Université de Montréal

Shared etiology between cancer and cardiovascular disease

Par

Maxine Sun

Programme de Sciences Biomédicales, Faculté de Médecine

Thèse présentée en vue de l'obtention du grade de Philosophiae Doctor (Ph.D)

en Sciences Biomédicales, option Médecine Expérimentale

Avril 2023

Université de Montréal

Faculté des études supérieures et postdoctorales

Cette thèse intitulée

Shared etiology between cancer and cardiovascular disease

Présentée par

Maxine Sun

A été évalué(e) par un jury composé des personnes suivantes

Anne-Marie Mes-Masson
Président-rapporteur

Marie-Pierre Dubé
Directeur de recherche

Jean-Claude Tardif
Codirecteur

Josée Hébert
Membre du jury

Kelly Bolton
Examinateur externe

Anne-Marie Laberge
Représentant du doyen

Résumé

Le cancer et les maladies cardiovasculaires sont deux problèmes de santé majeurs. Les survivants du cancer ont un risque plus élevé de développer et de mourir d'une maladie cardiovasculaire par rapport à la population générale. De même, les patients atteints de maladies cardiovasculaires sont également plus susceptibles de recevoir un diagnostic de cancer. Bien que les deux conditions partagent de nombreux facteurs de risque, le cancer et les maladies cardiovasculaires ont traditionnellement été étudiés séparément. Récemment, une vague de recherches en cours a mis en évidence des points communs biologiques entre le cancer et les maladies cardiovasculaires. Ces découvertes ont été rendues possibles grâce à des avancées scientifiques importantes, notamment la caractérisation moléculaire élargie des modulateurs de l'inflammation, l'élucidation des structures du protéasome et de l'inflammasome, ainsi que des contributions visant à renforcer nos connaissances sur l'immunité cellulaire.

Grâce à ces avancées, il est désormais possible d'obtenir des informations précieuses sur les circuits complexes de la réponse immunitaire innée et adaptative, ainsi que sur les mécanismes de défense de l'hôte. Les progrès notables réalisés dans ces domaines ont posé les bases solides de la biologie de l'inflammation. L'application des principes appris dans ce domaine à la maladie humaine n'a que récemment commencé à donner des résultats fructueux, conduisant à la croissance exponentielle du domaine de l'inflammation dans le cancer et les maladies cardiovasculaires. Pour le cancer comme pour les maladies cardiovasculaires, l'immunomodulation est récemment apparue comme un facteur clé dans leur traitement et leur prise en charge. Avec le temps, il est également devenu évident que cibler une voie inflammatoire particulière pour traiter une condition peut avoir des implications importantes pour l'autre, reflétant l'interaction complexe entre ces deux processus pathologiques.

Les avancées réalisées ont permis de reconnaître que l'inflammation joue un rôle actif dans le développement et la progression physiopathologique du cancer et des maladies cardiovasculaires. Toutefois, la compréhension de l'inflammation sur le cancer et les maladies cardiovasculaires est en constante évolution, et des recherches en cours sont nécessaires pour découvrir de nouvelles informations sur les mécanismes complexes de leurs relations. Dans ce

contexte, nous avons entrepris des analyses distinctes pour approfondir notre compréhension et contribuer aux efforts de recherche innovants en cours, afin de réduire le fardeau croissant du cancer et des maladies cardiovasculaires.

Premièrement, nous avons effectué une méta-analyse génétique sur le risque de cancer incident chez les patients atteints de maladie coronarienne prenant des statines. Dans cette étude, nous avons pu identifier et répliquer une variation génétique associée à un risque plus élevé de diagnostic incident de cancer chez les femmes utilisatrices de statines. La variation génétique se trouvait dans la région de l'antigène leucocytaire humain qui est largement responsable de l'activation des lymphocytes T et de la régulation des réponses immunitaires. Ces résultats réitèrent l'implication active des processus inflammatoires sous-jacents à la maladie coronarienne et au cancer.

Deuxièmement, nous avons cherché à explorer l'effet de l'hématopoïèse clonale, une condition dans laquelle certaines cellules sanguines sont produites à partir d'une seule cellule mutée plutôt qu'à partir de processus normaux de production de cellules sanguines, sur le risque de décès par causes cardiovasculaires chez les patients diagnostiqués avec un cancer en utilisant une grande cohorte prospective. Dans une première analyse, nos résultats ont montré que les porteurs de modifications chromosomiques en mosaïque, un type distinct d'hématopoïèse clonale, présentaient un risque plus élevé de décès par maladie coronarienne que les non-porteurs. Dans une analyse de suivi, nos résultats ont montré que les porteurs d'hématopoïèse clonale de potentiel indéterminé étaient plus susceptibles de succomber à la mort due à des causes de maladies cardiovasculaires que les non-porteurs. Bien que la présence combinée des deux types d'hématopoïèse clonale n'était pas associée à un effet additif sur les critères d'évaluation liés aux maladies cardiovasculaires, elle a conféré un risque plus élevé de mortalité globale par rapport à ceux ayant un seul type d'hématopoïèse clonale. D'autres contributions scientifiques apportées au cours de mes études doctorales mettent l'accent sur le potentiel de l'utilisation de thérapies anti-inflammatoires pour réguler l'inflammation, et donc modifier le développement de l'hématopoïèse clonale, réduisant efficacement les événements cardiovasculaires et éventuellement les effets liés au cancer.

Dans l'ensemble, les résultats de la thèse actuelle confirment que les processus inflammatoires sous-tendent le développement et la progression des maladies cardiovasculaires et du cancer. Cependant, les résultats de cette thèse soulignent la nécessité d'une évaluation continue, compte tenu des variations dans les liens entre l'inflammation, le cancer et les maladies cardiovasculaires à travers diverses voies inflammatoires et types de cancers.

Mots-clés : maladies cardiovasculaires, cancers, inflammation, cardio-oncologie, biomarqueur

Abstract

Cancer and cardiovascular diseases are two major health threats to humanity. Survivors of cancer have a higher risk of developing and dying from cardiovascular disease compared to the general population. Similarly, patients with cardiovascular disease are also more likely to be diagnosed with cancer. While both conditions share many risk factors, cancer and cardiovascular disease have traditionally been studied in isolation. Recently, ongoing research has revealed biological commonalities between the two diseases. These findings are owed to significant scientific advances that have allowed for an expanded molecular characterization of inflammatory modulators, the elucidation of the structures of proteasome and inflammasome, and the advancement of knowledge in cellular immunity.

Thanks to these advancements, valuable insights can now be gained into the complex circuity of the innate and adaptive immune response, as well as the mechanisms of the host defenses. The remarkable progress made in these areas has laid a strong foundation for inflammation biology, and the application of the principles learned in this field to human disease has only recently begun to yield fruitful results, leading to the exponential growth of the field of inflammation in cancer and cardiovascular disease. For both cancer and cardiovascular disease, immuno-modulation has recently emerged as a key factor in their treatment and management. Moreover, it has become apparent that targeting a particular inflammatory pathway to treat one condition can have significant implications for the other, reflecting the complex interplay between these two disease processes.

These developments have enabled the recognition that inflammation is an active participant in the pathophysiological development and progression of both cancer and cardiovascular disease. However, our understanding of the role of inflammation in cancer and cardiovascular disease is still evolving, as ongoing research is necessary to uncover new insights into the complex mechanisms of their triangular relationship. In the present PhD thesis, we sought to conduct distinct analyses that would further our understanding and contribute to the current efforts in innovative research that could help reduce the growing burden of cancer and cardiovascular disease.

First, we performed a genetic meta-analysis on the risk of incident cancer in patients with coronary artery disease taking statins. In that study, we were able to identify and replicate a genetic variant that was associated with a higher risk of incident cancer diagnosis in women statin users. The genetic variant was located in the human leukocyte antigen region that is largely responsible for the activation of T cells and the regulation of immune responses. These findings underscore the active involvement of inflammatory processes in the development of coronary artery disease and cancer.

Second, we sought to explore the effect of clonal hematopoiesis, a condition in which certain blood cells are produced from a single mutated cell rather than from normal blood cell production processes, on the risk of death from cardiovascular-related causes in a large prospective cohort of patients diagnosed with cancer. In a first analysis, our findings showed that carriers of mosaic chromosomal alterations, a distinct type of clonal hematopoiesis, had a higher risk of death from coronary artery disease than non-carriers. In a follow-up analysis, our findings showed that carriers of clonal hematopoiesis of indeterminate potential were more likely to succumb to death from cardiovascular disease causes than non-carriers, and while the combined presence of both types of clonal hematopoiesis was not associated with an additive effect on cardiovascular-related endpoints, it did confer a greater risk of overall mortality compared to those with either type of clonal hematopoiesis alone. Other scientific contributions made during the course of my doctoral studies emphasize on the potential use of anti-inflammatory therapeutics to alter the course of clonal evolution, for the purpose of reducing cardiovascular-related outcomes, and simultaneously cancer-related endpoints.

Overall, the results of this thesis support the notion that inflammatory processes underlie both diseases' development and progression. However, the results of this thesis also highlight the need for continuous evaluation, taking into account the variations in the connections among inflammation, cancer, and cardiovascular disease across various inflammatory pathways and types of cancers.

Keywords: cardiovascular disease, cancer, inflammation, cardio-oncology, biomarker

Table of contents

List of Tables

List of Figures

Abbreviations

ACE: angiotensin-converting enzyme
ACC/AHA: American College of Cardiology/American Heart Association
ASCVD: ACC/AHA atherosclerotic cardiovascular disease
ApoB: apolipoprotein B
BAF: B-allele frequency
BMI: body mass index
CABG: coronary artery bypass graft
CAD: coronary artery disease
CARE: Cholesterol and Recurrent Events Study
CAR-T: chimeric antigen receptor T
CH: clonal hematopoiesis
CHIP: clonal hematopoiesis of indeterminate potential
Chr: chromosome
CI: confidence interval
COLCOT: Colchicine Cardiovascular Outcomes Trial
CN-LOH: copy number neutral loss of heterozygosity
CRP: C-reactive protein
CTLs: cytotoxic T cells
CVD: cardiovascular disease
DDR: DNA damage response
DNA: deoxyribonucleic acid
EAF: effect allele frequency
EMT: epithelial to mesenchymal transition
EPIC: European Prospective Investigation into Cancer and Nutrition
eQTL: expression quantitative trait loci
GLM: generalized linear regression
GWAS: genome wide association studies
HDL: high-density lipoprotein
HDL-c: high-density lipoprotein cholesterol
HLA: human leukocyte antigen
HR: hazard ratio
hs-CRP: high-sensitivity C-reactive protein
HSPCs: hematopoietic stem and progenitor cells
IFN: interferon
IFNγ: interferon-gamma
IGF: insulin-like growth factor
IgM: immunoglobulin M
IL: interleukin
IQR: interquartile range
LDL: low-density lipoprotein
LDL-c: low-density lipoprotein cholesterol

LIPID: Long-Term Intervention with Pravastatin in Ischemic Disease
LRR: $\log_2$R ratio
mCAs: mosaic chromosomal alterations
MAF: minor allele frequency
MHC-II: major histocompatibility complex class II
MHI: Montreal Heart Institute
MI: myocardial infarction
mLoX: mosaic loss of the X chromosomal
mLoY: mosaic loss of the Y chromosome
MMPs: matrix metalloproteinases
NETs: neutrophil extracellular traps
NKT: natural killer T cells
NLRP3: NOD-like receptor family, pyrin domain-containing protein 3
NSAIDs: non-steroidal anti-inflammatory drugs
NSTEMI: non-ST-elevation myocardial infarction
OR: odds ratio
OxLDL: oxidized low-density lipoprotein
PC: principal components
PCI: percutaneous coronary intervention
PDGF: platelet-derived growth factor
PROSPER: PRospective Study of Pravastatin in the Elderly At Risk
RNA: ribonucleic acid
ROC: receiver operating characteristic
RR: relative ratio
SE: standard error
SEER: Surveillance, Epidemiology, and End Results
SNP: single-nucleotide polymorphism
STD: standard deviation
STEMI: ST-elevation myocardial infarction
TAM: tissue-associated macrophages
Tbd: to be determined
Th: helper T
TIA: transient ischemic attack
TME: tumor microenvironment
TGFβ: transforming growth factor beta
TNF-α: tumor necrosis factor alpha
Tregs: T regulatory cells
VAF: variant allele frequency
VSMCs: vascular smooth muscle cells
WES: whole-exome sequencing
WHO: world health organization

Pour Ludovic et Alexandre

Acknowledgements

I would like to express my sincere gratitude to the following individuals and institutions for their invaluable support and contributions to the completion of my thesis.

First and foremost, I would like to express my deep gratitude and appreciation to my thesis director, Marie-Pierre Dubé, for her guidance, support, and mentorship throughout my graduate studies.

I am also extremely grateful to my thesis co-director, Jean-Claude Tardif, for his generosity, valuable insights, and input to my studies.

I would like to extend my thanks to the lab members of the StatGen team, including Amina Barhdadi, Essaïd Oussaid, Géraldine Asselin, Johanna Sandoval, Louis-Philippe Lemieux Perreault, Marie-Christyne Cyr, and Sylvie Provost for their collaboration, technical support, moral support, and valuable feedback on my work. I also wish to express my appreciation to past members Audrey Lemaçon, Marc-André Legault, and Yassamin Feroz Zada.

I am grateful to the Pharmacogenomics support team, steered by Ian Mongrain, for providing me encouragement, support, and resources for my research. Additionally, I would like to thank the Montreal Heart Institute for providing access to their platform and facilities.

I would like to express my thanks to Lambert Busque, an internationally recognized pioneer of research in clonal hematopoiesis, for his continuous support and interest in the work that I have produced during the course of my thesis.

I would like to acknowledge the support of the FRQS for funding my studies, and the Université de Montréal for providing additional opportunities for professional development and networking.

I am grateful to the reviewing committee for overseeing my thesis and providing valuable feedback and suggestions.

Finally, I wish to express my heartfelt thankfulness to my family, especially Paul Soucie, for his unwavering support and encouragement. Paul has always shown interest in my work and been my biggest fan. I am forever grateful for his love and care.

Chapter 1

Introduction

Cancer and cardiovascular disease (CVD) are two of the leading causes of death worldwide, and while they are distinct diseases, recent research has highlighted a surprising connection between the two. The link between cancer and CVD is recognized as a consequence of shared risk factors, such as age, obesity, and smoking, but it is becoming apparent that inflammation may play a crucial role in the pathogenesis of both diseases. Chronic inflammation is a common denominator in many diseases, and an area of growing research interest is the shared mechanisms underlying the inflammatory processes of both conditions. The aim of this thesis is to explore the shared etiology of inflammation in cancer and CVD, with a focus on the inflammatory pathways involved and the potential implications for the prevention and treatment of these diseases.

1.1 – Intersection of Cancer and Atherosclerosis

1.1.1 Common Links Between Cancer and CVD

Without a doubt, cancer and CVD are the world's largest contributors to the burden of chronic disease, and will remain so in the upcoming years.[1] In 2022, approximately 18 million and 20 million individuals with a history of cancer and coronary artery disease (CAD), respectively, were estimated.[2,3] These numbers will unquestionably continue to rise given the ageing population and therapies prolonging life expectancy.

