Familial Risk of Hematologic Malignancies

a Twin Study

Clemmensen, Signe Bedsted

DOI:
10.21996/day7-6z07

Publication date:
2024

Document version:
Final published version

Citation for pulished version (APA):
Clemmensen, S. B. (2024). *Familial Risk of Hematologic Malignancies: a Twin Study* . [Ph.D. thesis, SDU]. Syddansk Universitet. Det Sundhedsvidenskabelige Fakultet. https://doi.org/10.21996/day7-6z07

Department of Public Health

Familial Risk of Hematologic Malignancies – a Twin Study

Signe Bedsted Clemmensen

PhD Thesis

SDU

May 2024

Academic advisors

Professor Jacob von Bornemann Hjelmborg
Epidemiology, Biostatistics and Biodemography
Department of Public Health
University of Southern Denmark

Associate professor Jonas Mengel-From
Epidemiology, Biostatistics and Biodemography
Department of Public Health
University of Southern Denmark

Professor Henrik Frederiksen
Department of Clinical Research
University of Southern Denmark
Department of Haematology
Odense University Hospital

Assessment committee

Associate professor Christel Nielsen
Division of Occupational and Environmental Medicine
Lund University

Professor Cecilia Ramlau-Hansen
Department of Epidemiology
Department of Public Health
Aarhus University

Professor Mso Gabriele Berg-Beckhof (Chair)
Health Promotion
Department of Public Health
University of Southern Denmark

Preface and acknowledgements

The work of this thesis was carried out October 2019 – May 2024 at the unit of Epidemiology, Biostatistics, and Biodemography, University of Southern Denmark and as part of the Danish Twin Research Center and the Nordic Twin Study of Cancer.

The project was supported by grants from the Nordic Cancer Union, the Department of Public Health, University of Southern Denmark, and the Faculty of Health Sciences, University of Southern Denmark.

I would like to express my gratitude to a number of people without whom the project would not have been possible.

First and foremost, to my main supervisor, Jacob von Bornemann Hjelmborg, thank you for taking me under your wing, for sharing your profound knowledge, and for providing me with the opportunity to become a researcher. I appreciate that I can always count on your support to make my own choices and your willingness to listen to whatever is on my mind.

To my co-supervisors, Jonas Mengel-From and Henrik Frederiksen, thank you for providing valuable inputs from epidemiologic and medical points of view. Likewise, to Jaakko Kaprio and Jennifer Harris, I am grateful for the opportunity to collaborate and learn from you.

I am also grateful for the thorough review and helpful comments provided by the assessment committee.

A special thanks to Ulrich Halekoh for your perpetual inquisitiveness and excellent humor.

To Lisbeth Aagaard Larsen and Kirsten Gauthier, thank you for the countless hours you have patiently spent providing guidance and making sure everything was in order.

To my colleagues at EBB, thank you for making every day at the office enjoyable. It has been a privilege to work with you and, more importantly, to relax and be a part of such a welcoming environment.

Last, but not least, thanks to my family for endless support.
Mikkel and Sofie, thank you for reminding me what is important in life.

Signe Bedsted Clemmensen, May 2024

List of papers in thesis

Paper 1 (Publication)

Clemmensen SB, Harris JR, Mengel-From J, Bonat WH, Frederiksen H, Kaprio J, Hjelmborg, JvB. Familial Risk and Heritability of Hematologic Malignancies in the Nordic Twin Study of Cancer. Cancers. 2021;13(12).

Paper 2 (Submitted manuscript)

Clemmensen SB, Frederiksen H, Mengel-From J, Heikkinen A, Kaprio J, Hjelmborg JvB. Novel epigenetic biomarkers for hematopoietic cancer found in twins. Submitted to Acta Oncologica, May 2024.

Paper 3 (Submitted manuscript)

Clemmensen SB, Mengel-From J, Kaprio J, Harris JR, Frederiksen H, Hjelmborg JvB. Tattooing is mainly cultural. A representative twin study of tattooing determinants. Submitted to Behaviour Genetics, May 2024.

Paper 4 (Unsubmitted manuscript)

Clemmensen SB, Mengel-From J, Kaprio J, Frederiksen H, Hjelmborg JvB. Tattoo ink exposure is associated with lymphoma and skin cancers – a Danish study of twins.

Table of contents

Introduction

This thesis is based on the work of four papers. An overview is provided aiming to describe the project on a more general level and cover elements of background material and statistical methods that were studied during the project, but not included in any of the papers.

The fundamental question that initiated this work stems from cancer epidemiology. Does tattoo ink increase risk of cancer? Some types of tattoo ink have been shown to contain suspected and known carcinogens. Tattoo ink is known to traverse from the skin through the lymphatic system to regional lymph nodes. We are concerned that it may induce hematopoietic cell abnormalities.

We conjecture that tattoo ink induce inflammation at deposit site that may eventually become chronic and increase risk of abnormal hematopoietic cell proliferation, especially lymphoma.

This is not straightforward to study since processes may be complex and initiate long before symptoms arise. We seek to approach this topic as follows:

1. To study etiology of hematologic malignancies through the Nordic Twin Cancer Study.
2. Epigenome-wide association twin study of hematologic malignancies.
3. To study underlying determinants of becoming tattooed using the Danish Twin Tattoo Cohort.
4. To study if tattoo ink is a risk factor for certain cancers including hematologic malignancies.

We believe having access to twin cohort data can provide invaluable insights: Firstly, twins may reveal underlying sources of variation, for instance familial factors. Secondly, the matched twin pair design has proven useful in identifying, with high validity, important risk factor associations.

We set out to understand underlying factors for hematologic malignancies, that is, to obtain insights into the genetic and environmental sources of variation of hematological malignancy risk using twins.

Further, to improve knowledge on feasible mechanisms bridging genetical sources and environmental exposures to development of these cancers, we set out to identify epigenetic markers. The matched twin pair design is key for this.

We then delve into the exposure of tattoo ink. The popularity of tattooing has increased greatly over the past decades, but which factors influence tattooing behavior? We establish the Danish Twin Tattoo Cohort to approach this question from an empirical point of view.

Finally, from the Danish Twin Tattoo Cohort we examine if there is an association between tattoo ink exposure and risk of hematologic malignancy and other types of cancer. We expect to obtain novel insights but recognize that detection requires both long follow-up time and many events. This will be an important contribution to a field of limited information. We have a duty to report any results hinting towards association so measures can be taken to obtain further knowledge.

Background

Biology of hematologic malignancies and related diseases

Human blood has four major components: 1) Plasma. It is mostly water but also contains various proteins, sugars, and fat particles. 2) Red blood cells (erythrocytes). Their essential role is to transport oxygen from the lungs to the rest of the body. 3) White blood cells (leukocytes). There are various types categorized as lymphocytes, monocytes, neutrophils, eosinophils, and basophils – all part of the immune response. 4) Platelets (thrombocytes). They help the blood to clot (1). All blood cells originate from the bone marrow. The process of blood cell formation is called hematopoiesis. A simplified overview of the hierarchical process of hematopoietic stem cell differentiation giving rise to all types of blood cells is depicted in Figure 1 (prepared with inspiration from Blecua et al. (2)).

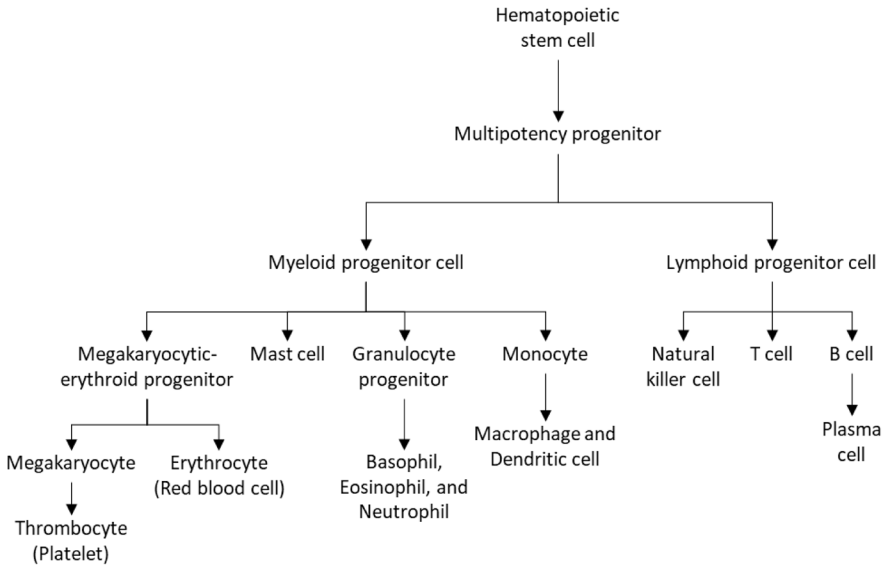


Figure 1. The classical hematopoietic hierarchy (simplified version). Development from hematopoietic stem cell to blood cells (2).

Disruption and misregulation of the processes governing hematopoiesis have the potential to lead to life-threatening hematological disorders including cancers (3). The main types of hematologic malignancies are lymphoma, leukemia, and multiple myeloma. Lymphomas develop from lymphocytes (T cells, B cells, and natural killer cells). There are a large number of subtypes based on biological features, but the overall classification into Hodgkin lymphoma (HL) and non-Hodgkin lymphoma (NHL), remains. Leukemia is characterized by an excessive production of abnormal white blood cells. Categorization is based on which type of white blood cells are involved, myeloid or lymphoid, as well as clinical course, acute or chronic, resulting in four main types: acute lymphoblastic leukemia (ALL), chronic lymphocytic leukemia (CLL), acute myeloid leukemia (AML), and chronic myeloid leukemia (CML). There are also other, less common, types of leukemia. Multiple myeloma (MM) develops from neoplastic plasma cells that when normally functioning produces antibodies (3).

Besides the three main types of hematologic malignancies, there are other hematopoietic diseases such as myelodysplastic syndromes, characterized by disorderly and ineffective hematopoiesis, and myeloproliferative diseases characterized by overproduction of blood cells (3). *Hematopoietic malignancies* is a commonly used term to describe a combination group including any of these cancers.

Descriptive epidemiology of hematologic malignancies and related diseases

Hematopoietic malignancies constitute 8.4% (females) and 9.7% (males) of all cancers diagnosed in the Nordic countries (excl. Faroe Islands and Greenland) in 2017-2021 (4, 5) and the risk of a diagnosis before age 75 was 2.6% for females and 3.7% for males. The number of new patients per year in that period was 7,078 for females and 9,207 for males. The overall incidence of hematopoietic malignancies has increased since the 1960s (Figure 2), mainly due to NHL, but seems to have reached a steady state (6). The estimated annual change in age-standardized incidence rate from 2011 to 2021 was 0.1% for males and 0.0% for females (4, 5).

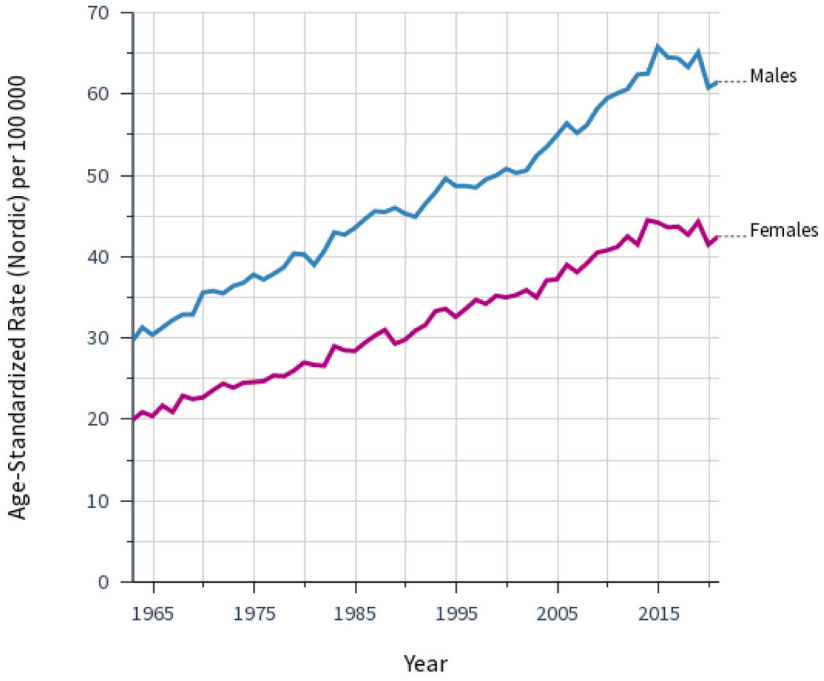


Figure 2. Age-Standardized Incidence Rate (Nordic) per 100,000 of malignant hematopoietic diseases by sex. Denmark, Finland, Norway, and Sweden. NORDCAN *(4, 5)*.

The most common hematopoietic malignancy in the Nordic countries is NHL which was diagnosed among 2,473 (females) and 3,234 (males) per year in 2017-2021. This is followed by leukemia with 1,556 (females) and 2,314 (males) new diagnoses per year and MM: 1,006 females and 1,331 males; myeloproliferative diseases: 1,097 females and 1,009 males; myelodysplastic syndrome: 407 females and 623 males; HL: 301 females and 416 males; other hematopoietic malignancies: 239 females and 279 males (4, 5). Most of the malignancies develop late in life but rarely some leukemias (ALL in particular) and especially HL are known to be pediatric.

The risk factors for these cancers remain to be fully elucidated and most leukemias and lymphomas are considered sporadic. Heterogeneity within hematologic malignancy subtypes seems to pose a challenge for epidemiological identification of risk factors. Additionally, the incidences vary with age, sex, geography, etc. (6). Known risk factors for lymphoma include immunodeficiency, autoimmune diseases, and viral infections such as Epstein-Barr virus, HIV and possibly hepatitis (7-9). For leukemia, some risk factors are chemo- and radiotherapy and other sources of ionizing radiation (AML) as well as Down's syndrome, common variable immunodeficiency and rare heritable stem cell deficiencies (ALL) (10). Radiation is also considered a risk factor of MM and AML. Besides age, sex, and geography there are only a few other established risk factors, one is obesity (11, 12).

Prognosis varies by subtype of malignancy, with poorer survival outcomes for adult ALL and better survival for HL (6). In the Nordic countries, hematopoietic malignancies constitute 8.5% and 9.9% of all cancer deaths for females and males, respectively (4, 5). Hemminki et al. (13) recently described survival trends in hematological malignancies (excluding ALL) in the Nordic countries from 1971 through 2020 using the NORDCAN database by the Association of the Nordic Cancer Registries. Improved treatment over the 50-year period is reflected in the relative survival that has improved with

20-50 percentage points. The lowest 1- and 5-year relative survival during 2015-2019 was found for myelodysplastic syndrome (approximately 20-85%) and AML (20-60%) while the highest was HL (85-95%), CLL (80-100%) and myeloproliferative disease (85-100%).

In this project, we focus on the main hematologic malignancies (HL, NHL, leukemia, and MM). When studying these malignancies using the Nordic Twin Study of Cancer, it was not possible to divide into further subcategories due to diagnostic inconsistencies between countries. Further details on disease coding and availability in the Nordic cancer registers can be found in Appendix 1 (in Danish only). Another reason it is difficult to use further subcategories in studies with long follow-up is that classification changed during the 1980s and 1990s with increasing understanding of cancer cell formation (13, 14). Besides, a study of Nordic twins has revealed that twin concordance for cancer often manifests across, rather than within, cancer types (15). It means that if one twin gets cancer, the co-twin has an increased risk of developing cancer, but not necessarily the same type of cancer. This suggests that, in spite of heterogeneity among subtypes, there are shared underlying factors that we may gain new information about through a twin-study of familial dependence in hematologic malignancies.

Twin studies of cancer

A recurring theme in this thesis is risk analysis in twins. Twins have contributed to essential understanding of disease occurrence in Denmark since the early 1950s. This through answering questions in surveys which have been repeated over longer periods of time. When the Danish Twin Registry was founded at University of Copenhagen in 1952, the original focus was to study causes of cancer (16). These studies have been continued and exists today in a Nordic context consisting of more than 350.000 twins having follow-up time since initiation of national cancer registries. A description of the Nordic twin cancer cohort along with study aims and main results is provided by Harris et al. (17).

Studying the patterns of cancer occurrence throughout life is a complex task as it is influenced by many varying factors such as biological inheritance and environmental exposures. There is variation in who gets cancer and when. In order to study this variation over time (spreading over several decades), it is important to consider censoring and competing events such as death. This will be explained further in the statistical methods section.

Studying twins can be particularly helpful in this endeavor, utilizing the inherent familial dependence structure of twins – namely, that monozygotic twins are genetically identical and dizygotic twins, on average, are as genetically alike as full siblings. The only requirement for standard methods of time-to-event analysis to apply in twin setting is that twin pairs are censored at the same time, which is rarely an issue in register-based studies. Through biometric modelling of variation in cancer risk within and between twin pairs, it is possible to assess relative influence of genes and environment (shared and non-shared) on said risk. This can provide hints on where to look for risk factors and, consequently, aid prevention as well as treatment. For instance, evidence from the Nordic Twin Study of Cancer points towards a very high heritability of prostate cancer: 58% (95% CI: 52-63) of variation in liability to develop prostate cancer can be explained by genetic effects (18). Another example is lung cancer. Biometric modelling in twins has enabled analysis of genetic contribution to risk of lung cancer and further, whether the genetic influence is modified by smoking and age. Evidence indicates decreasing familial influence with increasing age and does not support the long-standing hypothesis of interaction between genes and environment (19).

Besides studying sources of variation in risk, twins are optimal in case-control studies aiming to examine association between risk factors and disease. In pairs where one twin has been diagnosed with cancer and the other has not, who has been exposed to the risk factor of interest? The matched case-cotwin design (also referred to as discordant pair design) of monozygotic twins enables control for unobserved, shared confounding, including genetic background and shared environmental factors (e.g. childhood). A concern of this design is that only considers the pairs discordant for outcome event, e.g. cancer, but this discordance may be caused by non-shared confounders or because one twin simply has not experienced the event yet. The latter can be controlled for through time-to-event analysis, a relatively recent development in this context (20, 21). Both issues will be discussed further in the statistical methods section. An important example of also historic interest is where the matched case-cotwin design was applied, is the confirmation of smoking being a cause of lung cancer and other cancers (19).

In this project, we have applied such methods mainly to shed more light on potential risk factors of hematologic malignancies through time-to-event-analyses.

A brief introduction to epigenetics

Epigenetics can be described as cellular processes that affect gene expression or gene function without altering the underlying DNA and that are maintained through cell divisions (22). Epigenetic regulation of gene expression is mediated by a variety of molecular mechanisms such as histone modification and DNA methylation (23). DNA methylation is the addition of a methyl group to the fifth carbon of cytosine in a DNA strand. It regulates gene expression e.g. by repressing gene transcription. In the human cell, DNA methylation most often occur in cytosine guanine dinucleotides (CpGs) (22). Generally, increased methylation (hypermethylation) of CpGs, also referred to as CpG sites, in the gene body is associated with increased gene expression (24) while decreased methylation (hypomethylation) in promotor regions (a region of a gene where proteins bind to initiate transcription) is associated with gene repression (25).

Genome wide hypomethylation is a well-known hallmark of cancer, but also methylation levels of specific CpG sites have been associated with disease, e.g. NHL (26) and multiple myeloma (27), and lifestyle related phenotypes such as smoking (28) and alcohol consumption (29). Additionally, interaction of epigenetic markers leads to a complex network of signaling pathways (22).

Epigenetic variation between individuals is governed by environmental exposures, genes, and stochastic processes. MZ twins are practically genetically identical and share early life environment (8). Thus, analysis of epigenetic markers (e.g. DNA methylation) within MZ twin pairs provides a unique opportunity for investigating the influence of non-shared environmental factors occurring later in life since genetic and shared environment is controlled for. This is particularly useful when studying phenotypes with high heritability (22). Epigenetic studies of cancer etiology are of particular interest as abnormal cell proliferation might be traced to epigenetic changes. This would be an aim in our study.

An example of application of the matched case-cotwin design is a study analyzing DNA methylation at birth in 41 MZ twin pairs where one twin was diagnosed with pediatric acute lymphoblastic leukemia (pediatric ALL). In this study, a total of 240 CpG sites with FDR<0.05 were identified (30). Further, in a study by Svane et al. (31) DNA methylation CpG sites in association with human lifespan were studied. Certain markers and a predictor were proposed but could not be replicated in independent cohorts.

On tattoo ink and risk of cancer

In recent decades, tattoos have transitioned from being associated with low social status to being mainstream. The prevalence is up to 20-25% in some countries, even higher among the younger generations (32-34). There are several different types of tattoos. While the focus of this study is mainly the “classic” decorative tattoos, there is also permanent make-up (PMU) and medical tattoos used for reconstruction after surgery and for cancer treatment to ensure radiotherapy is applied to exactly the same area each time. Recently, tattoo ink has even been used as to assess patterns of vaccine drainage via the lymphatic system (35).

Many pigments used in tattoo ink were originally designed for industrial use or used to colorize food and consumer products. In the latter cases, toxicity is routinely tested concerning either oral or skin exposure, but never directly through injection into the skin. Until recently, substances banned from use in e.g. cosmetics, were not regulated in tattoo ink (36, 37). In an analysis conducted from 2008 to 2013, the Swiss health authorities identified 27 different organic colorants – 12 of which were tested for use in contact with the human body (38).

The chemical hazards of tattoos can be categorized as acute toxicity, carcinogenicity, genotoxicity, mutagenicity, immunotoxicity, reproductive toxicity, and irritant properties (39). This study revolves around the potential association between tattoo ink exposure and risk of cancer.

Tattoo carcinogenicity

When it comes to the question of tattoo carcinogenicity, opinions are widely varying. One extreme is not to worry about toxicity beyond immediate adverse reactions, and in some cases allergies, because tattoos are considered a single, one-time exposure of insoluble, inert substance (37, 40, 41). As Schreiver et al. puts it,

> “[…] the putatively low prevalence of severe side effects has rather led to a “show-me-the-dead-bodies” mind set in the general public and of healthcare officials alike” (42)(p. 1763).

One such example is the following comment by Serup:

> "Despite tattooing being practised for decennia, the general experience and the medical literature indicated no tattoo-associated cancer risk and no birth defects." (40)(p. 2)

Indeed, there is no empirical evidence suggesting that tattoo ink exposure could be associated with increased risk of cancer, however, it is important to point out, that until very recently, there were no studies in this field at all. In the literature, cancer occurrence (mainly skin cancer) among tattooed individuals is mostly considered coincidental due to low number of case reports (37, 43). When it comes to other cancers such as lymphoma, some researchers point out that since no such cases have been observed, there is nothing to worry about and argue that the body is able to encapsulate and render the foreign substances harmless (40). Given the long latency period of cancer and the possible combination of a variety of environmental exposures making it difficult to describe origin of disease development, the question remains: Is it fair to consider a lack of case reports evidence of no tattoo-associated cancer risk?

As opposed to beforementioned extreme, there are researchers arguing that a tattoo is a lifelong exposure. Indeed, reactions to tattoo ink have been observed to occur years after the tattoo was made (37). These investigators claim an urgent need for epidemiologic cohort studies (37, 42-49). This includes the International Agency for Research on Cancer (IARC), the European Commission, the Danish Environmental Protection Agency, and the Danish Cancer Society as well as researchers behind some of the most extensive studies in the area.

There is emphasis on the fact that ink particles with known or suspected carcinogenic properties have often been found to accumulate in regional lymph nodes (50-52) and they may be transported through the blood stream to other organs (43, 46, 49). For instance, a mice study found deposits of ink particles in the Kupffer cells of the liver (53). The question is whether they could cause any harm to internal organs – the potential seems to be there: Thus, it seems fitting to question the association to cancer, not only in the skin, but also elsewhere in the body (54).

The only published paper in the field so far is a Canadian study from 2020 reporting findings from two population-based case-control studies examining associations between tattoos and risk of non-Hodgkin lymphoma (NHL) and multiple myeloma (47). Among 1,518 participants (737 cases) in the NHL study and 742 (373 cases) in the multiple myeloma study, no statistically significant associations were found. They also reported no evidence of association when examining a "*broader association of tattoos and hematologic malignancies*" (47)(p. 2094).

In 2013, the German Federal Institute for Risk Assessment organized a conference about tattoo safety. One of the key papers within the field of tattoo research origins from a workshop held at this conference. It aimed to compose a review addressing, among other topics, the toxicological aspects of tattoo ink. The following quote nicely summarizes the duality of the question of tattoo ink carcinogenicity (as well as long-term tattoo safety in general):

> "Some have argued that the low solubility renders the respective pigments to be biologically unavailable; making them basically inert. Indeed, the persistence of tattoo colouring indicates that any metabolic processes are slow. Yet, low solubility is not a feature of all colourants and ink components and, with a lifelong deposit, even slow metabolism is relevant." (37)(p. 398)

Tattoo ink

Tattoo ink is mainly a mixture of pigments, suspensions (liquid carriers), preservatives and sometimes fragrances. Pigments are categorized as organic or inorganic. The former, typically azo- or polycyclic compounds, are most used as they generally have brighter colors and are cheaper to produce (36). Detailed product labels are required, but they are often lacking or downright incorrect as demonstrated by Wang et al. in 2021 (55). They reported that among 56 tattoo inks acquired in Sweden, 93% violated the European labelling requirements and 51% declared at least one pigment incorrectly. Besides the issue of mislabeling, the purity of tattoo ink is frequently less than 80%. The inorganic pigments may be polluted by heavy metals such as cadmium, lead, mercury and nickel, while the organic pigments may contain contaminants such as primary aromatic amines (PAAs) and polycyclic aromatic hydrocarbons (PAHs) (48). In addition to polluted ink as a source, beforementioned contaminants may also be formed inside the body through skin metabolism or light induced decomposition (36). Hauri and Hohl reports evidence of azo pigments releasing carcinogenic aromatic amines following exposure to sunlight or laser treatment tattoo removal (56). The toxicologic potential is increased by presence of nanoparticles as they can penetrate cell membranes, epithelium and cross the blood tissue barrier, and potentially enter different organs such as the liver, kidney, and lungs (36).

Background

The use of many hazardous heavy metals in tattoo ink has decreased in recent years but are still found in concentrations exceeding the recommended limits. A review by Kiszla et al. of studies quantifying restricted metals in tattoo ink reports that nearly every sample tested contained concentrations of chromium above allowed limits (57). Chromium-VI is classified as carcinogenic to humans by the International Agency for Research on Cancer (IARC) (58). A report on chemical substances in tattoo ink was performed on behalf of the Danish Environmental Protection Agency (DEPA) and includes results from chemical analysis of selected tattoo inks along with exposure scenarios and health risk assessments (43). In the following we provide a more in-depth description of PAHs, and refer to the study by Negi et al. (36) for an overview of other tattoo ink ingredients along with the DEPA report (43) and the IARC Monographs on the Identification of Carcinogenic Hazards to Humans for elaborate information on specific compounds (59).

The most frequently used tattoo color is black. Black tattoo inks typically contain soot products like carbon black which is listed as possibly carcinogenic to humans by IARC (60). As is the case with most concentration limits of various tattoo ink constituents, this suspicion is not based on human intradermal injection. Rather, it stems from the increased incidence of lung cancer after inhalation of carbon black and increased incidence of skin cancer among animals exposed to carbon black (39). Through the incomplete combustion used for carbon black production, PAHs are formed as byproducts. Part of the PAHs possibly stay lifelong in the skin where they can absorb UVA radiation from solar exposure and generate oxygen species, which may have adverse effects in the skin (61). Lehner it al. has demonstrated that substantial amounts of PAHs from black tattoo ink leaves the site of the tattoo and should thus be considered an additional effective source of PAH uptake in the human body besides inhalation, skin application and food (49).

The most dangerous PAH is benzo[a]pyrene (BaP) which is classified as carcinogenic to humans by the IARC (62). BaP, along with some other PAHs, can be metabolized by cytochrome P450 enzymes and in turn induce mutations of the KRAS oncogene and the TP53 tumor suppressor gene (36, 63). They can also have immunotoxic effects by impairing functional activation of lymphocytes (45). An acceptable limit for daily dose of BaP has been set to 0.6-5ng per kg body weight as that is estimated to correspond to cause 1 case of cancer per one million people. However, given the fluctuating amounts of BaPs and other PAHs in tattoo ink and the way an unknown proportion of it migrates from the tattoo site further into the body, it is extremely difficult to assess whether this limit is prone to be transgressed (43). It is unclear whether this “daily dose” refers to oral intake or other exposure routes as well, but most importantly, it may be different for tattoos where the ink particles accumulate in the lymphatic system over time.

The fading of tattoos is generally attributed to wound healing, chemical decomposition induced by exposure to sunlight or skin metabolism, and transport to other anatomical sites via blood and lymphatic system. It is difficult to find a good estimate for the decrease in pigment concentration. A study of tattoo ink in human and pig skin estimated the average amount of pigment injected to 2.53mg/cm^2, ranging from 0.60-9.42 mg/cm^2 (64). A mice study estimated that 30-60% of pigment leaves the tattoo within the first six weeks after tattooing (65). Another study claims a decrease estimate of 60-90%, however, they provide no explanation for the reasoning behind this number (46). In the following examples, we provide two scenarios estimating the amount of black tattoo ink injected in the body and PAH leaving the tattoo site. Similar calculations on amount of ink injected have been made by Sabboni et al. (66) and Jacobsen et al. (43). We have chosen 50% as a plausible value for the decrease in concentration.

