risk for complications and finding the balance of treatment intensity in the light of increasing survival in childhood cancers. Genetic predictors will improve individual counselling of patients and their families and help developing individualised follow-up guidelines.

Assessing multiple outcomes taking into account multiple covariates including treatment exposures will help identify particularly vulnerable patients. As identified in previous research, many patients suffer from several complications after childhood cancer treatment.[4] Finding genetic variants associated with increased risk for multiple health conditions will help identifying gene variants that contribute to several organ system complications. This approach might help identifying patients which could most benefit from treatment adaptation and preventive measures to reduce complications.

GECCOS provides a legal and organisational platform on how to use sensitive genetic data with clinical information in association studies. It establishes structures that can be used by researchers for national and international collaborative studies. Germline genetic sequencing data generated in the GECCOS substudies will be stored after completion in BISKIDS hosted within the BaHOP biobank, Geneva University Hospital. Clinical data will remain in the described databases and only be temporarily linked for research studies increasing data safety. The populations of interest will overlap between substudies and sequenced datasets generated from participants included in completed sub-studies will contribute with their germline genetic data to subsequent studies. This growing resource will reduce costs for future studies, where only DNA of a fraction of the participants will have to be sequenced. We will favour whole genome and whole exome sequencing to create datasets that are not restricted to a specific research question but can be used for further research. We will then be able to address different questions with the same datasets. A further strength of GECCOS is the availability of a large clinical dataset collected since 1976, with curated and regularly updated follow-up information and survival data.

GECCOS will be limited due to the fact that Switzerland is a small country with a limited number of possible participants, which we can recruit for our research despite the nationwide and population-based sampling. This will be counteracted by international collaborations. Collection of germline DNA in survivors was done in a first subset of participants. The collection is subject to survivor bias and omission of patients who died before they could be invited to germline DNA collection. This would lead to selection of patients with less severe phenotypes. We included a second stream of collection of samples from newly diagnosed patients through participating hospitals to include all childhood cancer patients early after diagnosis. Another issue with research on childhood cancer complications is that these health conditions are complex diseases with likely many mechanisms leading to a specific outcome. This makes identification of specific gene variants difficult. Many findings from studies were not replicated in independent datasets. Candidate-gene studies have particularly suffered from this. We will also explore novel methods to cluster and associate gene variants with clinical outcomes.

To overcome the issue of many previous studies using small sample sizes, the GECCOS study recruits participants from the nationwide SCCR with more than 8000 childhood cancer patients and survivors. This large base cohort will allow selection of specific treatment exposures to create homogenous samples for specific genotype–phenotype associations. Several studies are planned or ongoing to assess long-term complications in a standardised way in Switzerland (eg, cardiac outcomes)[69] which will improve quality of outcome assessments that can be used for the GECCOS study. Outcomes for health conditions that have been less investigated, like pulmonary complications, are also being collected and will be used in the GECCOS study.[70]

Our workflow combining a large dataset of clinical information with germline genetic data will enable genetic research on patient populations within Switzerland and facilitate collaborations with other research groups. As all childhood cancers are rare diseases by definition of the WHO with less than 1 in 2000 people being affected, patient numbers are generally small. Research on rare childhood cancer subtypes or specific rare outcomes is only possible through international collaborations. We will provide a platform for these collaborations with the GECCOS study.

PATIENT AND PUBLIC INVOLVEMENT

We did not formally involve patients and members of the public in the design of the study. We are collecting individual feedback through the available telephone hotline, and the dedicated study email address at various stages of the study, including the recruitment to the BISKIDS germline DNA sampling, presentations of the project at national meetings with patients, survivors and patient advocate groups, and presentations of results.

ETHICS AND DISSEMINATION

The Geneva Cantonal Commission for Research Ethics has approved the GECCOS study (approval 2020-01723), and the BaHOP biobank (approval PB_2017-00533).