Historically, cancer survivors were shown to have a higher risk of developing CVD and dying from it than the general population.[4–6] More recent studies have continued to support earlier findings. For example, using the Surveillance, Epidemiology, and End Results (SEER) program, Sturgeon et al.[7] showed among 3,234,256 patients with a history of cancer between 1973 and 2015, 38.0% died from cancer and 11.3% died from CVDs, defined as heart disease, hypertension, cerebrovascular disease, atherosclerosis, and aortic aneurysm/dissection. While the first year following cancer diagnosis was the peak of CVD mortality risk for all cancer sites, patients with a diagnosis of breast, melanoma, and prostate cancer experienced a continually elevated risk of

mortality from heart disease, even after adjusting for age at diagnosis and accounting for follow-up time following cancer diagnosis.[7]

In a similar fashion, patients with CVD are also more prone to being diagnosed with cancer.[8–12] In a prospective cohort study, the incident density rates for cancer diagnosis among residents of a local county in Minnesota were 33.7 per 1,000 person-years in patients who had heart failure compared to 15.6 per 1,000 person-years in residents without a previous heart failure (*P*=0.002).[8] In the Long-Term Intervention with Pravastatin in Ischemic Disease (LIPID) randomized trial which compared pravastatin to placebo in patients who experienced myocardial infarction (MI) or unstable angina, investigators sought to assess the role of D-dimer, a marker of hypercoagulability and thrombotic events, in the prediction of long-term vascular outcomes and incident cancers. Their findings showed that the highest D-dimer level quartile (>273 ng/mL) were significantly associated with cancer incidence (hazard ratio [HR]: 1.16, 95% confidence interval [CI]: 1.03 to 1.31) and cancer-specific mortality after 16 years (HR: 1.54, 95% CI: 1.25 to 1.91), compared to the lowest quartile (≤112 ng/mL, *P*<0.001).[10] Using a nationwide population-based Danish medical cohort, patients who were previously diagnosed with lower limb arterial thrombosis between 1994 and 2013 were identified and their standardized cancer incidence ratios computed. Among 6,600 patients, 772 subsequent cancers were observed. Within 6 months of follow-up, the standardized incidence ratio of any cancer was 3.28 (95% CI: 2.79 to 3.82), where a higher ratio was maintained in patients followed beyond 12 months (standardized incidence ratio: 1.14, 95% CI: 1.05 to 1.24).[11] In a recent study based on a retrospective Swedish cohort, patients hospitalized for their first MI between 2001 and 2014 were identified (n=175,146). Both prevalent and incident cancers were observed in 6.7% of patients between 2001 and 2002 and rose to 10.7% between 2013 and 2014. The presence of a cancer diagnosis was also shown to be significantly associated with recurrent MI (HR: 1.08, 95% CI: 1.04 to 1.12) and heart failure (HR: 1.10, 95% CI: 1.06 to 1.13).[12]

The co-occurrence of both cancer and CVD is not uncommon and is partly driven by their many shared risk factors, including age, smoking, obesity, and diabetes. Tobacco smoking produces many irritants, carcinogens, and pro-inflammatory stimuli that greatly impact both CVD and the development of multiple types of cancers.[13,14] Pro-inflammatory cytokines produced within

adipose tissue are more pronounced in the serum of individuals who are obese. One of the most common cytokines produced within adipose tissues is interleukin (IL)-6, which is associated with increased blood pressure and stimulates the production of C-reactive protein (CRP), a conventional marker of inflammation. Simultaneously, overexpression of IL-6 has been shown to inhibit cancer cell apoptosis and stimulate angiogenesis, leading to tumor progression.[15] Diabetes mellitus can exert an influence on CVD and the neoplastic process through mechanisms of hyperinsulinemia, hyperglycemia, insulin-like growth factor (IGF), and inflammation.[16–18]

Given these considerations, it may be less surprising that the control of CVD risk factors can simultaneously reduce the risk of cancer. In the European Prospective Investigation into Cancer and Nutrition (EPIC) study which followed 23,153 participants aged between 35 and 65 years old, those who adhered to four pre-determined healthy lifestyle factors, defined as never smokers, a body mass index (BMI) <30, physical activity of at least 3.5 hours per week, and a healthy diet had significantly lower risk of incident cancer (HR: 0.64, 95% CI: 0.43 to 0.95).[19] The study led to the development of a framework urging both the oncological and cardiovascular communities to recognize points of common occurrence between cancer and CVD, beyond the manifestation of cardio-toxicities as a result of chemo/radiotherapy.[20] Understanding the delicate interaction between cancer and CVD may eventually lead to better prevention and treatment strategies to reduce the growing burden of both diseases.

1.1.2 Inflammation: An Unifying Theme

Given the overlap in risk factors between cancer and CVD, there was premise of a common biological pathway. Historically, cancer and CVD have been viewed as two distinct disease entities. Over time, it became clear that there was more to the relationship between cancer and CVD than just the cardiotoxicities of chemotherapy/radiotherapy[21–23], with evidence suggesting a fundamental biological overlap between the two. Indeed, there was cumulating evidence that patients with cancer often had relative cardiac hypertrophy prior to receiving any treatments for cancer.[24] Others have also reported abnormal cardiac function at baseline, prior to the receipt of any treatment for their cancer.[25]

Recent years have witnessed remarkable progress in the fields of inflammation and immunity, with significant advancements in the understanding of oncology and CVD. One of the key developments in this field has been the identification of supramolecular intracellular structures such as the proteasome and the inflammasome. These structures have provided crucial insights into the immune (dys)-regulation and the resulting inflammatory responses generated by the body.[26,27] This enhanced understanding of innate and adaptive cellular immunity has led to the translation of such findings into clinical practice in both oncological and cardiovascular settings, enabling the development of novel therapeutic strategies that target the underlying mechanisms of these diseases.

Therefore, in addition to the treatment-induced CVD risks and shared risk factors, there is now a bourgeoning realization that inflammation participates in the pathogenesis of both cancer and CVD. The underlying clinical benefit in studying this unifying mechanistic link is the possibility that the inhibition of pro-inflammatory cytokines can simultaneously reduce events related to both CVD and cancer. Proof-of-concept data in agreement with this possibility already exist. For example, the frequent use of non-specific inhibition of inflammation, such as aspirin and statins, has often appeared to exert an influence on cancer development/progression. Additionally, the emerging concept of clonal hematopoiesis (CH), an age-related phenomenon driven by inflammation as a shared risk factor for both cancer and CVD, has further substantiated the idea of common patho-physiological processes underlying both diseases. While efforts to apply principles related to inflammatory processes in cancer and CVD have progressed, the fragmented nature of the field and the presence of many uncertainties insinuate that continuous work is needed. Nevertheless, ongoing research and a deeper understanding of the unifying mechanism of inflammatory processes in cancer and CVD have heralded the dawn of a new era in oncology and cardiology.

While the overarching objective of this new era is to characterize the synergistic inflammatory processes of cancer and atherosclerosis, it is essential to first understand the role of inflammation in each disease individually, as well as how our immune system works normally. The thesis begins with an overview of how the immune system self-regulates in a healthy state.

1.2 – The Field of Inflammation & Immunity

1.2.1 What is Inflammation and What Are the Causes of Inflammation?

Inflammation is an ancient, evolved process, which involves activation, recruitment, and action of cell of innate and adaptive immunity. More specifically, notions of this process emphasize on different aspects of inflammation: 1) a response to perturbation (e.g., infection); 2) a multistep process which seeks to mobilize defensive mechanisms to eliminate the source of perturbation; 3) or an altered state of the system that can be either protective or pathological. There is renewed interest in inflammation biology given the understanding that it is associated with nearly every human disease.[28] More and more so, inflammation is thought of as a *spectrum* where on one extreme there is acute inflammation and on the other end, cells and molecules that are typically involved in the acute inflammation process but also participate in normal homeostatic processes in the absence of a perturbation.[29] Acute inflammation is the most widely recognized type of inflammation in both the clinical and non-clinical settings and involves local changes in microvasculature, which results in inflammatory exudate giving rise to the cardinal signs of inflammation (e.g., swelling, pain, redness, heat). What has been uncovered more recently is that the so-called inflammatory signals (including cytokines, chemokines, bioactive amines) and cell types (including macrophages, neutrophils, mast cells, lymphocytes) are in fact involved in processes ascribed to both ends of the "inflammation spectrum" – the acute and the homeostatic phases.[30] There has been an increased appreciation that the best way to understand an acute inflammation is to understand the normal (homeostatic) immune system.

Truly, when one questions how inflammation can still be present in the absence of triggers, such as tissue injury or infection, it becomes clear that the traditional view on inflammation is limited and incomplete. But this is the type of inflammation that is typically associated with common health states, such as ageing, sleep deprivation, and obesity.[29,31–33] The common denominator, as it is for many causes of inflammation, is that there is a deviation of the biological system from the homeostatic "normal" state.[34,35] The elicited inflammation can be characterized as a physiological inflammatory response to environmental factors that threaten organismal homeostasis.

Homeostatic regulation is typically maintained through negative feedback mechanisms that work to correct any deviations of the immune system's state variables to an accepted range of values. However, when the deviations become significant, these mechanisms may no longer be sufficient to maintain stability, leading to the engagement of inflammatory signals to aid homeostatic regulation and restore the system to its normal state.[30] As an example, ageing and deteriorating health are characterized by the loss of cellular, tissue, and organismal homeostasis, with inflammation often being a consequence of the loss of homeostatic regulation. As the present thesis focuses on the relationship between inflammation and cancer and CVD, it is worth noting that the physiologic inflammation that occurs in the absence of infection or injury, and unlike inflammation caused by structural damage, is a distinct type of inflammatory response that requires further investigation in the context of these diseases.

1.2.2 Understanding the Inflammatory Pathways

Any inflammation is composed of dozens of molecular mediators and the involvement of various cell types. The general framework of any inflammatory response also depends on the cause of inflammation (i.e., whether it is a loss of structure or a loss of regulation, as discussed above). Regardless of the type of framework of the inflammatory response, four universal components encompass the inflammatory circuit (**Figure 1.1**).

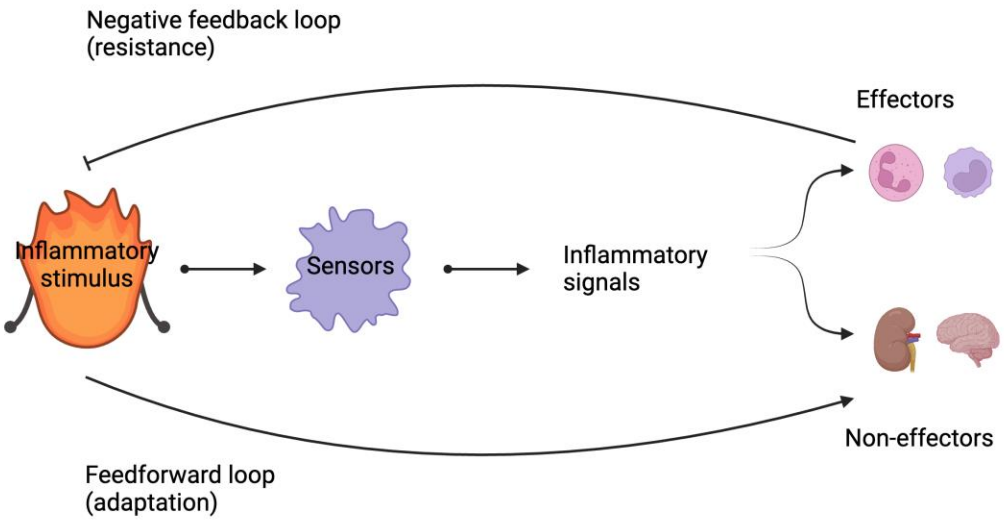


Figure 1.1 Proposed Inflammatory Circuit.

Adapted from Medzhitov[30]. *The circuit involved in inflammation includes an inflammatory stimulus, sensors, and inflammatory signals that act on two distinct types of targets. The first type, known as inflammatory effectors, aim to eliminate the cause of inflammation through a* ***negative feedback loop****, resulting in resistance. The second type, referred to as non-effectors, consists of tissues and organs that can alter their functions to adapt to the perturbation, seeking to offer support to the effector response through a* ***feedforward loop****, resulting in adaptation or tolerance to the inflammatory state of the organism.[1]*

Essentially, the inflammatory circuit involves an inflammatory stimulus, sensors, and inflammatory signals that act on two types of targets, one of which are inflammatory effectors, such as neutrophils, monocytes, eosinophils, mast cells, and basophils., and non-effectors. The effectors' goal is to eliminate the cause of the inflammation by creating a negative feedback loop as part of the circuit (resistance). In addition to signals emitted by the effector cells, other inflammatory signals will trigger tissues and organs that are not directly implicated in the elimination of pathogens. Any function that a tissue or organ will do under "normal" homeostatic state, inflammatory signals can alter these functions in accordance, or against homeostasis. Inflammatory signals from non-effectors do not explicitly eliminate threat (e.g., pathogen clearance) but undergo functional changes as directed by the inflammatory signals as an adjunct

[*1*] Reprinted from Medzhitov R. The spectrum of inflammatory responses. *Science*. 2021;374(6571):1070-1075. Reprinted with permission from the American Association for the Advancement of Science (AAAS, license #5518281450597).

to the effectors' response. When inflammatory signals target non-effectors, it represents a feedforward part of the inflammatory circuit as it seeks to promote adaptation or tolerance following infection.[36,37]

Understandably, when overstressed, both the resistance and adaptation parts of the inflammatory circuit can lead to pathological effects. Excessive resistance may result in collateral tissue damage and negative outcomes on the tissues and organs involved, while over-reliance on adaptation may result in extreme changes in the functional states of non-effector target tissues and organs. The challenge of inflammation biology lies in achieving a balance in the inflammatory circuit to optimize the magnitude and duration of an inflammatory response without over- or under-regulating itself. The inflammatory response can be characterized as acute when the resistance pathway dominates, while chronic inflammation may occur when the resistance pathway fails to remove the pathogen, and the adaption/tolerance pathway overtakes the circuit.[35]

1.2.3 Controlling Inflammation

Researchers have come to understand some of the ways that the inflammatory circuit operates to manage the trade-off between the magnitude of the response (i.e., pathogen elimination) without inflicting sustained inflammatory damage. Eliminating the cause of the inflammation is not enough and without inflammation regulation following the initial response, pathological outcomes are guaranteed. Therefore, one of the ways it does this is that vital tissues and organs more vulnerable to inflammatory damage have lower "reactiveness" to inflammatory signals.[36] Alternatively, counter-inflammatory signals may also be produced to reverse inflamed tissues back to its original homeostatic state. Such counter-inflammatory signals may include signals that harbor naturally homeostatic abilities (e.g., epinephrine), or specialized signals dedicated to counter inflammation, such as T regulatory cells[38] and signals involved in resolution, such as lipoxins and resolvins[39,40]. Experts have started to distinguish between anti-inflammatory and counter-inflammatory effects. Anti-inflammatory effects aim to inhibit the magnitude of the response, while counter-inflammatory effects focus on reducing the responsiveness of the target tissues and organs **(Figure 1.2).** In summary, an acute high-grade response triggered by an infection or tissue injury is the most known and well-characterized type of inflammation. A less

well-characterized form of inflammation is the chronic, low-grade variety, which entails a persistent cycle balancing the initial stress with homeostatic regulation. This type of inflammation has increasingly been linked with many patho-physiology of chronic diseases[28,29] and has also been associated with biological ageing [33,41].

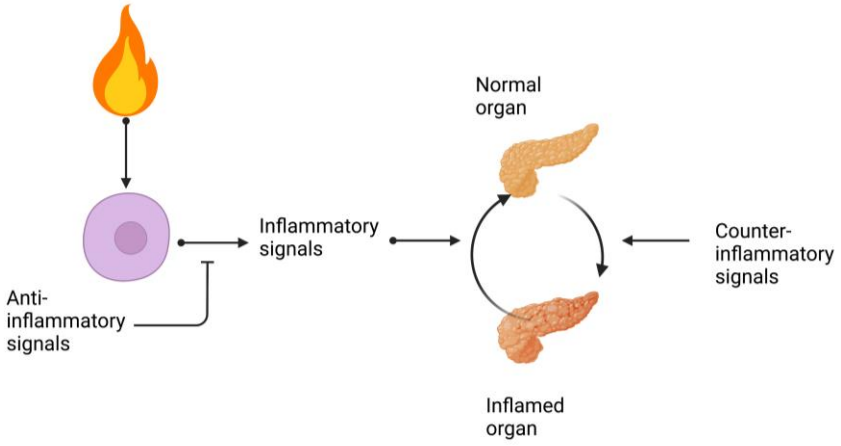


Figure 1.2 Distinction Between Anti-Inflammatory and Counter-Inflammatory Signals

Adapted from Medzhitov.[30] Anti-inflammatory signals control the magnitude of the inflammatory response, whereas counter-inflammatory effects rein back the inflammatory effects of targeted tissues and organs.[2]

As noted previously, inflammation serves as a unifying and mediator factor in many diseases, including cancer and CVD. The overlapping risk factors between these conditions, such as obesity, hyperglycemia, and hypertriglyceridemia can induce inflammation, underscoring the potential connection between the two.[42–45] Recognizing the profound patho-physiologic link between cancer and CVD through the common mechanism of inflammation biology opens up opportunities for innovative therapeutics in the emerging field of cardio-oncology. However, to investigate this field further, we must first comprehend how inflammation affects each condition individually.