A German study analyzed 19 commercially available tattoo inks and found the total concentration of PAH ranging from 0.14-201µg/g and 0.1-0.5µg/g for BaP. Note that these measures are based on ink suspensions, i.e., mixtures of pigments and liquid carriers. The average amount of pigment in ink is estimated to 54% (61).

Scenario 1: For a typical black, medium sized tattoo of 100 cm^2 with initial 2.5mg pigment per cm^2, the amount of ink injected is 460mg (corresponding to 250mg carbon black). Assuming a PAH concentration of 100µg/g and a decrease of 50%, the amount of PAH having left the tattoo is estimated to 23µg. Similarly, assuming a BaP concentration of 0.3µg/g, would amount to 69ng having left the tattoo site.

Scenario 2: For a larger black tattoo of 300 cm^2 (e.g. upper arm/shoulder) with initial 9.4mg pigment per cm^2, the amount of ink injected is 5.2g (corresponding to 2.8g carbon black). Assuming a PAH concentration of 200µg/g and a decrease of 50%, the amount of PAH having left the tattoo is estimated to 520µg. Similarly, assuming a BaP concentration of 0.5µg/g, would amount to 1.3µg having left the tattoo site.

The proportion of ink leaving the tattoo ending up in regional lymph nodes is uncertain. The distinct coloration of lymph nodes seen in case reports (50, 51) indicates it may be a significant proportion. Lehner et al. reported that the concentration

of 20 PAHs in the tattooed skin ranged from 0.1–0.6 mg/cm2 while the concentrations in the lymph nodes were 0.1–11.8 mg/g.

While PAH is most likely released in small amounts as a continuous exposure, regional lymph nodes function as accumulation points. Thus, it seems reasonable to contemplate the consequences of (at least a considerable proportion of) the estimated amounts of PAH in the lymph nodes.

As mentioned earlier, there are no guidelines for concentration limits of potentially harmful substances meant for intradermal injection. To put above estimated amounts of PAH in the body into perspective, we consider the main sources of exposure: smoking and diet. The estimated average amount of PAH in cigarette emissions is 474-1061ng per cigarette (67). A review of PAH occurrence in processed food reports of concentrations up to 1-2mg/kg in most extreme cases (smoked food) (68). While these measures are probably the best proxy we have for “uncommonly high” PAH exposure, it should be noted that they are not directly comparable to PAH from tattoos as the exposure routes and, likely, the effect of involved organs are different. For instance, how much of the PAH in cigarette emissions passes the air-blood barrier of the lungs?

Immunologic response

A mechanism behind the hypothesized association between tattoo ink and risk of cancer could be chronic immunologic response caused by tattoo ink deposited in regional lymph nodes ultimately leading to development of tumor cells. This is illustrated in Figure 3. Especially non-Hodgkin lymphoma seems a likely candidate as chronic immune stimulation is a known risk factor.

In analogy, one could consider breast implant-associated anaplastic large cell lymphoma (BIA-ALCL), a rare type of T-cell lymphoma, which was only recently discovered (first publications in 2008) even though breast implants have been used for decades. For this type of disease, clonal T-cell rearrangement and JAK-STAT signaling is often observed which is known to have the potential cause abnormal proliferation of tumor cells (69). The median time of BIA-ALCL diagnoses is 8-10 years after implantation. A similar latency period for tattoo ink related lymphatic malignancies seems plausible, supporting the claimed need for epidemiologic cohort studies.

Researchers are generally very cautious not to claim causal relations when studying association between tattoo ink and development of cancer. However, as far back as 1992, a paper was published on a so-called *tattoo-associated* case of non-Hodgkin lymphoma (NHL)(54). A 54-year-old male apparently developed a chronic immune response to a red tattoo. The authors postulate that, over a period of four years, a pseudolymphoma lead to formation of large B-cell monoclonal lymphoma (a type of NHL).

We are concerned that tattoo ink may induce hematopoietic cell abnormalities.

We conjecture that tattoo ink induce inflammation at deposit site that may eventually become chronic and increase risk of abnormal hematopoietic cell proliferation, especially lymphoma.

This is open for investigation. The present thesis may shed some light upon this.

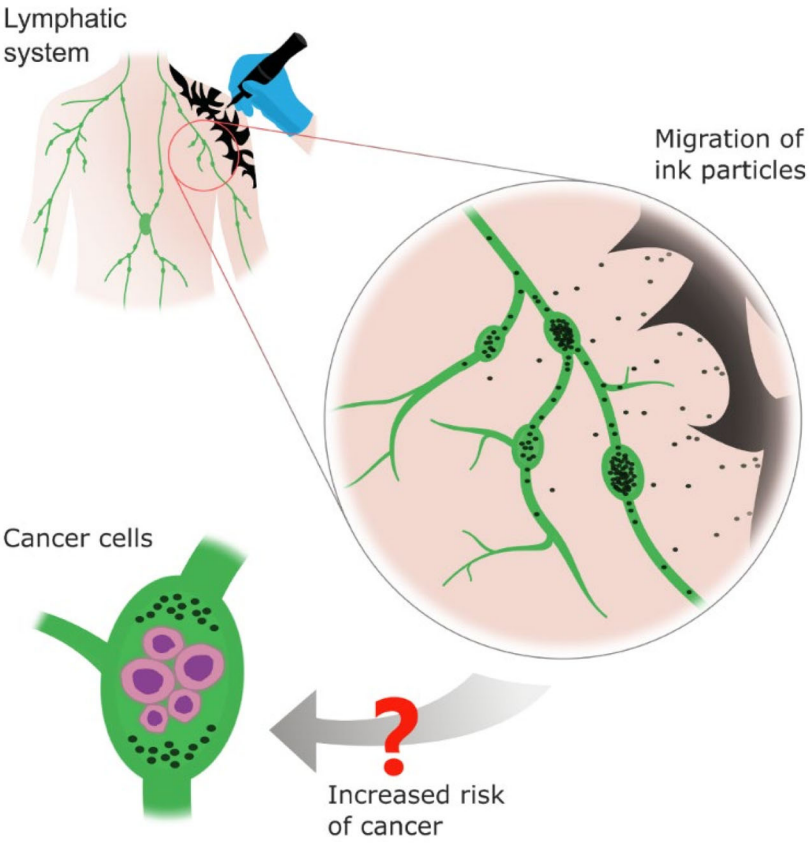


Figure 3. Illustration of hypothesis: Migration of tattoo ink particles to regional lymph nodes causing a chronic immunologic response which could ultimately lead to development of tumor cells.

Tattoo related complications

There are numerous, generally less severe, adverse reactions to tattoos. In the weeks following the breach of the skin, 1-5% are estimated to have infections ranging from relatively harmless to more serious cases such as blood-borne viruses like hepatitis B or C or HIV. Chronic inflammatory reactions have been observed to occur years after tattoo acquisition (37). Allergic reaction to red tattoo ink is one of the most frequent complications. Reviews and case reports provide information on findings such as squamous cell carcinoma, benign tumors, lymphoid conditions, and rare cases of malignant neoplasms occurring within the area of a tattoo (44, 70-74), indicating potential associations between tattooing and these conditions. A recent study by Bourgeois et al. showed unexpectedly high prevalence of abnormal lymphatic drainage under tattooed areas. Though the authors note that further research is required (75).

Hygiene plays an important role in this context. If the sanitary conditions are not adequate and the instruments used are not properly sterilized, there is a significant risk of being exposed to various infections. Informative, preventive initiatives have been made to reduce hygiene related hazards as well as help the costumer make an informed decision before getting a tattoo. There are various “Think Before You Ink” campaigns typically provided by health authorities (76, 77).

Another hazard of tattoo ink in the lymph nodes is misdiagnosis. Some examples are case reports of: i) a lymph node full of tattoo ink mistaken for malignant melanoma (52) and ii) tattoo pigment mimicking a positive sentinel lymph node in melanoma (78) and iii) lymph node calcification on mammography (79).

Delay in detection of suspicious birth mark alterations covered by tattoos is another issue. However, this is one of the well-known hazards that tattoo artists generally seem to be aware of and tend to preclude.

An extensive subject, only briefly touched upon in this paper, is laser tattoo removal. Through photo thermolysis, pigments are broken into smaller fragments that leave the site of the tattoo. The question is: where do they end up? Decreasing particle size often allows for greater migration potential. A review of degradation products formed when tattoo pigments

are irradiated with sun- or laser light is provided by Fraser et al (80). Some of the degradation products reported were hydrogen cyanide (81) and azo cleavage compounds such as 3,3'-dichlorobenzidine which has been shown to induce DNA strand break in human skin cells (82).

Tattoo ink regulations

Regulation of tattoo ink has been approached differently around the world. In most European countries, tattoo ink has been covered by the Regulation on the Registration, Evaluation, Authorization, and Restriction of Chemicals (REACH) and was, until recent years, regulated as any other chemical compound. It meant that there were no regulations specific to tattoo ink as known from the Cosmetic Product Regulation where e.g. azo pigments are banned for use in cosmetics because they may release carcinogenic aromatic amines (48). At the beginning of the 2000s, the safety of tattoo ink became a concern among various authorities. An overview of the development of the European regulatory framework is provided in the following list: (Sources: Negi et al. (36), Wang et al. (55))

2003 (revised in 2008): Council of Europe published resolution on requirements and criteria for safety of tattoos and permanent make-up (ResAP) (83, 84). Labeling of packages must adhere to chemical regulations set by the International Union of Pure and Applied Chemistry (IUPAC). Recommendations on maximum concentrations of certain impurities and ban on some potentially carcinogenic aromatic amines.

2015-2016: Extensive report by the Joint Research Centre of the European Commission stated that ResAP had been implemented to various degrees in ten European countries (48). Many of the colorants included in the negative lists by the European Council were found to be frequently used still.

2017-2019: Continued violation of regulations – ECHA (European Chemicals Agency) regularly found banned ingredients (85). A restriction proposal limiting more than 4,000 chemical substances used in tattoo ink and permanent make-up was submitted for evaluation by ECHA to the Committees for Risk Assessment and Socio-economic Analysis who passed the proposal on to the European Commission with only few modifications (86).

2020-2022: The European Commission adopted the new tattoo ink specific restrictions in REACH as of December 2020, allowing for a transition period so they came into force in January 2022. Applies to all member states of the European Union (87).

As the study in this thesis was based on Danish individuals, the more local regulatory initiatives should be mentioned: In 2010, the Swedish Chemicals Agency published a report finding that only 5 out of 31 tattoo inks were free of hazardous substances such as aromatic amines and metals above recommended limits (83). This led to Swedish legislation on tattoo ink covering, among other things, product information, importation, and usage of tattoo ink. A similar legislation proposal was made in Denmark in 2017, but it was rejected as it was decided to wait and see what came of the regulations proposed by ECHA (88).

It remains to be seen how well the latest legislation is adhered to. There have been complaints from tattoo artists finding the comprehensive ban of some of the most used pigments devastating as alternatives are not easy to find (89-91). Additionally, there is speculation as to whether it is even possible to check up on the new restrictions as few countries appear to have the equipment necessary for measuring concentrations on the small scales required (40).

Establishing the Danish Twin Tattoo Cohort

As part of this PhD project, the *Danish Twin Tattoo Cohort* was founded. The original aim of the cohort was twofold: 1) Development of a twin cohort aiming to study tattooing determinants and to characterize the prevalence of tattooing by various demographic- and lifestyle factors. 2) Analysis of the possible association between tattoo ink exposure and risk of certain types of cancer through a case-cotwin study. The process of compiling the cohort is described in the following.

Data collection: Questionnaire survey

The questionnaire was in Danish language and the translated title was *Risk factors of certain types of cancer diseases*. The first page of the questionnaire held a description of the project aims in layman's terms:

The aim of this survey is to study possible risk factors of certain types of cancer. We are, among other things, interested in knowing more about whether tattoo ink can be related to certain types of cancer such as skin and blood cancer. In the questionnaire, we ask about tattoos together with lifestyle and education.

The questionnaire included items about tattoo status and, when relevant, further tattoo details such as colors, size (measured in units of the size of the palm of one's hand), age at first tattoo, and medical issues such visits to the doctor/hospitals stays and swollen lymph nodes. In additions items about lifestyle factors; smoking (type and duration), physical exercise, alcohol consumption, and education. Finally, the questionnaire allowed written comments from participants. The total number of questions ranged from 11 to 23, depending on answers. The questionnaire (in Danish) and invitation letter is provided in Appendix 2.

The survey was initially developed for digital distribution via the public mail service e-Boks. Each participant would be directed to the SurveyXact webpage (an online tool for questionnaire surveys) through a unique link that could be used for identification by the data manager afterwards. This form of distribution allows the developer to guide the participant to a certain extent, e.g., only one question is shown at a time and certain restrictions can be made to allow single/multiple choices or only specific types of manual answers such as positive integers.

An overview of the survey timeline is provided in Figure 4. No pilot study was performed.

On January 21, 2021, the invitation was sent to 10,094 individual twins followed by a reminder on February 16 to 5,517 of those twins.

During the process when invitations were mailed using the public digital mailbox, e-Boks, we discovered that 1,033 (9%) of the survey population were unable to receive public digital mail. As this was a significant proportion mainly composed of the older twins, we decided to print and send invitations and questionnaires by letter mail. The participants could either fill out the printed version of the questionnaire and return by an enclosed envelope or use a link and unique code to access the online version. The invitation was nearly identical to the one distributed digitally. The letters were mailed over a period of a week starting on March 3, 2021, to 1,005 individuals (28 could not be reached). A reminder was sent to 688 twins on April 12 and the days following.

Distributing invitations by letter mail was not only time-consuming and costly, but also provided less of a guide for those participants replying through the printed questionnaire – a significant proportion either skipped questions or answered "incorrectly", e.g., by writing their own answers instead of using the ones provided in multiple choice questions.

The survey was closed on July 8, 2021.

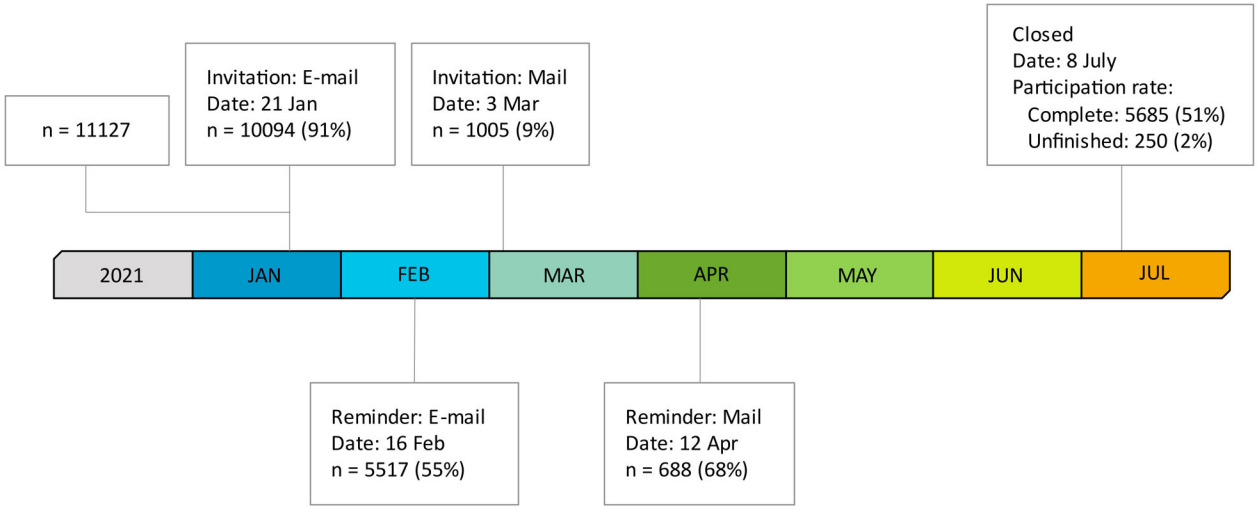


Figure 4. Survey timeline. Percentages in the top row refer to the initial 11,127 twins. The percentages in the bottom row refer to the proportions of the respective distribution forms.

Study population:

A flowchart of the survey study population is depicted in Figure 5.

For the case-cotwin study, we selected all twin pairs complying with the following conditions: 1) At least one twin could be contacted, i.e., was alive, not emigrated, and had not requested no contact. 2) At least one twin had one (or more) of the following cancer diagnoses: Hodgkin or non-Hodgkin lymphoma, multiple myeloma, liver, bladder/urinary tract, melanoma of skin or skin (non-melanoma). These cancers were selected based on suggestions from current literature and was discussed in the background section on tattoo carcinogenicity.

A total of individual 3,022 twins complied with these criteria out of which 2,246 individuals (including 735 complete pairs) could be contacted and were invited to participate in the survey.

A reference sample of 5,000 twin pairs was selected randomly among all pairs that could be contacted – after excluding the 3,022 pairs from the case-cotwin sample. 8,881 individual twins (including 3,881 complete pairs) were invited to participate in the survey.

The twin cohort was formed by combining the reference sample with pre-determined proportions of twins selected randomly from the case-cotwin study to make sure the cohort was representative of the general population. The proportions were defined by cancer type and sex and were based on readily available data on number of individuals alive with a cancer diagnosis per 100,000 in 2020 by sex. These data were provided by the *Association of the Nordic Cancer Registries* (ANCR) through the NORDCAN database (4, 5). The construction of the cohort is explained step-by-step in Appendix 3. The resulting cohort compiled 9,173 individual twins – 8,881 from the reference sample and 292 from the case-cotwin sample.

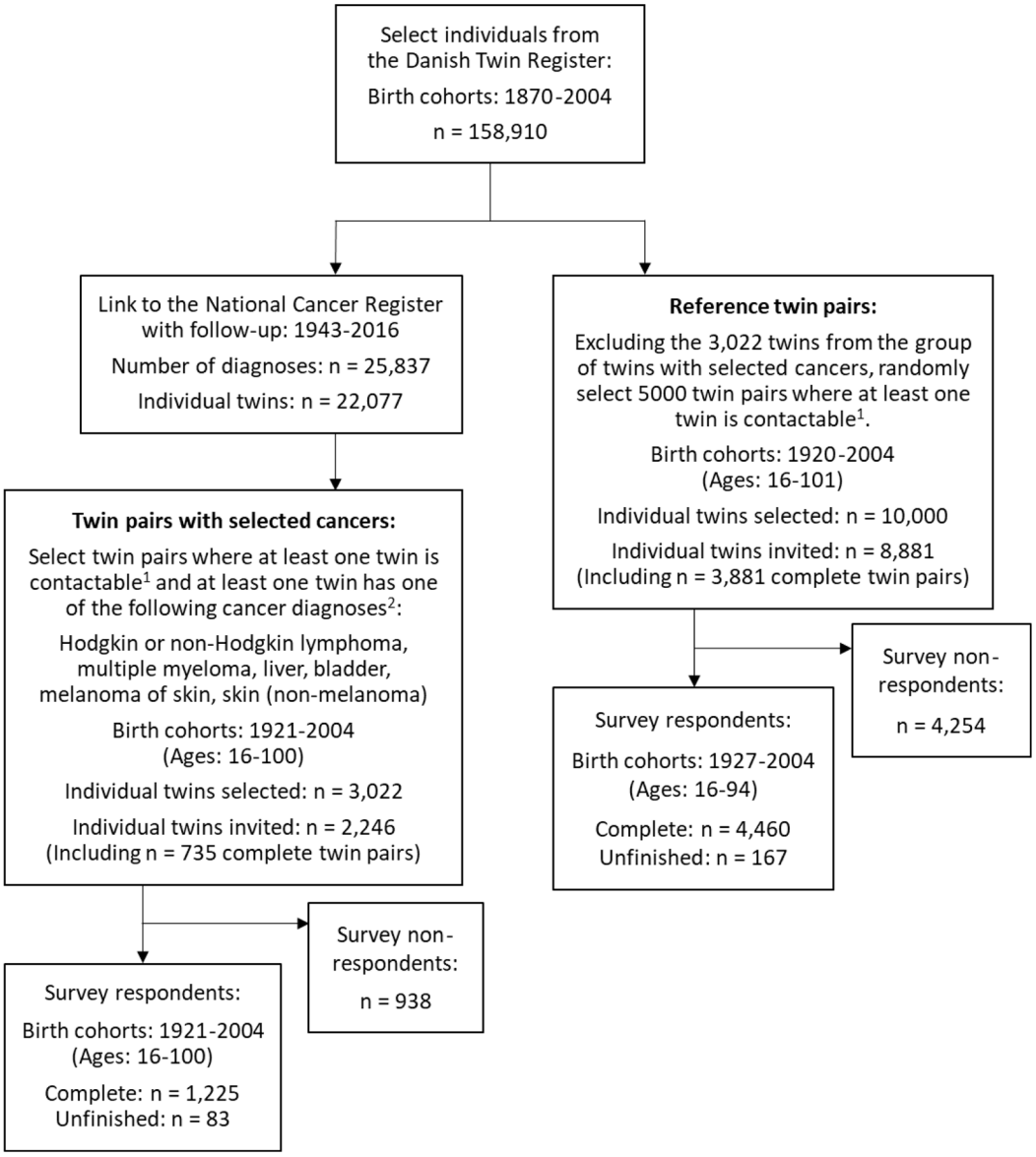


Figure 5. Selecting the study population for the questionnaire survey. Left: Case-cotwin sample. Right: Reference sample. [1]Alive, not emigrated, and has not requested no contact. [2]NORDCAN classification (4, 5).

Ethical statements and considerations

Details on statements of ethics and informed consent are described in each paper of the thesis. According to legislation in the Nordic countries, ethical review and approval are not needed for purely register-based studies. The appropriate register authorities have given permission for record linkages. Informed consents were obtained from all participants whose blood samples were used to obtain the DNA methylation data studied in Paper 2.

The Regional Committees on Health Research and Ethics for Southern Denmark were contacted to find out whether ethical approval was needed for the survey in the Danish Twin Tattoo Cohort, more specifically, for using information from the Danish Cancer Register to identify study subjects. They responded that no ethical was needed. Further, inviting a pair of twins to participate in a survey on the basis of one of them being diagnosed with cancer, would be a violation of the General Data Protection Regulation (GDPR), and unethical since the healthy twin may not be aware of the cotwin's diagnose. This issue was prevented by the inclusion of a random twin sample. The survey invitation letter included information on handling of personal data in accordance with the GDPR.

Regarding ethical considerations, this project involves no direct risks for participants. The only risk is illegitimate access to the databases. This risk is minimized since personal identification numbers will be removed before the data is accessed by the research group. All data studied in this project were stored on a secure server at SDU, and no files kept on personal computers. Furthermore, data are always presented at group level according to regulations and no individual register data will be transferred between other registers or linked beyond given approvals of involved registers.

In the development of our statistical models, bias and fairness is another ethical issue that must be addressed. A model can inherit biases from unobserved confounding factors, collider biases or simply mistakes by the statistician or from other sources. To mitigate this risk, we will use representative data with high coverage and quality, ensuring that all groups are included in the model development. Additionally, were possible, we will externally validate our models to ensure their generalizability.

Finally, one may question how to communicate potential findings. Particularly in relation to the study of tattoo ink exposure and risk of cancer. Should the results hint towards association, we have a duty to report this so measures can be taken to obtain further knowledge.

Aims

In this project, set out with the hypothesis that tattoo ink in lymph nodes may induce inflammation and eventually become chronic and increase risk of hematologic tumor cell development. This is not straightforward to study since processes may be complex and initiate long before symptoms arise. We therefore sought to approach the topic as follows:

Firstly, we aimed to study familial dependence among twins to get more knowledge on the increasing risk of hematologic malignancies. To shed more light on potential risk factors, we first set out to characterize hematologic malignancies in terms of genetic, environmental, and cross-cancer risk relationships varying in time using a population-based cohort of Nordic twins. The cross-cancer risk relationships could reveal mutual relations between cancer types and further describe the genetic and shared environmental influence.

Secondly, to improve knowledge on feasible mechanisms bridging genetical sources and environmental exposures to development of these cancers, we conducted an epigenome-wide association study of CpG methylation levels and their association with hematopoietic malignancies. DNA methylation is a mechanism that affects gene expression, and it is influenced by genes and environmental exposures. We aimed to identify epigenetic determinants of hematopoietic malignancies and, through application of the matched case-control design on monozygotic twins, to be able to describe to which extend genes and environment influences the methylation levels of the identified CpG sites.

Thirdly, set out to study the exposure of tattoo ink. The popularity of tattooing has increased greatly over the past decades. We aimed to explore this increase and identify underlying determinants of becoming tattooed. To approach this question from an empirical point of view, we set out to collect information on tattoo exposure and potential confounding lifestyle factors in a survey among Danish twins leading to establishment of the *Danish Twin Tattoo Cohort*.

Finally, tattoo ink has been hypothesized as a potential risk factor of certain cancer types, especially lymphoma, due to accumulation of ink particles with known or suspected carcinogenic properties in the lymphatic system. However, empirical evidence pertaining to carcinogenicity was extremely limited. Therefore, we aimed to study the association between tattoo ink exposure and risk of potentially related cancer types using the *Danish Twin Tattoo Cohort.*

Motivation

Utilizing twin zygosity provides a golden opportunity to study variation in risk of hematologic malignancies in terms of genetic and environmental influences and using the largest ever population-based cohort of twins allows for assessment of variation over time. Combined with epigenetic findings, this will expectedly be a substantial contribution to the field of these types of cancer.

We suspect that hematologic malignancies as well as other cancer types can arise due to certain tattoo ink exposure. Collecting and analyzing information on tattooing will be an original contribution to the field providing updated information on the increasing popularity of tattooing as well as empirical knowledge pertaining to the carcinogenicity of tattoo ink which has been called for by various health organizations.

Materials

The Danish Twin Register and the Danish Cancer Register

The Danish Twin Register (DTR) holds more than 175,000 twins born in 1870-2009 making it the oldest nationwide twin register in the world with complete ascertainment since 1968. As described by Pedersen et al. (92), comprehensive surveys conducted since 1994 on subsets of the DTR have provided detailed information on a wide range of health-related subjects such as physical performance, mental capability, lifestyle, and sociodemographic factors. Some surveys have also carried out clinical examinations providing biological material used to form the DTR biobank. This enables examination of genome-wide molecular data such as single nucleotide polymorphisms (SNPs), DNA methylation and gene expression (92).

Since 1968, each Danish citizen has been assigned a unique personal identification number (the Danish Civil Registration System) used in national administrative- and health registers.

Throughout this project, twin data from the DTR linked to the Danish Cancer Register (DCR) has been utilized in various ways to explore variation in the occurrence of hematologic malignancies. This provided decades of follow-up – from opening of the DCR in 1943 until 2016. The diagnoses were based on the International Statistical Classification of Diseases and Related Health Problems, Tenth Revision (ICD-10) and were grouped according to the NORDCAN classification system (5). When identifying lymphomas, the International Classification of Diseases for Oncology, Third Edition (ICD-O-3) was used along with ICD-10. Figure 6 provides an overview, including ICD-10 values, of the main cancer types studied in this project.

The Danish National Hematology Database holds specialized registers for various hematologic malignancies such as lymphoma, acute leukemia, and multiple myeloma. However, since the lymphoma registry did not become nationwide until 2000 and the other registers were founded afterwards, the DCR with its much longer follow-up time was preferable. Also, access to the DCR was readily available and while the specialized hematologic data bases hold more information on each patient, that level of detail was not needed.

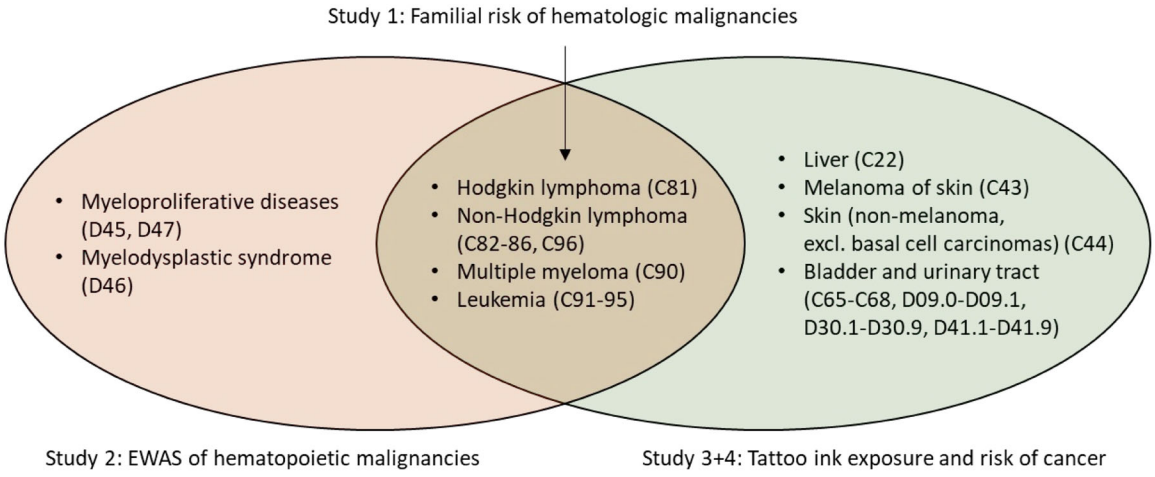


Figure 6. Overview of main cancer types studied in this project following the NORDCAN and ICD-10 classifications.