Research findings will be disseminated via national and international conferences and publication in peer-reviewed journals. For the lay audience, patients and survivors, we will translate research findings into lay language and publish them on freely available websites of the research groups involved in the research, websites of patients advocacy groups and social media.

PROTOCOL AND DATA AVAILABILITY

The protocol, as approved by the ethics committee (currently V.1.0, 28 May 2020), is available on request to the corresponding author of this manuscript. We plan to publish results in open access journals. Data will be available on request to the principal investigator due to the rarity of certain diagnoses and information that could compromise the privacy of research participants.

Author affiliations
[1]CANSEARCH Research Platform for Paediatric Oncology and Haematology, Department of Pediatrics, Gynecology and Obstetrics, Faculty of Medicine, University of Geneva, Geneva, Switzerland
[2]Childhood Cancer Research Group, Institute of Social and Preventive Medicine, University of Bern, Bern, Switzerland
[3]Graduate School for Cellular and Biomedical Sciences (GCB), University of Bern, Bern, Switzerland
[4]Division of Paediatric Oncology and Haematology, Department of Paediatrics, Inselspital, Bern University Hospital, Bern, Switzerland
[5]Graduate School for Health Sciences (GHS), University of Bern, Bern, Switzerland

[6]Division of Paediatric Oncology and Haematology, Department of Women, Children, and Adolescents, University Hospitals of Geneva, Geneve, Switzerland
[7]Division of Oncology-Hematology, Department of Pediatrics, Kantonsspital Aarau AG, Aarau, Switzerland
[8]SwissRDL - Medical Registries and Data Linkage, Institute of Social and Preventive Medicine, Universitat Bern, Bern, Switzerland

Acknowledgements We thank all childhood cancer patients, survivors and families for participating in our study. We thank the study team of the Childhood Cancer Research Group, Institute of Social and Preventive Medicine, University of Bern, and SCCSS (Fabien N. Belle, Christina Schindera, Tomas Slama), the data managers of the Swiss paediatric oncology clinics (Claudia Althaus, Nadine Assbichler, Pamela Balestra, Heike Baumeler, Nadine Beusch, Sarah Blanc, Pierluigi Brazzola, Susann Drerup, Janine Garibay, Franziska Hochreutener, Monika Imbach, Friedgard Julmy, Eléna Lemmel, Heike Markiewicz, Annette Reinberg, Renate Siegenthaler, Astrid Schiltknecht, Beate Schwenke, and Verena Stahel) and the data managers and administrative staff of the SCCR (Meltem Altun, Erika Brantschen, Katharina Flandera, Elisabeth Kiraly, Verena Pfeiffer, Julia Ruppel, Ursina Roder and Nadine Lötscher). We thank the study team of the Research Platform for Paediatric Oncology and Haematology at the Geneva medical school (Khalil Ben Hassine, Simona Jurkovic Mlakar, Fanny Muet, Mary Khoshbeen-Boudal, Laurence Lesne, Vid Mlakar, Shannon Robin, Yoann Sarmiento) and the Onco-Haematology Unit of the HUG (Fanette Bernard, Laurent Cimasoni, Violaine Guignon, Nelly Hafner-Bénichou, Rodolfo Lo Piccolo).

Contributors NW: conceptualisation, design, methodology, writing-all stages, visualisation. SS: design, methodology, writing-reviewing and editing. TN and CRSU: methodology, writing-reviewing and editing. DM and VM: design, writing-reviewing and editing. MO, FGP, AOVB, FB and LM: writing-reviewing and editing. CK and AS: design, methodology, writing-reviewing and editing. MA: supervision, conceptualisation, design, methodology, writing-reviewing and editing. All authors have approved the submitted final version of the manuscript.