[2] Reprinted from Medzhitov R. The spectrum of inflammatory responses. *Science*. 2021;374(6571):1070-1075. Reprinted with permission from the American Association for the Advancement of Science (AAAS, license #5518281450597).

1.3 – Inflammation and Cancer

1.3.1 The Tumor Microenvironment

In 1863, after noticing that certain irritants and tissue injuries induced by inflammation increased cell proliferation, Virchow proposed that cancer possibly originated at sites of chronic inflammation in the lymphoreticular infiltrate.[46] In the last three decades, our understanding of the inflammatory microenvironment and malignant tissues has unanimously corroborated Virchow's original hypothesis. The implications between cancer and inflammation have since flourished into prevention strategies and therapeutics for patients.

Presently, cancer biology has progressively shifted from a cancer cell centric view to a more inclusive conception that places cancer cells as part of network of stromal cells, comprised of fibroblasts and vascular cells and inflammatory immune cells – all of which form the tumor microenvironment (TME).[47] Activation of inflammation, whether due to a chronic inflammatory disease or tumor-elicited smoldering inflammation, can significantly affect the composition of the TME.

Evolutionary pressure has enabled the ability to respond to infections, facilitate wound healing, and regenerate tissue, but not necessarily to prevent the development of tumors (**Figure 1.3**). The upper panel depicts normal and organized tissue, where cells separated from blood vessels by a membrane. When there is an injury or wound, platelets are activated to form a clot and promote healing. Other cells are recruited to help with healing and restoring the veins. Ironically, the inflammatory response indispensable to wound healing inconveniently promotes tumor development. Specifically, the lower panel illustrates a cancerous tumor, which is less organized than normal tissue. The cancer cells interact with other cell types and blood vessels, creating a disorganized and chaotic network. Although the vascular network is not engaged in the same way during neoplastic progression as it is during wounding, many reciprocal interactions between cells and blood vessels end up fueling the growth and spread of the cancer cells throughout the body.

During a typical inflammatory response, tissue-associated macrophages (TAMs) respond to tissue changes by eliminating dying cells, producing chemotactic molecules to recruit other cell types if

necessary, regulating immune signals, and supporting the stem cell niche. Tissue inflammatory responses are maintained through three primary inter-related mechanisms:

1) Self-regulation of local tissue macrophages and dendritic cells through local proliferation;
2) Induction of immune cell recruitment from the bone marrow, including monocytes, neutrophils, monocyte-derived cells, and secondary lymphoid cells, in response to signals that threaten tissue homeostasis, and;
3) The recruitment of amplified inflammatory cells that undergo local activation, differentiation, and polarization as influenced by the microenvironment.[29,48]

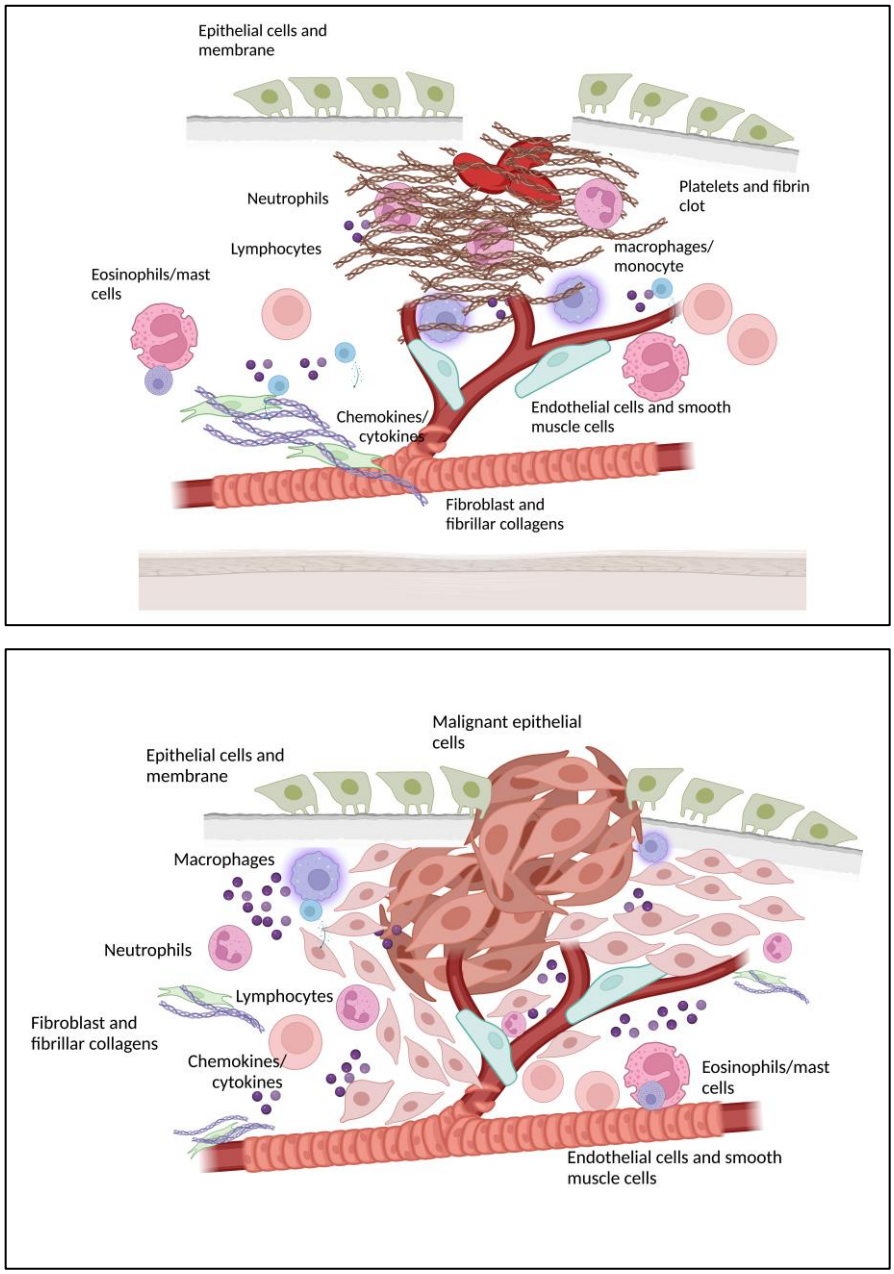


Figure 1.3 Visual Comparisons of a Typical Wound Healing Process Versus an Invasive Tumor Growth

Adapted from Coussens & Werb[49]. *The **top panel** shows organized normal tissues with segregated architecture. The epithelial cells are located at the top, separated from the vascularized stromal compartment by a membrane. In case of a wound or tissue injury, platelets become activated, forming a homeostatic plug and releasing vasoactive mediators that promote vascular permeability and the influx of serum fibrinogen, resulting in the formation of a fibrin clot. Subsequently, chemotactic factors such as tumor necrosis factor alpha (TNF-α) and platelet-derived growth factor (PDGF) initiate granulation tissue formation, activate fibroblasts and recruit proteolytic enzymes*

necessary for extracellular matrix remodeling. Recruitment of granulocytes, monocytes, and fibroblasts restores the venous network, and re-epithelialization across the wound takes place. During this process, there is reciprocal signaling dialogue between epithelial and stromal cells to facilitate healing. In contrast, invasive carcinomas in the ***bottom panel*** *demonstrate less organization. Neoplastic cells interact with other cell types, such as mesenchymal, hematopoietic, and lymphoid cells, as well as remodeled extracellular matrix, following neoplasm-induced angiogenesis and lymph-angiogenesis. The authors note that the vascular network is as perturbed during neoplastic progression as it is during wound healing. Dual interactions occur in parallel, where neoplastic cells produce cytokines and chemokines that are mitogenic and/or chemo-attractant for mast cells, granulocytes, monocytes and macrophages, fibroblasts, and endothelial cells. Activated fibroblasts and infiltrating inflammatory cells also secrete proteolytic enzymes, cytokines, and chemokines, which are mitogenic for neoplastic and endothelial cells involved in neo-angiogenesis and lymph-angiogenesis. These factors promote tumor growth, fuel angiogenesis, induce fibroblast migration and maturation, and facilitate metastatic spread through the venous or lymphatic networks.[3]*

In parallel, inflammation in cancer can be characterized by a hyperproliferation of epithelial cells which also prompt homeostatic responses for the purpose of increasing the numbers of macrophages and fibroblasts (tissue blocks). This happens by signaling circuits of chemokines and growth factors.[50] In contrast to wound healing, which resolves after immune cell recruitment and epithelial cell proliferation, growing tumors show persisting oncogene-derived stress, cell apoptosis, and microbial signals that altogether flourish into a feed-forward loop of inflammation-induced signaling (as described in **Figure 1.1**, section 1.2.2) and inflammatory cell recruitment. In this context, the enhanced inflammation and cytokine-driven proliferation is forced to accommodate tumor growth instead of restoring normal epithelial homeostasis.[27]

The infiltration of cells and inflammation within tumors can have a paradoxical effect on tumor growth, metastasis, and patient outcomes. Depending on the specific mechanisms involved, it can either promote or hinder tumor invasion, creating a paradox akin to Dr. Jekyll and My. Hyde.[51–53] Specifically, tumor cells produce different cytokines and chemokines attracting a diverse population of cells including neutrophils, dendritic cells, macrophages, eosinophils and mast cells. This leads to an array of cytokines, cytotoxic mediators, serine and cysteine proteases, matrix metalloproteinases (MMPs) and membrane-perforating agents then act on soluble

[3] Reprinted from Coussens LM, Werb Z. Inflammation and cancer. *Nature*. 2002;420(6917):860-867. Reprinted with permission from Springer Nature (license #5518290197283).

mediators of cell killing, such as TNF-α, ILs and interferons.[54,55] Monocytes are a type of white blood cell that can differentiate into different types of immune cells based on the environment they are in. When stimulated with granulocyte-macrophage colony-stimulating factor and IL-4, monocytes can differentiate into subsets of dendritic cells. These dendritic cells can then travel to areas of inflammation in the body, capture antigens (molecules that trigger an immune response) and bring them to lymph nodes. Once in the lymph nodes, the dendritic cells mature and present the antigens to T-lymphocytes (another type of immune cell), which activates them to fight off the infection or cause inflammation.

Returning to the subject of TAMs, they are the most found immune cells of the TME (**Table 1.1**). TAMs are recruited to the tumor site by chemokines and are mainly responsible for promoting tumor growth as well as angiogenesis (the formation of new blood vessels) and metastasis (the spread of cancer to other parts of the body). However, TAMs also have a dual role, as they are capable of killing cancerous cells when triggered by certain cytokines, including IL-2, interferon, and IL-12.[56,57] Despite this, TAMs also produce several growth factors, cytokines, and proteases that promote the progression of cancer.[58] Additionally, TAMs and tumor cells also produce IL-10, which can inhibit the immune system's anti-tumor response generated by cytotoxic T cells.

Mature T cells are commonly found in the TME and can be divided into two groups based on the T cell receptors they express: γδ and αβ. The latter is further sub-classified into $CD8^+$ cytotoxic T cells (CTLs) and $CD4^+$ helper T (Th) cells (including Th1, Th2, Th17, and T regulatory cells or Tregs), and natural killer T (NKT) cells. T cells have been found to exhibit both tumor-suppressing and promoting functions, depending on their effector functions.[59–61] For instance, activated CTLs and Th1 cells are associated with improved survival in patients with colon cancer, melanoma, and pancreatic cancer.[62,63] However, many T cell subsets found in solid malignancies are involved in tumor promotion, progression, and metastasis, including $CD8^+$ T cells[64] and Th cells[65,66]. Like TAMs, T lymphocytes can also promote tumor growth by releasing cytokines[67]. Nevertheless T lymphocytes, including Tregs, which are generally considered pro-tumorigenesis, have also demonstrated anti-tumorigenic effects in certain contexts.[68]

Aside from the immune cell content of the TME, the cytokine and chemokine profiles also induce various effects on tumor promotion and tumor inhibition.[67] Cytokines control the immune and inflammatory environment to either favor anti-tumor immunity (IFNγ, IL-12) or promote tumor development (IL-6, IL-17). TAMs, which are crucial producers of cytokines[69], can be classified into M1 and M2 types. M1 macrophages express high levels of pro-inflammatory cytokines (TNF-α, IL-1, IL-6, IL-12, IL-23), major histocompatibility complex molecules, and can kill pathogens and prime anti-tumor immune responses. In comparison, M2 macrophages downregulate major histocompatibility complex class II molecules and IL-12 expression, and are associated with a higher expression of anti-inflammatory cytokine IL-10, scavenger receptor A, and arginase.

To summarize, the TME contains innate immune cells (including macrophages, neutrophils, mast cells, myeloid-derived suppressor cells, dendritic cells and natural killer cells) and adaptive immune cells (T and B lymphocytes) in addition to the cancer cells and their surrounding tumor stroma (fibroblasts, endothelial, mesenchymal cells)[70] (**Table 1.1**). The balance between anti-tumor immunity and cancer-promoting inflammation within the TME is determined by various factors such as the expression level of immune mediators and modulators, as well as the abundance and activation state of different cell types present in the TME.[61,67]

While there are limited studies that have specifically evaluated how the balance between immunity and inflammation shifts during early tumor development, the landscape of evidence in established tumors indicates a tilt toward pro-tumor inflammation. Historically, inflammation's role in cancer has been a prevailing topic, but a turning point occurred when researchers sequencing aged normal tissues discovered 'micro-clones' carrying cancer driver mutations like *p53* and *KRAS*.[71] Surprisingly, these mutations existed without culminating into tumors, challenging the conventional belief that driver mutations are sufficient to instigate cancer. That revelation posited a 'missing component', which prompted researchers to explore triggers that may transform latent clones into active tumors. For instance, recent work has illuminated how environmental factors, such as air pollution, drive IL-β-mediated inflammation, transforming existing epidermal growth factor receptor (EGFR)-mutation clones into tumors through macrophage reprogramming.[72] Another study found that chronic inflammation provides a proliferative advantage to *p53* mutant cells, thereby facilitating their transformation into

tumors.[73] These studies underscore the critical influence of immune dysregulation and inflammation in early tumor initiation.

The question of immune regulation's role in cancer development has also been a compelling question in genome-wide association studies (GWAS). Previously, a study showed that variants in loci NKG2D responsible for coding immune regulatory factors were also associated with a higher risk of cervical cancer.[74] Another study noted that while the immune system generally hampers cancer development, GWAS results have identified variants that are linked to immune cell infiltration within tumors.[75] Nongenic-cancer-risk single-nucleotide polymorphisms (SNPs) have also been shown to act as facilitators in cancer development by directly altering immune functions.[76] In a separate study, investigators found that germline genetic determinants of host immune function and autoimmune diseases might elucidate the risk of toxicity induced by immune checkpoint inhibitors.[77] Collectively, these studies affirm that GWASs have previously identified germline variants linked with immune-mediated diseases and immune cell infiltration in tumors, reiterating the significance of immune-related processes in cancer development. However, it is essential to note that shared heritability does not necessarily translate to the same phenotypes. The complex interplay of genetics, environment, and other factors could lead to divergent outcomes, despite genetic overlap. Comprehensive research is warranted to also understand how such variants manifest in different cancer subtypes.

Table 1.1 Roles of Different Subtypes of Immune and Inflammatory Cells in Anti-Tumor Immunity and Pro-Tumor Inflammatory Response.