Besides Denmark, also Finland, Norway and Sweden have similar nationwide registers which have been combined to constitute the Nordic Twin Study of Cancer (NorTwinCan) (15, 17, 93, 94). Following the latest linkage update in 2018, the study included around 315,000 twins with known zygosity. The Lexis diagram in Figure 7 illustrates a rough outline of the NorTwinCan cohorts by country and cancer follow-up. The first study of this project utilized the NorTwinCan data to examine familial risk of the main hematologic malignancies: Hodgkin- and non-Hodgkin lymphoma, multiple myeloma, and leukemia. During the follow-up period, 3,459 hematologic malignancies were diagnosed.

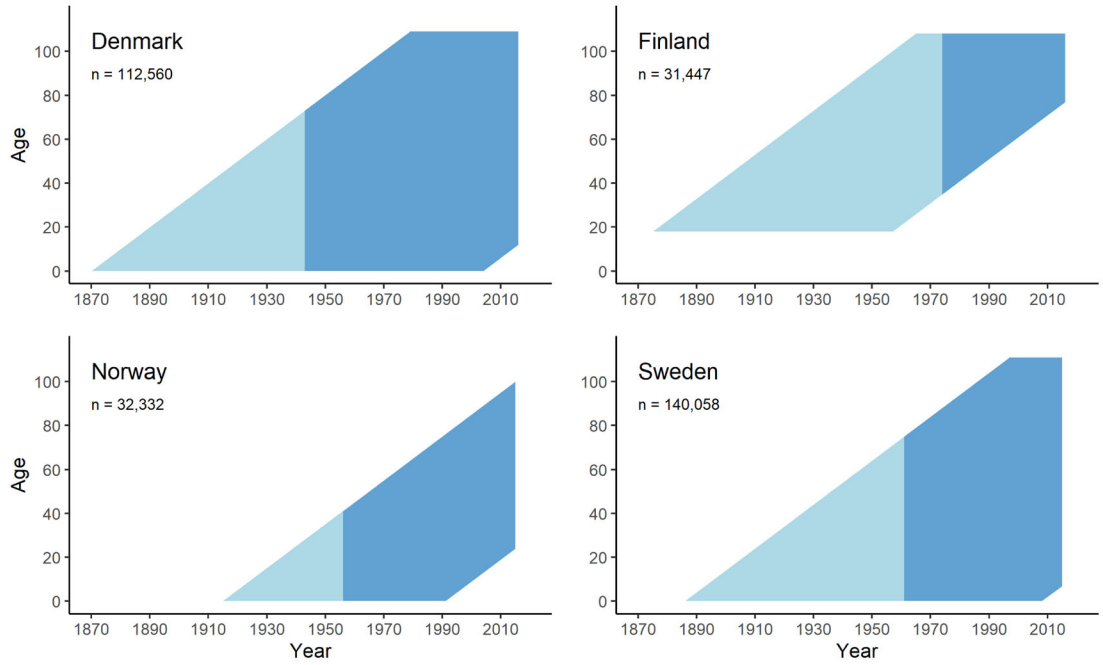


Figure 7. Lexis diagram of the NorTwinCan cohorts by country. Dark shaded areas indicate cancer follow-up.

Before any NorTwinCan related analyses could begin, data from the 2018-update had to be prepared. Prior to transfer of data, a translational program (R script) was created by the writer of this thesis and sent to colleagues in Finland for categorization of ICD-10 diagnoses according to the NORDCAN classification. The Norwegian data received was partly translated, the rest was done using the same program. Finally, the NORDCAN groups in Danish and Swedish data were scrutinized to ensure alignment between all four countries, that is, it was verified that the same mapping between diseases coding and NORDCAN groups was applied. The remaining preparations and the extensive set-up for statistical analyses laying the foundations for numerous NorTwinCan studies, were done by Wagner H. Bonat, Jacob Hjelmborg, and the writer of this thesis.

Secondly, epigenome-wide methylation data from 1,085 Danish twins, collected through three surveys by the DTR, was used to study association of methylation levels of 450,154 CpG sites to occurrence of hematopoietic malignancy. There were 31 such diagnoses in the data. In this study, the outcome of interest was hematopoietic malignancies. That is, besides the hematologic malignancies, also myeloproliferative diseases and myelodysplastic syndrome were included. There are other hematopoietic malignancies, but none were present in the analyzed data. This extended grouping was chosen to optimize conditions for analysis within twin pairs aiming to investigate genetic and environmental influences of the association between DNA methylation level and occurrence of disease (hematopoietic malignancy). Further, following the previous description of hematopoiesis (Background), this grouping is expectedly better suited for detection of DNA methylation aberrations arising early in the process of blood cell formation.

In the third and fourth studies on tattooing, we analyzed subsets of the Danish Twin Tattoo Cohort (DTTC) described previously. The study on tattooing determinants (Paper 3) included 9,173 randomly selected Danish twins born 1920-2004 (ages 16-101 at time of survey in 2021) among which 4,790 (52%) participated, thus providing information on tattoo exposure (e.g. tattoo status, colors, size, and age at first tattoo), lifestyle factors: smoking, physical exercise, alcohol consumption and education.

The study on association between tattoo ink exposure and risk of certain types of cancer consisted of two sub-studies for which data were available from the DTTC, i.e. survey data linked to information from the Danish Cancer Register with end of follow-up 1 January 2017. First, a case-cotwin study including all twin pairs born in Denmark from 1960 to 1996, that is, they have reached age 20 years at end of follow-up. Inclusion criteria were: 1) At least one twin had been diagnosed with one (or more) of the following cancers after reaching age 20 years: Hodgkin or non-Hodgkin lymphoma (ICD-10 was used along with ICD-O-3 by NORDCAN when identifying lymphoma incident cases), skin cancer (melanoma and non-melanoma – not including basal cell carcinoma) and bladder/urinary tract cancer. 2) At least one twin could be

contacted, i.e., was alive, not emigrated, and had not waivered contact by researchers. A total of 504 individual twins were invited to participate in the survey out of which 316 (56%) responded. Second, a twin cohort study of lymphoma, skin cancer and basal cell carcinoma. This sample was a subset of the twin cohort analyzed in Paper 3, restricting to twins born 1960-1996 Also, the lymphoma and skin cancer cases and their cotwins were a subset of the case-cotwin study. Among the 4,532 individuals invited to participate in the survey, 2,367 (52%) responded. We stress that twins invited to the survey were a mix of the pairs from the case-cotwin study and the cohort of randomly selected twins, thus, no personal information on cancer diagnosis was dispensed to potentially unknowing co-twins.

Methods

Study design and bias

In this thesis, two of the most common observational study designs, cohort and case-control studies, were applied with some modifications to adapt to the twin setting. In the following, they are described on a general level along with potential sources of bias relevant for the thesis manuscripts.

Cohort

In a cohort study, a group of people is followed over time to ascertain incidence of a health outcome. It is useful for assessing disease development over a longer period in a population through cumulative incidence. In a prospective cohort, the study population is identified at the beginning of the study, sometimes classified by exposure status, and subjects are followed through time. A retrospective (or historical) cohort is based on past exposures and diseases that have already occurred are evaluated. Prospective studies have less potential for recall bias, while retrospective studies are useful e.g. when studying diseases with long latency periods (95).

Case-control

In a case-control study, cases (e.g. subjects with disease) and suitable controls (typically disease-free) are identified, and their exposure status is examined – often retrospectively. It is a useful design when studying association between exposure and disease, especially when the disease is rare. There are different ways to identify controls, generally they must be drawn from the same population as the cases and be comparable in many ways except for disease status. In a matched case-control study, one or more controls are identified per case matching on potential confounders such as sex or age to ensure equality among cases and controls (95).

Twin cohort

A cohort of twins is no different from a classic cohort, in terms of study design. It does, however, require certain adjustments when analyzing as described in the statistical methods section. Besides estimation of cumulative incidence (risk), twin studies have the ability to provide information on genetic and environmental influences on variation in risk (described in detail later).

When studying nationwide twin cohorts (or representative samples from these), the aim is often to obtain results that applies to the general population. This generalizability between studies in twins and singletons is often not a matter of concern but must be addressed in relation to each specific research question. A comprehensive discussion of differences between twins and singletons, for instance in relation to health and behavior is provided by Christensen and McGue (96). Arguments for generalization of results will be discussed as relevant with each manuscript of the thesis.

Case-cotwin

The case-cotwin design, also known as discordant twin design, is similar to a matched case-control study where the control is the cotwin, which has the advantage of matching for age, shared environment (e.g. upbringing), and shared genetics. Monozygotic twins are genetically identical while dizygotic twins share on average 50% of their genetic material (97). As non-shared confounding may lead to bias, it is custom to do a parallel analysis where the twins are considered as individuals instead of pairs and the within-pair dependence is taken into account. This way observed confounders can be included in the analysis.

Bias

A register-based study has the advantage of reduced risk of selection bias compared to a study where subjects are recruited by advertisement. Also, information bias such as recall- and misclassification bias is limited. However, the latter could be questionable when the register has long follow-up since classification systems are prone to change over time in step with improving knowledge.

A survey-based study may be influenced by selection bias causing the sample of survey participants to be non-representative. Examples are participation bias – some are more willing to participate than others – and survivorship bias – some are unable to participate (95). Statistical sources of bias relating to analysis of time-to-event data, such as immortal time bias, will be covered in the statistical methods section.

A confounder is a variable that is associated with both exposure and outcome and distorts the estimated measure of association between exposure and outcome. There are several ways to control for confounding – each of different validity. Those may be by design, such as randomization, matching, stratification, or through analytic modelling approaches. Further, it can be of interest to investigate dose-response relationships, that is, if there is an association between dose of an exposure and the effect it has on an outcome.

The statistical field of causal inference aims for confounder control through assumptions met by the randomized design of study, those being exchangeability, positivity and consistency (98). It turns out that in some cases, observational studies may be analyzed as if they were randomized, however, the underlying assumptions will be untestable. This approach to analyzing data is developing and will expectedly become standard remedy for observational studies.

In this thesis, we apply the notion of average causal treatment effect, which has the interpretation of a counterfactual effect, that is, the effect if everyone were treated compared to if everyone were untreated. Being treated would in general mean being exposed for this property. Above assumptions would relate this measure to actual available data if fulfilled. In general, the causal inference is not fully developed in case of twin analysis having clustered data, however, we propose in this thesis a way to address the causal interpretation in relation to the matched case-cotwin design that exploits the ability to effectively control for shared confounding, that is negating effects of confounders that are shared for the pair. A reference to consider for this reducing confounder influence by matching is the smoking-lung cancer study with twins in which shared confounders are controlled for (19). Other main sources of bias will occur for colliders and is named collider-bias. A collider is a variable that is influence by both exposure and outcome. Stratifying or conditioning by a collider of two variables (say exposure and outcome) will induce an association. An artifact association is obtained and the general approach to avoid collider-bias is to avoid conditioning or stratification by such (99). This will be the approach taken in the present thesis having most validity in general. Another issue to consider is effect modification. It occurs when the effect measure of one explanatory variable varies with different levels of another explanatory variable. In this thesis, it will be taken into account when relevant.

When it comes to variable selection, there are various methods. A common approach when doing regression modelling is to start with the most credible (minimal) model based on prior information such as previous studies. The complexity of the model is then increased gradually while keeping an eye on the influence each added covariate has as well as statistical model diagnostics. This approach extends readily to the case of modelling twin data. The aim is to find a model which minimizes confounding according to above considerations on confounders and colliders, but at the same time not including overly many covariates as it might increase the bias rather than decrease it or simply lower the precision (without inducing bias).

Statistical methods

In this section, the statistical methods applied throughout the thesis are discussed on a general level along with methodological considerations not included in the manuscripts. The overall topic is sources of variation in risk assessed through time to event analysis of twin data.

Time to event analysis

When analyzing the occurrence of an event of interest, consideration regarding the timing is required. If the follow-up period is incomplete, it is not enough to simply count the number of subjects experiencing said event as it may occur at a later point in time. This is an example of *right censoring*. Incomplete follow-up could also be due to delayed entry, meaning that it is unknown whether the event took place before beginning of follow-up. This is called *left truncation*. Failing to consider the timing of events could lead to bias if the time one is at risk of experiencing an event is not defined correctly. Further, the estimates derived would depend on end of follow-up.

Another important aspect to consider is the presence of *competing events* preventing a subject from experiencing the event of interest. This is often mistaken as censoring; however, it is important to distinguish – is it possible to experience the event later or not? Ignoring competing events would likely lead to overestimation of the risk.

In our study, the main event of interest is cancer. Study entry is determined by date of register initiation and status at follow-up may be right-censored. In this setup, death is a competing risk of cancer diagnosis before follow-up. We note that death would not be a censoring event at follow-up. The classic alive-illness-death model shown in Figure 8 is a typical example of a competing event setting: One may go from being alive (and healthy) directly to death or develop cancer before dying. It is not possible, however, to die and then develop cancer.

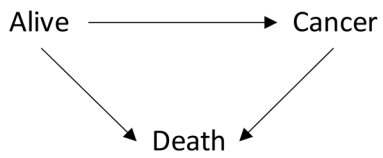


Figure 8. Alive-illness-death model with cancer as illness of interest.

Cumulative incidence function

Assuming independent censoring and in the presence of competing events, the cumulative incidence function can be obtained from the non-parametric Aalen-Johansen estimator. For cluster-correlated data, such as twin pairs, generalized estimating equations (GEE) type standard errors are applied. In the presence of covariates, the cumulative incidence function is estimated using semi-parametric counting process modelling accounting for censoring by inverse probability of weighting methods as described in Scheike et al. (100). Test for comparison of cumulative incidence functions in this setting can be done with Gray's test.

In our study, we choose age as timescale and hence consider events at various ages, e.g., age at study entry and age at diagnosis. This ensures that effect of age is part of the model and in particular non-linear effects by age are comprehended by this approach in the analysis. The cumulative incidence function at time t will then have the interpretation of the probability (risk) of obtaining a cancer diagnosis before age t, where t is an age typically between zero and a hundred. The *lifetime risk* of diagnosis is then estimated as the risk of cancer diagnosis before age 100.

Familial risk

Studying sources of variation in risk is a complex task that we address by examining various measures of familial dependence. One such measure is the *familial risk*, also called casewise concordance. It is defined as the conditional risk of observing an event in a twin before some time t given that the co-twin experienced the event before that time. This is equivalent to the ratio of the concordance (will be defined shortly) to the cumulative incidence (i.e., the individual risk) assuming equal marginals, that is, equal cumulative incidence among monozygotic (MZ) and dizygotic (DZ) twins – an assumption that is most often fulfilled as twins rarely differ by zygosity as singletons. An elevated familial risk compared to the cumulative incidence is an indication of familial effects of some kind; genetic or environmental – or both. Further,

assuming environmental influence is independent of zygosity, a higher familial risk among MZ twins compared to DZ twins indicates shared genetic influence. As a direct measure for these relative differences, we define the *relative recurrence risk* as the ratio of the familial risk to the cumulative incidence.
When assessing familial dependence, it is important to keep in mind that the proportion of shared genes between the DZ twins are the same as among full siblings, as it enables generalization of results.

The concordance function is a bivariate probability expressing the risk of observing an event in both twins in a pair before some time *t*. Several approaches were considered for obtaining this concordance risk, as described in Appendix 4. We chose to proceed with non-parametric counting process modelling (semi-parametric in the presence of covariates) accounting for censoring by inverse probability of weighting methods as explained by Scheike et al. (100). The key assumption here is same censoring within twin pairs meaning that both twins must be censored at the same time. For the register data used in this project that is not a problem as the time of censoring is the same for nearly everyone, namely end of follow-up, the only exception being emigration. If one twin emigrates, both twins in the pair will have to be censored at that time.

Inference of these measures of familial dependence can be obtained by tests of differences using a Pepe–Mori type test derived in a way similar to that described by Scheike et al. (100) assuming approximate normal distribution of estimators (94).

Familial risk of multivariate events

As part of this project, we have looked into extending above methods to apply to multivariate events described by a multidimensional random variable, e.g., representing different cancer types. We assess familial dependence by looking at pairwise combinations of events, referred to as *cross-events*, i.e., one event occurring in one twin and another (non-competing) event occurring in the cotwin. In this situation, one must be aware that two familial risks can be defined for each pair of events – one for each event being conditioned upon. For instance, is the risk of cancer type *A* increased if the cotwin has been diagnosed with cancer type *B*? The relative recurrence risk is defined as the risk of observing a set of cross-events in a twin pair before time t relative to observing the same set of cross-events in a pair of unrelated individuals before that time (94). This is estimated by the bivariate concordance function relative to the product of the cumulative incidence functions for the pair of events. As a result of this extension to twin methodology, it was necessary to adapt the available statistical software used to compute inference. The R function *casewise.test* from the mets package by Scheike et al. (100) is based on a Pepe-Mori type test and can be used to test equality of the cumulative incidence function and MZ- or DZ- concordance- or casewise concordance functions for the same event in a pair of twins. In order to apply to cross-events, the source code of this R function was adapted to test the hypothesis of equal MZ and DZ concordance functions. That way, it could also be used to test equality of MZ and DZ casewise concordance (familial risk) or MZ and DZ relative recurrence risk since both relations can be reduced to concordance equality.

Biometric modelling

Another approach to assess variation within and between twin pairs is by modelling the overall magnitudes of the genetic and environmental contributions to the variation in risk (94). We apply the *polygenic ADCE model* from quantitative genetics containing the following variance components: additive genetic (A), dominant genetic (D), shared environmental (C), and unique environmental (E) effects. Information on the underlying assumptions of this model is provided by Sham (97). We derive further measures of familial dependence based on this setting: 1) The *heritability,* or the shared genetic effects, defined as the proportion of variation in risk attributed to genetic factors, i.e., the sum of the A and D components. 2) The *shared environmental effects*, which is simply the contribution of the C component to the total variation in risk.

As described by Holst et al. (101), when using the classic twin design, one cannot identify all four variance components simultaneously. We typically go through a development process applying model selection criteria to determine the most parsimonious model (101). For now, we choose the ACE-model as it is the most biologically interesting of the largest possible three component models. With this as the starting point, we estimate the heritability as the ratio of twice the difference in MZ and DZ concordance risk to the total variance in risk. Assuming equal marginals, i.e., the same cumulative incidence functions for MZ and DZ twins, the total variance in risk is given by the product of the cumulative incidence function and one minus the cumulative incidence function. The shared environmental effects are estimated as

the difference between twice the DZ concordance and the MZ concordance relative to the total variation in risk (102). Derivations of both measures are provided in Appendix 5.

The classic liability threshold model

An alternative approach to biometric modelling is the classic *liability threshold model*. Instead of working directly on the risk scale, it models the probability of experiencing an event as a function of the latent unmeasured liability assuming normal distribution. It can be used to provide measures of familial dependence analogous to those described above and the canonical *tetrachoric correlation*, which is defined as the Pearson product moment correlation of the latent Gaussian random variables – the liabilities. Estimates are provided through a bivariate biprobit model assuming equal marginals and using inverse probability weighting to account for censoring (15, 101). It is not yet feasible to account for delayed entry or competing risk with this approach. A more comprehensive presentation of the liability threshold model is provided by Holst et al. (101).

Risk versus liability scale

It should be noted that the heritabilities and shared environmental effects in our approach stems from two different scales: The risk scale and the liability scale. From the liability threshold model, we may predict the risk scale measures above, casewise concordance etc., however, the direct non-parametric modelling on risk scale is an advantage in our study. There are various pros and cons for both methods: The risk scale approach based on counting process modelling requires few assumptions and takes into account delayed entry and competing risk of death. It also has the ability to apply to cross-events. However, there are several ways of estimating risk scale concordance, and as demonstrated in Appendix 4, it is yet uncertain how well they perform under extreme circumstances such as very few twin pairs concordant for an event, e.g., a rare cancer. The parametric liability threshold model requires more assumptions, does not apply to cross-events, and is perhaps rendered less intuitive by its latency. Nonetheless, with the additional assumptions, it often provides more accurate inference, and the heritability of liability is (yet) by far the most common in current literature.

Individual level regression analysis

We will now move the focus from measures of familial dependence providing information on where to look for risk factors (genes or environment) to various regression methods, still in the time to event setting, aiming to provide measures of associations between specific exposures and outcome.

We utilize the hazard function, that is, the short-term event rate considering at time t only individuals who have not yet experienced the event of interest. Covariate effects are given by hazard ratios (HR's) estimated by a regression model such as the Cox proportional hazards model. However, assuming that a HR does not vary over time, as is the case for the standard Cox PH model, may not be reasonable. A common way to deal with this is through estimation of *time-dependent coefficients*, in other words, HR's can be estimated for specific time periods. However, caution must be taken when dealing with time-dependent coefficients: The interpretation of the HR for the first time-period is straightforward, but the HR's of the following periods come with a selection bias since they only apply to those who made it until that time (103).

Assuming independent censoring, models based on the hazard function allows for dynamic modelling which naturally handles delayed entry. Additionally, it may allow for studying effects of both covariates measured at baseline and covariates that change during follow-up, referred to as *time-dependent covariates*. Two common errors, when dealing with an exposure that changes during follow-up, are: 1) using time of (first) exposure as starting point or 2) correctly using a time-scale independent of exposure, but then assume exposure since study entry. The first would lead to immortal time bias since it corresponds to assuming that every exposed individual is "immortal" (event free) until study entry. The second would underestimate the HR because the event-free time since baseline until the actual time of exposure would be misattributed to the exposed group instead of the unexposed group (104).

Another measure of interest, when assessing association between exposure and outcome, is the average treatment effect. It is a measure of risk difference (absolute or relative) in a setting where we consider the effect if everybody were exposed compared to if none were exposed. That is, the risk difference in a causal setting assuming exchangeability, positivity, and consistency (98). It can be derived in accordance with the censored competing risks framework previously described e.g., by using G-computation on a Cox proportional hazards model or a Fine-Gray subdistribution-hazards model.

All regression methods described in this section may account for cluster-correlated data, such as twin pairs, by applying robust variance estimation.

The discordant twin design

Besides polygenic biometric modelling of time-to-event twin cohort data used for examination of familial dependence, the discordant twin design is another powerful approach demonstrating the advantages of twin studies. It has previously been applied for the Nordic twin cancer data for instance in smoking exposure studies (19). The matched analysis enables control for unobserved, shared confounding, including genetic background and shared environmental exposures (e.g. upbringing) and reduces the impact of several limitations common in standard observational approaches. For example, the classic Cox proportional hazards regression model is easily adapted to a stratified model fit for the matched setting by using twin pair specific baseline hazards (105). A matched case-cotwin design is powerful in the sense that few informative pairs, that is, pairs discordant for exposure and outcome, are needed to detect a significant effect. As non-shared confounding may impact this design, standard analysis using the twins as singletons (as described previously) while correcting for within-pair dependence and observed risk factors is carried out as a parallel analysis, presumably less powerful, but using all available information (106, 107).

Prediction modelling

Finally, we have attempted to create a prediction model selecting predictors from a high number of covariates. In our study, this followed an epigenome wide association study (EWAS) aiming to identify possible epigenetic markers of cancer through application of previously described individual level regression analyses, applying time-to-event modelling, and taking into account timing of events and competing risk of death. A specific form of stability selection particularly well suited for high dimensional data (108) was used on a subset of the most promising covariates (CpG sites) from the EWAS. To identify candidates for a predictor we computed a number of replications applying the least absolute shrinkage and selection operator (109) in a Cox regression model. Evaluation of the prediction model was done through estimation of cross-validated Harrel's C, a concordance index used for censored time-to-event data, followed by predictive ability over time via the cross-validated, time-varying area under the curve (AUC).

To assess whether the effect of the predictor was independent of underlying genetic effects, the discordant twin design was applied using a stratified Cox model on the monozygotic twins.

Results

In this section, the main results of the four studies included in this thesis are summarized. Further details can be found in the enclosed papers.

Paper 1: Familial risk and heritability of hematologic malignancies in the Nordic Twin Study of Cancer

- Among 316,397 Nordic twins with a median follow-up time of 42 years, there were 3,459 hematologic malignancies. The distribution was: 44% non-Hodgkin lymphoma (NHL), 35% leukemia, 15% multiple myeloma (MM), and 7% Hodgkin lymphoma (HL).
- Cumulative incidence functions by age were estimated for each of the four hematologic malignancies. The lifetime risks (in percentages with 95% CI) were estimated to: NHL: 1.1 (1.0–1.2), leukemia: 0.9 (0.8–1.0), MM: 0.4 (0.3–0.4), and HL: 0.1 (0.1–0.1). The overall risk was 2.5 (2.4-2.6).
- The familial risks by age for overall hematologic malignancy was significantly higher than that of the general population at all ages (from age 40) with the DZ familial risk being nearly twice as high. The MZ familial risk was suggestively higher than DZ at all ages, but not significantly so.
- Lifetime estimates of genetic and environmental contributions to variation in liability to develop cancer for each of the subtypes, except for HL, were very similar with low heritability, high non-shared environmentability, and no shared environmentability. For HL, the contribution from genetics and non-shared environment were approximately evenly shared.
- Assessing the same measures over age for overall hematologic malignancy revealed even contributions from genetics and non-shared environment around age 40 years, which then diverged until stabilization around age 55, where the heritability was around 20-25% and the remaining variation was attributed to non-shared environmentability. There was no indication that shared environment had an influence at any age.
- Lifetime relative recurrence risks revealed significantly increased risk of co-occurrence of NHL and leukemia in DZ twin pairs suggesting familial predisposition between these cancer types.

Paper 2: Novel epigenetic biomarkers for hematopoietic cancer found in twins

- The study included 1,085 Danish twins with 31 hematopoietic malignancies. DNA methylation levels were available for 450,154 CpG sites.
- An epigenome wide association study of DNA methylation levels and hematopoietic malignancy identified 15,432 CpG sites with a false discovery rate below 0.05.
- Among these CpG sites were 67 significant epigenetic markers – 12 of which were found to be linked to genes that were associated with hematologic malignancies in the FinnGen study and further 16 sites linked to genes associated with other immune disorders or diseases of the blood.
- A matched case-cotwin study, providing strong confounder control, yielded similar measures of association, but only 2 out of 12 were statistically significant. Within pair correlations of monozygotic twins were significantly different from zero for 7 sites, suggesting a moderate upper bound of genetic influence on liability of hematopoietic malignancy at 0.34.
- We further identified 12 candidates for a predictor of hematopoietic malignancies. The predictor performed well on the Danish data under cross-validation with a prediction performance of 92% three years after blood sampling and persistent performance above 70% up to six years after blood sampling.
- The findings of associated and predicted CpG sites were validated externally to a satisfactory extent in a population representative Finnish twin sample.

Paper 3: Tattooing is mainly cultural. A representative twin study of tattooing determinants

- The study included 4,790 twins among which 1,061 (22%) had a tattoo.
- Cumulative incidence of age at first tattoo by sex was estimated for three birth cohorts and the cumulative incidences at age 25 years were: 5.8% (95% CI: 4.2-7.4%) for males and 0.4% (0.0-0.8%) for females born in the years 1925-1960, 15.4% (12.5-18.3%) for males and 11.5% (9.1-13.9%) for females born in the years 1961-1980, and 30.2% (25.4-35.0%) for males and 41.3% (37.0-45.5%) for females born in the years 1981-2004.
- More than half of the tattooed males and females born in the years 1981-2004 were younger than 20 years old when they had their first tattoo.
- The average causal risk difference (ATE) of having a tattoo before age 25 years among those born in 1981-2004 was 9.9% (4.4-15.4%) when comparing females to males.
- Likewise, the ATE of tattooing was 35.9% (29.2-42.7%) when comparing ever-smokers to never-smokers born in 1981-2004. No significant ATE could be detected for alcohol consumption or physical activity.
- The familial risk of tattooing by age was nearly twice as high as the risk in the background population at ages 20-80 years suggesting a familial predisposition governed by environmental effects.
- The shared environmental influence of variation in risk of having a tattoo by age was around 65% from age 30 while the genetic influence was close to zero at all ages.

Paper 4: Tattoo ink exposure is associated with lymphoma and skin cancers – a Danish study of twins

- The 316 twins in the case-cotwin study were split into three (slightly overlapping) samples. The first sample included 32 lymphoma cases and 34 cotwin controls. (The numbers are not equal because incomplete pairs were included). The second sample included 119 skin cancer cases and 104 cotwin controls. The third sample included 8 cases with cancer of the bladder and urinary tract and 6 cotwin controls.
- Approximately 30% of both lymphoma cases and controls were tattooed and around 25% of the skin cancer cases and control were tattooed. In the sample with bladder cancer there were fewer than five tattooed individuals.
- Individual level analysis revealed increased hazard of lymphoma among twins with tattoos larger than the palm of one's hand (HR: 2.73 (95% CI: 1.33-5.60)). Also, the HR of skin cancer was 1.62 (95% CI: 1.08-2.41) comparing tattooed to non-tattooed, while the HR was 2.37 (95% CI: 1.11-5.06) when considering large tattoos. There was no evidence for a confounding effect of smoking when estimating the hazard of skin cancer. It was not possible to fit any models estimating hazard of bladder cancer, modelling the confounding effect of smoking on lymphoma hazard, or doing matched analysis for lymphoma or bladder cancer.
- There were 2,367 individual twins in the twin cohort. Among these were 6 cases with lymphoma (fewer than five of them tattooed), 16 with skin cancer (half of them were tattooed), and 29 with basal cell carcinoma (11 of them were tattooed).
- Cumulative incidence functions for skin cancer and basal cell carcinoma by age were not significantly different among tattooed and non-tattooed. The risk (cumulative incidence) of being diagnosed with skin cancer before age 57 years was 1.1% (95% CI: 0.3%-1.8%) among tattooed and 0.4% (95% CI: 0.2%-0.7%) among non-tattooed. Likewise, for basal cell carcinoma the risks were 1.5% (95% CI: 0.6%-2.4%) for tattooed and 0.9% (95% CI: 0.5%-1.4%) for non-tattooed.
- Individual level analysis revealed increased hazard of skin cancer among tattooed twins (HR: 3.90 (95% CI: 1.42-10.78)). Likewise, for basal cell carcinoma the HR was 2.83 (95% CI: 1.30-6.16). There was no evidence of a confounding effect of smoking when studying skin cancer outcome, but for basal cell carcinoma, the effect of tattooing was increased with a HR of 3.52 (95% CI: 1.63-7.61) when adjusting for smoking.