Funding This work was supported by the CANSEARCH Foundation for BISKIDS, the host biobank BaHOP, the research study GECCOS and salary support to Nicolas Waespe and Sven Strebel. Further funding comes from the Swiss National Science Foundation (31BL30_185396), and Swiss Cancer Research (KFS-4722-02-2019, KLS/KFS-4825-01-2019). The sponsor of the GECCOS study is the University Hospital Geneva, Rue Willy-Donzé 6, 1211 Geneva 4, Switzerland.

Disclaimer The sponsor bodies had no role in the design of the study and collection, analysis, and interpretation of data and in writing this manuscript.

Competing interests None declared.

Patient consent for publication Not applicable.

Provenance and peer review Not commissioned; externally peer reviewed.

ORCID iD
Nicolas Waespe http://orcid.org/0000-0002-2271-8959

REFERENCES

1 Bhatia S, Armenian SH, Armstrong GT, *et al*. Collaborative research in childhood cancer survivorship: the current landscape. *JCO* 2015;33:3055–64.
2 Nurgali K, Jagoe RT, Abalo R. Editorial: adverse effects of cancer chemotherapy: anything new to improve tolerance and reduce sequelae? *Front Pharmacol* 2018;9.
3 Hudson MM, Ness KK, Gurney JG, *et al*. Clinical ascertainment of health outcomes among adults treated for childhood cancer. *JAMA* 2013;309:2371–81.
4 Bhakta N, Liu Q, Ness KK, *et al*. The cumulative burden of surviving childhood cancer: an initial report from the ST Jude lifetime cohort study (SJLIFE). *The Lancet* 2017;390:2569–82.
5 Gibson TM, Mostoufi-Moab S, Stratton KL, *et al*. Temporal patterns in the risk of chronic health conditions in survivors of childhood cancer diagnosed 1970–99: a report from the childhood cancer Survivor study cohort. *Lancet Oncol* 2018;19:1590–601.
6 Schindler M, Spycher BD, Ammann RA, *et al*. Cause-specific long-term mortality in survivors of childhood cancer in S witzerland: A population-based study. *Int. J. Cancer* 2016;139:322–33.
7 Fidler-Benaoudia MM, Oeffinger KC, Yasui Y, *et al*. A comparison of late mortality among survivors of childhood cancer in the United States and United Kingdom. *J Natl Cancer Inst* 2021;113:562–71.
8 Winther JF, Kenborg L, Byrne J, *et al*. Childhood cancer Survivor cohorts in Europe. *Acta Oncol* 2015;54:655–68.
9 Guy GP, Berkowitz Z, Ekwueme DU, *et al*. Annual economic burden of productivity losses among adult survivors of childhood cancers. *Pediatrics* 2016;138:S15–21.
10 McLeod HL, Coulthard S, Thomas AE, *et al*. Analysis of thiopurine methyltransferase variant alleles in childhood acute lymphoblastic leukaemia. *Br J Haematol* 1999;105:696–700.
11 Zhou H, Li L, Yang P, *et al*. Optimal predictor for 6-mercaptopurine intolerance in Chinese children with acute lymphoblastic leukemia: NUDT15, TPMT, or ITPA genetic variants? *BMC Cancer* 2018;18:516.
12 Schaeffeler E, Jaeger SU, Klumpp V, *et al*. Impact of NUDT15 genetics on severe thiopurine-related hematotoxicity in patients with European ancestry. *Genetics in Medicine* 2019;21:2145–50.