Cell types	Anti-tumor	Pro-tumor
Macrophages, dendritic cells, myeloid-derived suppressor cells	Antigen presentation; production of cytokines (IL-12 and type I IFN)	Immunosuppression, production of cytokines, chemokines, growth factors, proteases, and angiogenic factors
Mast cells		Production of cytokines
B cells	Production of tumor-specific antibodies	Production of cytokines and antibodies, activation of mast cells, immunosuppression
$CD8^+$ T cells	Production of cytotoxic cytokines	Production of cytokines (tbd)
$CD4^+$ Th2 cells		Education of macrophages, production of cytokines, B cell activation
$CD4^+$ Th1 cells	Help to CTLs in tumor rejection, production of cytokine (IFNγ)	Production of cytokines

$CD4^+$ Th17 cells	Activation of CTLs	Production of cytokines
$CD4^+$ Treg cells	Suppression of inflammation (cytokines and other suppressive mechanisms)	Immunosuppression, production of cytokines
Natural killer cells	Direct cytotoxicity toward cancer cells; production of cytotoxic cytokines	
Natural killer T cells	Direct cytotoxicity toward cancer cells; production of cytotoxic cytokines	
Neutrophils	Direct cytotoxicity, regulation of CTL responses	Production of cytokines, proteases, ROS

Adapted from Grivennikov et al.[27] *CTL: cytotoxic T lymphocytes, IFN: interferon, IFNγ: interferon gamma, IL-12: interleukin-12, ROS: reactive oxygen species, tbd: to be determined.[4]*

In corroboration to that theory, a large body of evidence has consistently shown that individuals with autoimmune disease appear to have a higher risk of cancer likely due to an underlying immune dysregulation. Specifically, the presence of autoimmune diseases including rheumatoid arthritis, lupus erythematosus, psoriasis, celiac disease, and inflammatory bowel disease have been associated with the development of cancer.[78–82] The co-relationship between autoimmune disease and cancer susceptibility is further supported by the increased risk of cancer onset in patients undergoing hormonal therapy (i.e., androgens, estrogens, progesterone) for immune response regulation.

Furthermore, as another proof-of-concept, many studies have shown that non-specific inhibition of inflammation as seen with non-steroidal anti-inflammatory drugs, such as aspirin, has effectively reduced the incidence and mortality of several types of cancers.[83,84] Similarly, statins, which in addition to their lipid-lowering abilities also harbor anti-inflammatory abilities, have shown to be associated with a lower risk of cancer-related death in patients diagnosed with cancer.[85]

1.3.2 Types of Inflammation in Cancer

What is also important to note is that there are highly heterogenous TMEs for different tumors and metastatic lesions as demonstrated by cell-to-cell transcriptomics which have clearly

[*4*] Reprinted from Grivennikov SI, Greten FR, Karin M. Immunity, inflammation, and cancer. *Cell*. 2010;140(6):883-899. Reprinted with permission from Elsevier (license #5518290612440).

indicated distinct inflammatory cytokine and chemokine profiles with notable qualitative and quantitative differences in the way that inflammatory cells are being recruited.[47,86] Furthermore, it was shown that some cancers that were previously considered as 'non-inflammatory' can actually recruit immune cells and augment the expression of inflammatory mediators to sustain tumor growth, all the while re-formatting their TME to better accommodate the pro-tumor milieu.[27,87]

While a better understanding of the TME has elicited significant progress in deconstructing the roles and mechanisms of action of the immune system in cancer and inflammatory responses associated with them, it is also important to note how inflammation in cancer is induced and how it is sustained, in terms of timing and the cause (**Figure 1.4**). As proposed by experts in cancer immunology, different types of inflammation – divergent in mechanism, outcome, and intensity – can promote tumor development. For example, infections resulting in a persisting inflammation such as from schistosoma or bacteroides species can increase the risk of bladder and colon cancers, respectively.[88,89] The inflammatory response following an infection occurs before tumor development and is part of the normal host-defense response, which aims to remove the pathogen. However, tumorigenic pathogens can subvert host immunity and impose repeated infections associated with low-grade, chronic inflammation. Another example of this is inflammatory bowel disease, considered an autoimmune disease, which has been shown to greatly increase the risk of colorectal cancer.[90]

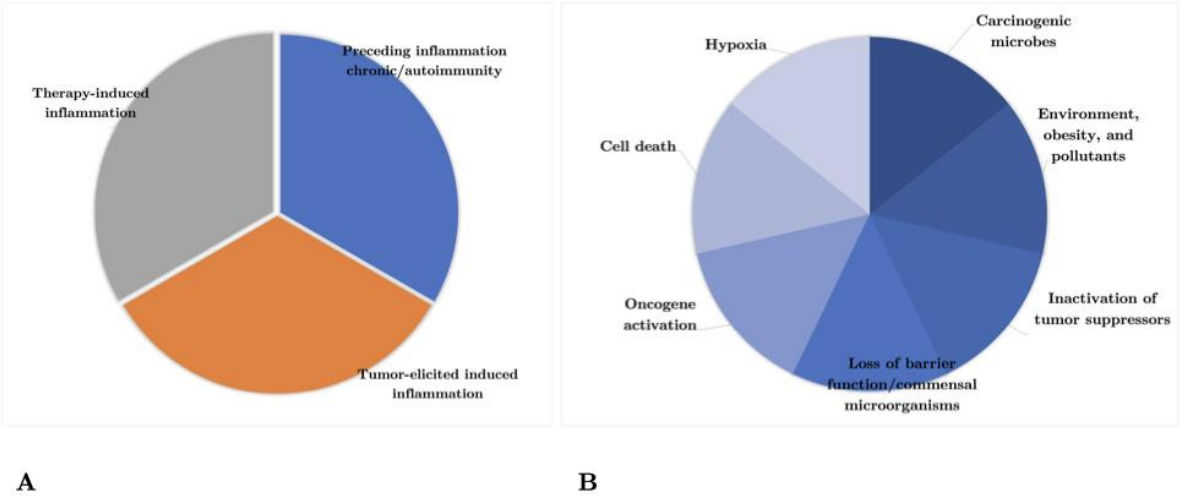


Figure 1.4 Types of Inflammation Based on Timing and Stimulus.

Adapted from Greten & Grivennikov[91]. ***Panel A*** *shows how inflammation related to cancer can be induced at different time points of tumor development. It may start as a form of autoimmune reaction or infection, or induced from malignant cells, or stimulated by anti-tumor therapy.* ***Panel B*** *depicts various sources of inflammation that can aggravate tumor development and progression.[5]*

Some environmental factors can also promote tumor development by induction of chronic inflammation, whether it be low-grade or low-intensity. This could include obstructive pulmonary disease which occurs as a result of repeated exposure to particulate materials from tobacco, which then increases individuals' risk of lung cancer.[92] Alternatively, systemic low-grade inflammation induced by obesity[93] or lipid accumulation has also been shown to increase the risk of various cancers[94–96].

1.3.3 Inflammation and Tumor Initiation, Progression, and Metastasis

Tumor initiation is a process in which normal cells acquire a mutation that sends them on a tumorigenic track by supporting growth and survival advantages over other cells. A single mutation is typically not sufficient for tumor development. Each oncogenic mutation must occur in long-lived stem cells or transient amplifying cells instead of within differentiated cells, as the

[*5*] Reprinted from Greten FR, Grivennikov SI. Inflammation and Cancer: Triggers, Mechanisms, and Consequences. *Immunity*. 2019;51(1):27-41. Reprinted with permission from Elsevier (license #5518290770278)

latter are depleted quickly before the next mutation can spawn.[97] An inflammatory microenvironment will not only accelerate mutations rates, but it will also enhance the proliferation of mutated cells. The activated inflammatory cells are sources of reactive oxygen species and reactive nitrogen intermediates able to induce DNA damage and genomic instability. The inflammatory responses also lead to various epigenetic changes in adjacent epithelial cells.

Within the intestine tract, sustained inflammation elevates the number of mutations in *TP53* and additional oncogenic genes within the epithelial cells[98,99], which facilitates the onset of of tumorigenesis[100]. It is noteworthy that inflammation has the potential to cause DNA damage and mutations. However, certain inflammatory cytokines can induce the activation of DNA damage response genes, which can counteract the genotoxicity caused by inflammation. This presents an interesting and complex relationship between inflammation and DNA damage response.[101] In response to inflammation, cytokines such as IL-6 and IL-β can activate the epigenetic system in epithelial cells including DNA and histone modifications (via DNA methyl transferases DNMT1 and DNMT3), as well as microRNA and long non-coding RNA, ultimately leading to changes in the expression levels of oncogenes and tumor suppressors.[102] When a tissue experiences injury or infection, the immune system responds by triggering an inflammatory response. One of the effects of this response is the expansion of stem cell pool, either by inducing the proliferation of existing stem cells or by de-differentiating other types of cells into stem cells, as the goal is to restore the tissue to its normal state. However, if there already are stem cells in the tissue that have cancer-causing mutations, the expansion of the stem cell pool can lead to the growth and spread of cancerous cells throughout the body (metastasis). The presence of these stem cells which are mediated by many inflammatory stimuli will ultimately affect regeneration, thereby influencing tumor progression.[103]

Inflammatory mechanisms hold a prominent role in the regulation of metastasis. More than 90% of cancer-related deaths are caused by tumor metastasis.[104] The metastatic formation is a culmination of the Darwinian evolutionary process within the tumor, where the competition of multiple subclones favors tumor dissemination. At the core of tumor evolution is the cancer stem cell, which has the potential for self-renewal, clonogenic properties, and genomic instability. Cancer cells that gain the capacity to colonize distant organs (metastatic formation) have not only

the traits of cancer stem cells, but are also able to initiate tumor development in distal tissues undergoing adverse microenvironmental conditions.[105] Cancer stem cells are transcriptionally and functionally closer to mesenchymal cells than regular epithelial or bulk tumor cells.[106]

As mentioned earlier, the lower overall rate of metastatic gastrointestinal cancer deaths in long-term aspirin users illustrates well the role of inflammation in distant disease. Experimental studies have also shown that inducing the resolution of inflammation or blocking inflammatory responses results in poor tumor colonization and helps destroy micrometastases.[107] This is likely because metastasis is an inefficient process, as most cells released from the primary tumor die and do not form distant metastases. Drug targets that can pass through the bottleneck would significantly increase the chances of success and have a positive impact on patient survival.

Metastasis is a process by which cancerous cells spread from their original location to other parts of the body. The initiation of metastasis involves cancer cells breaking through the outermost layer of the organ or tissue where they are located, the epithelial cell layer, allowing them to invade neighboring tissues. The process of initiation is facilitated by the acquisition of a phenotype called epithelial to mesenchymal transition (EMT), which allows the cancer cells to become more mobile and invasive. Essentially, cancer cells lose their normal characteristics and gain new traits that help them spread and invade other parts of the body, including the lymphatics or blood vessels for further dissemination. Inflammatory processes act on cancer invasion, the EMT and cell migration. Recruited cytokines including TNF and IL-1β can directly enhance expression of EMT-inducing transcription factors.[108] Recruited myeloid cells results in the production of matrix metalloproteinases and other enzymes facilitating cell migration.[109] Myeloid cells of granulocytic and monocytic lineage, often referred to as myeloid-derived suppressor cells also contribute to suppression of anti-tumor responses.[110,111] Evidence suggests that a higher number of cytotoxic T and NKT cells in combination with a decreased number of myeloid cells correlates with better prognosis.[112]

In summary, research on the mechanisms that drive cancer initiation and progression has revealed that the same inflammation pathways involved in promoting immunity and tissue homeostasis following an infection are used by tumors to their advantage. The timing of

inflammation-initiated TME varies, but it can happen before the development of a tumor (chronic infection), or it might sporadically occur at later stages of tumorigenesis. This timing is strongly dependent on the occurrence of triggers, as it can be induced by environmental factors or cell intrinsic factors, and even therapy related. In recent developments, the field of cancer immunology has successfully implemented approaches that redirect or hyperactivate the immune system to recognize and kill cancer cells. Such therapeutics include immunological checkpoint blockade, the immunization of cancer vaccines, neutralization of immunosuppressive cells, treatment with oncolytic viruses, or the use of synthetic biology with bi-specific antibodies or cells with "chimeric antigen receptors" (CAR-T cells).[47,113,114] This segment has highlighted numerous fundamental principles and mechanisms through which inflammation contributes to the progression of cancer. Moving forward, we will now explore the parallels of how the knowledge and complexities governing molecular and cellular pathways of tumor-promoting inflammation are applicable to atherosclerosis.

1.4 Inflammation and Atherosclerosis

The standard of care for atherosclerosis treatment is the prevention of cardiovascular events by targeting modifiable risk factors (e.g., diet, exercise, smoking cessation), the use of intensive lowering of low-density lipoprotein (LDL), and the re-establishment of arterial flow by percutaneous or surgical interventions. However, the cardiology community has recognized that the therapeutic effectiveness of these approaches have its limitations.[115]

Historically, high plasma concentrations of cholesterol, particularly those of LDL, were considered a primary risk factor of atherosclerosis. Consequently, many believed that the process of atherogenesis was primarily a result of lipid accumulation within the artery wall. Over time, it became more widely understood that atherosclerosis is not solely a disease of cholesterol deposits, but that lesions of atherosclerosis represent a series of highly specific cellular and molecular responses with inflammatory origins.

The functional role of inflammation in atherosclerosis was proposed as early as in the 1970s by Ross & Glomset[116] who argued that lesions of atherosclerosis arose as a result of some form of

injury to arterial endothelium. Stating that when unabated and excessive, the lesion would advance to a more complicated lesion. The endothelium dysfunction is thus the result from the injury which lead to compensatory properties of the endothelium.[117] As referred to in previous sections on the immune system's self-regulatory state (section 1.2.3), if the inflammatory response does not effectively eliminate the offending pathogens, it can continue indefinitely. In this process, it will provoke the migration and proliferation of smooth-muscle cells forming an intermediate lesion. If the inflammatory response is sustained, the number of macrophages and lymphocytes will then multiply at the lesion site.

In 1994, Attilio Maseri conducted a study to assess the bivariate associations of established markers of inflammation such as CRP and serum amyloid A protein with outcomes following hospitalization in a small group of patients with CAD, reporting that high CRP and serum amyloid A protein levels were strongly correlated with worse outcomes in patients with unstable angina.[118] In a follow-up study by the same group, patients with unstable angina were shown to have higher levels of IL-6 compared to patients with stable angina, which was significantly correlated with CRP.[119]

The current understanding of atherogenesis involves the complex interplay between cellular immunity, lipid retention, and the interaction of immune and non-immune cells.[120] Similar to tumor development, atherosclerosis development is determined by the balance between inflammatory and anti-inflammatory processes. Atherosclerosis mostly initiates by entrapment of LDL in the intimal layer of medium and large arteries. As atherosclerotic plaques begin to form, the endothelial injury provides suitable circumstances for circulating monocytes to access the subendothelial layer. As the plaque grows, other types of the immune cells, such as mast cells and T lymphocytes are gradually introduced. Each of those cells release various cytokines, causing the recruitment of other immune cells to the atherosclerotic plaque sites.[121]

The macrophages take up lipoproteins present in the plaque and become lipid-laden foam cells, setting the foundation for the development of the plaque necrotic core. Antigen-presenting cells, such as macrophages, dendritic cells, and B cells, present lipid antigens to NKT cells and peptide antigens to T cells at the immune synapse. This interaction leads to the engagement of adaptive

T and B cell responses. Antigen presentation occurs in the plaque and in the lymph nodes.[122] The combination of these processes contribute to endothelial dysfunction, which further perpetuates inflammation as the recruitment of monocytes and macrophages continues, which then increases the uptake of lipoproteins, further increasing the plaque lipid burden, and finally leading to fibroblast migration which contributes to the fibrous cap formation. Almost all components of the immune system are involved in atherosclerosis development, with sometimes contradictory consequences at different stages of disease development and progression. This brings us to a discussion of the functions of the most impactful immune cells in the formation and development of atherosclerosis.