Discussion

In this project, we set out to study the hypothesis that tattoo ink induce inflammation at deposit site which may eventually become chronic and increase risk of abnormal hematopoietic cell proliferation. The first step of the project was to initiate the survey which would later become the foundation for the Danish Twin Tattoo Cohort. While this was in the making, we studied etiology of hematologic malignancies through the Nordic Twin Cancer Study and conducted an epigenome-wide association twin study of hematopoietic malignancies. We then examined the hypothesized risk factor, tattooing, finding it to be a cultural behavior with limited genetic influence, and provided novel evidence associating tattoo exposure with increased risk of lymphoma and skin cancer.

In greater detail, we have established the Danish Twin Tattoo Cohort (DTTC) for which more than eleven thousand Danish twins were invited and about half participated. The DTTC holds a population-based cohort of 4,790 participating twins randomly selected among Danish twins born 1920-2004 and case-cotwin samples of twins with various cancers, including lymphoma and skin cancer. All twin pairs in the Danish Twin Register, where at least one twin had been diagnosed with a cancer of interest and at least one twin in a pair could be contacted, were invited, leading to a total of 1,328 participants.

Meanwhile, in the first paper, we verified previous findings of familial predisposition to hematologic malignancies and co-occurrence with other cancers such as prostate and lung cancer. The estimation of familial risk by age and further the influence of genetic and environmental factors by age was a novel contribution. The discovery of decreasing familial predisposition with increasing age is what could be expected from a disease influenced by non-genetic factors since more environmental exposures will have had opportunity to exert an influence. As there were few diagnoses before age 40 years in the cohort, we were unable to study familial influence at younger ages. This may also explain our finding of no shared environmental influence on risk of hematologic malignancies since this is expectedly mainly governed by early life environmental exposures (upbringing).

As a link between environmental exposures and the effect they have on our genes, we have studied epigenetics. In the study of DNA methylation and development of hematopoietic malignancy, we identified 67 epigenetic quantitative trait loci of which 12 are linked to genes associated with hematologic malignancy in FinnGen. For instance, the CpG sites cg00984696 and cg05575733 which are both related to NHL. Additionally, these two sites presented moderate genetic influence. To the best of our knowledge, we are the first to present estimates of average causal CpG effects comparing risk of hematopoietic malignancy five years after individual blood sampling among those with high versus low methylation levels. Moreover, the formation of a predictor performing persistently at a high level up to six years after blood sampling could, with further exploration, aid prevention and early detection of such diseases.

In the study of tattooing determinants, we obtained several new insights. First, we provided empirical evidence of the dramatic increase in cumulative incidence of tattooing over the past decade. Second, we found a strong familial influence on tattooing behavior, similar among MZ and DZ twins, this mainly due to environmental factors. In other words, tattooing is a cultural behavior. Third, a strong association between smoking and tattooing was detected, and fourth, findings indicated that females of the younger generations are more likely to have a tattoo than males.

Finally, the twin study of tattoo ink exposure and risk of certain cancers lead to novel evidence associating tattoos with risk of lymphoma and skin cancer. This is a finding of great importance seeing as more than a fifth of the Danish adult population are estimated to have at least one tattoo. In a case-cotwin study, hazards of skin cancer, basal cell carcinoma, and lymphoma were about 2-3 times higher among tattooed. Further, the cohort study suggested risks for skin cancer being around twice as high for tattooed individuals.

Previous knowledge

As explained in Paper 1, numerous studies have provided evidence of familial predisposition to hematologic malignancies (110-122). For example, in a study of NHL in the Nordic countries, Fallah et al. found evidence of increased familial risks for various familial relationships (including twins) by sex compared to the general population (122). A similar study of HL has also demonstrated increased familial risks (121). Both studies found the highest familial influence on development of hematologic malignancy among twin siblings, however, zygosity was ignored in both cases.

Most epigenetic studies of hematopoietic malignancies concern specific subtypes. However, a Swedish study of familial risk of hematopoietic malignancies, including more than 150,000 patients, provided evidence for shared etiological factors across subtypes (117). To the best of our knowledge, the CpG sites identified in our study (Paper 2) has not been published previously in relation to association between DNA methylation level and risk of hematopoietic malignancies. This may indicate that our findings are novel, but there is also a possibility that some are false positive findings resulting from large scale testing. We conjecture that some of the identified sites stem from DNA methylation aberrations arising early in the process of hematopoiesis where hematopoietic stem cells have yet to be restricted to a certain lineage (3).

Previous studies have suggested modification of tattoo prevalence by age, with a higher proportion of young females being tattooed compared to males of similar age (32, 123). This was not directly tested in our study, but it is consistent with tattoos being more common among males in the oldest birth cohorts and more common among females in the youngest birth cohorts. Moreover, there is limited literature about the decision of getting a tattoo. A study among American college students found a relationship between tattooing and having tattooed peers while limited familial influence was suggested (124). In a review of tattoo studies, Kluger points out that peer influence is known to play an important role, in line with our findings of environmental influence, and further, that smoking was increased among tattooed individuals, which is also consistent with our results (125).

To this date (May, 2024), only one study has been published regarding tattoo ink exposure and risk of cancer: A Canadian study from 2020 (47). It considered two population-based case-control studies holding 1,518 participants including 737 cases with non-Hodgkin lymphoma and 742 participants including 373 cases with multiple myeloma. No significant associations could be detected. We note a considerable limitation of the study, namely that logistic regression was used instead of time-to-event modelling. Consequently, two types of bias in the direction of underestimation of risk are likely to occur: One due to incomplete follow-up – controls are practically assumed to stay cancer free, and another since the time-varying nature of tattooing is ignored, exposed individuals are assumed to be tattooed since birth.

Someone without prior experience in twin research might be tempted to question whether results of twin studies have any use beyond twins. It should be clear by now that it indeed is the case for the topics covered in this project. Cancer risk has previously been shown to mirror that of the general population (15, 93). That was verified for hematologic malignancies in Paper 1. There is no reason to believe that merely being a twin influences risk of cancer (or tattooing). When we condition on cotwin it is a different matter entirely – and a different kind of risk. In a book chapter by Christensen & McGue, twin-singleton differences are discussed based on literature and the authors conclude that while twins differ from singletons in early life, twin-singleton differences are generally small after childhood (96).

Through Statistics Denmark (126) it is possible to compare number of singleton- and twin births in the younger birth cohorts of the DTTC, that is, those born in the years 1990-2004. While the number of singleton births per year appears relatively stable around 61 to 68 thousand, the number of twin births steadily increased to from around 700 to 1400 in that period. The increasing use of fertility treatment may be a contributing factor to this development. This would explain the discrepancy in age distributions for the 16–25-year-olds in the tattoo cohort compared to the background population (figure not shown, but the patterns were very similar to that of Supplementary Figure 1 and Supplementary Figure 2 in Paper 3 depicting age distribution among participants in the cohort study of tattoo exposure).

Strength and weaknesses

A strength of twin studies, which has been exploited in the first three papers of this thesis, is the use of zygosity to provide information on relative genetic and environmental influences on occurrence of hematologic (Paper 1) and hematopoietic malignancy (Paper 2) and tattooing (Paper 3). In the cancer studies, the use of broader disease classification involves both limitations and strengths: While the study of disease subgroups likely yields information more applicable for the individual, e.g., specific information on (epi)genetics is useful in classification and treatment of hematologic malignancies, the use of broader groupings provides more power which is needed for assessment of beforementioned genetic and environmental influences and how they vary by age. This information can in turn aid in the search of risk factors which can be useful in prevention and early detection of disease.

Variation in disease occurrence among twins has been studied through both classical twin modelling and state-of-the-art risk-based methodology throughout this project. Pros and cons of both liability and risk scale modelling have been discussed. While the former is well-known in the field of twin research, its lack of intuitiveness is a drawback. Also, having

started to look into co-occurrence of cancers have revealed variation in risk shared between cancer types typically studied separately (ref: Thesis Paper 1 and yet unpublished results from the Nordic Twin Study of Cancer on more than 40 cancer sites). Such analyses of cross-events would not have been possible on the liability scale. However, there is still work to be done in evaluating the performance of such methods under extreme circumstances such as very few twin pairs concordant for an event, e.g., a rare cancer.

Matched analysis for average treatment effect was not possible in the study of tattooing determinants (Paper 3) as we did not yet have the methods required to do so while considering the time-varying nature of tattoo exposure. We hope to derive the means to do so soon. It is problematic because modelling within- and between pair associations of tattoo exposure and cancer outcome may introduce competing events when enforcing pairwise censoring (if one twin develops cancer, the other twin is censored at that point in time). Should both twins acquire tattoos before one (potentially) develops cancer, that would be a competing event as it is no longer possible for the pair to become exposure and outcome discordant. Careful consideration must be given to this issue as competing events complicate the interpretation of results estimated through a Fine-Gray subdistribution hazards model (127). Given the finding of low heritability of tattoo exposure (Paper 3) and low heritability of hematologic malignancies (Paper 1), one might question whether a matched twin study would provide much more information than the individual level analysis. At least, the match for genetic background would expectedly have little influence.

An advantage of studying cancer in the Nordic countries is the nationwide registers. This limits the risk of selection bias compared to a study where subjects are recruited by advertisement. Also, information bias such as recall- and misclassification bias is limited. However, the latter could be questionable in the NorTwinCan study (which Paper 1 is part of) since disease classification has changed over time in step with improving knowledge. For the biometric analysis in Paper 1, such changes would be regarded as random noise and be contained in the E-component (non-shared environment) of the ADCE model. Another source of misclassification bias is that of tattooing since the survey did not specify type of tattoo. The intention was to study decorative tattoos, but participants with permanent make-up and medical tattoos may also have responded as being tattooed.

Since the DTTC is a survey-based cohort, the studies described in Paper 3 and 4 may be influenced by selection bias causing the sample of survey participants to be non-representative. For instance, survivorship bias is present since one had to be alive in 2021 to participate in the survey. In particular, the case-cotwin sample consists of cancer survivor and their cotwins. Another example is the higher participation rates among those with higher levels of education which could bias the data towards lower tattoo rates (32, 125). It is also possible that some based their decision of whether to participate on their personal opinion on tattoos or thinking it was related to cancer. Survivorship bias may also be present in the epigenetic study since the survey participants, whose blood samples were used to extract DNA methylation data, are unlikely to have included the frailest individuals.

Finally, in each study, the expected measured confounders have been taken into account (e.g. age/cohort, sex, zygosity). One additional factor that could provide important information, mainly in the study of tattooing determinants, is socioeconomic background. That information was not available; however, the twin design expectedly reduces such bias due to familial components.

Perspectives

Multiple ideas for potential future studies have arisen during the project. First, we believe there is more knowledge to be learned regarding manifestation of hematologic cancer across cancer types within twin pairs (or families) as was studied in Paper 1. That is, the increased risk of cancer in a twin given another cancer in the co-twin. A study of familial dependence including not only the hematologic malignancies, but all major cancer types, could reveal hitherto unknown links and provide important insights into shared etiology. The NorTwinCan study would provide an optimal setting for this. Our paper can be seen as an initiation for a cross cancer relationship study, and it would be further interesting to study the impact of risk factors, in particular tattooing as it has been done for smoking. This perspective requires extensive survey data and collection of these have just begun.

The epigenetic study of hematopoietic malignancies (Paper 2) has led to a vast number of potential further studies. For instance, a more in-depth exploration of the 67 identified CpG sites including search of linked genes in other databases besides FinnGen. Also, as opposed to an epigenome-wide analysis, one could perform a site-specific study of previously

identified CpGs related to hematopoietic malignancy using the matched case-cotwin design to provide novel information on genetic and environmental influences on DNA methylation level. Another perspective for the identified CpGs would be to study their pathways through omics layers 'beyond and above', that is, their further role in genomics as already started in Paper 2 and similarly for proteomics and mRNA expression or transcriptomics. Finally, combining DNA methylation data from the three cohorts included in the epigenetic study with the DTTC, we could search for CpG sites associated with tattoo exposure through the matched case-cotwin design among twin pairs discordant for tattoo exposure and further compare these with established CpG sites related to development of hematopoietic malignancies. This could reveal more information about the biological pathway from tattoo exposure to development of cancer.

In addition, there is more information to be utilized in the DTTC. The study of tattoo ink and cancer (Paper 4) was restricted to individuals born since 1960, but information is available dating back to opening of the Danish Cancer Register in 1943. This would require different modelling approaches but would be possible with e.g. an extended Poisson model allowing for two timescales: age and chronological time.

Another perspective is a matched study comparing known cancer biomarkers or inflammatory markers measured in blood samples among healthy tattooed and non-tattooed individuals. Through linkage to some of the other national health registers available in Denmark, the twin cohort established during this project could be used to study association between tattoo ink exposure (along with smoking habits) and other immune-related diseases. It would for instance be possible to estimate and compare cumulative incidence functions by age of sarcoidosis among tattooed and non-tattooed individuals. Besides, current knowledge on migration of ink from the site of the tattoo to somewhere beyond the regional lymph nodes is based on animal studies. We propose a review of medical and forensic pathologic reports to look for observations of ink in other organs.

Finally, the European laws from 2021 regulating use of many known and suspected carcinogens in tattoo ink will expectedly aid the prevention of tattoo-related cancers to a certain extent, but it may take years for the payoff to show. We note, however, that according to our main hypothesis, the presence of ink of any sort in the lymph nodes may induce increased cell proliferation due to an immunologic response and we do not expect this to be affected by the stricter regulations.

Conclusion

In the exploration of determinants influencing familial dependence of hematologic malignancies, twin studies play an important role and has benefitted from our analysis of the nationwide registers in the Nordic countries. Furthermore, in a world where tattooing has become immensely popular and knowledge on the long-term safety has been found sorely lacking, this thesis is considered a contribution of great importance.

Our contribution consists of; i) The establishment of The Danish Twin Tattoo Cohort beneficial for public health studies. ii) Results shed light on etiology of hematologic malignancies through the Nordic Twin Cancer Study. iii) Epigenetic biomarkers identified through an epigenome-wide association twin study of hematopoietic malignancies. iv) We examined the hypothesized risk factor, tattooing, finding it to be a cultural group clustering phenomenon of limited genetic influence. v) Conjectured and supported novel evidence associating tattoo exposure with increased risk of lymphoma and skin cancer. It is the hope that further knowledge will be pursued for the benefit of public health and health science.

English summary

In this project, we set out to study the hypothesis that tattoo ink, which is known to migrate to regional lymph nodes, induce inflammation at deposit site which may eventually become chronic and increase risk of hematopoietic cancer cell formation. First, we study etiology of hematologic malignancies through the Nordic Twin Cancer Study. We verify previous findings of familial predisposition and provide novel information on the influence of genetic and environmental factors by age. As a link between environmental exposures and the effect they have on our genes, we conduct an epigenome-wide association twin study of DNA methylation and hematopoietic malignancies identifying 67 CpG biomarkers of which 12 are linked to genes associated with hematologic malignancy. Further, we set up a predictor performing persistently at a high level up to six years after blood sampling.

Moreover, we establish the Danish Twin Tattoo Cohort which has so far enabled two studies and may become an important source of knowledge in the field. First, in a study of tattooing determinants, we provide empirical evidence of the dramatic increase in cumulative incidence of tattooing over the past decade, find that tattooing behavior is a cultural group clustering phenomenon with limited genetic influence, and detect a strong association between smoking and tattooing. Second, we examine the hypothesized risk factor, tattooing, and provide novel evidence associating tattoo exposure with increased risk of lymphoma and skin cancer. These findings are of great importance in a world where tattooing has become immensely popular and knowledge on the long-term safety has been found sorely lacking. It is our hope that further knowledge will be pursued for the benefit of public health and health science.

Danish summary (dansk resumé)

Dette projekt tager udgangspunkt i hypotesen om at tatoveringsblæk, som er kendt for at kunne migrere til omkringliggende lymfeknuder, inducerer inflammation på aflejringsstedet. Vi mistænker at inflammationen kan blive kronisk og dermed øge risikoen for dannelse af hæmatopoietiske kræftceller. Vi starter med at undersøge ætiologien af hæmatologiske cancere ved brug af det Nordiske Tvilling Cancer Studie. Resultaterne bekræfter tidligere fund af familiær indflydelse og finder ny information om indflydelsen af genetiske og miljømæssige faktorers variation med alder. For at forbinde miljøeksponeringer og den effekt, de har på vores gener, udfører vi et tvilling-associationsstudie af hele epigenomet. Vi identificerer 67 CpG-biomarkører for hæmatopoietisk cancer, hvoraf 12 er forbundet med gener, som er forbundet med hæmatologisk cancer. Vi opstiller også en prædiktor, der præsterer på et vedvarende højt niveau op til seks år efter blodprøvetagning.

Vi etablerer desuden den Danske Tvilling Tatoveringskohorte, som indtil videre har givet mulighed for to studier, men som kan gå hen og blive en vigtig informationskilde indenfor feltet. Det første studie handler om faktorer, der har indflydelse på tatoveringsadfærd. Her finder vi ud af, at tatoveringsadfærd er et kulturelt gruppefænomen med begrænset genetisk indflydelse. Vi finder desuden empirisk evidens for den dramatiske stigning i kumulativ incidens af tatoverede individer i løbet af det sidste årti og opdager en stærk sammenhæng mellem rygning og tatoveringer. Det andet studie handler om den hypoteserede risikofaktor, dét at være tatoveret, og fremlægger ny evidens, der forbinder tatoveringseksponering med øget risiko for lymfe- og hudcancer. Disse resultater er af stor betydning i en verden, hvor tatoveringer er blevet enormt populære, og viden om langsigtet sikkerhed har vist sig at være meget mangelfuld. Det er vores håb, at yderligere viden vil blive efterstræbt til gavn for folkesundhed og sundhedsvidenskab.

References

1. Dean L. Chapter 1, Blood and the cells it contains. 2005. In: Blood Groups and Red Cell Antigens [Internet]. Bethesda (MD): National Center for Biotechnology Information (US). Available from: https://www.ncbi.nlm.nih.gov/books/NBK2263/.
2. Blecua P, Martinez-Verbo L, Esteller M. The DNA methylation landscape of hematological malignancies: an update. Mol Oncol. 2020;14(8):1616-39.
3. Hu D, Shilatifard A. Epigenetics of hematopoiesis and hematological malignancies. Genes Dev. 2016;30(18):2021-41.
4. NORDCAN: Cancer Incidence, Mortality, Prevalence and Survival in the Nordic Countries, Version 9.3 (02.10.2023). Association of the Nordic Cancer Registries. Cancer Registry of Norway. [Internet]. 2023 [cited 16 November 2023]. Available from: https://nordcan.iarc.fr/.
5. Engholm G, Ferlay J, Christensen N, Bray F, Gjerstorff ML, Klint A, et al. NORDCAN-a Nordic tool for cancer information, planning, quality control and research. Acta Oncol. 2010;49(5):725-36.
6. Bispo JAB, Pinheiro PS, Kobetz EK. Epidemiology and Etiology of Leukemia and Lymphoma. Cold Spring Harb Perspect Med. 2020;10(6).
7. Skrabek P, Turner D, Seftel M. Epidemiology of Non-Hodgkin Lymphoma. Transfusion and Apheresis Science. 2013;49(2):133-8.
8. Huh J. Epidemiologic overview of malignant lymphoma. Korean J Hematol. 2012;47(2):92-104.
9. Hjalgrim H. On the aetiology of Hodgkin lymphoma. Danish medical journal. 2012;59(7):B4485.
10. Juliusson G, Hough R. Leukemia. Prog Tumor Res. 2016;43:87-100.
11. Huang J, Chan SC, Lok V, Zhang L, Lucero-Prisno DE, Xu W, et al. The epidemiological landscape of multiple myeloma: a global cancer registry estimate of disease burden, risk factors, and temporal trends. The Lancet Haematology. 2022;9(9):e670-e7.
12. Marques-Mourlet C, Di Iorio R, Fairfield H, Reagan MR. Obesity and myeloma: Clinical and mechanistic contributions to disease progression. Front Endocrinol (Lausanne). 2023;14:1118691.
13. Hemminki K, Hemminki J, Försti A, Sud A. Survival trends in hematological malignancies in the Nordic countries through 50 years. Blood Cancer Journal. 2022;12(11):150.
14. Smith A, Howell D, Patmore R, Jack A, Roman E. Incidence of haematological malignancy by sub-type: a report from the Haematological Malignancy Research Network. Br J Cancer. 2011;105(11):1684-92.
15. Mucci LA, Hjelmborg JB, Harris JR, Czene K, Havelick DJ, Scheike T, et al. Familial risk and heritability of cancer among twins in nordic countries. JAMA. 2016;315(1):68-76.
16. Hjelmborg JVB, Clemmensen SB. Twin studies on cancer. Ugeskr Laeger. 2024;186(4).
17. Harris JR, Hjelmborg J, Adami HO, Czene K, Mucci L, Kaprio J, et al. The Nordic Twin Study on Cancer - NorTwinCan. Twin Res Hum Genet. 2019:1-7.
18. Hjelmborg JB, Scheike T, Holst K, Skytthe A, Penney KL, Graff RE, et al. The heritability of prostate cancer in the Nordic Twin Study of Cancer. Cancer Epidemiol Biomarkers Prev. 2014;23(11):2303-10.
19. Hjelmborg J, Korhonen T, Holst K, Skytthe A, Pukkala E, Kutschke J, et al. Lung cancer, genetic predisposition and smoking: the Nordic Twin Study of Cancer. Thorax. 2017;72(11):1021-7.
20. Gerster M, Madsen M, Andersen PK. Matched survival data in a co-twin control design. Lifetime data analysis. 2014;20(1):38-50.
21. Petersen AH, Lange T. What Is the Causal Interpretation of Sibling Comparison Designs? Epidemiology. 2020;31(1):75-81.
22. Heikkinen A, Lundgren S, Ollikainen M. Chapter 33 - Epigenetics. In: Tarnoki A, Tarnoki D, Harris J, Segal N, editors. Twin Research for Everyone: Academic Press; 2022. p. 585-616.
23. Moore LD, Le T, Fan G. DNA Methylation and Its Basic Function. Neuropsychopharmacology. 2013;38(1):23-38.
24. Jjingo D, Conley AB, Yi SV, Lunyak VV, Jordan IK. On the presence and role of human gene-body DNA methylation. Oncotarget. 2012;3(4):462-74.
25. Zemach A, McDaniel IE, Silva P, Zilberman D. Genome-wide evolutionary analysis of eukaryotic DNA methylation. Science. 2010;328(5980):916-9.
26. Espín-Pérez A, Brennan K, Ediriwickrema AS, Gevaert O, Lossos IS, Gentles AJ. Peripheral blood DNA methylation profiles predict future development of B-cell Non-Hodgkin Lymphoma. NPJ Precis Oncol. 2022;6(1):53.
27. Kaiser MF, Johnson DC, Wu P, Walker BA, Brioli A, Mirabella F, et al. Global methylation analysis identifies prognostically important epigenetically inactivated tumor suppressor genes in multiple myeloma. Blood. 2013;122(2):219-26.

28. Joehanes R, Just AC, Marioni RE, Pilling LC, Reynolds LM, Mandaviya PR, et al. Epigenetic Signatures of Cigarette Smoking. Circ Cardiovasc Genet. 2016;9(5):436-47.
29. Liu C, Marioni RE, Hedman Å K, Pfeiffer L, Tsai PC, Reynolds LM, et al. A DNA methylation biomarker of alcohol consumption. Mol Psychiatry. 2018;23(2):422-33.
30. Nickels EM, Li S, Myint SS, Arroyo K, Feng Q, Siegmund KD, et al. DNA methylation at birth in monozygotic twins discordant for pediatric acute lymphoblastic leukemia. Nat Commun. 2022;13(1):6077.
31. Svane AM, Soerensen M, Lund J, Tan Q, Jylhava J, Wang Y, et al. DNA Methylation and All-Cause Mortality in Middle-Aged and Elderly Danish Twins. Genes (Basel). 2018;9(2).
32. Sagoe D, Pallesen S, Andreassen CS. Prevalence and correlates of tattooing in Norway: A large-scale cross-sectional study. Scand J Psychol. 2017;58(6):562-70.
33. Nielsen C, Andreasson K, Olsson H, Engfeldt M, Joud A. Cohort profile: The Swedish Tattoo and Body Modifications Cohort (TABOO). BMJ Open. 2023;13(5):e069664.
34. Paprottka FJ, Krezdorn N, Narwan M, Turk M, Sorg H, Noah EM, et al. Trendy Tattoos-Maybe a Serious Health Risk? Aesthetic Plast Surg. 2018;42(1):310-21.
35. Barber-Axthelm IM, Kelly HG, Esterbauer R, Wragg KM, Gibbon AM, Lee WS, et al. Coformulation with Tattoo Ink for Immunological Assessment of Vaccine Immunogenicity in the Draining Lymph Node. J Immunol. 2021;207(2):735-44.
36. Negi S, Bala L, Shukla S, Chopra D. Tattoo inks are toxicological risks to human health: A systematic review of their ingredients, fate inside skin, toxicity due to polycyclic aromatic hydrocarbons, primary aromatic amines, metals, and overview of regulatory frameworks. Toxicol Ind Health. 2022;38(7):417-34.
37. Laux P, Tralau T, Tentschert J, Blume A, Dahouk SA, Bäumler W, et al. A medical-toxicological view of tattooing. The Lancet. 2016;387(10016):395-402.
38. Hauri U, Amberg J, Lüthi K-W, Baumgartner A, Gautsch S, Brüschweiler B. Konformität von Tätowier- und Permanent-Make-up Farben nicht zufriedenstellend. BAG Bulletin. 2009;29.
39. Weiss KT, Schreiver I, Siewert K, Luch A, Haslbock B, Berneburg M, et al. Tattoos - more than just colored skin? Searching for tattoo allergens. J Dtsch Dermatol Ges. 2021;19(5):657-69.
40. Serup J. Chaotic Tattoo Ink Market and No Improved Costumer Safety after New EU Regulation. Dermatology. 2023;239(1):1-4.
41. Pazos MD, Hu Y, Elani Y, Browning KL, Jiang N, Yetisen AK. Tattoo Inks for Optical Biosensing in Interstitial Fluid. Adv Healthc Mater. 2021;10(21):e2101238.
42. Schreiver I, Luch A. At the dark end of the rainbow: data gaps in tattoo toxicology. Arch Toxicol. 2016;90(7):1763-5.
43. Jacobsen E, Tønning K, Pedersen E, Bernth N, Serup J, Høgsberg T, et al. Kemiske Stoffer i Tatoveringsfarver. 2012.
44. Kluger N. An update on cutaneous complications of permanent tattooing. Expert Rev Clin Immunol. 2019;15(11):1135-43.
45. Baumler W. Absorption, distribution, metabolism and excretion of tattoo colorants and ingredients in mouse and man: the known and the unknown. Curr Probl Dermatol. 2015;48:176-84.
46. Foerster M, Schreiver I, Luch A, Schuz J. Tattoo inks and cancer. Cancer Epidemiol. 2020;65:101655.
47. Warner FM, Darvishian M, Boyle T, Brooks-Wilson AR, Connors JM, Lai AS, et al. Tattoos and Hematologic Malignancies in British Columbia, Canada. Cancer Epidemiol Biomarkers Prev. 2020;29(10):2093-5.
48. Piccinini P, Pakalin S, Contor L, Bianchi I, Senaldi C. Safety of Tattoos and Permanent Make-Up: Final Report. Joint Research Centre; 2016. Report No.: EUR 27947 EN.
49. Lehner K, Santarelli F, Vasold R, Penning R, Sidoroff A, Konig B, et al. Black tattoos entail substantial uptake of genotoxic polycyclic aromatic hydrocarbons (PAH) in human skin and regional lymph nodes. PLoS One. 2014;9(3):e92787.
50. Schreiver I, Hesse B, Seim C, Castillo-Michel H, Villanova J, Laux P, et al. Synchrotron-based nu-XRF mapping and mu-FTIR microscopy enable to look into the fate and effects of tattoo pigments in human skin. Sci Rep. 2017;7(1):11395.
51. Dominguez E, Alegre V, Garcia-Melgares ML, Laguna C, Martin B, Sanchez JL, et al. Tattoo pigment in two lymph nodes in a patient with melanoma. J Eur Acad Dermatol Venereol. 2008;22(1):101-2.
52. Dissing J, Bille C, Jylling AMB. Tatoveringsblæk gav mistanke om malignt melanom i lymfeknude. Ugeskrift for Læger. 2022;184.
53. Sepehri M, Sejersen T, Qvortrup K, Lerche CM, Serup J. Tattoo Pigments Are Observed in the Kupffer Cells of the Liver Indicating Blood-Borne Distribution of Tattoo Ink. Dermatology. 2017;233(1):86-93.
54. Sangueza OP, Yadav S, White CRJ, Braziel RM. Evolution of B-Cell Lymphoma from Pseudolymphoma A Multidisciplinary Approach Using Histology, Immunohistochemistry, and Southern Blot Analysis. The American Journal of Dermatopathology. 1992;14(5):408-13.