13 Dulucq S, St-Onge G, Gagné V, *et al*. Dna variants in the dihydrofolate reductase gene and outcome in childhood all. *Blood* 2008;111:3692–700.
14 Ceppi F, Gagné V, Douyon L, *et al*. DNA variants in *DHFR* gene and response to treatment in children with childhood B ALL: revisited in AIEOP-BFM protocol. *Pharmacogenomics* 2018;19:105–12.
15 Ansari M, Curtis PH-D, Uppugunduri CRS, *et al*. Gsta1 diplotypes affect busulfan clearance and toxicity in children undergoing allogeneic hematopoietic stem cell transplantation: a multicenter study. *Oncotarget* 2017;8:90852–67.
16 Bhatia S. Role of genetic susceptibility in development of treatment-related adverse outcomes in cancer survivors. *Cancer Epidemiol Biomarkers Prev* 2011;20:2048–67.
17 Clemens E, van der Kooi ALF, Broer L, *et al*. The influence of genetic variation on late toxicities in childhood cancer survivors: a review. *Crit Rev Oncol Hematol* 2018;126:154–67.
18 Xu H, Robinson GW, Huang J, *et al*. Common variants in ACYP2 influence susceptibility to cisplatin-induced hearing loss. *Nat Genet* 2015;47:263–6.
19 Thiesen S, Yin P, Jorgensen AL, *et al*. Tpmt, COMT and ACYP2 genetic variants in paediatric cancer patients with cisplatin-induced ototoxicity. *Pharmacogenet Genomics* 2017;27:213–22.
20 Vos HI, Guchelaar H-J, Gelderblom H, *et al*. Replication of a genetic variant in ACYP2 associated with cisplatin-induced hearing loss in patients with osteosarcoma. *Pharmacogenetics and Genomics*2016;26:243–7.
21 Clemens E, Broer L, Langer T, *et al*. Genetic variation of cisplatin-induced ototoxicity in non-cranial-irradiated pediatric patients using a candidate gene approach: the International PanCareLIFE study. *Pharmacogenomics J* 2020;20:294–305.
22 Singh P, Wang X, Hageman L, *et al*. Association of *GSTM1* null variant with anthracycline-related cardiomyopathy after childhood cancer—A Children's Oncology Group ALTE03N1 report. *Cancer* 2020;126:4051–8.
23 Wang X, Sun C-L, Quiñones-Lombraña A, *et al*. *CELF4* Variant and Anthracycline-Related Cardiomyopathy: A Children's Oncology Group Genome-Wide Association Study. *JCO* 2016;34:863–70.
24 Wilson CL, Liu W, Yang JJ. Genetic and clinical factors associated with obesity among adult survivors of childhood cancer: a report from the ST. *Jude Lifetime Cohort Cancer* 2015;121:2262–70.
25 van Waas M, Neggers SJCMM, Uitterlinden AG, *et al*. Treatment factors rather than genetic variation determine metabolic syndrome in childhood cancer survivors. *Eur J Cancer* 2013;49:668–75.
26 Romerius P, Giwercman A, Moëll C, *et al*. Estrogen receptor α single nucleotide polymorphism modifies the risk of azoospermia in childhood cancer survivors. *Pharmacogenet Genomics* 2011;21:263–9.
27 van Dorp W, van den Heuvel-Eibrink MM, Stolk L, *et al*. Genetic variation may modify ovarian reserve in female childhood cancer survivors. *Hum Reprod* 2013;28:1069–76.