1.4.1 Innate and Adaptive Immune Cells in Atherosclerosis

1.4.1.1 Monocytes and macrophages

Circulating monocytes and macrophages are the first immune cells to appear in the atherosclerotic lesions.[123] In humans, monocytes can be classified into two main subtypes based on the expression of surface markers: classical monocytes (CD14+CD16-), the most abundant, and non-classical monocytes (CD14+CD16+). They are involved in various atherosclerotic processes including plaque formation, progression, and rupture.[124] During atherosclerosis, classical monocytes are recruited into plaque lesions following engagement with the chemokine receptors CCR2, CCR5, CX3CR1.[125,126] Non-classical monocytes are also associated with CAD progression, contributing to plaque reduction via their anti-inflammatory abilities.[127] Monocytes are present in the blood, bone marrow, and spleen during homeostasis. In atherosclerosis, a specific type of monocytes (Ly6C^{high}, lymphocyte antigen 6 complex) derived from hematopoietic stem and progenitor cells (HSPCs), are produced through extramedullary hematopoiesis in the spleen.[128] Ly6C^{high} monocytes can infiltrate the atherosclerotic lesions and exert inflammatory properties. Hyperlipidemia and low levels of high-density lipoprotein (HDL) accelerate HSPCs proliferation and monocytosis, worsening atherosclerosis progression.[129]

Macrophages play a balancing role in the orientation of plaque fate. They can affect growth and rupture by maintaining, or attenuating inflammation. In atherogenesis, monocytes rebuild the population of resident macrophages in the arterial intima at the early stages of atherosclerosis.[130]

In later stages of the disease, the proliferation of macrophages within the lesion site contribute to the accumulation of macrophages.[131]

Recent studies using single-cell analyses in humans[132] and mouse[133,134] have identified three distinct macrophage populations with specific inflammatory properties. These findings challenge the traditional M1/M2 macrophage polarization paradigm.[135] One of these macrophage populations express high levels of IL-1β[132,133], which has been implicated as an immune target in atherosclerosis. Another population of macrophages was found to be involved in antigen presentation and endocytosis.[133,136] A third group of macrophages were shown to have a role in lipid handling in the plaques.[133,136,137] Additionally, macrophages were found to play a role in plaque rupture and thrombosis through the production of MMPs and tissue factor[138], and the coordination of intraplaque efferocytosis.

1.4.1.2 Dendritic Cells

Another type of immune cells that play a crucial role in driving atherosclerotic plaque inflammation in both the innate and adaptive immune responses are dendritic cells. Plasmacytoid dendritic cells (cDC1s) are located in the blood and lymphoid tissues. In the presence of pathogens, these cells produce large quantities of interferon-α and interferon-β (IFN-α and IFN-β), both of which hold important pro-atherogenic functions.[139] Conventional dendritic cells (cDC2s) are found in lymphoid and non-lymphoid sites. cDC1s are involved in the cross-presentation of antigens and regulate cytotoxic immune responses, whereas cDC2s are involved in T cell priming.[140]

Dendritic cells hold important roles in lipid uptake and lipid metabolism and are considered crucial mediators in early accumulation of lipids in atherosclerotic lesions.[141] Plaque dendritic cell numbers have been shown to be positively correlated with plaque vulnerability.[142] Mouse models studies suggest that dendritic cells bear both pro-atherogenic and anti-atherogenic functions.[122] Specifically, dendritic cells maturation can have two effects on atherosclerosis: they secrete pro-inflammatory cytokines and also simultaneously activate T cells.[121,142] $CD103^+$ dendritic cells induce the development of Tregs, which suppress the activation of endothelial cells[143] and macrophages via secretion of anti-inflammatory cytokines such as TGF-β and IL-10.[144,145], inciting

athero-protection. On the other hand, dendritic cells expressing CCL17 (a protein that functions as a chemokine which binds to CCR4, expressed on the surface of various immune cells, including T cells and dendritic cells) have a pro-atherogenic role in mice.[146] $CCL17^{+}$ dendritic cells are known suppressors of Treg development and encourage apoptosis of Tregs.[146]

1.4.1.3 Neutrophils

Neutrophils are involved in all stages of atherosclerosis. Activated platelets release specific chemokines that trigger the recruitment of neutrophils to the atherosclerotic lesions.[147] Lipid-rich plaques that have a neovascular base are prone to the effect of injury, leaving them vulnerable for neutrophil infiltration.[148] Monocytes are attracted to neutrophils via chemotactic molecules secreted by the neutrophils, and they can activate macrophages by releasing nuclear material in the form of neutrophil extracellular traps (NETs)[149] NETs extrusions contribute to different pathological conditions including atherosclerosis and thrombosis.[150] In mice, the depletion of neutrophils reduces atherosclerosis, whereas increasing neutrophil levels aggravate plaque formation.[151] NETs carry histone H4 that can bind to vascular smooth muscle cells (VSMCs) and cause their lysis, leading to destabilization of the plaque.[152] Moreover, NETs induce plaque erosion and platelet accumulation, which lead to thrombosis.[153] Therefore, neutrophils are generally pro-atherogenic. Ironically however, they also defer to reparative functions during a thrombotic event by promoting endothelial repair and angiogenesis.[147]

1.4.1.4 T Cells

T cells are one of the most important immune cell type for atherosclerosis initiation and progression.[154,155] In human carotid artery plaques, a mass cytometry analysis showed that T cells outnumber macrophages[136], as opposed to mice whose T cells number is lower. T cells in atherosclerotic plaques also show more activation- and exhaustion-related gene expression compared to peripheral blood T cells. The high concentration of PD-1 inhibitory molecule can result in inefficient T cell effector function and the dysregulation of the immune response within plaques.[132,136] The activation of T cells directly regulate effector functions in the arterial wall and may help B cells produce antibodies.

The most prevalent T cells in mouse atherosclerotic plaques are $CD4^+$. $CD4^+$ T cells are mostly polarized towards pro-inflammatory phenotype (Th1).[155]. However, their role in promoting or protecting against atherogenesis depends on the subset involved. Th1 cells have been shown to have pro-atherogenic roles, whereas Treg cells are suspected to have athero-protective roles through IL-10 and TGFβ secretion.[154] Using single-cell RNA sequencing of $CD4^+$ phenotype T cells in mouse models with atherosclerosis showed that a specific $CD4^+$ T cell population bore transcriptional similarities with apolipoprotein B (ApoB)-reactive $CD4^+$ T cells.[156] As atherosclerosis advances, ApoB-reactive $CD4^+$ T cells transition from a Treg cell into a pro-inflammatory phenotype, which might explain their contribution towards disease progression.

$CD8^+$ T cells in mice have been shown to drive plaque inflammation and apoptosis, supporting unstable plaque phenotypes and erosion.[157,158] In atherosclerosis, CD8+ T cells hold dual functions, one with pro-atherogenic effects mediated by the IFNγ production and macrophage activation, and two with athero-protective effects via B cell modulation.[154] Proportionally, $CD8^+$ T cells outnumber $CD4^+$ T cells in individuals with advanced atherosclerotic plaques[136,157], where higher $CD8^+$ T cell numbers were linked to CAD.[159,160]

Invariant NKT cells are a distinct subset of T cells that express unique invariant T cell receptors and natural killer cell surface molecules. Invariant NKT cells are a CD1d-restricted T cell population that can respond to lipid antigenic stimulation rapidly by secreting a wide variety of cytokines.[161] In mice, the invariant NKT cells show pro-atherogenic properties owing to their production of pro-inflammatory cytokines.[162] Human studies show that plaques with high vulnerability have higher levels of invariant NKT cells compared to stable plaques.[163]

1.4.1.5 B Cells

B cells present in atherosclerotic lesions play diverse roles in the development of atherosclerosis. Some B cell subpopulations can produce cytokines and antibodies against plaque antigens, which can help suppress inflammation and exhibit anti-atherosclerotic activity. B cells are classified as B1 cells and B2 cells. The former is mainly produced in the fetal liver, whereas the latter originates from the bone marrow. While the secretion of antibodies by B cells show athero-protective effects, B1 and B2 cells are not similar regarding their influence on atherosclerosis development.

B1 cells in mice have been considered protective against atherosclerosis due to the IgM antibodies they produce, which can hinder the absorption of oxidized-LDL by macrophages located in lesions.[164,165] On the other hand, B2 cells have demonstrated pro-atherogenic traits through antibody responses formed via germinal center B cell reactions that enhance adaptive immunity.[166] Interestingly, when fed a high-cholesterol diet, B2 cells with athero-protective functions arise in secondary lymphoid organs, such as the lymph node or the spleen.[167] B cells are also involved in atherosclerosis in the form of cellular immunity. In a complex involvement, T and B cells interact to regulate T cell activation via antigen activation, cytokine production and co-stimulation.

1.4.2 Rebalancing the Immune System in CVD

The immune system's primary goal is to maintain a balance between pro-inflammatory and anti-inflammatory processes. Failure to do so can lead to chronic inflammation, as seen in both atherosclerosis and cancer. In atherosclerosis, a critical mechanism for resolving inflammation is efferocytosis, which restores the balance of pro-inflammatory lipid mediators by promoting the production of specialized pro-resolving mediators.

1.4.2.1 Efferocytosis

Efferocytosis, the process by which phagocytic cells remove apoptotic cells is an essential process in the maintenance of tissue homeostasis, often referred to as the 'burying of dead cells', in cell biology. Defective efferocytosis initiates various chronic inflammatory diseases, including atherosclerosis and cancer.[168] Atherosclerotic plaque inflammation and necrotic core formation are attributed to a defective efferocytosis mechanisms and insufficient immunomodulation. This leads to the accumulation of apoptotic cells and a subsequent increase in necrosis, which can result in plaque destabilization.[169] Efferocytosis is mediated through phagocytic receptors such as tyrosine-protein kinase MER (MerTK) or LDL-receptor (Ldlr) related protein 1 (LRP1), and apoptotic cell ligands.[170–172] In atherosclerosis, defective efferocytosis can be attributed to the downregulation of such regulators. Indeed, the downregulation of efferocytosis receptors[170–172] and dysregulated expression of "eat me" signals[173,174] lay the foundation of an impaired efferocytosis. In experimental studies, mice with high MerTK levels have increased levels of

efferocytosis and less necrotic core formation compared to Ldlr$^{-/-}$ mice.[170,171] Similarly, the loss of LRP1 in macrophages or hematopoietic stem cells in athero-sensitive mice was shown to lead to larger lesion sites and necrotic core size.[175] Neutralizing antibodies that block CD47, improve efferocytosis and reduce atherosclerosis in Apoe$^{-/-}$ mice by masking “don’t eat me” signals.[174] The demonstrated efficacy in pre-clinical trials prompted investigators to assess the potential therapeutic benefit of CD47 targets (Hu5F9-G4 and TTI-621) as cancer therapies in randomized clinical trials.[176,177] The concomitant inhibition of CD47 and TNF appears to provide a combined benefit in eliminating apoptotic cells in mice.[174] This suggests that there may be a rationale to consider the combination of anti-inflammatory and pro-efferocytic therapies for the management of advanced atherosclerosis. Furthermore, it is worth noting that anti-TNF therapy has already been shown to reduce the risk of cardiovascular outcomes in patients with prevalent rheumatoid arthritis, further supporting the potential benefit of this combination therapy.[178] In this regard, a group of researchers pondered on the outsized benefit of statins that was solely attributed to their influence on LDL-c. These investigators have recently confirmed a novel pleiotropic effect of statins, which is that statins have lipid-independent anti-atherosclerotic effects mediated by increasing the rate of macrophage efferocytosis through suppression of the expression of the "don’t-eat-me" molecule CD47 on plaque apoptotic cells.[179] It had already been established that statins may possess anti-inflammatory properties, in addition to their lipid-lowering effects. However, the exact mechanism by which they exert these anti-inflammatory effects had remained unclear until then. Using RNA-sequencing to investigate gene expression profiles in macrophages after inhibition of CD47-SIRPα axis, responsible in reactivating efferocytosis, they found that inhibition of the axis led to gene expression changes comparable to those expected to occur after treatment with statins. In mouse models, the combined use of CD47-SIRPα blockade and atorvastatin induced additive anti-atherosclerotic effects, such as a reduction in lesion size and the necrotic core area compared to mice given atorvastatin alone. Both in vitro and in vivo, the combination regimen also demonstrated additive effects on the rate of efferocytosis. In carotid endarterectomy samples obtained from the Munich Vascular Biobank, the authors showed that patients receiving statins had lower CD47 expression compared to propensity-matched patients not receiving statins.[179]

1.4.3 Anti-Inflammatory Therapy in the Prevention of CVD

Historically, the idea that inflammation was responsible for atherosclerosis was not widely accepted. To challenge the conventional belief in the early 1990s, which held that atherosclerosis was solely caused by the build-up of fatty deposits known as plaques in blood vessels, leading to heart attacks and strokes due to excess cholesterol, researchers had to present concrete clinical data in a stepwise fashion.

A cornerstone study was a report led by Ridker et al. which demonstrated that middle-aged men with high levels of high-sensitivity CRP (hs-CRP), a marker for systemic inflammation, were at higher risk of cardiovascular events, regardless of their cholesterol level (n=1,086).[180] Specifically, those within the highest quartile group of hs-CRP had nearly 3.0-fold higher risk of MI (P<0.001) and 2.0-fold higher risk of ischemic stroke (P=0.02) than their lowest quartile counterparts. Furthermore, the study showed that the influence of hs-CRP on the events of interest were maintained over time, and remained independent of traditional cardiovascular risk factors, such as obesity, hypertension, smoking habits, type 2 diabetes, and lipid factors. As other studies were then able to replicate similar risk associations between hs-CRP levels and incident CVD for men and women[181,182], there began a growing collective enthusiasm for the role of inflammation in CVD.

As researchers accepted the notion that inflammation did play an important role in atherosclerosis development, the next step involved proving that there would be a potential therapeutic benefit in preventing atherosclerosis by inhibiting inflammation. For this objective, researchers utilized statins as a therapeutic strategy. First, studies had to demonstrate that in addition to its cholesterol-lowering abilities, statins were also efficient at reducing inflammation. Within the Cholesterol and Recurrent Events Study (CARE) study, a randomized, placebo-controlled trial which compared pravastatin to placebo among 4,159 individuals with previous MI and elevated levels of cholesterol (LDL-C 115-175 mg/dL)[183], 472 individuals chosen at random demonstrated reduced mean in hs-CRP levels (–21.6%) when treated with pravastatin compared to placebo over a period of five years[184]. For patients without any history of CVD, among 1,702 participants randomized to pravastatin at 40 mg daily experienced up to 17% reduction of hs-CRP levels after 24 weeks compared to none in the control group treated with placebo.[184] What was

particularly noteworthy of that trial was the observation that no significant association was found between baseline hs-CRP levels and baseline LDL-c levels, as well as end-of-study hs-CRP and end-of-study LDL-c levels. Univariable analyses appeared to suggest that only treatment with pravastatin was associated with change in hs-CRP levels over the 24-week period (*P*<0.001).

More potent statins also showed a dose-response effect on hs-CRP levels. The JUPITER study randomly assigned 17,802 participants without baseline CVD and with low levels of LDL-c at study entry (criteria set at <130 mg/dL), yet high levels of hs-CRP (≥2 mg/L), to either rosuvastatin at 20 mg or placebo.[85] Median LDL-c levels at the 12-, 24-, 36- and 48-month mark were 55, 54, 53, and 55 mg/dL for those treated with rosuvastatin, respectively. Meanwhile, their median hs-CRP levels for the same time points were 2.2, 2.2, 2.0 and 1.8 mg/L, respectively, compared to 3.5, 3.5, 3.5 and 3.3 mg/L for the placebo control population, respectively (all between-group comparisons were *P*<0.001). Overall, rosuvastatin led to a median 37% decrease in hs-CRP compared to placebo (*P*<0.001). Along the same lines, another study randomized 3,745 individuals with acute coronary syndrome to high-dose statin (atorvastatin 80 mg) or low-dose statin (pravastatin 40 mg) to assess the risk of recurrent MI or coronary-related death (PROVE-IT TIMI 22).[185] After 30 days of treatment, 58% of individuals treated with high-dose statins saw their hs-CRP levels dip below 2.0 mg/L vs. 46% of individuals treated with low-dose statins.[186]

The lowering of hs-CRP levels was also concordant with lower cardiovascular event rates. In the JUPITER trial, rosuvastatin was shown to effectively reduce the risk of MI (HR: 0.46, 95% CI: 0.30 to 0.70, *P*<0.001), stroke (HR: 0.52, 95% CI: 0.34 to 0.79, *P*=0.002), revascularization or unstable angina (HR: 0.53, 95% CI: 0.40 to 0.70, *P*<0.001), and the composite endpoint of MI, stroke, or death from cardiovascular causes (HR: 0.53, 95% CI: 0.40 to 0.69, *P*=0.02).[85] In the PROVE-IT TIMI 22 study, individuals who reached hs-CRP levels below 2.0 mg/L had fewer event rates irrespective of LDL-c lowering.[187] Furthermore, the event rates were comparable between groups who reached LDL-c levels <70 mg/dL or hs-CRP <2 mg/L.