55. Wang X, Josefsson L, Meschnark S, Lind ML, Emmer A, Goessler W, et al. Analytical survey of tattoo inks-A chemical and legal perspective with focus on sensitizing substances. Contact Dermatitis. 2021;85(3):340-53.
56. Hauri U, Hohl C. Photostability and breakdown products of pigments currently used in tattoo inks. Curr Probl Dermatol. 2015;48:164-9.
57. Kiszla BM, Elmets CA, Mayo TT. Quantitative analysis of restricted metals and metalloids in tattoo inks: A systematic review and meta-analysis. Chemosphere. 2023;313:137291.
58. IARC. IARC monographs on the evaluation of carcinogenic risks to humans - arsenic, metals, fibres, and dusts. 2012;100C.
59. IARC Monographs on the Identification of Carcinogenic Hazards to Humans. International Agency for Research on Cancer. World Health Organization.; 1972-onwards.
60. IARC. IARC monographs on the evaluation of carcinogenic risks to humans - carbon black, titanuim dioxide, and talc. 2010;93.
61. Regensburger J, Lehner K, Maisch T, Vasold R, Santarelli F, Engel E, et al. Tattoo inks contain polycyclic aromatic hydrocarbons that additionally generate deleterious singlet oxygen. Exp Dermatol. 2010;19(8):e275-81.
62. IARC. IARC monographs on the evaluation of carcinogenic risks to humans - chemical agents and related occupants. 2012;100F.
63. Jacobsen NR, Clausen PA. Carbon black nanoparticles and other problematic constituents of black ink and their potential to harm tattooed humans. Curr Probl Dermatol. 2015;48:170-5.
64. Engel E, Santarelli F, Vasold R, Maisch T, Ulrich H, Prantl L, et al. Modern tattoos cause high concentrations of hazardous pigments in skin. Contact Dermatitis. 2008;58(4):228-33.
65. Engel E, Vasold R, Santarelli F, Maisch T, Gopee NV, Howard PC, et al. Tattooing of skin results in transportation and light-induced decomposition of tattoo pigments--a first quantification in vivo using a mouse model. Exp Dermatol. 2010;19(1):54-60.
66. Sabboni G, Hauri U. Carcinogenic Tattoos. Epidemiology Biostatistics and Public Health. 2016;13(4).
67. Haleem A, Amin S, Mahmood U. Heavy metal and polycyclic aromatic hydrocarbons in cigarettes: An analytical assessment. Population Medicine. 2020;2(June).
68. Singh L, Varshney JG, Agarwal T. Polycyclic aromatic hydrocarbons' formation and occurrence in processed food. Food Chem. 2016;199:768-81.
69. Lisbeth Rosenkrantz Hölmich SJH-D, Francesco Annibale d'Amore. Brystimplantatassocieret anaplastisk storcellet lymfom. Ugeskrift for Læger. 2021(183).
70. Cohen PR. Squamous Cell Carcinoma and Tattoo: A Man With a Tattoo-Associated Squamous Cell Carcinoma and Review of Benign Tumors, Lymphoid Conditions, and Malignant Neoplasms Occurring Within a Tattoo. Cureus. 2022;14(1):e21083.
71. van der Bent SAS, Rauwerdink D, Oyen EMM, Maijer KI, Rustemeyer T, Wolkerstorfer A. Complications of tattoos and permanent makeup: overview and analysis of 308 cases. J Cosmet Dermatol. 2021;20(11):3630-41.
72. Serup J. Atlas of Illustrative Cases of Tattoo Complications. Curr Probl Dermatol. 2017;52:139-229.
73. Wenzel SM, Rittmann I, Landthaler M, Baumler W. Adverse reactions after tattooing: review of the literature and comparison to results of a survey. Dermatology. 2013;226(2):138-47.
74. Kluger N. Cutaneous complications after tattooing in Finland from 2016 to 2021. J Eur Acad Dermatol Venereol. 2022;36(1):e72-e3.
75. Bourgeois P, Roman MM, Schweicher J, Lavoisier P, Maquet P, Karler C, et al. Lymphatic Alterations Under Tattoos: Preliminary Reports of One Observational Study. Clin Cosmet Investig Dermatol. 2023;16:257-65.
76. U. S. Food and Drug Administration (FDA). Think Before You Ink: Tattoo Safety: United States Government; 2023 [Available from: https://www.fda.gov/consumers/consumer-updates/think-you-ink-tattoo-safety.
77. Miljøstyrelsen. Think Before You Ink: Miljøministeriet; 2023 [Available from: https://mst.dk/borger/sundhed-og-kemi/personlig-pleje/think-before-you-ink.
78. Moehrle M, Blaheta HJ, Ruck P. Tattoo Pigment Mimics Positive Sentinel Lymph Node in Melanoma. Dematology. 2001;203(4):342-4.
79. Paul Litton T, Vijay Ghate S. Tattoo pigment mimicking axillary lymph node calcifications on mammography. Radiol Case Rep. 2020;15(8):1194-6.
80. Fraser TR, Ross KE, Alexander U, Lenehan CE. Current knowledge of the degradation products of tattoo pigments by sunlight, laser irradiation and metabolism: a systematic review. J Expo Sci Environ Epidemiol. 2022;32(3):343-55.
81. Schreiver I, Hutzler C, Laux P, Berlien HP, Luch A. Formation of highly toxic hydrogen cyanide upon ruby laser irradiation of the tattoo pigment phthalocyanine blue. Sci Rep. 2015;5:12915.

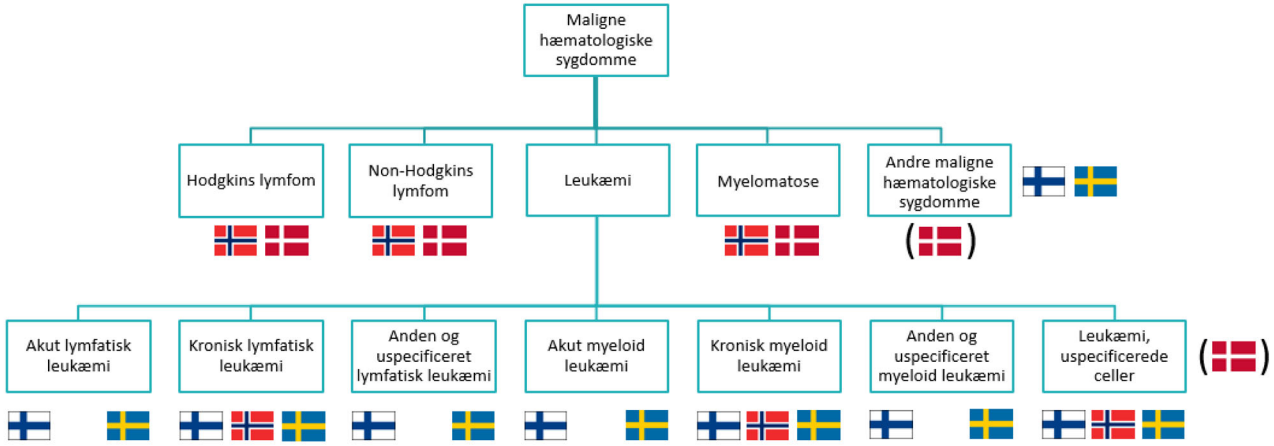
82. Hering H, Sung AY, Roder N, Hutzler C, Berlien HP, Laux P, et al. Laser Irradiation of Organic Tattoo Pigments Releases Carcinogens with 3,3'-Dichlorobenzidine Inducing DNA Strand Breaks in Human Skin Cells. J Invest Dermatol. 2018;138(12):2687-90.
83. Kemikalieinspektionen. Farliga ämnen i tatueringsfärger. Rapport 3/10.; 2010.
84. Council of Europe. Resolution ResAP (2008)1 on requirements and criteria for the safety of tattoos and permanent make-up (superseding Resolution ResAP (2003)2 on tattoos and permanent make-up). 2008.
85. European Chemicals Agency (ECHA). Annex XV Restriction Report. Proposal for eestriction, substances in tattoo inks and permanent make up. Version 1.2.; 2017.
86. European Chemicals Agency (ECHA). Committee for Risk Assessment (RAC) Committee for Socio-Economic Analysis (SEAC) - Opinion on an Annex XV Dossier Proposing Restrictions on Substances Used in Tattoo Inks and Permanent Make-Up.; 2019.
87. European Commission. Chemicals: EU takes action for safer tattooing inks and permanent make-up 2020 [Available from: https://single-market-economy.ec.europa.eu/news/chemicals-eu-takes-action-safer-tattooing-inks-and-permanent-make-2020-12-14_en.
88. Søndergaard S, Brix S, Gjerding MR, Villumsen N. B 39 Forslag til folketingsbeslutning om forbud mod giftstoffer i tatoveringsblæk. 2017.
89. McDonagh S. How the new EU ink ban may be permanently scarring the tattoo industry. Euronews Culture.5 January 2022. [accessed 27 September 2023]. Available from: https://www.euronews.com/culture/2022/01/05/how-the-new-eu-ink-ban-may-be-permanently-scarring-the-tattoo-industry.
90. Kaadt T. Fyns mest rutinerede tatovør risikerer at gå fra hus og hjem. Fyns Amts Avis. 2021 26 October 2021;Sect. Fyn.
91. Kaadt T. Efter møde mellem tatovør og Miljøstyrelsen: Fortsat ingen afklaring i tatoveringskrise. Fyens Stiftstidende. 2021 29 October 2021;Sect. Nyheder.
92. Pedersen DA, Larsen LA, Nygaard M, Mengel-From J, McGue M, Dalgard C, et al. The Danish Twin Registry: An Updated Overview. Twin Res Hum Genet. 2019;22(6):499-507.
93. Skytthe A, Harris JR, Czene K, Mucci L, Adami HO, Christensen K, et al. Cancer incidence and mortality in 260,000 Nordic twins with 30,000 prospective cancers. Twin Res Hum Genet. 2019;22:99-107.
94. Clemmensen SB, Harris JR, Mengel-From J, Bonat WH, Frederiksen H, Kaprio J, et al. Familial Risk and Heritability of Hematologic Malignancies in the Nordic Twin Study of Cancer. Cancers (Basel). 2021;13(12).
95. Rothman K, Greenland S. Modern Epidemiology. 2nd ed. Philadelphia: Lippincott-Raven; 1998.
96. Christensen K, McGue M. Chapter 26 - Twin-singleton differences. In: Tarnoki A, Tarnoki D, Harris J, Segal N, editors. Twin Research for Everyone: Academic Press; 2022. p. 439-56.
97. Sham P. Statistics in Human Genetics: Arnold; 1998.
98. Pearl J, Glymour M, Jewell NP. Causal inference in statistics : a primer. Chichester, West Sussex, UK: John Wiley & Sons Ltd Chichester, West Sussex, UK; 2016. Available from: http://site.ebrary.com/id/11151663.
99. Tönnies T, Kahl S, Kuss O. Collider Bias in Observational Studies. Dtsch Arztebl Int. 2022;119(7):107-22.
100. Scheike TH, Holst KK, Hjelmborg JB. Estimating twin concordance for bivariate competing risks twin data. Stat Med. 2014;33(7):1193-204.
101. Holst KK, Scheike TH, Hjelmborg JB. The liability threshold model for censored twin data. Computational Statistics & Data Analysis. 2016;93:324-35.
102. Scheike TH, Holst KK, Hjelmborg JB. Estimating heritability for cause specific mortality based on twin studies. Lifetime data analysis. 2014;20(2):210-33.
103. Hernan MA. The hazards of hazard ratios. Epidemiology. 2010;21(1):13-5.
104. Kragh Andersen P, Pohar Perme M, van Houwelingen HC, Cook RJ, Joly P, Martinussen T, et al. Analysis of time-to-event for observational studies: Guidance to the use of intensity models. Stat Med. 2021;40(1):185-211.
105. Holt JD, Prentice RL. Survival analyses in twin studies and matched pair experiments. Biometrika. 1974;61(1):17-30.
106. Rasmussen SHR, Ludeke S, Hjelmborg JVB. A Major Limitation of the Direction of Causation Model: Non-Shared Environmental Confounding. Twin Research and Human Genetics. 2019;22(1):14-26.
107. Li W, Christiansen L, Hjelmborg J, Baumbach J, Tan Q. On the power of epigenome-wide association studies using a disease-discordant twin design. Bioinformatics. 2018;34(23):4073-8.
108. Meinshausen N, Bühlmann P. Stability selection. Journal of the Royal Statistical Society: Series B (Statistical Methodology). 2010;72(4):417-73.
109. Friedman J, Hastie T, Tibshirani R. Regularization Paths for Generalized Linear Models via Coordinate Descent. Journal of Statistical Software. 2010;33(1).
110. Zhang Y, Wang R, Holford TR, Leaderer B, Zahm SH, Boyle P, et al. Family history of hematopoietic and non-hematopoietic malignancies and risk of non-Hodgkin lymphoma. Cancer Causes Control. 2007;18(4):351-9.

References

111. Chang ET, Smedby KE, Hjalgrim H, Porwit-MacDonald A, Roos G, Glimelius B, et al. Family history of hematopoietic malignancy and risk of lymphoma. J Natl Cancer Inst. 2005;97(19):1466-74.
112. Altieri A, Bermejo JL, Hemminki K. Familial risk for non-Hodgkin lymphoma and other lymphoproliferative malignancies by histopathologic subtype: the Swedish Family-Cancer Database. Blood. 2005;106(2):668-72.
113. Wang SS, Slager SL, Brennan P, Holly EA, De Sanjose S, Bernstein L, et al. Family history of hematopoietic malignancies and risk of non-Hodgkin lymphoma (NHL): a pooled analysis of 10 211 cases and 11 905 controls from the International Lymphoma Epidemiology Consortium (InterLymph). Blood. 2007;109(8):3479-88.
114. Schinasi LH, Brown EE, Camp NJ, Wang SS, Hofmann JN, Chiu BC, et al. Multiple myeloma and family history of lymphohaematopoietic cancers: Results from the International Multiple Myeloma Consortium. Br J Haematol. 2016;175(1):87-101.
115. VanValkenburg ME, Pruitt GI, Brill IK, Costa L, Ehtsham M, Justement IT, et al. Family history of hematologic malignancies and risk of multiple myeloma: differences by race and clinical features. Cancer Causes Control. 2016;27(1):81-91.
116. Cerhan JR, Slager SL. Familial predisposition and genetic risk factors for lymphoma. Blood. 2015;126(20):2265-73.
117. Sud A, Chattopadhyay S, Thomsen H, Sundquist K, Sundquist J, Houlston RS, et al. Analysis of 153115 patients with hematological malignancies refines the spectrum of familial risk. Blood. 2019;134(12).
118. Mack TM, Cozen W, Shibata DK, Weiss LM, Nathwani BN, Hernandez AM, et al. Concordance for Hodgkin's disease in identical twins suggesting genetic susceptibility to the young-adult form of the disease. The New England journal of medicine. 1995;332(7):413-8.
119. Mensah FK, Willett EV, Ansell P, Adamson PJ, Roman E. Non-Hodgkin's lymphoma and family history of hematologic malignancy. Am J Epidemiol. 2007;165(2):126-33.
120. Goldgar DE, Easton DF, Cannon-Albright LA, Skolnick MH. Systematic Population-Based Assessment of Cancer Risk in First-Degree Relatives of Cancer Probands. JNCI. 1994;86(21).
121. Kharazmi E, Fallah M, Pukkala E, Olsen JH, Tryggvadottir L, Sundquist K, et al. Risk of familial classical Hodgkin lymphoma by relationship, histology, age, and sex: a joint study from five Nordic countries. Blood. 2015;126(17):1990-5.
122. Fallah M, Kharazmi E, Pukkala E, Tretli S, Olsen JH, Tryggvadottir L, et al. Familial risk of non-Hodgkin lymphoma by sex, relationship, age at diagnosis and histology: a joint study from five Nordic countries. Leukemia. 2016;30(2):373-8.
123. Heywood W, Patrick K, Smith AM, Simpson JM, Pitts MK, Richters J, et al. Who gets tattoos? Demographic and behavioral correlates of ever being tattooed in a representative sample of men and women. Ann Epidemiol. 2012;22(1):51-6.
124. Armstrong ML, Owen DC, Roberts AE, Koch JR. College students and tattoos. Influence of image, identity, family, and friends. Journal of Psychosocial Nursing and Mental Health Services. 2002;40(10).
125. Kluger N. Epidemiology of tattoos in industrialized countries. Curr Probl Dermatol. 2015;48:6-20.
126. Single and multiple deliveries by type of delivery (1850-2018) [Internet]. Statistikbanken - Danmark Statistik. 2023 [cited 12 November 2023]. Available from: https://www.statbank.dk/20017.
127. Austin PC, Latouche A, Fine JP. A review of the use of time-varying covariates in the Fine-Gray subdistribution hazard competing risk regression model. Stat Med. 2020;39(2):103-13.
128. Sundhedsdatastyrelsen. Cancerregistret [Available from: https://sundhedsdatastyrelsen.dk/da/registre-og-services/om-de-nationale-sundhedsregistre/sygdomme-laegemidler-og-behandlinger/cancerregisteret.
129. Scheike TH, Sun Y, Zhang MJ, Jensen TK. A semiparametric random effects model for multivariate competing risks data. Biometrika. 2010;97(1):133-45.

Appendix 1: Disease classification of hematologic malignancies in the Nordic countries

Oversigt over klassifikation af hæmatologiske cancere i det nordiske tvilling cancer studie. Udarbejdet i forbindelse med artikel 1 i denne afhandling.



Tilgængelige variable:

Sverige: ICDO3, ICDO10, icd7, ICD9

Finland: Nordcan (baseret på icd10/Nordcan liste)

Norge: Tum_icd7, Tum_icd10 (intervaller), Tum_topo_O3, Morf_O2, Tum_hist_morf (før 1993), Morf_O3

Danmark: C_ICD10, C_MORFO3, C_ICD7 (+ en masse andet, se nedenfor)

Udsnit af variable fra det danske cancerregister – fra "*Dokumentation fra cancerregistret*": (128)

Variabel	årsinterval	Indhold
K_TUMORNR	1943-2016	Tumor løbenr
C_ICD10	1978-2016	ICD10 koden
C_MORFO3	1978-2016	Histologi, Morfologi iflg. ICDO3 klassifikationen
C_GRAD	1943-2008	Gradering (Kun urinvejstumorer)
C_LATERAL	1943-2016	Lateralitet for parrede organer
C_UDBRED_KLASSIFIKATION	2004-2016	Anvendt udbredelsesklassifikation
C_UDBRED	1943-2003	Tumorens Udbredelse
C_MAKROGRUNDLAG	1943-2016	Makroskopisk grundlag
C_MIKROGRUNDLAG	1943-2016	Mikroskopisk grundlag
C_BEHANDLING	1943-2003	Behandling
C_ICD7	1943-1977	Den gamle nordiske ICD7-klassifikation
C_DIAGGR	1943-1977	ICD7 diagnosegruppering - undergrupper
C_ORGGR	1943-1977	ICD7 diagnosegruppering - overordnede grupper
C_SARC	1943-1977	Er sarcom eller ej ?
C_DIAGGR_ICD10	1978-2016	ICD10 - diagnosegruppering - undergrupper
C_ORGGR_ICD10	1978-2016	ICD10 - diagnosegruppering - overordnede grupper
c_DIAGGR_Nordcan	1943-2016	NordCan grupper

Appendix 1: Disease classification of hematologic malignancies in the Nordic countries

Sammenfatning af informationer fra klinisk professor i hæmatologi, Henrik Frederiksen, og Nordcan (4, 5).

	Henrik Frederiksen	**Nordcan**
Hematological malignancies	C81-C96	C81-86,C88,C90-96,D45-47.0-1,3-9
Hodgkin lymphoma	C81	C81
Non-Hodgkin lymphomas	C82.0-C85.9, C86, C88.0-9, C91.1b	C82-86
Leukemia		C91-95
Acute myeloid leukemia	C92.0 (excl. C921, C922, c927b), C93.0, C93.7, C93.9 (except C94.6 and C931), C95.0	C92.0+C93.0+C94.0+C94.2+C94.4-5
Acute lymphoid leukemia	C91.0	C91.0
Chronic myeloid leukemia	C92.1, C92.2	C92.1+C93.1+C94.1
Multiple myeloma	C90.0	C90
Chronic lymphocytic leukemia	C91.1 excl C91.1b	C91.1
Myelodysplastic syndromes / chronic myelomonocytic leukemia	D46, C931	
Other hematological malignancy	All others codes between C81-C96 not included in the above mentioned classifications	C88,C96,D47.0,7-9
Chronic myeloproliferative neoplasms (not in cancer registries until recent years)		
Essential thrombocythaemia (ET)	D473; D752	
Polycythaemia vera (PV)	D459	
Myelofibrosis (MF) / Myeloproliferative neoplasm unspecified (MPN-U)	D471; D474	

Appendix 2: Questionnaire and invitation letter

Full questionnaire: (7 pages)

Risikofaktorerfor visse typer af kræftsygdomme

Formålet med dette spørgeskema er at undersøge mulige risikofaktorer for visse typer af kræft. Vi er blandt andet interesserede i at vide mere om, hvorvidt tatoveringsblæk kan relateres til visse kræftformer såsom hud- og blodkræft.
I undersøgelsen spørges til tatoveringer samt livsstil og uddannelse.

Du kan læse mere om baggrunden for undersøgelsen på følgende hjemmeside:
www.sdu.dk/dtr.

Dine svar vil være nyttige, uanset om du har en tatovering eller ej. Information om hvordan vi behandler personoplysninger kan findes i invitationsbrevet.

Når du har svaret på spørgeskemaet, bedes du sende det tilbage i vedlagte returkuvert.

På forhånd tak for hjælpen!

Navn: «Navn»
Respondent-ID: «Kode» 1

1. Har du en eller flere tatoveringer? (Sæt ét kryds)

- ❑ Ja, en
- ❑ Ja, flere
- ❑ Nej

Hvis du svarede ”Ja, en” eller ”Ja, flere” til spm. 1, gå venligst til spm. 3.

2. Overvejer du at få lavet en tatovering? (Sæt ét kryds)

- ❑ Ja
- ❑ Nej
- ❑ Ved ikke
- ❑ Vil ikke oplyse

3. Har din medtvilling en eller flere tatoveringer? (Sæt ét kryds)

- ❑ Ja, en
- ❑ Ja, flere
- ❑ Nej
- ❑ Ved ikke

Hvis du svarede ”Nej” til spm. 1, gå venligst til spm. 13.

4. Hvilke farver er du tatoveret med? (Sæt ét eller flere krydser)

- ❑ Sort
- ❑ Rød
- ❑ Blå
- ❑ Hvid
- ❑ Grå
- ❑ Gul
- ❑ Orange
- ❑ Lilla
- ❑ Grøn
- ❑ Andre

Respondent-ID: «Kode» 2

5. Er du bekendt med hvilket blækprodukt, du er blevet tatoveret med? (Sæt ét kryds)

- ❑ Ja
- ❑ Nej

Hvis du svarede ”Nej” til spm. 5, gå venligst til spm. 7.

6. Skriv venligst navnet (mærket) på blækproduktet

_____

7. Hvor stor er din(e) tatovering(er) cirka målt i håndflader? (Ved flere tatoveringer angives den samlede størrelse) (Sæt ét kryds)

- ❑ Mindre end 1 håndflade
- ❑ 1-5 håndflader
- ❑ Mere end 5 håndflader
- ❑ Ved ikke

8. Hvor gammel var du, da du fik din første tatovering? (år)

_____

9. Har du eller din læge bemærket hævede lymfekirtler i nærheden af tatoveringen? (Sæt ét kryds)

- ❑ Ja
- ❑ Nej
- ❑ Ved ikke

Hvis du svarede ”Nej” eller ”Ved ikke” til spm. 9, gå venligst til spm. 12.

10. Hvor havde du hævede lymfekirtler? (Sæt ét eller flere krydser)

- ❑ Lyske
- ❑ Armhule
- ❑ Hals
- ❑ Andet
- ❑ Ved ikke

11. Hvor var den nærmeste tatovering til hævelsen placeret? (Sæt ét eller flere krydser)

- ❑ Hovedet eller halsen
- ❑ Armene eller hænderne
- ❑ Overkroppen (over bæltet)
- ❑ Underkroppen (under bæltet)
- ❑ Benene eller fødderne
- ❑ Ved ikke

12. Har du nogensinde været ved læge eller på hospitalet, fordi du havde gener af en tatovering? (Sæt ét kryds)

- ❑ Ja
- ❑ Nej
- ❑ Ved ikke
- ❑ Vil ikke oplyse

13. Har du fået fjernet en tatovering? (Sæt ét kryds)

- ❑ Ja
- ❑ Nej
- ❑ Ved ikke
- ❑ Vil ikke oplyse

Hvis du svarede ”Nej”, ”Ved ikke” eller ”Vil ikke oplyse” til spm. 13, gå venligst til spm. 15.

Respondent-ID: «Kode» 4

14. Hvad var årsagen til, at tatoveringen blev fjernet? (Sæt ét eller flere krydser)

- ❑ Var ikke længere tilfreds med udseendet
- ❑ Var ikke længere interesseret i at have en tatovering
- ❑ Medicinske gener som f.eks. kløe eller betændelse
- ❑ I forbindelse med hudkræft eller mistanke herom
- ❑ Andet: _____

15. Har du fået fjernet hud, der IKKE var tatoveret, i forbindelse med hudkræft eller mistanke herom? (Sæt ét kryds)

- ❑ Ja
- ❑ Nej
- ❑ Ved ikke
- ❑ Vil ikke oplyse

16. Ryger du? (cigaret, e-cigaret, cigar, cerut, pibe eller andet) (Sæt ét kryds)

- ❑ Ja
- ❑ Nej, jeg er holdt op
- ❑ Nej

Hvis du svarede ”Nej” til spm. 16, gå venligst til spm. 21.

Hvis du svarede ”Ja” eller ”Nej, jeg er holdt op” til spm. 16, gå venligst til spm. 17

17. Hvad ryger/røg du? (Sæt ét eller flere krydser)

- ❑ Cigaret
- ❑ E-cigaret
- ❑ Cigar
- ❑ Cerut
- ❑ Pibe
- ❑ Andet

18. Hvor mange gange ryger/røg du gennemsnitligt om dagen?
(Eksempel: 1 gang = 1 cigaret)

_____gange

19. Hvor gammel var du, da du begyndte at ryge?

_____år

Hvis du svarede "Ja" til spm. 16, gå venligst til spm. 21.

20. Hvor mange år af dit liv har du røget?

_____år

21. På en typisk uge, hvor meget tid bruger du på moderat og hård fysisk aktivitet?
Medtag kun aktiviteter, der varer i mindst 10 minutter ad gangen.

	Fysisk aktivitet (antal minutter)
Moderat fysisk aktivitet, hvor du kan mærke, at pulsen og vejrtrækningen øges (det kan f.eks. være rask gang, cykling som transport eller motion, tungt havearbejde, løb eller motionsidræt)	_____minutter
Hård fysisk aktivitet, som øger pulsen væsentligt, får dig til at svede og gør dig så forpustet, at det er svært at tale (det kan f.eks. være svømning, løb, cykling i højt tempo, konditionstræning, hård styrketræning eller boldspil)	_____minutter

22. Hvor mange genstande alkohol drikker du typisk om ugen? (Sæt ét kryds)

- ❑ 0 genstande
- ❑ 1-7 genstande
- ❑ 8-14 genstande
- ❑ 15-21 genstande
- ❑ 22 eller flere genstande
- ❑ Ved ikke
- ❑ Vil ikke oplyse

23. Hvad er din senest afsluttede uddannelse? (Sæt ét kryds)

- ❑ Jeg har ingen formel uddannelse
- ❑ Folkeskole
- ❑ Gymnasial uddannelse
- ❑ Erhvervsuddannelse/faglært (fx social- og sundhedsassistent, tømrer, frisør)
- ❑ Kort videregående uddannelse (fx tandplejer, politibetjent, markedsføringsøkonom)
- ❑ Mellemlang videregående uddannelse (fx pædagog, sygeplejerske, bachelor)
- ❑ Lang videregående uddannelse (fx civilingeniør, gymnasielærer, læge)
- ❑ Anden uddannelse
- ❑ Ved ikke
- ❑ Vil ikke oplyse

24. Har du yderligere kommentarer til spørgeskemaet?

_____

_____

_____

_____

_____

_____

Respondent-ID: «Kode» 7

Survey invitation

Page 1/3:

SyddanskUniversitet
Institut for Sundheds-
tjenesteforskning,
DanskCenterfor
Tvillingforskning,
J.B. WinsløwsVej 9B
5000 OdenseC

www.sdu.dk/dtr

Dato: Januar2021

Kære Signe Bedsted Clemmensen

Vi er en gruppe forskere tilknyttet Dansk Center for Tvillingforskning ved Syddansk Universitet, som gerne vil undersøge hvilke faktorer, der kan medvirke til at udvikle visse typer af kræft.

Vi vil gerne invitere dig til at deltage i studiet og vil derfor høre, om du vil besvare et kort spørgeskema. Dine oplysninger vil udelukkende blive brugt til forskning og vil blive opbevaret forsvarligt i henhold til databeskyttelsesloven (se nedenfor).

Vi er blandt andet interesserede i at vide mere om, hvorvidt tatoveringsblæk kan relateres til visse kræftformer såsom hud- og blodkræft. Du kan læse mere om baggrunden for undersøgelsen på følgende hjemmeside: www.sdu.dk/dtr

I undersøgelsen spørger vi ind til tatoveringer samt livsstil og uddannelse. Dine svar vil være nyttige, uanset om du har en tatovering eller ej.