28 Karol SE, Mattano LA, Yang W, *et al*. Genetic risk factors for the development of osteonecrosis in children under age 10 treated for acute lymphoblastic leukemia. *Blood* 2016;127:558–64.
29 Fernandez CA, Smith C, Yang W, *et al*. Genome-Wide analysis links NFATc2 with asparaginase hypersensitivity. *Blood* 2015;126:69–75.
30 Chen S-H, Pei D, Yang W, *et al*. Genetic variations in GRIA1 on chromosome 5q33 related to asparaginase hypersensitivity. *Clin Pharmacol Ther* 2010;88:191–6.
31 Diouf B, Crews KR, Lew G, *et al*. Association of an inherited genetic variant with Vincristine-Related peripheral neuropathy in children with acute lymphoblastic leukemia. *JAMA* 2015;313:815–23.
32 Ansari M, Petrykey K, Rezgui MA, *et al*. Genetic susceptibility to hepatic sinusoidal obstruction syndrome in pediatric patients undergoing hematopoietic stem cell transplantation. *Biol Blood Marrow Transplantation* 2020;26:920–7.
33 Langer T, Clemens E, Broer L, *et al*. Usefulness of current candidate genetic markers to identify childhood cancer patients at risk for platinum-induced ototoxicity: results of the European PanCareLIFE cohort study. *Eur J Cancer* 2020;138:212–24.
34 Morton LM, Sampson JN, Armstrong GT, *et al*. Genome-Wide association study to identify susceptibility loci that modify radiation-related risk for breast cancer after childhood cancer. *J Natl Cancer Inst* 2017;109.
35 Hosking FJ, Feldman D, Bruchim R, *et al*. Search for inherited susceptibility to radiation-associated meningioma by genomewide SNP linkage disequilibrium mapping. *Br J Cancer* 2011;104:1049–54.
36 Knight JA, Skol AD, Shinde A, *et al*. Genome-wide association study to identify novel loci associated with therapy-related myeloid leukemia susceptibility. *Blood* 2009;113:5575–82.
37 Hawkins M, Bhatia S, Henderson TO. Subsequent primary neoplasms: risks, risk factors, surveillance, and future research. *Pediatr Clin North Am* 2020;67:1135–54.
38 Zidane M, Cazier J-B, Chevillard S, *et al*. Genetic susceptibility to radiation-related differentiated thyroid cancers: a systematic review of literature. *Endocr Relat Cancer* 2019;26:R583–96.
39 Howell CR, Bjornard KL, Ness KK, *et al*. Cohort profile: the St. Jude lifetime cohort study (SJLIFE) for paediatric cancer survivors. *Int J Epidemiol* 2021;50:39–49.
40 Berbis J, Michel G, Baruchel A, *et al*. Cohort profile: the French childhood cancer Survivor study for leukaemia (Lea cohort). *Int J Epidemiol* 2015;44:49–57.
41 Opstal-van Winden AWJ, de Haan HG, Hauptmann M, *et al*. Genetic susceptibility to radiation-induced breast cancer after Hodgkin lymphoma. *Blood* 2019;133:1130–9.
42 Wang B, Mezlini AM, Demir F, *et al*. Similarity network fusion for aggregating data types on a genomic scale. *Nat Methods* 2014;11:333–7.
43 Mezlini AM, Goldenberg A. Incorporating networks in a probabilistic graphical model to find drivers for complex human diseases. *PLoS Comput Biol* 2017;13:e1005580.
44 Michel G, von der Weid NX, Zwahlen M. The Swiss childhood cancer registry: rationale, organisation and results for the years 2001-2005. *Swiss Med Wkly* 2007;137:502–9.
45 Schindler M, Mitter V, Bergstraesser E. Death certificate notifications in the Swiss childhood cancer registry: assessing completeness and registration procedures. *Swiss Med Wkly* 2015;145:w14225.
46 Le Conseil fédéral CS. RS 818.33 Loi fédérale du 18 mars 2016 sur l'enregistrement des maladies oncologiques (LEMO). Loi fédérale sur l'enregistrement des maladies oncologiques, 2020. Available: https://www.admin.ch/opc/fr/classified-compilation/20121618/index.html [Accessed 20 Nov 2020].
47 Kuehni CE, Rueegg CS, Michel G, *et al*. Cohort profile: the Swiss childhood cancer Survivor study. *Int J Epidemiol* 2012;41:1553–64.
48 Allen CE, Merad M, McClain KL. Langerhans-Cell histiocytosis. *N Engl J Med* 2018;379:856–68.