The main criticism of these data was whether the reduction of inflammation, in the absence of cholesterol lowering, would still result in fewer vascular-related events. There were concerns that the observed reduction in inflammation may have merely been a secondary outcome associated

with the LDL reduction following treatment with statins. To formally test the hypothesis that anti-inflammatory therapies may effectively reduce cardiovascular-related outcomes, regardless of LDL lowering, randomized trials targeting IL-1 through 6 in the CRP signaling pathway without any consequence on atherogenic lipids would have to be conducted. Several clinical trials have been designed in the last five years to directly target inflammatory pathways in the context of CVD (**Figure 1.5**).

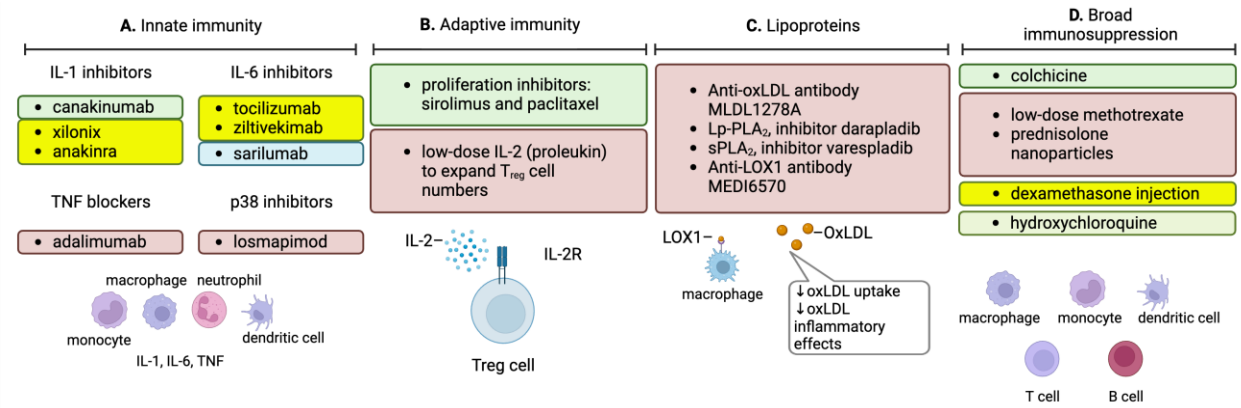


Figure 1.5. Visual Diagram Depicting Trials that Target the Immune System in Atherosclerosis.

> *Adapted from* Engelen et al.[188]. *Therapies for the treatment of atherosclerosis that showed a clinical benefit (green), no benefit (red), or potential benefit (yellow) that have been tested in trials or are ongoing (blue) are illustrated. Therapies that target innate immunity include IL-1 inhibitors, IL-6 inhibitors, TNF blockers and p38 inhibitors (see **panel A**). Therapies that target adaptive immunity include local proliferation inhibitors in drug-eluting stents and low-dose IL-2 targeting Treg cells (see **panel B**). Therapies that target lipoproteins to reduce inflammation include antibodies against oxidized LDL (oxLDL), lipoprotein-associated phospholipase A2 (Lp-PLA$_2$), secretory phospholipase A2 (sPLA$_2$) and lectin-like oxidized LDL receptor 1 (LOX1) (see **panel C**). Therapies with broad immunosuppressive effects include colchicine, low-dose methotrexate, glucocorticoids, and hydroxychloroquine (see **panel D**).[6]*

To date, two molecules have shown clinical efficacy in improving cardiovascular outcomes in patients with CVD (2nd line therapy): canakinumab and colchicine.

[6] Reprinted from Engelen SE, Robinson AJB, Zurke YX, Monaco C. Therapeutic strategies targeting inflammation and immunity in atherosclerosis: how to proceed? *Nat Rev Cardiol*. 2022;19(8):522-542. Reprinted with permission from Springer Nature (license #5518291072021).

1.4.3.1 Canakinumab

Canakinumab is a fully human IL-1β neutralizing monoclonal antibody. For CVD prevention, canakinumab was attractive because atherosclerosis risk factors upregulate IL-1β via the NLRP3 inflammasome. In a pilot study, canakinumab showed a significant decrease in hs-CRP, fibrinogen, and IL-6 with no effect on LDL-c or other established lipid measures.[189] The CANTOS trial was a double-blind, randomized, placebo-controlled study which investigated the effects of canakinumab in patients who recently experienced a MI.[190] To focus on patients with a "residual inflammatory risk", as opposed to a "residual cholesterol risk", specific inclusion criteria included baseline hs-CRP of >2 mg/L. Subcutaneous injection of canakinumab was performed at doses of 50 mg, 150 mg, or 300 mg every 90 days (median follow-up 3.7 years). Investigators focused on a composite primary endpoint that consisted of non-fatal MI, non-fatal stroke, or cardiovascular death. Overall, average age was 61 years, with 75% who were male. Baseline diabetes was observed in 40% of patients, and 81% of patients had undergone previous percutaneous or surgical coronary revascularization. Of note, at least 93% had ongoing LDL lowering therapy, with a mean LDL-c of 92 mg/dL. Median hs-CRP at baseline was 4.2 mg/L.

Overall, patients who received 150 mg of canakinumab experienced lower rates of recurrent cardiovascular events compared to placebo, independent of lipid-lowering (HR: 0.85, 95% CI: 0.74 to 0.98, *P*=0.021). As expected, there was a dose-dependent reduction of approximately 26–41% in median hs-CRP after 48 months. Canakinumab had no effect on LDL-c or HDL cholesterol (HDL-c). In pre-specified analysis, the efficacy of canakinumab depended considerably on the magnitude of inflammation reduction for trial participants.[191] Specifically, those with hs-CRP levels <2mg/L after the 1st dose experienced –25% reduction in major cardiovascular events (*P*<0.001) compared to only 5% reduction in those with on-treatment hs-CRP levels ≥2 mg/L (non-significant). Patients who achieved hs-CRP levels of <2 mg/L also experienced lower risks of cardiovascular mortality (HR: 0.69, 95% CI: 0.56 to 0.85, *P*<0.001) and lower risks of overall mortality (HR: 0.69, 95% CI: 0.58 to 0.81, *P*<0.001), whereas patients who did not sustain an hs-CRP level less than the median did not experience such benefits from treatment. The differing benefits persisted despite adjustment for clinical confounders, including baseline hs-CRP and LDL-c, age, sex, smoking status, hypertension, diabetes, and body mass index.

Median IL-6 also reduced by 19–38% after 12 months. A post-hoc analysis showed that the effects of canakinumab extended beyond hs-CRP where 4,833 stable atherosclerosis patients had their IL-6 measures taken before and after treatment.[192] For those with on-treatment IL-6 below the median level of 1.65 ng/L, canakinumab resulted in lower risks of major adverse cardiovascular events (HR: 0.68, 95% CI: 0.56 to 0.82, *P*<0.001), lower risks of cardiovascular mortality (HR: 0.48, 95% CI: 0.34 to 0.68, *P*<0.001), and lower risks of overall mortality (HR: 0.52, 95% CI: 0.40 to 0.68, *P*<0.001). In contrast, patients with IL-6 levels above the median did not experience any treatment-related benefit with respect to any of the pre-determined endpoints. These additional analyses provide evidence that modulation of the IL-6 signaling pathway has potential to reduce cardiovascular event rates, independent of lipid lowering.

Unfortunately, due to higher sepsis-related deaths in the intervention arm relative to placebo (incidence rate 0.31 vs. 0.18 person-years, *P*=0.02), canakinumab was not approved for secondary prevention of CVD. However, the CANTOS trial did serve as a first-time proof-of-concept that therapeutics targeting the immune system can result in beneficial cardiovascular outcomes in patients.

1.4.3.2 Colchicine

Previously widely used for the treatment of gout and pericarditis, colchicine decreases inflammation by inhibiting cytoskeletal microtubule formation.[193,194] In contrast to canakinumab, colchicine has demonstrated a wide range of cellular effects, including the reduction of monocytes and neutrophil motility and inhibition of the inflammasome assembly in vitro[195] and alteration of leukocyte responsiveness[196–198]. As stated in another section (section 1.4.1.3), lipid-rich plaques with a neovascular base are susceptible to the effect of cardiac injury, which may leave them vulnerable to neutrophil infiltration.[148] As neutrophils entering the interstitial space become activated upon exposure to the plaque contents, it incites an aggressive inflammation response that drives plaque instability, thereby increasing the risk of plaque enlargement and rupture, aggravating the risk of clinical events.

The sequence of these events raises the possibility of inhibiting neutrophil function to lower the risk of plaque instability and cardiac disease progression. Given colchicine's anti-tubulin effect

inhibiting neutrophil, investigators first conducted a prospective, randomized, observer-blinded endpoint trial to assess the effect of colchicine (0.5 mg/day) combined with then-standard of care such as aspirin and high-dose statins on the risk of cardiovascular events in patients with stable coronary disease (n=532, LoDoCo).[199] The primary endpoint of interest was a composite of incident acute coronary syndrome, out-of-hospital cardiac arrest, or non-cardioembolic ischemic stroke. The investigators found that colchicine was associated with a significantly lower risk of the primary endpoint (HR: 0.33, 95% CI: 0.18 to 0.59, *P*<0.001) compared to those who did not get colchicine.

In an effort to circumvent the small sample size of the LoDoCo trial[199] and in light of the negative trial of an alternative anti-inflammatory agent that was published at the time[200] (to be discussed below), investigators conceptualized the Colchicine Cardiovascular Outcomes Trial (COLCOT). COLCOT was a randomized, double-blind, placebo-controlled study that enrolled participants who had experienced a MI <30 days prior and managed with guideline recommended standard of care.[201] Patients were assigned to either colchicine (0.5 mg/day) or placebo (n=4,745). The primary endpoint of interest was a composite of death from cardiovascular causes, resuscitated cardiac arrest, MI, stroke, or urgent hospitalization for angina leading to coronary revascularization. Each endpoint was then assessed individually as secondary endpoints of interest. Overall, the primary endpoint was observed in 5.5% of patients treated with colchicine vs. 7.1% of patients given placebo (HR: 0.77, 95% CI: 0.61 to 0.96, log-rank *P*=0.02). A striking protective effect was also observed for those treated with colchicine vs. placebo with respect to stroke (HR: 0.26, 95% CI: 0.10 to 0.70).

In COLCOT, data on the effects of colchicine on inflammatory markers were not readily obtained for all enrolled patients. Specifically, hs-CRP levels were measured only in a small subset of patients (n=207, colchicine was 99 and placebo was 108) at randomization and 6 months thereafter. In those patients, the adjusted geometric mean % changes in hs-CRP levels at the 6-month mark was –70% (95% CI: –74.6 to –64.5) for colchicine and –66.6% (95% CI: –71.5 to –60.8, not significant). While the clinical and patient characteristics of this subgroup was not very different from the overall population, the lack of a significant hs-CRP change was likely due to small sample size.

Encouraged by the cumulating evidence of the effective anti-inflammatory effects of colchicine on the reduction of cardiovascular events in patients with MI, investigators from the LoDoCo2 trial sought to assess its use in patients with chronic coronary disease. In a second randomized, placebo-controlled, double-blind study, patients with chronic CAD were assigned to 0.5 mg of colchicine per day or placebo (n=5,522).[202] Similarly, the primary endpoint was a composite of cardiovascular death, spontaneous (nonprocedural) MI, ischemic stroke, or ischemia-driven coronary revascularization. Patients who were given colchicine had significantly fewer primary events compared to placebo (HR: 0.69, 95% CI: 0.57 to 0.83, *P*<0.001).

There is now convincing data that hs-CRP holds prognostic value in the prediction of cardiovascular events, possibly even more so than LDL-c. In a meta-analysis of three randomized-controlled trials (n=31,245), the highest hs-CRP quartile was shown to be significantly associated with a larger risk of major acute cardiovascular events compared to the lowest hs-CRP quartile (HR: 1.31, 95% CI: 1.20 to 1.43, *P*<0.001). In contrast, the influence of residual cholesterol profile was non-statistically significant for major acute cardiovascular events (highest LDL-c vs. lower LDL-c HR: 1.07, 95% CI: 0.98 to 1.17, *P*=0.11).[203]

Altogether, the evidence generated from the CANTOS and the COLCOT/LoDoCo/2 trials provided actionable, concrete indication that anti-inflammatory therapies have the ability to effectively reduce cardiovascular events in patients with stable CVD. However, the immunomodulatory therapeutics to explore remain vast, and our understanding of the complexities of the functioning of the immune system and its effect on cardiovascular health is still growing.

1.4.3.3 Challenges and Potential Targets

Currently, several potential inflammatory targets have been, or are being tested in clinical trial settings. As strategies continue to expand, not all anti-inflammatory molecules have resulted in an efficacious reduction of inflammation and/or cardiovascular events. Noteworthy examples are methotrexate and p38 inhibitors, both of which failed to significantly reduce cardiovascular events in patients with CVD.[200,204] p38 is an intracellular kinase that is activated by specific triggers, such as oxLDL and hypertension, and is involved in the stabilization of mRNA encoding several inflammatory mediators.[204,205] Losmapimod was the first p38 blocker study, which

included patients with stable atherosclerosis. Losmapimod was not significantly associated with less vascular inflammation in the index vessel, as defined with lower uptake of fluorodeoxyglucose PET-CT imaging, compared with placebo, but losmapimod did result in lower inflammation in regions that showed the highest inflammation.[206] However, losmapimod failed to have any clinical benefit in patients with acute MI in a subsequent trial.[207,208]

In another negative study, CIRT, methotrexate was tested in a randomized, double-blind trial against placebo in patients with a history of MI or multivessel coronary disease, with either type 2 diabetes or metabolic syndrome. The trial was stopped after 2.3 years for futility, where no significant different with respect to the primary endpoint (composite of non-fatal MI, non-fatal stroke, or cardiovascular death) was observed between methotrexate and placebo (HR: 0.96, 95% CI: 0.79 to 1.16).[200]

A potential explanation from trials that did not show efficacy of the anti-inflammatory drug being tested is that they did not focus on a stringent selection criterion that required patients to have high inflammation at study entry (e.g., CANTOS required baseline hs-CRP of ≥2 mg/L and COLCOT required acute MI <30 days). However, LoDoCo2 also tested colchicine in a relatively unselected group of patients and did end up confirming a benefit in cardiovascular outcomes for colchicine.[202] An alternative explanation is that mechanism-specific inhibition of inflammation in CVD could be important, and so far, the correct target appears to be the IL-1β-IL-6-CRP pathway of innate immunity. The progressive understanding of the pathogenesis of chronic inflammation on atherosclerosis has highlighted a multitude of potential molecular therapeutic targets that have undergone several lines of testing and are on the verge to being translated into potential clinical therapies for patients (**Figures 1.5 & 1.6**).

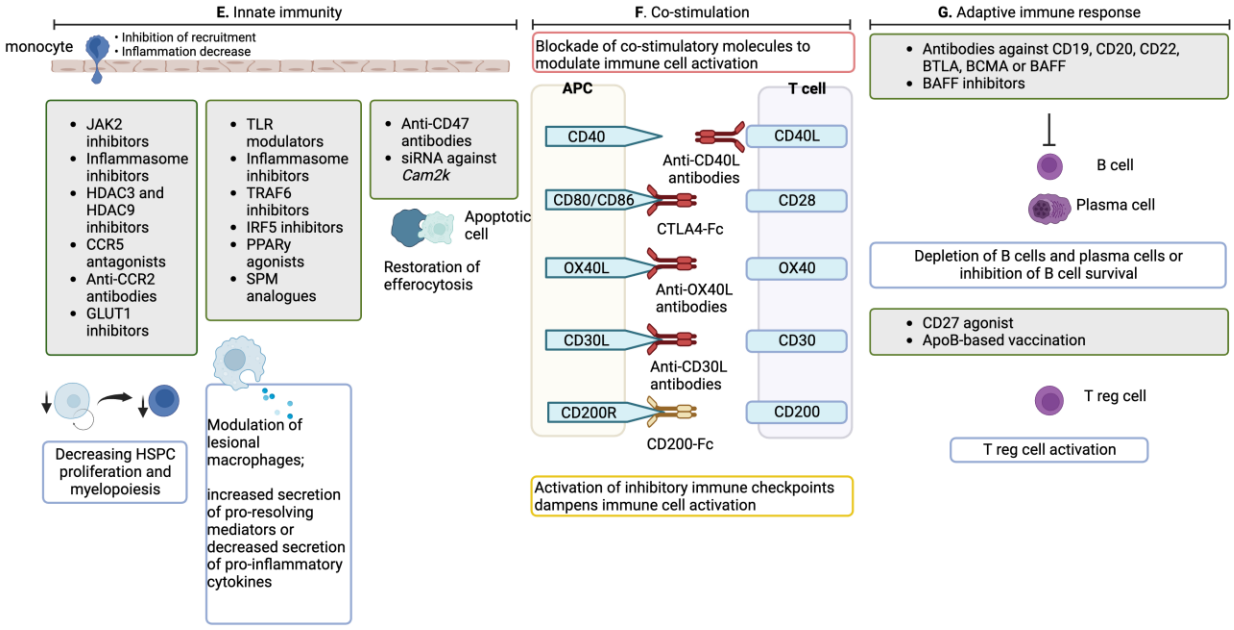


Figure 1.6 Visual Depiction of Pre-Clinical Trials That Target the Immune System in Atherosclerosis.