Undersøgelsen foregår via internettet og spørgeskemaet skal derfor besvares online. Nedenfor finder du det link, du skal benytte, hvis du beslutter dig for at deltage i undersøgelsen:

Link til spørgeskema:

https://www.survey-xact.dk/answer?key=U6MPKJI3A2Z8

181584373

Åbner spørgeskemaet ikke, når du trykker på linket, kan du gå ind på www.datafabrikken.dk og indtaste følgende kode: U6MPKJI3A2Z8

Det tager cirka 5 minutter at besvare de spørgsmål, der indgår i undersøgelsen.

Page 2/3:

Det er helt frivilligt, om du ønsker at deltage. De oplysninger, du giver os, er omfattet af tavshedspligt og derfor kun tilgængelige for de deltagende forskere. Vi udleverer ikke oplysninger til offentlige myndigheder, forsikringsselskaber eller arbejdsgivere, og de anvendes ikke til kommercielt brug.

Undersøgelsen gennemføres med støtte fra forskellige fonde. Ingen af de deltagende forskere har økonomisk udbytte af undersøgelsen, ligesom ingen af dem har tilknytning til private virksomheder eller fonde, som har direkte interesse i det pågældende projekt.

I henhold til databeskyttelsesloven er vi forpligtet til at fortælle, hvilke personoplysninger, vi behandler, samt på hvilken måde. Det kan du læse mere om på den sidste side.

Skulle du have nogle spørgsmål vedrørende undersøgelsen, kan du kontakte den ansvarlige for undersøgelsen:

Signe B. Clemmensen, e-mail: sbclemmensen@health.sdu.dk, tlf. 6550 3970

Vi håber, at du vil tage dig tid til at deltage i undersøgelsen ved at svare på spørgsmålene.

På forhånd tak for hjælpen!

Venlig hilsen

Signe B. Clemmensen
Ph.d.-studerende
Epidemiologi, Biostatistik og Biodemografi
Institut for Sundhedstjenesteforskning
Syddansk Universitet
Dansk Center for Tvillingforskning

Jacob v. B. Hjelmborg
Professor, ph.d.
Epidemiologi, Biostatistik og Biodemografi
Institut for Sundhedstjenesteforskning
Syddansk Universitet
Dansk Center for Tvillingforskning

Henrik Frederiksen
Professor, overlæge, ph.d.
Hæmatologisk afdeling X,
Odense Universitetshospital
Klinisk Institut, Syddansk Universitet

Page 3/3:

Oplysning om behandling af personoplysninger

I forbindelse med projektet *Tatoveringer og risiko for cancer*, ønsker SDU at indsamle oplysninger om dig. SDU er ansvarlig for beskyttelsen af dine personoplysninger til brug for forskningsprojekter. Det er frivilligt at deltage i projektet. Indsamlingen sker via Det Danske Tvillingregister ved brug af e-Boks.

Formål med behandlingen af personoplysninger
Undersøgelsen er en del af et studie, der har til formål at undersøge, hvorvidt tatoveringsblæk kan relateres til risikoen for visse kræftformer såsom hud- og blodkræft.

De oplysninger, der behandles er:
Navn, adresse, om du er enægget eller toægget tvilling, fødselsdato/alder, køn, om du har haft cancer og, hvis det er tilfældet, hvilken diagnose, samt de svar du giver i forbindelse med denne undersøgelse.

Sådan bruger vi personoplysningerne
SDU behandler personoplysningerne fortroligt - i overensstemmelse med gældende ret. Oplysningerne vil kun blive brugt til forskning mens formidling af forskningsresultater vil ske i anonymiseret form. Vi sørger for at opbevare dine oplysninger sikkert, så det kun er de relevante forskere på SDU, der har adgang til dem. Efter afslutning af projektet opbevares data ved Det Danske Tvillingregister til anvendelse indenfor fremtidig sundhedsvidenskabelig forskning og kan i den forbindelse blive videregivet til andre forskere ved andre anerkendte forskningsinstitutioner.

Ønske om at deltage
Når der skal bruges personoplysninger til forskningsprojektet, er der nogle særlige bestemmelser i lovgivningen, som giver mulighed for at indsamle og bruge personoplysninger uden samtykke fra deltageren. Disse bestemmelser findes i databeskyttelseslovens § 10 og databeskyttelsesforordningens art. 6, stk. 1, litra e.

Bestemmelsen giver os lov til at bruge dine personoplysninger til forskning uden dit samtykke, men forbyder også brug af oplysningerne til andre formål end forskning og statistik. Du risikerer derfor ikke, at dine oplysninger vil blive brugt til andre formål.

Databeskyttelseslovens § 10 og databeskyttelsesforordningens art. 6, stk. 1, litra e giver også mulighed for, at vi kan videregive oplysningerne til andre forskningsprojekter, men der vil fortsat være tale om epidemiologisk forskning.

Du skal derfor beslutte dig for, om du vil deltage i projektet, så vi kan anvende dine personoplysninger til forskning. Hvis du indvilliger i at deltage, vil vi indsamle de oplysninger, som er nævnt ovenfor, og anvende dem i projektet. Skulle du på et tidspunkt i forløbet ikke længere have lyst til at deltage, kan du trække dig fra projektet. Det medfører at vi ikke længere vil indsamle nye oplysninger om dig, men vi har fortsat lov til at bruge de oplysninger, vi allerede har fået.

Yderligere information
Hvis du har spørgsmål til undersøgelsen, kan du til enhver tid kontakte Signe B. Clemmensen på tlf. 6550 3970 eller via mail til sbclemmensen@health.sdu.dk.

Hvis du har spørgsmål omkring databeskyttelse og dine rettigheder, kan du kontakte vores databeskyttelsesrådgiver, Simon Kamber, på tlf. 6550 3906 eller via mail til dpo@sdu.dk

Appendix 3: Setting up the twin cohort

Introduction

In the case-cotwin sample, we selected all twin pairs where at least one twin has been diagnosed with a tattoo interest cancer and at least one twin could be contacted.

In the reference sample, we selected 5,000 twin pairs randomly among all twins in the Danish Twin Register (DTR), excluding those in the case-cotwin sample, where at least one twin could be contacted.

	Invited	Unreachable	Total
Reference	8,881	1,119	10,000
Case-control	2,246	776	3,022
Total	11,127	1,895	13,022

Among the 2,246 invited in the case-cotwin sample, the proportions of individuals with cancer were 48.6% for males and 44.1% for females. In the reference sample, there were 4,419 males and 4,462 females.

Proportions of cancers within the cohort

To form a cohort representative of the background population, we combined the reference sample with a pre-determined proportion of the case-cotwin sample randomly selected. The proportions were determined by sex using data from the NORDCAN database on number of individuals alive per 100,000 in 2020 (4, 5). The proportion were:

Cancer type	Male	Female
Skin, non-melanoma	0.5159 %	0.4318 %
Melanoma of skin	0.5260 %	0.7281 %
Liver	0.0205 %	0.0099 %
Bladder and urinary tract	0.5625 %	0.1999 %
Non-Hodgkin lymphoma	0.2676 %	0.2207 %
Hodgkin lymphoma	0.0652 %	0.0493 %
Multiple myeloma	0.0639 %	0.0529 %
Sum	2.0216 %	1.6926 %

It was possible for an individual to be diagnosed with more than one of the cancers listed. This was ignored in the sum. Hence, the expected total estimated below assumed at most one diagnosis per individual. That was incorrect, but the error was considered negligible.

Estimating the expected total number of twins in the cohort

We now have the proportions of cancer by sex in the cohort, but we need to find the total number of males and females to extract the number of individuals with each cancer type.

For males, 2.02% of the cohort were expected to have an interest cancer. The number of males in the cohort with interest cancer, C, is expected to be 48.6% of T, the (unknown) total number of males in the cohort, minus 4,419 (number of males in reference sample):

$$C = 0.0202 \cdot T$$

$$C = (T - 4419) \cdot 0.486$$

That gave an expected total number of males in the cohort of $T = 4{,}611$.

Similarly, for females, 1.69% of the cohort was expected to have an interest cancer. The number of females in the cohort with interest cancer, C, was expected to be 44.1% of T, the (unknown) total number of females in the cohort, minus 4,462 (number of females in reference sample):

$$\mathrm{C} = 0.0169 \cdot T$$

$$C = (T - 4462) \cdot 0.441$$

That gave an expected total number of females in the cohort of $T = 4{,}640$.

Number of males and females from to case-cotwin sample to be included in the cohort

The proportions provided in the table above are individuals alive in 2020. We assumed the proportions were also valid when survey invitations were sent in January/March 2021. We selected the estimated numbers of males and females with each cancer type among the invited individuals of the case-cotwin sample and then added co-twins afterwards. The co-twin may or may not have been among the invited case-cotwin pairs - that would correspond to the procedure in the reference sample where we chose 5,000 reachable twins and added the cotwin (who may not be reachable).

Number of individuals to select from case-control group:

Cancer type	Male	Female	n Male	n Female
Skin, non-melanoma	0.5159 %	0.4318 %	24	20
Melanoma of skin	0.5260 %	0.7281 %	24	34
Liver	0.0205 %	0.0099 %	1	0
Bladder and urinary tract	0.5625 %	0.1999 %	26	9
Non-Hodgkin lymphoma	0.2676 %	0.2207 %	12	10
Hodgkin lymphoma	0.0652 %	0.0493 %	3	2
Multiple myeloma	0.0639 %	0.0529 %	3	2
Sum	2.0216 %	1.6926 %	93	79

Forming the cohort

We then selected the estimated numbers of individuals from the case-cotwin sample and combined with the reference sample. It was possible to select the same individual more than once due to multiple diagnosis and it was possible to select both twins in a pair by chance.

Random selection (with a specified seed to ensure replication) resulted in 172 twin pairs out of which 292 twin individuals could be contacted. The resulting cohort then held 9,173 individual twins - all 8,881 individuals from the reference sample and 292 individuals from the case-cotwin sample.

Appendix 4: Estimation of the covariance function

The concordance function is a bivariate probability expressing the risk of observing an event in both twins in a pair before some time *t*. In the following, we briefly present and compare three methods of estimation.

Counting process modelling

The standard non-parametric product-limit estimator (Aalen-Johansen) applies to bivariate competing risks data when assuming that both twins in a pair are censored at the same time. When adjusting for covariates, it is necessary to assume a standard semi-parametric regression structure and thus rely on inverse probability weights to account for delayed entry and right-censoring as described by Scheike et al. (100).

Semi-parametric random effects modelling

The concordance function of competing risk data can also be modelled as a semi-parametric random effects model combining the marginal cumulative incidence functions on copula form and parameters from the distribution of random effects as proposed by Scheike et al (129) and (102). It is still necessary to assume same censoring within twin pairs and care must be taken to correctly specify the dependence structure over time (101).

Parametric liability threshold modelling

The parametric liability threshold model estimates the probability of experiencing an event as a function of the latent unmeasured liability assuming normal distribution. It is a bivariate biprobit model assuming equal cumulative incidence functions for MZ and DZ twins and using inverse probability weighting to account for censoring (15, 101). This model only applies to bivariate events of the same type, i.e., not cross-events. It is not yet feasible to account for delayed entry or competing risk with this approach. A more comprehensive presentation of the liability threshold model is provided by Holst et al. (101).

Comparison of practical applications

All three approaches were tested on the data of hematologic malignancies described in Paper 1. Fitting the random effects model and the liability threshold model when there were only few twin pairs concordant for an event (around five or less) was not feasible as they did not yield converging estimates. That was not an issue with counting process modelling suggesting this would be the preferred method when dealing with rare events or events with low familial dependence.

When concordant events are more frequent, all three models are expected to provide similar estimates. Generally, the variability of inference is expected to be lowest among parametric models provided that assumptions are met.

Holst et al. (101) provides an example of breast cancer heritability estimation (which relies upon the concordance functions) applying the counting process and random effects models. The estimates are similar between the two methods, but the variability of the counting process is a lot higher.

For concordance function estimations carried out throughout this thesis, we decided to rely on 1) non-parametric counting-process modelling, as it provides the most reliable estimates with as few assumptions as possible, and 2) the liability threshold model, mainly to allow for comparison with other studies of the heritability and shared environmental effects on the liability scale. Further comparison of the selected models is given in the methods section in the main text.

Appendix 5: Shared genetic and environmental effects

We are interested in expressing the genetic and shared environmental contributions to the total variation in risk in terms of cumulative incidence and concordance functions.

Based on the ACE model, the within pair covariances for MZ and DZ twin pairs are

$$Cov_{MZ}(t) = \sigma_A^2(t) + \sigma_C^2(t)$$
$$Cov_{DZ}(t) = \frac{1}{2}\sigma_A^2(t) + \sigma_C^2(t)$$

The covariance of the E component is zero by definition as unique environmental effects are non-shared.

In this setting, the MZ and DZ covariances can be substituted by the respective concordance functions given by the bivariate probability of event in twin 1 and 2 before time t

$$C(t) = P(T_1 \leq t, T_2 \leq t,\ \epsilon_1 = 1,\ \epsilon_2 = 1)$$

assuming twin indices 1 and 2 are randomly assigned.

Expressions of variation for genetic and shared environmental contribution are obtained by simple rewriting:

$$\sigma_A^2(t) = C_{MZ}(t) - \sigma_C^2(t) = C_{MZ}(t) - C_{DZ}(t) + \frac{1}{2}\sigma_A^2(t) \Leftrightarrow$$
$$\sigma_A^2(t) = 2\big(C_{MZ}(t) - C_{DZ}(t)\big)$$
$$\sigma_C^2(t) = C_{MZ}(t) - 2\big(C_{MZ}(t) - C_{DZ}(t)\big) = 2C_{DZ}(t) - C_{MZ}(t)$$

The total variation is expressed as $F(t)\big(1 - F(t)\big)$ where $F(t)$ is the cumulative incidence function is given by the probability of event before time t

$$F(t) \ = \ P(T \leq t, \epsilon = 1)$$

Combining these expressions and assuming equal marginals, i.e., the same cumulative incidence functions for MZ and DZ twins, yields the measures of interest: Heritability, $h^2(t)$, and shared environmental effects, $c^2(t)$:

$$h^2(t) = \frac{2\big(C_{MZ}(t) - C_{DZ}(t)\big)}{F(t)\big(1 - F(t)\big)}$$

$$c^2(t) = \frac{2C_{DZ}(t) - C_{MZ}(t)}{F(t)\big(1 - F(t)\big)}$$

More information on underlying assumptions can be found in Scheike et al. (102) and Sham (97).

Paper 1

Adapted from published version to make sure everything is readable.

cancers

MDPI

Article

Familial Risk and Heritability of Hematologic Malignancies in the Nordic Twin Study of Cancer

Signe B. Clemmensen [1,2,*], Jennifer R. Harris [3], Jonas Mengel-From [1,2,4], Wagner H. Bonat [5], Henrik Frederiksen [6,7], Jaakko Kaprio [8] and Jacob v. B. Hjelmborg [1,2]

check for updates

Citation: Clemmensen, S.B.; Harris, J.R.; Mengel-From, J.; Bonat, W.H.; Frederiksen, H.; Kaprio, J.; Hjelmborg, J.v.B. Familial Risk and Heritability of Hematologic Malignancies in the Nordic Twin Study of Cancer. *Cancers* 2021, *13*, 3023. https://doi.org/10.3390/cancers13123023

Academic Editor: Francesco Bertolini

Received: 12 May 2021

Accepted: 11 June 2021

Published: 16 June 2021

[1] Department of Epidemiology, Biostatistics, and Biodemography, Institute of Public Health, University of Southern Denmark, 5000 Odense C, Denmark; jmengel-from@health.sdu.dk (J.M.-F.); jhjelmborg@health.sdu.dk (J.v.B.H.)
[2] Danish Twin Registry, Institute of Public Health, University of Southern Denmark, 5000 Odense C, Denmark
[3] Division of Health Data and Digitalisation, Norwegian Institute of Public Health, 0213 Oslo, Norway; jenniferruth.harris@fhi.no
[4] Department of Clinical Genetics, Odense University Hospital, 5000 Odense C, Denmark
[5] Department of Statistics, Paraná Federal University, Curitiba 81531-980, Brazil; wbonat@ufpr.br
[6] Department of Haematology, Odense University Hospital, 5000 Odense C, Denmark; Henrik.Frederiksen@rsyd.dk
[7] Department of Clinical Research, Institute of Public Health, University of Southern Denmark, 5000 Odense C, Denmark
[8] Department of Public Health and Institute for Molecular Medicine Finland, University of Helsinki, 00014 Helsinki, Finland; jaakko.kaprio@helsinki.fi
* Correspondence: sbclemmensen@health.sdu.dk

Simple Summary: Hematologic malignancies account for 8–9% of all incident cancers. Both genetic and environmental risk factors contribute to cancer development, but it is unclear if there is shared heritability between hematologic malignancies. This study aimed to investigate familial predisposition to hematologic malignancies using the largest twin study of cancer in the world.
We compared individual risk in the general population and the risk of cancer in one twin before some age given that the other twin had (another) cancer before that age. Furthermore, by analyzing information about whether the twins were identical or fraternal, we could estimate the relative importance of genetic and environmental influences on the risk for developing hematologic cancers. This study confirmed previous findings of familial predisposition to hematologic malignancies and provides novel evidence that familial predisposition decreases with increasing age. The latter points to the importance of taking age into account in the surveillance of hematological cancers.

Abstract: We aimed to explore the genetic and environmental contributions to variation in the risk of hematologic malignancies and characterize familial dependence within and across hematologic malignancies. The study base included 316,397 individual twins from the Nordic Twin Study of Cancer with a median of 41 years of follow-up: 88,618 (28%) of the twins were monozygotic, and 3459 hematologic malignancies were reported. We estimated the cumulative incidence by age, familial risk, and genetic and environmental variance components of hematologic malignancies accounting for competing risk of death. The lifetime risk of any hematologic malignancy was 2.5% (95% CI 2.4–2.6%), as in the background population. This risk was elevated to 4.5% (95% CI 3.1–6.5%) conditional on hematologic malignancy in a dizygotic co-twin and was even greater at 7.6% (95% CI 4.8–11.8%) if a monozygotic co-twin had a hematologic malignancy. Heritability of the liability to develop any hematologic malignancy was 24% (95% CI 14–33%). This estimate decreased across age, from approximately 55% at age 40 to about 20–25% after age 55, when it seems to stabilize. In this largest ever studied twin cohort with the longest follow-up, we found evidence for familial risk of hematologic malignancies. The discovery of decreasing familial predisposition with increasing age underscores the importance of cancer surveillance in families with hematological malignancies.

Keywords: twin study; cumulative risk; familial risk; risk between different cancers; heritability; biometric modelling; hematologic malignancy

1. Introduction

Hematologic malignancies account for 8–9% of all incident cancers, and the incidence has generally increased since the 1960s [1,2]. The major hematologic malignancies are Hodgkin lymphoma (HL), non-Hodgkin lymphoma (NHL), leukemia, and multiple myeloma (MM). The epidemiology and risk factors of these cancers have been studied for decades, but their etiology is still not entirely clarified. Heterogeneity within hematologic malignancy subtypes seems to pose a challenge for epidemiological identification of risk factors. Additionally, the incidence rates vary by age, sex, and geography [3–9].

A recent study of twins in the Nordic countries revealed that twin concordance for cancer often manifests across, rather than within, cancer types [10,11]. Namely, if one twin develops cancer, the co-twin has an increased risk of developing cancer but not necessarily at the same site. Previous epidemiological studies of hematologic malignancies suggest that this is indeed the case; they report evidence of familial predisposition across both hematologic and non-hematologic malignancies [12–21]. For instance, among women from Connecticut, lung cancer in first-degree relatives as well as breast and ovary cancer in siblings are associated with a significantly increased risk of NHL [12]. Moreover, the results indicate a stronger association between siblings than parents [9,12–15]. The impact of sex on familial predisposition has also been assessed, though the results differ by type of malignancy [9,14–16]. While the body of research indicates that familial effects for hematologic malignancies influence the risks, they are not able to further identify the source of these effects into those explained by shared genes and environment.

Twin studies allow for delving into the nature of this familial predisposition by partitioning the variation in risk into components explained by genetic and environmental influences. In 1996, Mack et al. [22] published findings from an American twin study suggesting genetic susceptibility to HL in young adulthood. A few years later, Lichtenstein et al. [23] conducted a study of twins from Denmark, Finland, and Sweden and reported familial influences for leukemia but did not take timing of events into account. A more recent Nordic study by Mucci et al. [10] also included Norwegian twins and had a longer follow-up time; they categorized leukemia as "acute" or "other" and provided an uncertain estimate of heritability for other leukemia: 57% (95% CI: 0–100%). Both of these Nordic studies analyzed multiple cancer types, including hematologic malignancies, but neither were able to estimate genetic and environmental variance for the remaining hematologic malignancies as no or too few concordant pairs were observed.

In this study, we utilize the world's most comprehensive database of twins for studying cancer. We aim to explore the genetic and environmental contributions to variation in risk of hematologic malignancies and to characterize familial dependence within and across hematologic malignancies. We take into account timing of events, censoring, and competing risk of death.

2. Materials and Methods

The Nordic Twin Study of Cancer (NorTwinCan) is an international, multidisciplinary collaboration of researchers working to investigate the genetic and environmental underpinnings of cancer. The NorTwinCan cohort is based on nationwide twin registers linked to national cancer and mortality registers in Denmark, Finland, Norway, and Sweden [10,11,23]. Details about the twin registers in each of the four participating countries were described previously by Skytthe et al. [24]. Following the latest linkage update in 2018, the study now includes around 400,000 individual twins from four countries, out of which around 315,000 have known zygosity [11]. The analyses presented here are based on these 315,000 monozygotic (MZ) and dizygotic (DZ) twins. The DZ twin pairs can be of the same or opposite sex unless otherwise specified.

Cancer diagnoses were based on the International Statistical Classification of Diseases and Related Health Problems, Tenth Revision (ICD-10) and were grouped according to the NORDCAN classification system [25]. This study includes the main hematologic malignancies, i.e., Hodgkin lymphoma (HL) (C81), non-Hodgkin lymphoma (NHL) (C82–86), multiple myeloma (MM) (C90), and leukemia (C91–95). Due to inconsistencies among the classification of leukemia subtypes between Nordic countries, a sub-classification of leukemia was not carried out. In the Danish data, ICD-10 was used along with the International Classification of Diseases for Oncology, Third Edition (ICD-O-3) when identifying lymphoma incident cases, which might cause the cumulative incidences of HL and NHL to diverge from those of other Nordic countries. The incidence of cancer among twins has been demonstrated to mirror that of the general population [10,24].

2.1. *Cumulative Incidence and Measures of Cross-Cancer Familial Dependence*

For hematologic malignancies overall and separately, we estimated the lifetime risk, defined as risk of cancer diagnosis before age 100. The lifetime risks were obtained from the cumulative incidence function by age estimated using nonparametric counting process modelling, as described by Scheike et al. [26]. The lifetime risks were stratified by sex, zygosity, and cohort for representativeness comparison. Delayed entry determined by the date of register initiation and right-censoring caused by end of follow-up or emigration were accounted for by this approach. Likewise, competing risk of death was accounted for.

The nonparametric risk scale approach, described by Scheike et al. [27], allows for the estimation of cross-cancer familial dependence in terms of case-wise concordance and relative recurrence risk. The case-wise concordance is the conditional risk of cancer in a twin before some age given that the other twin developed (another) cancer before that age and was estimated using nonparametric counting process modelling, as described by Scheike et al. [27]. This is referred to as familial risk. Familial risks for DZ twins also apply to full siblings, as they are as genetically alike. Familial risk by age was estimated for overall hematologic malignancies, and lifetime familial risks were estimated for hematologic malignancies, overall and separately. The relative recurrence risk of cross- cancers was defined as the concordance of cross-cancers, that is, the risk of both twins in a pair having cancer (different cancers) before some age, relative to the risk of the two cancers occurring independently. The latter is given by the product of cumulative incidences. Cross-cancer relative recurrence risks were estimated within the hematologic malignancies and for overall hematologic malignancy vs. other cancer sites. Familial predisposition to cross-cancer is indicated by an increased familial risk among DZ twins compared to the cumulative incidence or a relative recurrence risk for DZ twins greater than one. Furthermore, the influence of shared genetic factors is indicated by a higher familial dependence among MZ twins compared to DZ twins [26]. Inference of the measures assume approximate normality in distribution of estimators and test of differences was performed using a Pepe–Mori type test derived in a way similar to that described by Scheike et al. [27].

2.2. *Biometric Modelling*

The overall magnitude of the genetic and environmental contributions to the variation in risk for each hematologic cancer, taking the timing of events into account, can be assessed by modelling the difference in MZ and DZ pair covariances to the total variance in cancer risk assuming equal margins in MZ and DZ twins. This was achieved parametrically using the bivariate biprobit model, that is, the classic liability-threshold approach, taking censoring into account by inverse probability weighting [28]. Tetrachoric correlations for MZ and DZ pairs and further variance components of the polygenic biometric model from quantitative genetics, the "ADCE model" [28,29], were obtained from this model. The components that contribute to variance are additive genetic (A), dominant genetic (D), shared environmental (C), and unique environmental (E) effects. Heritability is then defined as the proportion of variation in liability to develop cancer attributed to genetic factors, that is, the sum of A and D. Shared environmental effects are those experiences and exposures that affect both twins in a pair. Unique environmental effects influence one twin but not the other, but it also includes measurement error as well as random effects and genetic differences even between MZ pairs. Hence, environmental effects in this setting refer to external non-genetic influences and should not be confused with environmental risk factors, i.e., environmental exposures and lifestyle-related behaviors [30]. This model also provides estimates of cumulative incidence by age as well as familial risk by age, though not for different cancers. The parametric estimates of familial risk were compared to those from the nonparametric model and in turn also adjusted for competing risk of death.

Furthermore, the heritability estimated nonparametrically on the risk scale was obtained in line with the "ACE" sub-model as the ratio of twice the difference in MZ and DZ concordance risk to the total variance in risk based on nonparametric estimates [26,31]. This is not directly comparable to the heritability modelled parametrically on the liability scale [26]. The heritability in risk was estimated by age for overall hematologic malignancies and at age 100 for overall hematologic malignancy vs. other cancer sites.

When illustrating risks and genetic and environmental effects by age, we focused on age 40 and older but did not restrict occurrences by age.

Model sensitivity was assessed by comparing the parametric and nonparametric estimates of cumulative incidence functions and familial risks by age. If there were no substantial differences, we relied on the parametric estimates as they allow for more accurate inference.

All analyses were performed using the statistical software R version 3.5.2 [32] with mets package version 1.2.6 [26,28]. All tests were two-sided and performed at the 5% significance level.

3. Results

The study base included 316,397 individual twins from the NorTwinCan cohort; 88,618 (28%) of these were monozygotic (Table 1). A total of 58,449 cancer diagnoses were reported; out of these, 3459 (6%) were hematologic malignancies. The most common types, NHL and leukemia, constituted 44% and 35% of the hematologic malignancies, respectively. MM and HL accounted for 15% and 7%, respectively.

Table 1. Characteristics of the NorTwinCan cohort of 316,397 twins and a total of 3459 incident hematologic malignancies.

Country	Denmark	Finland	Norway	Sweden	Total
Birth Cohort	1870–2004	1875–1957	1915–1991	1886–2008	
n individual twins	112,560	31,447	32,332	140,058	316,397
n (%) MZ twins	25,120 (22)	8368 (27)	13,100 (41)	42,030 (30)	88,618 (28)
n (%) female twins	55,652 (49)	15,717 (50)	17,473 (54)	73,640 (53)	162,482 (51)
First date follow-up	January 1943	February 1974	April 1956	April 1961	
End of follow-up	December 2016	December 2016	December 2015	December 2015	
Median follow-up time (IQR), years	46.6 (26.5–62.1)	41.3 (31.7–41.3)	52.9 (39.4–59.9)	42.0 (22.2–58.2)	41.8 (26.4–58.2)
Median entry age (IQR), years	0 (0–0)	31.8 (25.6–43.1)	2.4 (0–16.3)	1.7 (0–23.3)	0 (0–24.6)
Number of incident cancers					
Any cancer site	19,581	7044	5465	26,359	58,449
Non-Hodgkin lymphoma	424	261	184	645	1514
Hodgkin lymphoma	80	29	22	98	229
Multiple myeloma	144	72	53	253	522
Leukemia	406	141	102	545	1194

The cumulative incidences by age of hematologic malignancies are shown in Figure 1. There was no substantial variation among cumulative incidences by age of leukemia, MM, and HL across the Nordic countries. However, the cumulative incidences of NHL diverged between the Nordic countries, mainly at high ages (Figure S1). Stratification by sex revealed that males had slightly higher risks for all four cancer types (Figure S2). No difference between cumulative incidences by zygosity was seen (Figure S3).