49 Working Group Report&NA;. International rules for multiple primary cancers (ICD-0 third edition). *Eur J Cancer Prev* 2005;14:307–8.
50 Ng A, Brock KK, Sharpe MB, *et al*. Individualized 3D reconstruction of normal tissue dose for patients with long-term follow-up: a step toward understanding dose risk for late toxicity. *Int J Radiat Oncol Biol Phys* 2012;84:e557–63.
51 Maraldo MV, Jørgensen M, Brodin NP, *et al*. The impact of involved node, involved field and mantle field radiotherapy on estimated radiation doses and risk of late effects for pediatric patients with Hodgkin lymphoma. *Pediatr Blood Cancer* 2014;61:717–22.
52 Gasic D, Rosenschöld PMaf, Vogelius IR, *et al*. Retrospective estimation of heart and lung doses in pediatric patients treated with spinal irradiation. *Radiother Oncol* 2018;128:209–13.
53 Waespe N, Strebel S, Marino D, *et al*. Predictors for participation in DNA self-sampling of childhood cancer survivors in Switzerland. *BMC Med Res Methodol* 2021;21.
54 Coman Schmid D, Crameri K, Oesterle S. SPHN - The BioMedIT Network: A Secure IT Platform for Research with Sensitive Human Data. *Stud Health Technol Inform* 2020;270:1170–4.
55 McKenna A, Hanna M, Banks E, *et al*. The genome analysis toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res* 2010;20:1297–303.
56 Roy S, Coldren C, Karunamurthy A. Standards and guidelines for validating next-generation sequencing bioinformatics pipelines: a joint recommendation of the association for molecular pathology and the College of American pathologists. *J Mol Diagn* 2018;20:4–27.
57 Purcell S, Neale B, Todd-Brown K, *et al*. PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet* 2007;81:559–75.
58 Marees AT, de Kluiver H, Stringer S, *et al*. A tutorial on conducting genome-wide association studies: quality control and statistical analysis. *Int J Methods Psychiatr Res* 2018;27:e1608.
59 Skol AD, Scott LJ, Abecasis GR, *et al*. Joint analysis is more efficient than replication-based analysis for two-stage genome-wide association studies. *Nat Genet* 2006;38:209–13.
60 Kumar P, Henikoff S, Ng PC. Predicting the effects of coding non-synonymous variants on protein function using the SIFT algorithm. *Nat Protoc* 2009;4:1073–81.
61 Adzhubei I, Jordan DM, Sunyaev SR. Predicting functional effect of human missense mutations using PolyPhen-2. *Curr Protoc Hum Genet* 2013;76.
62 Desmet F-O, Hamroun D, Lalande M, *et al*. Human splicing finder: an online bioinformatics tool to predict splicing signals. *Nucleic Acids Res* 2009;37:e67.
63 Cartharius K, Frech K, Grote K, *et al*. MatInspector and beyond: promoter analysis based on transcription factor binding sites. *Bioinformatics* 2005;21:2933–42.
64 Nakka P, Raphael BJ, Ramachandran S. Gene and network analysis of common variants reveals novel associations in multiple complex diseases. *Genetics* 2016;204:783–98.
65 Kalinin AA, Higgins GA, Reamaroon N, *et al*. Deep learning in pharmacogenomics: from gene regulation to patient stratification. *Pharmacogenomics* 2018;19:629–50.
66 Riley RD, Jackson D, Salanti G, *et al*. Multivariate and network meta-analysis of multiple outcomes and multiple treatments: rationale, concepts, and examples. *BMJ* 2017;358:j3932.
67 Ristl R, Urach S, Rosenkranz G, *et al*. Methods for the analysis of multiple endpoints in small populations: a review. *J Biopharm Stat* 2019;29:1–29.
68 Landier W, Skinner R, Wallace WH, *et al*. Surveillance for late effects in childhood cancer survivors. *JCO* 2018;36:2216–22.
69 Schindera C, Kuehni CE, Pavlovic M, *et al*. Diagnosing preclinical cardiac dysfunction in Swiss childhood cancer survivors: protocol for a single-center cohort study. *JMIR Res Protoc* 2020;9:e17724.
70 Otth M, Yammine S, Usemann J, *et al*. Longitudinal lung function in childhood cancer survivors after hematopoietic stem cell transplantation. *Bone Marrow Transplant* 2021;47.