Adapted from Engelen et al.[188] *Therapeutics in development targeting the innate immunity are shown in **panel E**, co-stimulation pathways in **panel F**, and B cell and T cell regulation in **panel G**.[7]*

The acknowledgement that chronic inflammation holds a significant role in atherosclerosis and the understanding of its effects has undoubtedly helped translate various inflammatory targets into human cardiovascular therapy. Throughout the process, there is a better acceptance that the pathogenesis of CVD is multifactorial, and different disease settings have distinct immune signatures. Similarly with cancer, the identification of specific disease settings can materialize into pinpointing which specific therapeutic strategy may be more suitable. Although the community is better equipped in its comprehension of the role of inflammation in atherosclerosis, the nuanced balance between pro-inflammatory and anti-inflammatory responses required for homeostasis remains intangible.

[7] Reprinted from Engelen SE, Robinson AJB, Zurke YX, Monaco C. Therapeutic strategies targeting inflammation and immunity in atherosclerosis: how to proceed? *Nat Rev Cardiol*. 2022;19(8):522-542. Reprinted with permission from Springer Nature (license #5518291072021).

1.5 – Inflammation: A Shared Patho-Physiological Mechanism Between Cancer and CVD

1.5.1 Modulation of Inflammatory Pathways Can Alter the Risk of Either Cancer or CVD

As discussed in previous sections, immune dysregulation, characterized by elevated levels of pro-inflammatory markers detected in the blood or tissues even after the resolution of a foreign pathogen, directly contributes to both cancer and CVD. Over the past two decades, both fields have increasingly focused on testing and employing immune-modulatory therapies in pre-clinical and clinical trials for the effective management of these diseases. In both settings, the challenge ahead lies in optimizing modulators of disease-specific inflammation considering the plethora of inflammatory processes while simultaneously sparing the host defense. In this context, therapeutics in the treatment of CVD predominantly strive to attenuate inflammation, whereas novel cancer treatments attempt to stimulate the immune system, which can either re-activate or sustain inflammation.

In a simplified perspective, a targeted interruption of a specific inflammatory pathway can have simultaneous benefits in reducing both atherosclerotic events and cancer-related events. For example, in the CANTOS study, canakinumab administration was associated with a notable reduction in the incidence of fatal and non-fatal lung cancer compared to placebo.[209] Notably, this effect was most pronounced in patients who were in the higher-dose group, demonstrating a relative risk reduction of –67% for total lung cancer and –77% for fatal lung cancer, which may imply a potential benefit for more aggressive disease phenotype. This unexpected finding has motivated a number of subsequent clinical trials (CANOPY study program) investigating the safety and efficacy of canakinumab on overall survival in metastatic non-small cell lung cancer.[210] So far, however, two trials (CANOPY-1 and CANOPY-2) have reported negative results, where canakinumab failed to significantly improve overall survival and progression-free survival in patients with non-small cell lung cancer.[211,212]

On the other hand, anti-cancer treatments with immune checkpoint inhibitors (e.g., CTLA-4, PD-1/PD-L1) modulate the immune system's response so that healthy cells in the body are not destroyed. The checkpoints are activated when proteins on the surface of immune T cells recognize and bind to partner proteins on other tumor cells. Under normal circumstances, when

checkpoint and partner proteins bind together, it sends a signal to T cells to disengage, prompting the immune system to allow cancer progression and development. Immune checkpoint inhibitors block the proteins from binding with such partner proteins, which bypasses the "off" signaling, allowing the T cells to effectively strike out cancer cells. Immune checkpoint inhibitors have now been widely shown to effectively improve cancer prognosis by leveraging patients' own immune system across many cancer types. However, because immune checkpoint inhibitors can induce widespread inflammation, they have also paradoxically been associated with an increased risk of CVD-related events.[213–215] Using a retrospective database of individual case safety reports provided by the WHO, investigators were able to correlate the increased use of immune checkpoint blockade with severe, fatal cases of myocarditis.[213] The mechanism underlying the increased risk of myocarditis following use of immunotherapy in patients diagnosed with advanced cancer identified cardiac-specific protein α-myosin as the cognate antigen source for major histocompatibility complex class I restricted T cell receptors derived from mice with myocarditis.[216,217] In humans with immune checkpoint inhibitors-induced myocarditis, peripheral blood T cells were expanded by α-myosin peptides. Myocarditis samples in mice also indicated a high degree of clonal T cell receptors (63%) compared to healthy mice controls. They also found that non-specific immunosuppressive modulations were insufficient in attenuating myocarditis. However, the depletion of $CD8^+$ cells, but not $CD4^+$ cells rescued mice who would have succumbed to myocarditis.

While the finding that canakinumab use lowered lung cancer incidence may have been unexpected, the proof-of-concept that inflammation, defined as a general upregulation of inflammatory mediators and active recruitment of cells with tumor-promoting properties (i.e., macrophages, neutrophils, etc.), stimulates enhanced tumor incidence, growth, and progression was already evident in a vast body of epidemiological studies. As mentioned, meta-analytic analyses of clinical trials on "non-specific" inhibition of inflammation with non-steroidal anti-inflammatory drugs (NSAIDs), such as aspirin[83,84], or statins[85,218] have been associated with reduced incidence and mortality for many cancers.

Naturally, simply lowering inflammation in patients with CVD does not consistently result in lower incidence of cancer altogether. The inhibition of adaptive immunity can also result in a higher risk

of cancer through a dysregulation in antitumor immunity.[219] For example, in the CIRT study, methotrexate was associated with a small, albeit higher increase in the incidence of skin cancer relative to those treated with placebo.[91,200] With regard to non-specific anti-inflammatory agents such as statins, the clinical evidence regarding its use and cancer incidence is also inconsistent. For example, post-hoc analyses from two previous trials that tested simvastatin to placebo showed that melanoma diagnoses were more common in the intervention arms.[220,221] The CARE investigators reported on a double-blind phase III trial that tested pravastatin vs. placebo in patients who experienced a MI.[183] They observed a total of 161 fatal or nonfatal primary cancers in the placebo group compared to 172 in the statin group. Similarly, the PRospective Study of Pravastatin in the Elderly At Risk (PROSPER) phase III trial assigned participants aged >70 years old to pravastatin or placebo.[222] After a follow-up of 3.2 years, they found that pravastatin-treated patients had a higher risk of incident cancer compared to placebo (HR: 1.25, 95% CI: 1.04 to 1.51, *P*=0.020).

Large population-based cohorts have also been used to examine the association between statins and cancer. Within a Danish cohort of patients diagnosed with cancer between 1995 and 2007, statin users were compared to non-statin users in a nested matched-analysis and it was reported that statin users were significantly less likely to die from cancer causes than non-statin users (HR: 0.85, 95% CI: 0.81 to 0.87, *P*<0.001).[223] In a Quebec population, Blais et al. performed a nested case-control analysis using administrative health data to assess the association between statin and cancer incidence.[224] Beneficiaries aged ≥65 years old without prevalent cancer at study entry who used statins were less likely to be diagnosed with any cancer during the study period compared to their bile acid-binding resin controls (relative risk [RR]: 0.72, 95% CI: 0.57 to 0.92).

Other retrospective studies also confirmed the protective effect of statin use on incident cancers. In a matched case-control study, Graaf et al. identified 3,129 cancer cases who were matched to 16,976 controls. Statin users were found to be less likely to be diagnosed with incident cancer compared to a control population who used other CVD-related medications (odds ratio [OR]: 0.80, 95% CI: 0.66 to 0.96).[225] Friis et al., relying on the Prescription Database of North Jutland County and the Danish Cancer Registry, compared overall and site-specific cancer incidence among

12,251 statin users (≥2 prescriptions) with cancer incidence among non-users and users of other lipid-lowering drugs (n=1,257).[226] The authors found that the risk of overall cancer among statin users were lower compared to nonusers (RR: 0.86, 95% CI: 0.78 to 0.95) and to users of other types of lipid-lowering agents (RR: 0.73, 95% CI: 0.55 to 0.98).

However, retrospective observational studies have not always been consistent. Kaye & Jick[227] found that within the General Practice Research Database, statin use was not associated with the incidence of 13 cancers (RR: 1.0, 95% CI: 0.9 to 1.2). In fact, the authors observed that statin utilization for greater than 5 years was associated with a significantly higher risk of colon (RR: 3.5, 95% CI: 1.1 to 10.9) and rectal cancers (RR: 4.2, 95% CI: 1.0 to 16.6). Subsequently, the observed increased risk between statin use and colon/rectal cancer was then reversed in a large population-based matched case- control study.[228] In that study, long-term use of statins (≥5 years) actually reduced the risk of colorectal cancer compared to non-users, even after adjusting for the use or non-use of aspirin or other nonsteroidal anti- inflammatory drugs, physical activity status, hypercholesterolemia, family history of colorectal cancer, ethnic group, and level of vegetable consumption (OR: 0.53, 95% CI: 0.38 to 0.74). Within a comprehensive meta-analysis of 19 studies totaling over 1.5 million patients, Bonovas et al.[229] then reported no evidence of an association between statin use and the risk of colorectal cancer in randomized-controlled trials (RR: 0.95, 95% CI: 0.80 to 1.13, *k*=6) or among cohort studies (RR: 0.96, 95% CI: 0.84 to 1.11, *k*=3). However, the authors did observe that statin use was associated with a slight reduction in the risk of colorectal cancer when focusing on case-control studies only (RR: 0.91, 95% CI: 0.87 to 0.96, *k*=9).

Since the CARE trial suggested that statin use may be related to an increased risk of breast cancer, several other epidemiological studies sought to confirm the association. Using a Saskatchewan population health services database[230], the authors identified women with at least one prescription of statin between the years 1989 to 1997. Following an age and sex-matched non-exposed group, the authors found the risks of breast cancer in those aged >55 years old to be higher for statin users (RR: 1.15, 95% CI: 0.97 to 1.37). An interaction with hormonal therapy was also observed, where for women aged >55 years old with hormone replacement therapy duration

of more than 6 years (i.e., ≥37 prescriptions), statin use was associated with an increased rate of breast cancer (RR: 2.04, 95% CI: 1.20 to 3.46).

Opposing these findings, Cauley et al.[231] found that women statin users were less likely to be diagnosed with breast cancer compared to women using other types of lipid-lowering drugs (RR: 0.28, 95% CI: 0.09 to 0.86), as well as non-users (RR: 0.37, 95% CI: 0.14 to 0.99) within a prospective cohort study comprising of community-based clinical centers. Finally, Coogan et al.[232] found that among more than 1,000 breast cancer cases and matched clinic controls, statin users were not significantly more likely to develop invasive breast cancer compared to non-statin users (RR: 1.20, 95% CI: 0.70 to 2.0). Similarly, the lack of a significant association was also observed in a population- based case-control study that comprised of female residents within Washington counties (OR: 0.90, 95% CI: 0.70 to 1.20).[233]

1.5.2 Clonal Hematopoiesis: A Novel Driver for Both Cancer and CVD

Recently, the emergence of clonal hematopoiesis (CH) as a common, age-related risk factor for both cancer and CVD has provided yet another novel link between the two disease phenotypes. Previously established as a risk factor in cancer genomics, somatic or acquired mutations in a variety of tissues have now also shown unequivocal associations with the development of CVD, some of which with a substantial impact on its severity. Previous genetic investigations of CVD have focused on inherited genetic mutations; although individuals acquire mutations throughout their lifetime. Improved technological advances have allowed the appreciation of acquired mutations beyond the scope of cancer. Such technologies comprise error-corrected deep sequencing and single-cell multi-omics sequencing, with increasingly higher throughput over time. The application of such innovative technologies onto large tissue- and blood sample repositories allowed the sensitive detection of mutations at low variant allele frequencies, which led to a greater understanding of somatic mutations in CVD.

CH can be defined as any clonal expansion of hematopoietic cells, with somatic mutations in otherwise normal white blood cells.[234,235] Such mutations occur throughout an individual's lifespan due to various biological mechanisms[236] and robustly occur as early as the first cellular division of embryogenesis.[237,238] Ageing hematopoietic stem cells have a decreased capacity to

prevent and repair DNA replication errors.[239] A common mechanism for somatic mutation formation is spontaneous deamination of 5-methylcytosine to thymine, primarily at methylated CpG dinucleotides.[240] If left unrepaired, the alteration is then passed to progeny cells over a lifetime. The recurrent observation of expanded single-nucleotide polymorphisms (SNPs) and short insertions or deletions in the blood in individuals without other cytological abnormalities is termed CH of indeterminate potential (CHIP), and is relatively common in the general population.[241] In whole-exome sequencing data of DNA samples in the peripheral-blood cells across 17,182 individuals, detectable somatic mutations were observed in nearly 20% of those aged >90 years old. Mutations commonly observed occur in genes *DNMT3A*, *TET2*, *ASXL1*, *JAK2*, and *TP53*. The analyses of haplotype imbalance can also represent the presence of large structural variants, indicative of large genomic gains, losses, or copy number neutral loss of heterozygosity, not necessarily indicative of cancer mutations, termed mosaic chromosomal alterations (mCAs).[242,243]

1.5.2.1 CH and Atherosclerosis

Research from over 50 years ago have shown an enrichment of monoclonality in atherosclerotic plaques.[244] However, the underlying drivers of clonality and mechanisms of mutational burden in peripheral blood lymphocytes in patients with atherosclerosis were unknown. Recently, a number of case-control studies has revealed associations between CHIP and CVD (**Figure 1.7**).

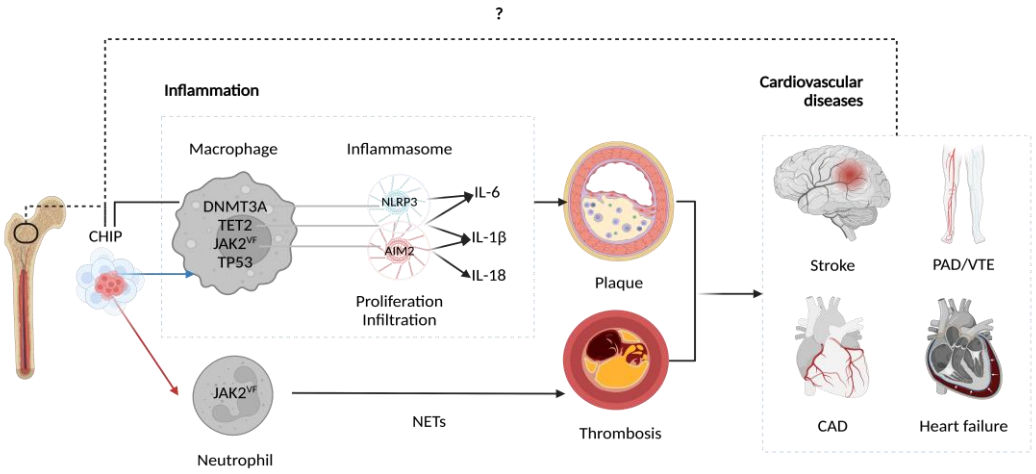


Figure 1.7 Proposed Mechanisms of Action Between CHIP and CVD.