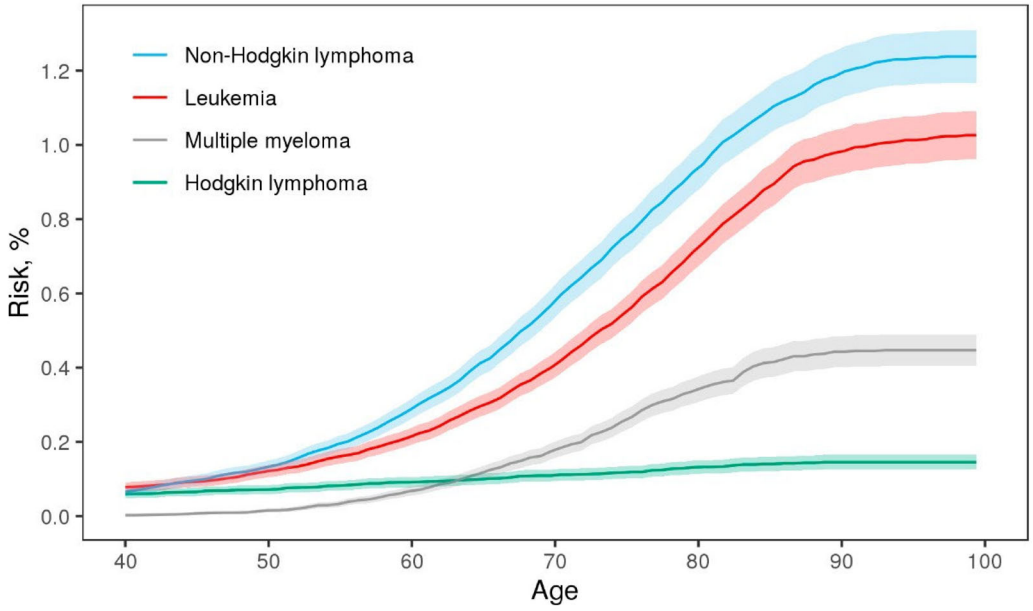


Figure 1. Cumulative incidence and 95% confidence intervals for hematologic malignancies by age in the NorTwinCan cohort, adjusted for censoring and competing risk of death. Incidence before age 40 was very low.

The lifetime risk (in %, 95% CI) of overall hematologic malignancy, given by cumulative incidence at age 100, was 2.5 (2.4–2.6) (Table 2). For individuals diagnosed with more than one hematologic malignancy during follow-up, time to first diagnosis was used. The lifetime risks for each cancer site were NHL: 1.1 (1.0–1.2), leukemia: 0.9 (0.8–1.0), MM: 0.4 (0.3–0.4), and HL: 0.1 (0.1–0.1). The number of concordant and discordant pairs by zygosity is also provided in Table 2. Discordance, in this case, refers to pairs in which one twin was diagnosed with a hematologic malignancy, and (i) the co-twin was diagnosed with a different cancer or (ii) none at all. Only complete pairs, i.e., pairs where both twins were in the cohort, were included in this table and in the following concordance-related estimates.

Familial risks of hematologic malignancies conditional on the same cancer in the co-twin before age 100 are shown in Table 2. They were elevated for DZ pairs in comparison to the corresponding individual lifetime risks, but the difference was only significant for overall hematologic malignancy, leukemia, and NHL. Furthermore, all familial risks were suggestively higher among MZ than DZ pairs; however, the differences were not significant. There were not enough DZ pairs concordant for HL to provide an estimate of familial risk.

Estimates of tetrachoric correlation (from the biprobit model), and genetic and environmental contributions to variation in liability to hematologic malignancies (from the "ACE" sub-model) are shown in Table 3. The tetrachoric correlations were higher for MZ than DZ pairs, but none of the differences were statistically significant. There were not enough DZ pairs concordant for HL to estimate the tetrachoric correlation.

Table 2. Estimates of lifetime and familial risk, and the number of concordant and discordant pairs for hematologic malignancies among twins in the NorTwinCan cohort, adjusted for censoring.

Cancer Site	Lifetime Risk, % (95% CI) [1]	Number of Twin Pairs						Familial Risk, % (95% CI) [1]	
		MZ			DZ				
		Conc.	Disc.		Conc.	Disc		MZ	DZ
			Other [3]	None [4]		Other [3]	None [4]		
Overall hematologic [2]	2.5 (2.4-2.6)	21	232	591	30	592	1653	7.6 (4.8–11.8)	4.5 (3.1–6.5)
Non-Hodgkin lymphoma	1.1 (1.0–1.2)	4	115	246	8	279	732	5.2 (1.8–13.9)	2.1 (1.0–4.3)
Hodgkin lymphoma	0.1 (0.1–0.1)	1–3	14	52	0	33	110	7.4 (1.8–25.6)	-
Multiple myeloma	0.4 (0.3–0.4)	1–3	49	94	1–3	92	244	2.6 (0.6–10.0)	1.4 (0.2–9.2)
Leukemia	0.9 (0.8–1.0)	6	72	207	8	219	582	5.4 (2.1–13.3)	3.6 (1.7–7.3)

[1] Estimated at age 100. [2] For individuals who were diagnosed with more than one hematologic malignancy during follow-up (n = 34), time to first diagnosis was used. [3] Discordant for cancer type, i.e., one twin was diagnosed with the specified cancer and the co-twin was diagnosed with another type of cancer. [4] Discordant for cancer, i.e., one twin was diagnosed with the specified cancer and the co-twin did not have cancer.

Table 3. Estimates of tetrachoric correlation (biprobit model), and genetic and environmental contribution (ACE model) to variation in liability to hematologic malignancies among twins in the NorTwinCan cohort, adjusted for censoring.

Cancer Site	Tetrachoric Correlations, (95% CI)		Heritability, (95% CI)	Shared Environment, (95% CI)	Unique Environment, (95% CI)
	MZ	DZ			
Overall hematologic [1]	0.24 (0.12–0.35)	0.12 (0.03–0.20)	0.24 (0.14–0.33)	0.00 (NA–NA)	0.76 (0.67–0.86)
Non-Hodgkin lymphoma	0.28 (0.05–0.48)	0.10 (−0.03–0.22)	0.25 (0.08–0.42)	0.00 (NA–NA)	0.75 (0.58–0.92)
Hodgkin lymphoma	0.57 (0.29–0.76)	-	0.56 (0.32–0.80)	0.00 (NA–NA)	0.44 (0.20–0.68)
Multiple myeloma	0.26 (0.02–0.48)	0.17 (−0.12–0.43)	0.19 (−0.54–0.93)	0.07 (−0.54–0.69)	0.74 (0.50–0.97)
Leukemia	0.31 (0.10–0.49)	0.23 (0.09–0.36)	0.16 (−0.31–0.64)	0.14 (−0.19–0.48)	0.69 (0.50–0.89)

[1] For individuals who were diagnosed with more than one hematologic cancer during follow-up (n = 34), time to first diagnosis was used.

The genetic and environmental contributions to variation in liability to overall hematologic malignancy and to NHL, MM, and leukemia individually were very similar. Generally, the heritability was moderate, while the contribution from shared environment was practically zero and the remaining variation was attributed to unique environmental effects (Table 3). For overall hematologic malignancy, the estimates and 95% CI were heritability: 0.24 (0.14–0.33), shared environment: 0 (NA–NA), and unique environment: 0.76 (0.67–0.86). The respective values for the specific hematological malignancies are also listed in Table 3. For HL, the heritability was higher, 0.56 (0.32–0.80); the unique environmental contribution equivalently lower, 0.44 (0.20–0.68); and there was no shared environmental contribution.

Relative recurrence (RR) risks and 95% confidence intervals of cross-cancers at age 100 are shown in Figure 2. For example, the risk for leukemia and NHL to cooccur in DZ twin pairs was 2.3 times higher than the risk of these two cancers occurring independently. The RR for DZ twins were generally higher than one indicating a familial predisposition, though only leukemia/NHL and leukemia/leukemia were statistically significant. Furthermore, all of the ratios were higher among MZ than DZ twins, suggesting the influence of a shared genetic component, though none of these differences were statistically significant.

Cumulative incidence and familial risk by age for overall hematological malignancy are depicted in Figure 3. The DZ familial risk lying above the cumulative incidence, especially at old age, indicates a familial predisposition. The MZ familial risk was somewhat higher than DZ, suggesting a genetic component, but this difference was not statistically significant. The curves were estimated using the "AE" submodel as it allowed for more accurate inference than the nonparametric counting process modelling (Figure S4). Curves based on the biprobit model and "ACE" sub-model were almost identical to Figure 3 but with wider confidence intervals.

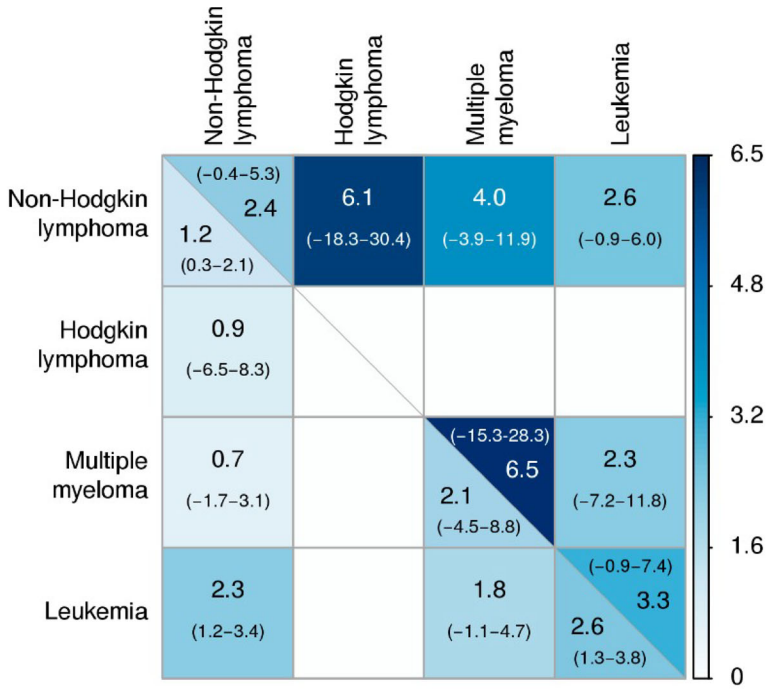


Figure 2. Lifetime cross-cancer relative recurrence risk (concordance risk in pairs relative to independent occurrence risk) and 95% confidence intervals for hematologic malignancies in the NorTwinCan cohort, adjusted for censoring and competing risk of death. Monozygotic pairs are in upper triangle, and dizygotic pairs are in lower triangle. The color scale indicates the magnitude of estimates.

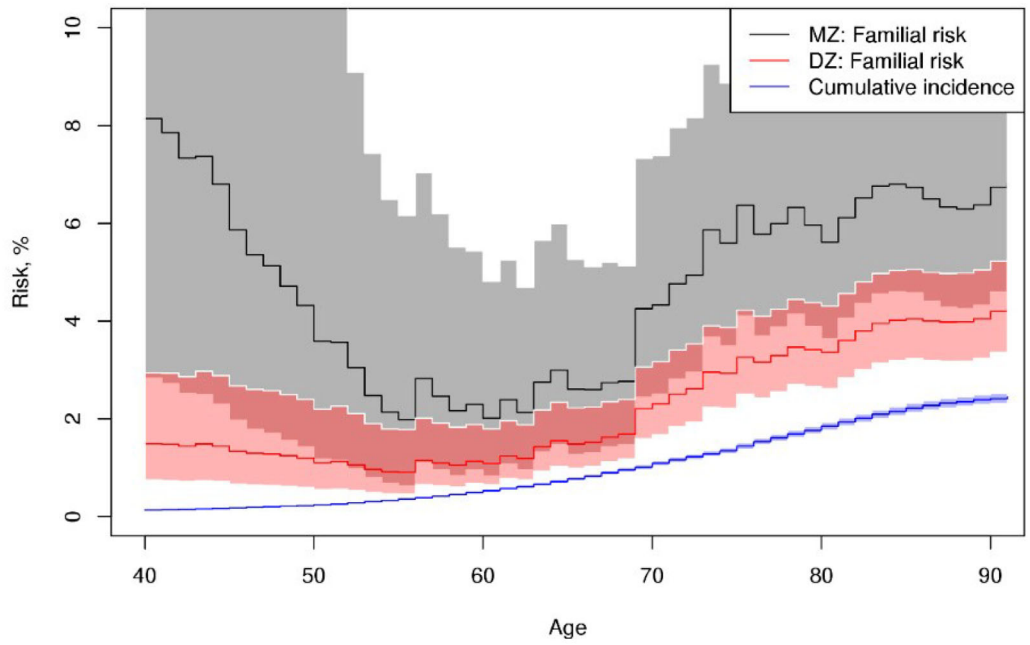


Figure 3. Cumulative incidence and familial risk for monozygotic (MZ) and dizygotic (DZ) twins by age and 95% confidence intervals for overall hematologic malignancy in the NorTwinCan cohort, adjusted for censoring.

The genetic (heritability) and unique environmental contributions to variation in liability to overall hematologic malignancy by age are shown in Figure 4. The heritability decreases from around 55% at age 40 and stabilizes around 20–25% at age 55. The remaining variation is accounted for by the unique environmental contribution as shared environmental contribution (C) was not different from zero at any age.

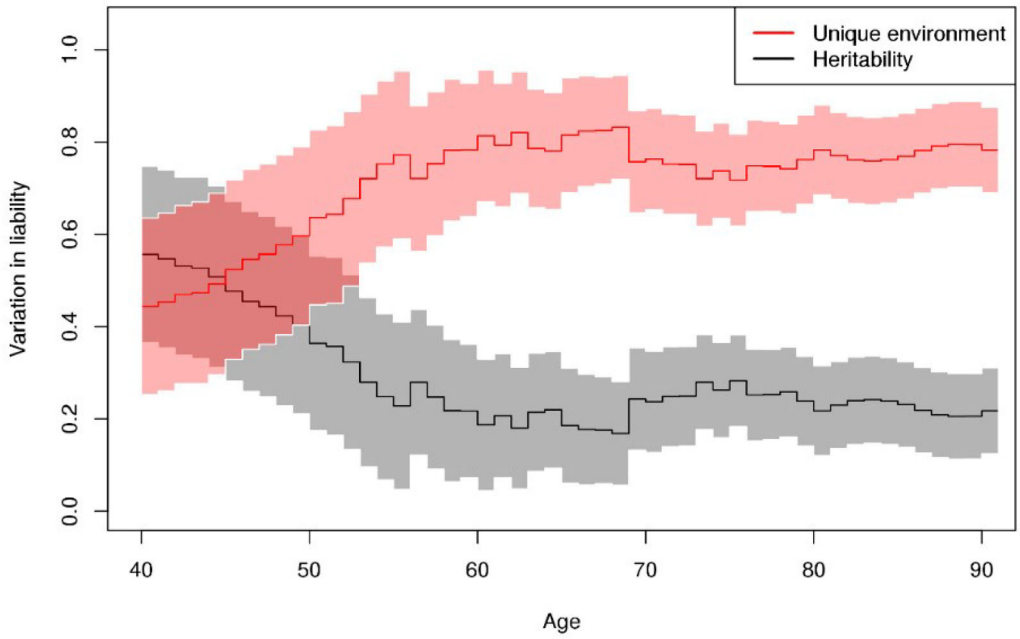


Figure 4. Shared genetic (heritability) and unique environmental contribution to variation in liability to hematologic malignancy by age and 95% confidence intervals in the NorTwinCan cohort, adjusted for censoring.

The heritability in risk of overall hematologic malignancy by age estimated using the nonparametric approach was fairly stable over age and not significantly different from zero (Figure S5). The shared environmental contribution was zero at all ages.

Familial risks for each hematologic malignancy given any hematologic malignancy in the co-twin were estimated for MZ and DZ (in %, 95% CI): NHL: 3.4 (1.1–5.6) and 1.9 (1.0–2.8); HL: 0.7 (−0.2–1.6) and 0.1 (−0.1–0.2); MM: 1.4 (0.3–2.6) and 0.6 (0–1.2); and leukemia: 2.4 (0.7–4.1) and 2.3 (1.2–3.4).

Finally, to analyze the concordance of overall hematologic malignancy and non- hematologic cancers, we counted the number of occurrences by zygosity and estimated lifetime cross-cancer relative recurrence risk and heritability on the risk scale (Table S1). The relative recurrence risk of overall hematologic malignancy and the following cancers were borderline significantly higher than one among DZ twins, indicating a familial predisposition for colon, lung, prostate, kidney, bladder, skin (melanoma and non-melanoma), brain/central nervous system, and lip/oral cavity/pharynx cancers. However, none of the ratios were significantly higher among MZ twins than DZ and none of the heritabilities were significantly different from zero, providing no evidence of shared genetic influence.

4. Discussion

The results from this prospective cohort study of Nordic twins with the longest follow-up ever provide evidence of increased familial risk of hematologic malignancies. For a DZ twin with an affected co-twin, the overall risk (4.5%) of developing a hematologic malignancy was almost twice the risk in the general population (2.5%). Notably, this DZ familial risk was significantly higher than that of the general population already from 40 years of age. The lifetime risks of leukemia and NHL were about twice as high if a DZ co-twin had been diagnosed with any hematologic malignancy. The estimates of relative recurrence risk indicated familial predisposition, not only within the hematologic malignancies but also between overall hematologic malignancy

and other types of cancer (as defined by the NORDCAN classification). Moreover, about a quarter (24%) of the variation in liability to develop a hematological malignancy was attributed to genetic effects (heritability) and the remaining variation was explained by effects not shared by twin siblings. The heritability varied across age and was around 55% at age 40 but decreased thereafter to 20–25% and stayed constant at the older ages.

The lifetime risks for individuals were not entirely consistent with the estimates reported by NORDCAN. They provide the following risks of developing cancer in the absence of competing risks of death before age 75 in the Nordic countries for male/female (%); NHL: 1.3/0.9, HL: 0.2/0.2, MM: 0.5/0.3, and leukemia: 0.9/0.6 [1]. The lifetime risks for the corresponding age and sex estimated in this study (Figure S2) were all lower than those from NORDCAN. The differences between the estimates from NORDCAN and our findings might reflect the fact that we adjusted for competing risk of death; omitting this will likely bias estimates upwards. There is a variable age-at-entry of the twins into the cohorts, which may be associated with absence of hematological cancers diagnosed early in life in some cohorts. However, this is taken into account by adjusting for delayed entry, providing valid estimates at all ages.

Several studies have reported evidence of familial predisposition to hematologic malignancies [9,12–22]. One is a study by Fallah et al. [7], who found significantly increased familial risks of NHL for various familial relationships by sex relative to the general population in the Nordic countries. Not surprisingly, the highest familial risk of NHL was found among twins; however, the zygosity was unfortunately not specified. Similar results were seen in a study of familial risk of HL by Kharazmi et al. [9], who also ignored zygosity. The distinction between MZ and DZ twins is considered crucial to the analysis presented in this paper as it allows for the partitioning of genetic and environmental influences.

Another approach was taken by Sud et al. [19], who published results from an extensive study including 153,115 Swedish patients with hematologic malignancies. They utilized the large sample size by examining the interrelationships between familial relative risks (corresponding to the relative recurrence risks of this study) of 22 different hematologic malignancies. To the extent that they are comparable, the results are similar to the ones we provide. A weakness of our study is the lack of power in the sense that we were limited by the low number of concordant twin pairs. Therefore, we consider only the four main hematologic malignancies when assessing familial dependence measures and when estimating familial predisposition by age hematologic malignancy as an overall group. However, the fact that we were able not only to assess the course of familial risk by age but also to partition the variation in liability to hematologic malignancy by age into genetic and environmental factors is a novel contribution.

A common problem regarding studies of familial risk of hematologic malignancies, especially the subtypes, is the questionable validity of self-reported family history [20,33]. As this is a multi-country study based on four nationwide registries with very long follow- up times, we avoided issues of misclassification and recall bias.

The estimated genetic and environmental influences for NHL, HL, and MM based on the NorTwinCan cohort are reported for the first time in this paper. The heritability estimates for NHL, MM, and leukemia were generally low (16–25%), while the shared environmental effects were estimated to be zero. The remaining variance was attributed to exposures and experiences not shared by twin siblings, which includes a wide range of possible factors from occupational exposures and lifestyle factors to somatic mutations and random effects at the cellular level. The estimates of genetic and environmental influences on liability to leukemia were similar to those from Lichtenstein et al. [23] despite using methods that take into account the timing of events and censoring [26].

In contrast to other hematological malignancies, HL showed considerably more heritable variation, with a heritability estimate (56%) that was more than twice as high as the other hematologic malignancies. This finding is congruent with the study by Mack et al. [22]. Excluding HL from overall hematologic malignancies in our analyses did not result in noteworthy differences in familial risk by age or in genetic and environmental effects by age, though this might not have been the case if we focused on childhood and young adulthood.

Though it could be argued that the hematologic malignancies are fundamentally different and should be studied separately, we considered it relevant, from an epidemiological point of view, to look at overall hematologic malignancies when estimating genetic and environmental effects. By doing so we have, for the first time, gained insight into the time-varying influence of genetic and environmental influences on the liability to develop a hematologic malignancy.

We found evidence of familial predisposition to cross-cancer occurrences when com- paring overall hematologic malignancy and other cancers, including prostate, lung, and non-melanoma skin cancer. For these cross-cancer associations, the relative recurrence risks were higher among MZ than DZ twins, indicating a possible shared genetic influence, though it was not statistically significant. These are verifications of previous findings; for instance, Goldgar et al. [21] reported evidence of familial predisposition for, e.g., breast cancer/NHL, prostate cancer/NHL, and leukemia/colon cancer and Zhang et al. [12] found evidence of familial predisposition to NHL/lung cancer.

Etiological studies of hematologic malignancies indicate that there are numerous risk factors and that they influence the subtypes to various degrees. For instance, immunodeficiency is deemed a relevant risk factor of NHL along with radiation exposure, alcohol intake, and certain infectious organisms, but we have yet to fully explain the increase of NHL incidence rate that has been seen in developed countries [4,8]. One should keep in mind that many risk factors tend to be aggregated within families, representing shared environmental effects and heritability. Additionally, there are (rare) congenital syndromes, such as Downs syndrome, Fanconi anemia, Diamond Blackfan anemia, and other rare bone marrow failure syndromes, that increase the risk of hematologic malignancy [34]. While an underlying syndromic predisposition could also explain familial clustering of these cancers, it seems unlikely that these known rare syndromes explain our findings. Though the influence of many environmental risk factors has already been studied, it seems there are more to be found. A more recent trend that could influence cancer risk, especially cancers of the lymphatic system, is tattooing. Pathologists have reported cases of pigments from tattoo ink found in lymph nodes [35–37], and it has been hypothesized that ink particles may also reach the blood stream and become distributed to organs. In a mouse study, the authors reported that tattoo pigments were detected in the liver but not in other internal organs [38]. However, the (cancer) risk of long-term exposure to tattoo ink has yet to be examined. The matched co-twin design provides the perfect setting for such studies.

5. Conclusions

In this largest ever studied twin cohort with the longest follow-up, we found evidence of increased familial risk of hematologic malignancies and indications of familial predisposition, not only within the hematologic malignancies but also between overall hematologic malignancy and other types of cancer. Heritable factors were most important among young and middle-aged adults and decreased across age after 55 years of age. The discovery of decreasing familial predisposition with increasing age underscores the importance of cancer surveillance in families with hematological malignancies.

Supplementary Materials: The following are available online at https://www.mdpi.com/article/10.3390/cancers13123023/s1, Figure S1: Cumulative incidence by age and country, Figure S2: Cumulative incidence by age and sex, Figure S3: Cumulative incidence by age and zygosity, Figure S4: Cumulative and familial risk by age, Figure S5: Heritability in risk, Table S1: Number of pairs where one twin has been diagnosed with a hematologic malignancy and the co-twin another type of cancer and estimates of lifetime relative recurrence risk and heritability in risk.

Author Contributions: Conceptualization, S.B.C. and J.v.B.H.; data curation, S.B.C., W.H.B. and J.v.B.H.; formal analysis, S.B.C., W.H.B. and J.v.B.H.; funding acquisition, J.R.H., J.K. and J.v.B.H.; methodology, S.B.C. and J.v.B.H.; project administration, J.v.B.H.; supervision, J.v.B.H.; visualization, S.B.C.; writing—original draft, S.B.C. and J.v.B.H.; writing—review and editing, S.B.C., J.R.H., J.M.-F., H.F., J.K. and J.v.B.H. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by grants from the Nordic Cancer Union grant number 195020, project entitled "Genetic Epidemiology and Familial Risk of Cross-Cancer Associations: A Nordic Twin Study-II", and from the AgeCare program of the Academy of Geriatric Cancer Research, Odense University Hospital, as well as by a scholarship from the Faculty of Health Sciences, University of Southern Denmark.

Institutional Review Board Statement: Ethical review and approval were waived for this study, since no ethics approval is needed for register-based studies according to legislation in the Nordic countries. The appropriate register authorities in each country have given permission for the record linkages.

Informed Consent Statement: Not applicable.

Data Availability Statement: Restrictions apply to the availability of these data. The data were obtained from the participating twin cohorts and national cancer registries. Requests to access additional data need to be made through the individual national cohorts and registers responsible for the data sets.

Acknowledgments: The authors are thankful to The Danish Twin Registry for hosting and managing the joint Nordic twin data. We also acknowledge the contribution of all of the researchers and staff of the twin and cancer registries contributing to the NorTwinCan project.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Danckert, B.; Ferlay, J.; Engholm, G.; Hansen, H.L.; Johannesen, T.B.; Khan, S.; Køtlum, J.E.; Ólafsdóttir, E.; Schmidt, L.K.H.; Virtanen, A.; et al. NORDCAN: Cancer Incidence, Mortality, Prevalence and Survival in the Nordic Countries. Available online: https://www-dep.iarc.fr/nordcan/english/frame.asp (accessed on 3 May 2021).
2. Global Burden of Disease Cancer Collaboration. Global, Regional, and National Cancer Incidence, Mortality, Years of Life Lost, Years Lived with Disability, and Disability-Adjusted Life-Years for 29 Cancer Groups, 1990 to 2016: A Systematic Analysis for the Global Burden of Disease Study. *JAMA Oncol.* **2018**, *4*, 1553–1568. [CrossRef]
3. Hjalgrim, H. On the aetiology of Hodgkin lymphoma. *Dan. Med. J.* **2012**, *59*, 4485.
4. Chiu, B.C.-H.; Weisenburger, D.D. An Update of the Epidemiology of Non-Hodgkin's Lymphoma. *Clin. Lymphoma* **2003**, *4*, 161–168. [CrossRef]
5. Juliusson, G.; Hough, R. Leukemia. *Neuronal Act. Tumor Tissue* **2016**, *43*, 87–100.
6. Alexander, D.D.; Mink, P.J.; Adami, H.-O.; Cole, P.; Mandel, J.S.; Oken, M.M.; Trichopoulos, D. Multiple myeloma: A review of the epidemiologic literature. *Int. J. Cancer* **2007**, *120* (Suppl. 12), 40–61. [CrossRef]
7. Fallah, M.; Kharazmi, E.; Pukkala, E.; Tretli, S.; Olsen, J.H.; Tryggvadottir, L.; Sundquist, K.; Hemminki, K. Familial risk of non-Hodgkin lymphoma by sex, relationship, age at diagnosis and histology: A joint study from five Nordic countries. *Leukemia* **2015**, *30*, 373–378. [CrossRef] [PubMed]
8. Alexander, D.D.; Mink, P.J.; Adami, H.-O.; Chang, E.T.; Cole, P.; Mandel, J.S.; Trichopoulos, D. The non-Hodgkin lymphomas: A review of the epidemiologic literature. *Int. J. Cancer* **2007**, *120*, 1–39. [CrossRef]
9. Kharazmi, E.; Fallah, M.; Pukkala, E.; Olsen, J.H.; Tryggvadottir, L.; Sundquist, K.; Tretli, S.; Hemminki, K. Risk of familial classical Hodgkin lymphoma by relationship, histology, age, and sex: A joint study from five Nordic countries. *Blood* **2015**, *126*, 1990–1995. [CrossRef] [PubMed]
10. Mucci, L.A.; Hjelmborg, J.; Harris, J.R.; Czene, K.; Havelick, D.J.; Scheike, T.; Graff, R.E.; Holst, K.; Möller, S.; Unger, R.; et al. Familial Risk and Heritability of Cancer Among Twins in Nordic Countries. *JAMA* **2016**, *315*, 68–76. [CrossRef]
11. Harris, J.R.; Hjelmborg, J.; Adami, H.-O.; Czene, K.; Mucci, L.; Kaprio, J. Nordic Twin Study of Cancer (NorTwinCan) Collaboration The Nordic Twin Study on Cancer—NorTwinCan. *Twin Res. Hum. Genet.* **2019**, *22*, 817–823. [CrossRef]
12. Zhang, Y.; Wang, R.; Holford, T.R.; Leaderer, B.; Zahm, S.H.; Boyle, P.; Zhu, Y.; Qin, Q.; Zheng, T. Family history of hema-topoietic and non-hematopoietic malignancies and risk of non-Hodgkin lymphoma. *Cancer Causes Control* **2007**, *18*, 351–359. [CrossRef] [PubMed]
13. Chang, E.T.; Smedby, K.E.; Hjalgrim, H.; Porwit-MacDonald, A.; Roos, G.; Glimelius, B.; Adami, H.-O. Family History of Hematopoietic Malignancy and Risk of Lymphoma. *J. Natl. Cancer Inst.* **2005**, *97*, 1466–1474. [CrossRef]
14. Altieri, A.; Bermejo, J.L.; Hemminki, K. Familial risk for non-Hodgkin lymphoma and other lymphoproliferative malignancies by histopathologic subtype: The Swedish Family-Cancer Database. *Blood* **2005**, *106*, 668–672. [CrossRef] [PubMed]

15. Wang, S.S.; Slager, S.L.; Brennan, P.; Holly, E.A.; De Sanjose, S.; Bernstein, L.; Boffetta, P.; Cerhan, J.R.; Maynadie, M.; Spinelli, J.J.; et al. Family history of hematopoietic malignancies and risk of non-Hodgkin lymphoma (NHL): A pooled analysis of 10 211 cases and 11 905 controls from the International Lymphoma Epidemiology Consortium (InterLymph). *Blood* **2006**, *109*, 3479–3488. [CrossRef]
16. Schinasi, L.H.; Brown, E.E.; Camp, N.J.; Wang, S.S.; Hofmann, J.N.; Chiu, B.C.; Miligi, L.; Freeman, L.E.B.; De Sanjose, S.; Bernstein, L.; et al. Multiple myeloma and family history of lymphohaematopoietic cancers: Results from the International Multiple Myeloma Consortium. *Br. J. Haematol.* **2016**, *175*, 87–101. [CrossRef] [PubMed]
17. VanValkenburg, M.E.; Pruitt, G.I.; Brill, I.K.; Costa, L.; Ehtsham, M.; Justement, I.T.; Innis-Shelton, R.D.; Salzman, D.; Reddy, E.S.P.; Godby, K.N.; et al. Family history of hematologic malignancies and risk of multiple myeloma: Differences by race and clinical features. *Cancer Causes Control.* **2015**, *27*, 81–91. [CrossRef] [PubMed]
18. Cerhan, J.R.; Slager, S.L. Familial predisposition and genetic risk factors for lymphoma. *Blood* **2015**, *126*, 2265–2273. [CrossRef] [PubMed]
19. Sud, A.; Chattopadhyay, S.; Thomsen, H.; Sundquist, K.; Sundquist, J.; Houlston, R.S.; Hemminki, K. Analysis of 153115 patients with hematological malignancies refines the spectrum of familial risk. *Blood* **2019**, *134*. [CrossRef]
20. Mensah, F.; Willett, E.; Ansell, P.; Adamson, P.; Roman, E. Non-Hodgkin's Lymphoma and Family History of Hematologic Malignancy. *Am. J. Epidemiol.* **2006**, *165*, 126–133. [CrossRef] [PubMed]
21. Goldgar, D.E.; Easton, D.F.; Cannon-Albright, L.A.; Skolnick, M.H. Systematic Population-Based Assessment of Cancer Risk in First-Degree Relatives of Cancer Probands. *J. Natl. Cancer Inst.* **1994**, *86*, 1600–1608. [CrossRef] [PubMed]
22. Mack, T.M.; Cozen, W.; Shibata, D.K.; Weiss, L.M.; Nathwani, B.N.; Hernandez, A.M.; Taylor, C.R.; Hamilton, A.S.; Deapen, D.M.; Rappaport, E.B. Concordance for Hodgkin's Disease in Identical Twins Suggesting Genetic Susceptibility to the Young-Adult Form of the Disease. *N. Engl. J. Med.* **1995**, *332*, 413–419. [CrossRef]
23. Lichtenstein, P.; Holm, N.V.; Verkasalo, P.K.; Iliadou, A.; Kaprio, J.; Koskenvuo, M.; Pukkala, E.; Skytthe, A.; Hemminki, K. Environmental and Heritable Factors in the Causation of Cancer—Analyses of Cohorts of Twins from Sweden, Denmark, and Finland. *N. Engl. J. Med.* **2000**, *343*, 78–85. [CrossRef] [PubMed]
24. Skytthe, A.; Harris, J.R.; Czene, K.; Mucci, L.; Adami, H.-O.; Christensen, K.; Hjelmborg, J.; Holm, N.V.; Nilsen, T.S.; Kaprio, J.; et al. Cancer Incidence and Mortality in 260,000 Nordic Twins With 30,000 Prospective Cancers. *Twin Res. Hum. Genet.* **2019**, *22*, 99–107. [CrossRef]
25. Engholm, G.; Ferlay, J.; Christensen, N.; Bray, F.; Gjerstorff, M.L.; Klint, Å.; Køtlum, J.E.; Ólafsdóttir, E.; Pukkala, E.; Storm, H. NORDCAN—A Nordic tool for cancer information, planning, quality control and research. *Acta Oncol.* **2010**, *49*, 725–736. [CrossRef]
26. Scheike, T.H.; Holst, K.K.; Hjelmborg, J.B. Estimating heritability for cause specific mortality based on twin studies. *Lifetime Data Anal.* **2013**, *20*, 210–233. [CrossRef] [PubMed]
27. Scheike, T.H.; Holst, K.K.; Hjelmborg, J. Estimating twin concordance for bivariate competing risks twin data. *Stat. Med.* **2013**, *33*, 1193–1204. [CrossRef] [PubMed]
28. Holst, K.K.; Scheike, T.H.; Hjelmborg, J. The liability threshold model for censored twin data. *Comput. Stat. Data Anal.* **2016**, *93*, 324–335. [CrossRef]
29. Sham, P. *Statistics in Human Genetics*; Wiley: London, UK, 1998.
30. Peters, A.; Nawrot, T.S.; Baccarelli, A.A. Hallmarks of environmental insults. *Cell* **2021**, *184*, 1455–1468. [CrossRef] [PubMed]
31. Hjelmborg, J.B.; Scheike, T.; Holst, K.; Skytthe, A.; Penney, K.L.; Graff, R.E.; Pukkala, E.; Christensen, K.; Adami, H.-O.; Holm, N.V.; et al. The Heritability of Prostate Cancer in the Nordic Twin Study of Cancer. *Cancer Epidemiology Biomarkers Prev.* **2014**, *23*, 2303–2310. [CrossRef]
32. R Core Team. R: A Language and Environment for Statistical Computing. Available online: https://www.R-project.org/ (accessed on 3 May 2021).
33. Chang, E.T.; Smedby, K.E.; Hjalgrim, H.; Glimelius, B.; Adami, H.-O. Reliability of Self-Reported Family History of Cancer in a Large Case–Control Study of Lymphoma. *J. Natl. Cancer Inst.* **2006**, *98*, 61–68. [CrossRef]
34. Godley, L.A.; Shimamura, A. Genetic predisposition to hematologic malignancies: Management and surveillance. *Blood* **2017**, *130*, 424–432. [CrossRef] [PubMed]
35. Dominguez, E.; Alegre, V.; García-Melgares, M.; Laguna, C.; Martin, B.; Sánchez, J.; Oliver, V. Tattoo pigment in two lymph nodes in a patient with melanoma. *J. Eur. Acad. Dermatol. Venereol.* **2007**, *22*, 101–102. [CrossRef] [PubMed]
36. Schreiver, I.; Hesse, B.; Seim, C.; Castillo-Michel, H.; Villanova, J.; Laux, P.; Dreiack, N.; Penning, R.; Tucoulou, R.; Cotte, M.; et al. Synchrotron-based ν-XRF mapping and μ-FTIR microscopy enable to look into the fate and effects of tattoo pigments in human skin. *Sci. Rep.* **2017**, *7*, 11395. [CrossRef]
37. Singh, R.S.; Diwan, A.H.; Prieto, V.G. Potential diagnostic pitfalls in melanoma arising in a cutaneous tattoo. *Histopathology* **2007**, *51*, 283–285. [CrossRef]
38. Sepehri, M.; Sejersen, T.; Qvortrup, K.; Lerche, C.M.; Serup, J. Tattoo Pigments Are Observed in the Kupffer Cells of the Liver Indicating Blood-Borne Distribution of Tattoo Ink. *Dermatology* **2017**, *233*, 86–93. [CrossRef] [PubMed]

cancers

Supplementary Materials: Familial Risk and Heritability of Hematologic Malignancies in the Nordic Twin Study of Cancer

Signe B. Clemmensen, Jennifer R. Harris, Jonas Mengel-From, Wagner H. Bonat, Henrik Frederiksen, Jaakko Kaprio and Jacob v. B. Hjelmborg

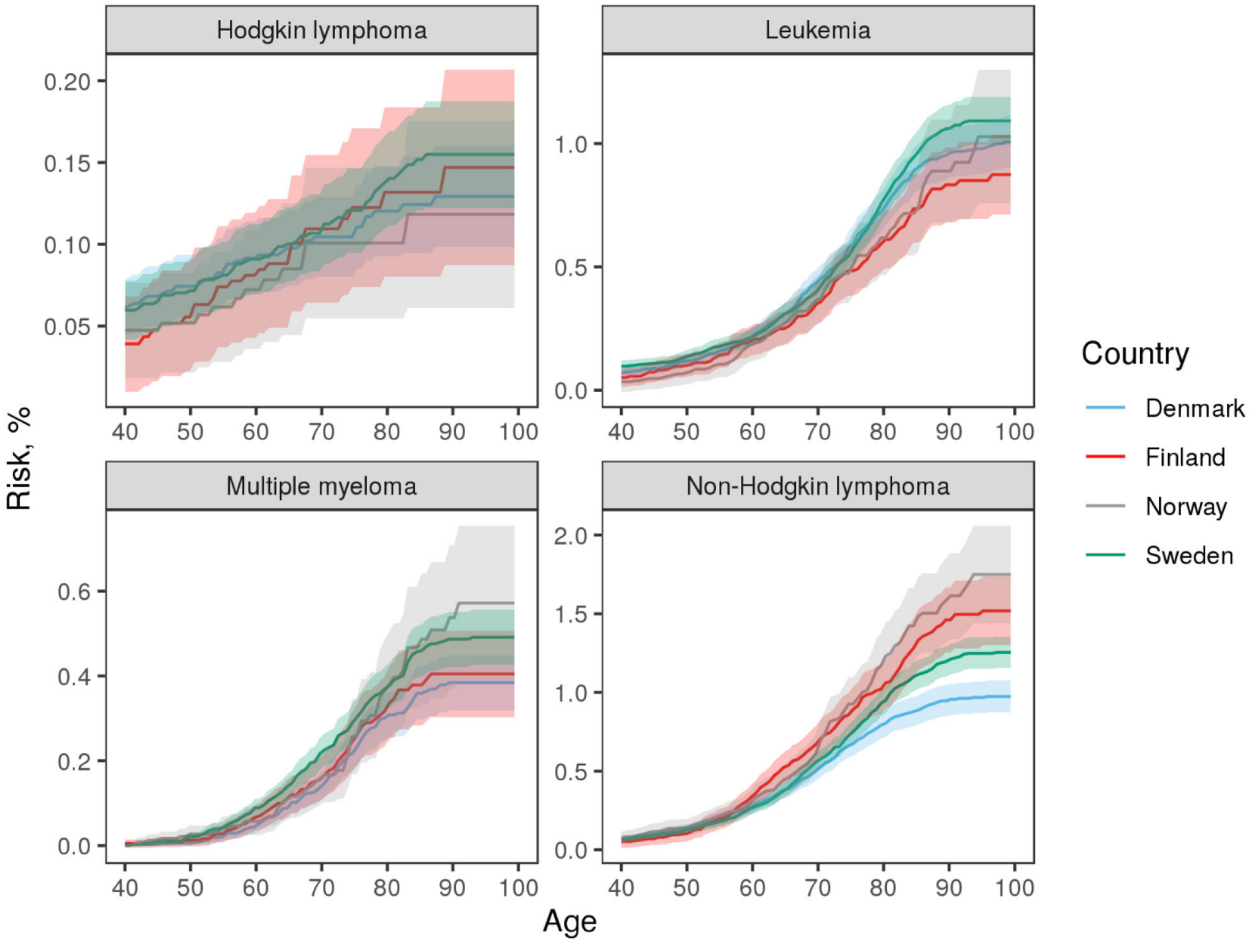


Figure S1. Cumulative incidence and 95% confidence intervals for hematologic malignancies by age in each of the Nordic countries, adjusted for censoring and competing risk of death.

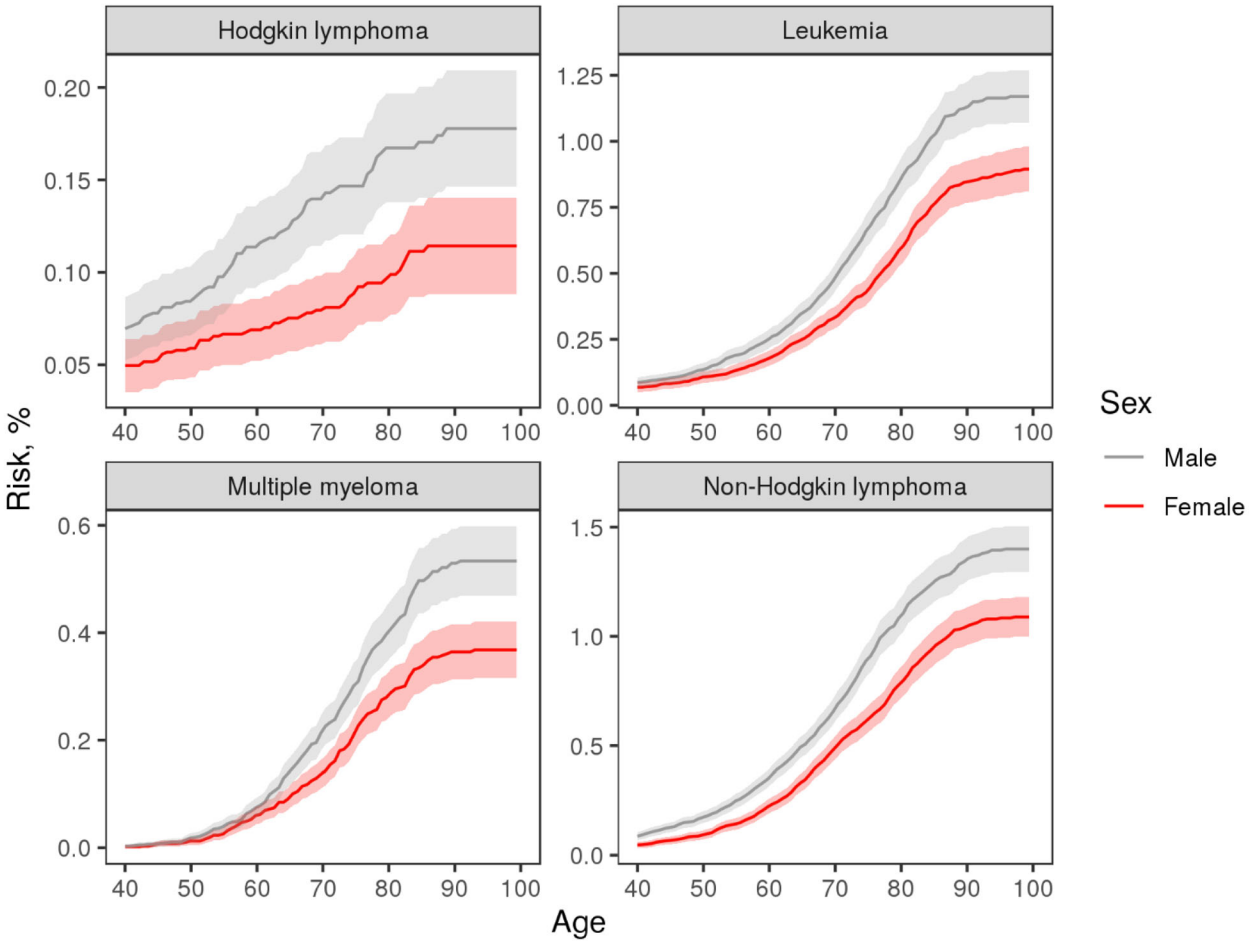


Figure S2. Cumulative incidence and 95% confidence intervals for hematologic malignancies by age and sex, adjusted for censoring and competing risk of death.

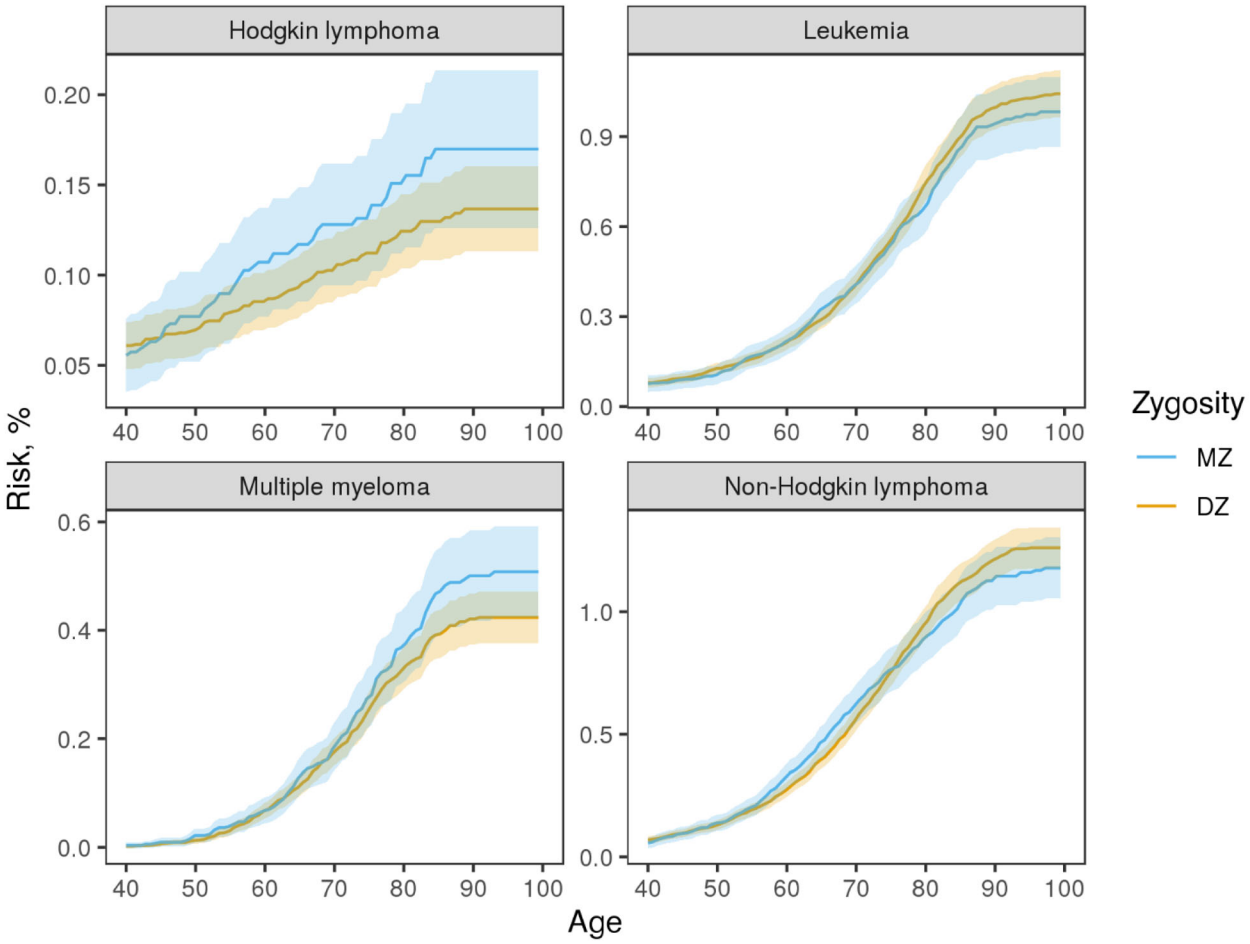


Figure S3. Cumulative incidence and 95% confidence intervals for hematologic malignancies by age and zygosity, adjusted for censoring and competing risk of death.

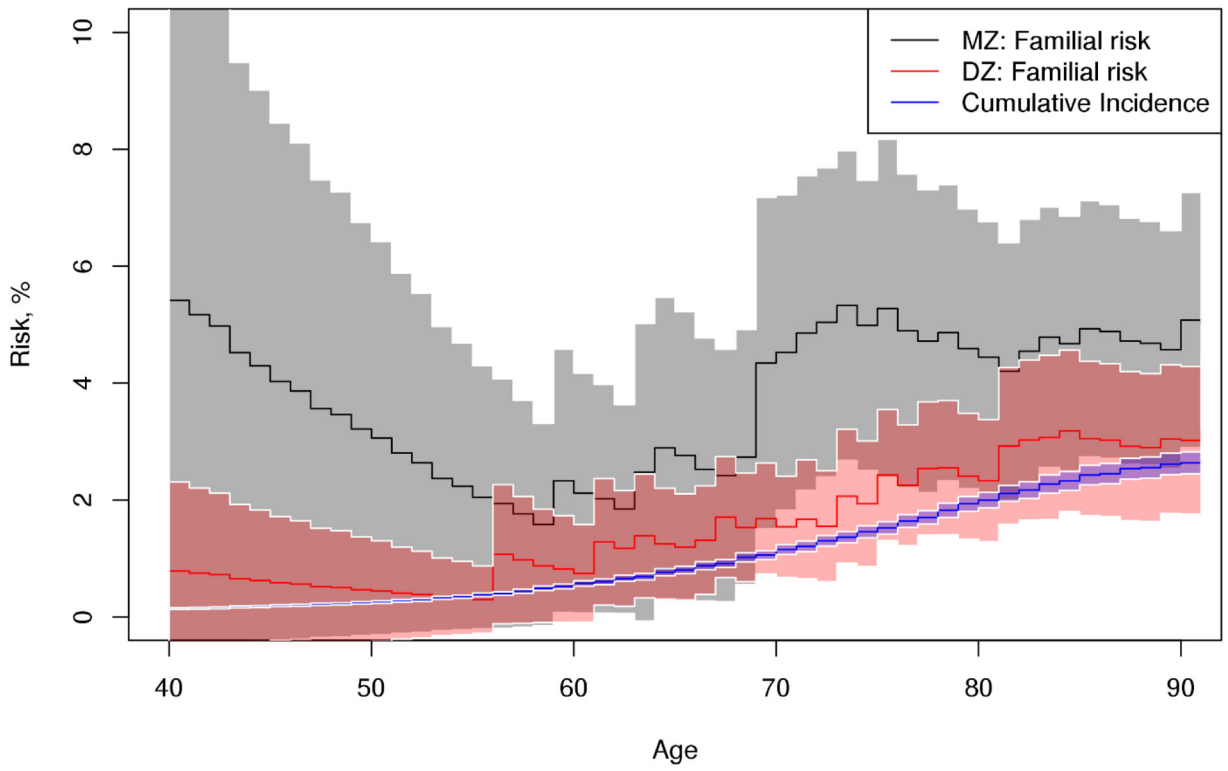


Figure S4. Cumulative incidence and familial risk for monozygotic (MZ) and dizygotic (DZ) twins by age and 95% confidence intervals for overall hematologic malignancy in the NorTwinCan cohort, adjusted for censoring and competing risk of death.

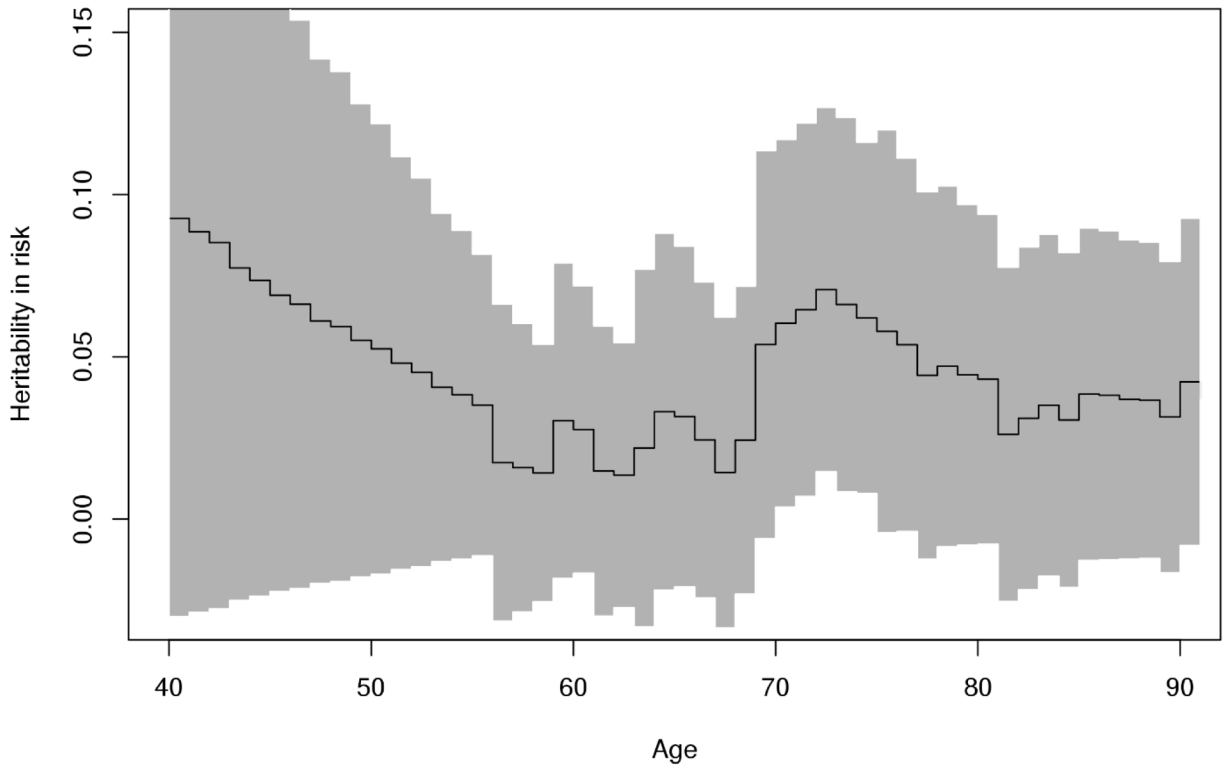


Figure S5. Heritability in risk by age and 95% confidence interval for overall hematologic malignancy in the NorTwinCan cohort, adjusted for censoring and competing risk of death.

Table S1. Number of pairs where one twin was diagnosed with a hematologic malignancy and the co-twin was diagnosed with another type of cancer (following the NORDCAN classification). Estimates of lifetime relative recurrence risk and heritability at age 100 among twins in the NorTwinCan cohort, adjusted for censoring and competing risk of death.

Hematologic Malignancy	Number of Twin Pairs		Relative Recurrence Risk (95% CI)		Heritability in Risk, % (95% CI)
	MZ	DZ	MZ	DZ	
Lip, oral cavity and pharynx	9	14	3.0 (1.0-5.1)	1.7 (1.2-2.1)	4.0 (-2.8-10.9)
Oesophagus	4	8	2.2 (0.0-4.5)	2.1 (0.6-3.6)	0.4 (-5.9-6.8)
Stomach	6	12	1.6 (0.2-3.1)	1.0 (0.4-1.7)	2.2 (-3.8-8.3)
Small intestine	0	4	-	-	-
Colon	20	46	2.1 (1.1-3.1)	1.6 (1.1-2.2)	2.7 (-3.9-9.4)
Rectum and anus	9	25	1.2 (0.3-2.0)	1.4 (0.8-2.0)	-1.2 (-6.1-3.7)
Liver	4	9	2.2 (0.0-4.5)	1.9 (0.6-3.2)	0.8 (-5.5-7.1)
Gallbladder	5	7	3.3 (0.3-6.2)	1.7 (0.4-3.0)	3.7 (-3.6-11.0)
Pancreas	8	12	1.9 (0.5-3.3)	1.4 (0.6-2.2)	1.9 (-3.9-7.7)
Nose, sinuses	0	0	-	-	-
Larynx	1-3	5	2.2 (-1.0-5.5)	1.7 (0.2-3.3)	0.9 (-5.3-7.1)
Lung	22	54	1.9 (1.1-2.8)	1.6 (1.2-2.1)	1.8 (-4.6-8.2)
Pleura	0	1-3	-	-	-
Bone	1-3	0	-	-	-
Melanoma of skin	9	28	1.4 (0.4-2.3)	1.8 (1.1-2.5)	-1.8 (-7.3-3.6)
Skin, non-melanoma	24	50	2.5 (1.5-3.6)	1.9 (1.3-2.4)	4.4 (-3.5-12.3)
Soft tissues	1-3	1-3	2.2 (-0.9-5.3)	1.1 (-0.7-2.8)	1.9 (-3.8-7.6)
Breast [1]	34	86	1.7 (1.1-2.4)	1.5 (1.0-1.9)	2.9 (-5.7-11.4)
Cervix uteri [1]	11	34	1.2 (0.4-2.0)	1.0 (0.4-1.6)	1.4 (-5.3-8.1)
Corpus uteri [1]	6	12	1.4 (0.2-2.6)	0.8 (0.1-1.5)	3.1 (-3.6-9.7)
Uterus, other [1]	0	1-3	-	-	-
Ovary [1]	5	15	1.6 (0.1-3.1)	1.1 (0.3-2.0)	2.0 (-5.7-9.7)
Other female genital organs [1]	0	1-3	-	-	-
Prostate [2]	45	106	2.5 (1.7-3.2)	2.3 (1.7-2.8)	2.5 (-8.9-13.9)
Testis [2]	1-3	4	5.0 (-2.3-12.3)	1.3 (-0.2-2.9)	8.8 (-9.0-26.7)
Penis [2]	0	1-3	-	-	-
Kidney	6	16	2.0 (0.4-3.7)	2.2 (1.1-3.4)	-0.7 (-6.9-5.5)
Bladder	18	36	2.3 (1.1-3.4)	2.0 (1.3-2.6)	1.6 (-5.3-8.5)
Eye	0	1-3	-	-	-
Brain, central nervous system	7	26	1.5 (0.2-2.8)	2.0 (1.2-2.8)	-1.9 (-7.8-4.1)
Thyroid	1-3	1-3	1.2 (-0.5-2.9)	1.0 (-0.2-2.3)	0.3 (-3.6-4.1)

[1] Females only.
[2] Males only.