Adapted from Nakao & Natarajan[245]*. Clonal hematopoiesis is the expansion of hematopoietic stem cells with somatic driver mutations that can result in clones of hematological cells in the body, including macrophages and neutrophils. CHIP refers to clonal hematopoiesis without hematological malignancy. Macrophages with CHIP-related mutations, such as DNMT3A, TET2, or JAK2V617F, can produce inflammatory cytokines that promote the fitness of these mutations through inflammasome activation. Preclinical studies have shown that JAK2V617F mutations promote the formation of NETs, which increase the risk of thrombosis, inflammation, and plaque formation. Carriers of CHIP are at higher risk of stroke, CAD, and PAD, VTE, and heart failure. CAD: coronary artery disease, CHIP: clonal hematopoiesis of indeterminate potential, CVD: cardiovascular disease, NETs: neutrophil extracellular traps, PAD: peripheral artery disease, VTE: venous thromboembolism[8]*

In a large retrospective cohort study, it was recently shown that patients who carried CHIP mutations had a 10-fold increased risk of developing hematological cancer. Unexpectedly, investigators found that although carriers of CHIP mutations have up to 40% higher risk of all-cause mortality, only a small fraction of excess deaths were explained by blood cancers. Rather, patients with CHIP mutations demonstrated significantly higher rates of incident ischemic stroke (HR: 2.6, 95% CI: 1.4 to 4.8) and coronary heart disease (HR: 2.0, 95% CI: 1.2 to 3.4).[241] Another important finding pertains to the association between CH and early-onset MI. In patients aged <50 years old, mutations in genes *TET2*, *JAK2*, and *ASXL1* were associated with higher odds of early-onset MI (OR: 4.0, 95% CI: 2.4 to 6.7, n=7,245).[246] In another study, it was shown that among

[8] Reprinted from Nakao T, Natarajan P. Clonal hematopoiesis, multi-omics and coronary artery disease. *Nat Cardiovasc Res*. 2022;1(11):965-967. Reprinted with permission from Sprinter Nature (license #5518291249797)

200 patients, those with CHIP mutations were more likely to experience heart failure progression and a poor prognosis from ischemic heart disease.[247]

CVD risks have also been shown to vary according to the most common CHIP mutations, where for example, carriers of *JAK2* mutations confer a 12-fold relative increased risk of incident CAD.[248] Differential effects between CHIP mutations and CVD may depend on the mutation's role. For example, *DNMT3A* and *TET2* are involved in DNA methylation, whereas ASXL1 is a chromatin regulator. Consequently, the loss of function mutations in *DNMT3A* and *TET2* have shown enhanced renewal capability in mice in vitro which led to the development of clonal populations.[249] Alternatively, *TP53* is involved in DNA damage response (DDR) and hematopoietic stem cells with loss of TP53 have a competitive advantage over their neighbors.[250] *PPM1D* is also part of the DDR pathway and interacts with p53-mediated apoptosis to gain proliferative advantage.[251]

While CHIP mutations occur in somatic cells, increasing evidence points towards the presence of a germline predilection for CHIP mutation development in hematopoietic stem cells.[252] In 2009, a genome-wide association study (GWAS) identified an allele at the *JAK2* locus that predisposed to the development of JAK2p. Val617Phe ($JAK2^{V617F}$)– derived myeloproliferative neoplasms was corroborated by several other subsequent studies, thereby substantiating the notion of inherited risk of clonality.[253–255] The first inherited variation linked to *JAK2* ^{V617F}–mutated myeloproliferative neoplasms was 46/1 or GGCC haplotype, a collection of SNPs that included the *JAK2* gene itself.[255,256] Interestingly, the haplotype may also drive JAK2 ^{V617F} clonal expansion. In a small study of 12 patients, the homozygosity for 46/1 preceding myeloproliferative neoplasm diagnosis was enriched and demonstrated high clonal growth rate.[257] That said, the genetic predisposition of *JAK2*-CH does not entirely overlap with JAK2-malignancy. This suggests that the inherited landscape of disease-negative and disease-positive CH settings can result in different phenotypic risk and magnitude.

A key enzyme in the maintenance of telomeres, *TERT*, is associated with the occurrence of CHIP and was identified in a whole-genome GWAS Icelandic cohort[235] and replicated in a larger cohort (100,000 genomes)[258]. In addition, several novel single nucleotide variants associated with the

occurrence of CHIP were also identified, confirming that germline variation is involved in the risk of developing CHIP mutations.[258] In particular, an intergenic region near *TET2* (rs144418061) was identified in individuals with African ancestry and shown to increase the risk of CHIP by 2.4-fold.

With respect to mCAs, a substantial fraction of risk for the mosaic loss of the Y chromosome (mLoY) appears to be genetically-based, with estimates of mLOY heritability ranging between 9% to 34%.[259–261] The first germline association with mLoY to be identified was rs2887399 near TCL1A, which encodes the protein T cell leukemia-lymphoma 1A.[262] Other studies conducted in population-based biobanks have shown significant germline associations with mCAs (**Table 1.2**).

Table 1.2 Cis-Acting and Trans-Acting Risk Variants For mCAs.

Inherited risk locus (chromosome)	**Risk variant frequency**	**Effect type**	**Associated mCAs**	**Reported OR (95% CI)**
FH(1q)	Rare	*Cis*	CN-LOH	28 (14-55)
NBN(8q)	Rare	*Cis*	CN-LOH	210 (92-484); 91 (52-159)
MRE11(11q)	Rare	*Cis*	CN-LOH	130 (50-338); 37 (17-84)
SH2B3(12q)	Rare	*Cis*	CN-LOH	11 (5.8-20)
MPL(1p)	Rare	*Cis*	CN-LOH	142 (111-184); 54 (30-100)
ATM(11q)	Rare	*Cis*	CN-LOH	96 (52-177)
TM2D3(15q)	Rare	*Cis*	CN-LOH	555 (425-724)
TCL1A(14q)	Common	*Cis*	CN-LOH	0.84 (0.75-0.94); 0.88 (0.79-0.98)
DLK1(14q)	Common	*Cis*	CN-LOH	1.24 (1.13-1.37); 1.38 (1.31-1.44)
JAK2(9p)	Common	*Cis*	CN-LOH	2.29 (1.99-2.63)
SP140(2q)	Common	*Trans*	Any autosomal mCAs	1.08 (1.05-1.10)
TERC(3q)	Common	*Trans*	Any autosomal mCAs	0.93 (0.91-0.96)
TERT(5p)	Common	*Trans*	Any autosomal mCA 14q CN-LOH	1.11 (1.08-1.14); 1.27 (1.21-1.33)
FRA10B(10q)	Common	*Cis*	Loss of 10q	18 (12-26)
DXZ1(X)	Common	*Cis*	Loss of X	1.09 (1.04-1.15)
DXZ4(X)	Common	*Cis*	Loss of X	1.10 (1.04-1.17)
HLA(6p)	Common	*Trans*	Loss of X	1.18 (1.12-1.25)
SP140L(2q)	Common	*Trans*	Loss of X	1.17 (1.12-1.20)
NEDD8–TINF2(14q)	Common	*Cis*	CN-LOH	1.62 (1.42-1.85)
CTU2(16q)	Rare	*Cis*	CN-LOH	28 (17-45)
MAD1L1(7p)	Common	*Trans*	Gain of 15	1.61 (1.46-1.77)

Adapted from Silver et al.[252] *Odds ratios (ORs) given for the variant with the most significant P value.[9] CI: confidence interval; CN-LOH: copy-neutral loss of heterozygosity, mCAs: mosaic chromosomal alterations*

Because epidemiological studies showed strong associations between CH and CVD, it raised the question as to whether both entities shared heritability. A study showed that on top of the risk of hematological malignancy, germline *TET2* mutations were associated with pulmonary arterial hypertension[263], believed to result from the overproduction of inflammatory cytokines in differentiated immune cells. A genetic variation in the gene *SH2B3* has also been shown to predispose to cardiovascular dysfunction, including hypertension, aortic dissection, atherosclerosis, and stroke.[264–266]

Contrary to CHIP, and prior to the work reported in the present thesis, mCAs alone had not been shown to be associated with CAD in the general population. However, mCAs had been linked to increased risks of severe infection diseases, including severe COVID-19.[267] In another study, the concomitant presence of both mCAs and CHIP was shown to further increased the risk of CVD mortality compared to those with either type alone in the general population.[268]

1.5.2.2 CH and Peripheral Artery Disease and Venous Thromboembolism

Within the UK and the Mass General Brigham Biobanks among 50,122 individuals with whole-exome sequencing, investigators reported that carriers of CHIP mutations were at higher risks of peripheral artery disease than their non-CHIP carrier counterparts (HR: 1.67, 95% CI: 1.32 to 2.11, $P=2.2\times10^{-5}$) where the risk was particularly high for carriers of DNA damage response (DDR) genes (*TP53*, *PPM1D*; HR: 2.72, 95% CI: 1.20 to 1.75, $P<0.001$), and notably high for carriers of *TP53* mutations (HR: 4.98, 95% CI: 1.23 to 20.09, $P=0.024$).[269] In the same study, a mouse model of atherosclerosis transplanted with 20% $Trp53^{-/-}$ bone marrow cells was sufficient to accelerate atherosclerosis development through a macrophage-driven process, independent of IL-1β and IL-6, thereby suggesting a distinct mechanism relating *TP53* CHIP and peripheral artery disease. Previously it has also been shown that venous thromboembolism is linked with $JAK2^{V617F}$ CHIP. In mice with hematopoietic $JAK2^{VF}$, higher rates of thrombosis have been observed owing to

[*9*] Reprinted from Silver AJ, Bick AG, Savona MR. Germline risk of clonal haematopoiesis. *Nat Rev Genet*. 2021;22(9):603-617. Reprinted with permission from Sprinter Nature (license #5518291460721)

increased NETs formation, an important factor of thrombosis. In vivo analyses showed that treatment with a JAK1/2 inhibitor was associated with a decrease in NET formation and decreased thrombosis in $JAK2^{VF}$ mice.[270]

1.5.2.3 Inflammatory Processes Underlie CH

Experimental studies have attempted to elucidate the mechanisms relating CH and CVD. It was known that macrophages with CHIP mutations had a tendency to overproduce inflammatory cytokines, such as IL-1β, IL-6, IL-18, through the NLRP3 inflammasome and AIM2 inflammasome activation in TET2 and $JAK2^{VF}$ CHIP, respectively. Specifically, in atherosclerotic-prone mice, transplanted bone marrow cells that contained 10% of $TET2^{-/-}$ cells showed clonal expansion, which led to a 60% increase in atherosclerotic plaque size, as well as an augmented nucleotide-binding oligomerization domain, leucine-rich repeat and pyrin domain-containing (NLRP3) inflammasome-mediated secretion of IL-1β by $TET2^{-/-}$macrophages.[271] This atherogenic effect was successfully blocked following the administration of an NLRP3 inhibitor. In another study, the investigators transplanted the bone marrow from TET2 heterozygous ($TET^{+/-}$) or homozygous ($TET2^{-/-}$) knockout mice or control mice into atherosclerosis-prone animals and noticed that mice who were transplanted with the homozygous mutation had 2.7-fold larger aortic root atherosclerotic plaques compared to their controls.[246] Myeloid lineage-specific TET2 knockout mice and macrophages revealed accelerated atherosclerosis development following an enhanced CXC chemokine expression followed by secretion of IL-1β and IL-6.[246,271] Other mechanistic studies have also previously showed that in primary cells, *TET2* was found to actively suppress IL-6 transcription during inflammation in innate myeloid cells, including dendritic cells and macrophages. Loss of *TET2* resulted in the upregulation of inflammatory mediators, such as IL-6.[272]

The relevance of these studies in the context of atherosclerosis involves the putative inflammatory mechanism implicating IL-6 in the development of CAD. A disruption of IL-6 signaling as a result of a common variant (IL-6Rp.Asp358Ala), a well-established germline allele associated with modest CAD reduction in the general population, was shown to be associated with a reduction of the risk of CAD in carriers of CHIP mutations.[273] Serum IL-6 has also been

shown to be more accurate in discriminating individuals at risk of CAD compared to the ACC/AHA atherosclerotic CVD risk score (ASCVD, receiver operating characteristic [ROC] curve: 0.72 vs. 0.54).[274] Comparable to *TET2*, CH with JAK2^{V617F} mutations modulate the risk of CAD. Mice with myeloid-specific JAK2^{V617F} mutations develop accelerated atherosclerosis in tandem with cellular proliferation reinforced by the AIM2 inflammasome activation and IL-1β production.[275] In mice with JAK2^{V617F} positive transplanted bone marrow after irradiation, a higher burden of inflammatory macrophages in atherosclerotic lesions were observed, as well as increased necrotic core formation and higher plaque instability. These effects were circumvented with the administration of an IL-1β inhibitor.[275,276]

To further explore whether the neutralization of IL-1β signaling could benefit patients with CHIP, the investigators of CANTOS[190] performed a post-hoc analysis comparing the risk of major adverse cardiovascular events between canakinumab and placebo exposed participants with and without CHIP. The results were inconclusive when all CHIP mutations were considered (HR: 1.32, 95% CI: 0.86 to 2.04, *P*=0.21)[277]. However, patients with *TET2* CHIP had significantly lower risk of incident major cardiovascular events (HR: 0.38, 95% CI: 0.15 to 0.96, stratified *P*=0.04), however, the effect was not specific to canakinumab treatment (interaction *P*=0.14). These analyses further support the underlying putative mechanism of inflammatory processes between CH and atherosclerosis.

1.6 – Proposed Framework to Thesis

1.6.1 Rationale & Clinical Issues Considered

Over the last twenty years, the field of inflammation and immunity has flourished in both CVD and tumor biology, independently. Despite the enormous strides and clinical translation of these findings in the practice of cardiology and oncology, both entities have remained separate, with the application of inflammation biology onto cardiovascular and cancer therapeutics remaining disintegrated. Recently, the medical community has increasingly appreciated the convergence of inflammation processes as a fundamental unifying mechanistic connection between the two diseases. Focusing on the commonality of inflammation biology in oncogenesis and atherogenesis provides a compelling rationale to align forces and enhance multi-disciplinary patient care, with

undeniable therapeutic implications. In consequence, the current doctorate aims at studying the coinciding pathogenic inflammatory mechanisms at the intersection of cardiology and oncology by concentrating on two issues where cancer and CVD have shown overlap, namely non-specific inhibition of inflammation with statin, and clonal hematopoiesis as a marker of CVD in patients diagnosed with cancer.

1.6.2 Objectives & Hypotheses

The overarching objective of this doctoral thesis is to enhance our understanding of inflammation as a shared biological mechanism underlying both cancer and CVD. Our central hypothesis posits that inflammation mechanisms, either directly or indirectly, contribute to both disease phenotypes, acting as a common risk factor for cancer and CVD outcomes. This implies that modulating the inflammatory environment in individuals with CVD can influence the risk of cancer, and vice versa.

Under this premise, we will investigate the contribution of inflammation to the risk of disease co-occurrences. To explore this hypothesis, we will rely on clinical data and extensive genotyping repositories, which have informed the development of specific objectives to guide our analyses.

Objective 1: Investigate the potential of non-specific inhibition of inflammation in reducing the risk of cancer.

- Statins are widely used cholesterol-lowering agents that have proven to be effective in lowering cardiovascular morbidity and mortality in first- and second-line settings. Satins also exhibit anti-inflammatory properties, which have been shown to slow down cancer cell growth, reduce metastasis, and promote apoptosis. However, evidence regarding statin effects on cancer is inconsistent. By using genome-wide analyses, this objective aims to uncover pathophysiological pathways mediating the potential effect of statins on cancer risk in patients with prevalent CAD.

Objective 2: Evaluate clonal hematopoiesis as a biomarker for CVD risk in cancer survivors.

- Clonal hematopoiesis, a known predictive biomarker of hematological cancer, has recently been shown to be linked with CVD, possibly through inflammatory mechanisms. With an expected increase in cancer survivors in the coming decades, cardiovascular morbidity

remains a significant concern. Identifying a biomarker that can more accurately detect subclinical cardiovascular risk in these patients is crucial for improving patient care. This objective focuses on the potential role of clonal hematopoiesis mutations as a CVD risk marker in cancer survivors, addressing the limited information available on the relationship between clonal hematopoiesis and cardiovascular events in the context of malignancy.

1.6.3 Assumptions & Limitations

In this thesis, we acknowledge some assumptions and limitations that provide context for interpreting our findings. We assume that inflammation is a shared biological mechanism underlying both cancer and CVD and that modulating inflammation in individuals with one condition may influence the risk of the other. We presume that individual genetic differences may account for variability in the effects of statins on cancer risk. Furthermore, we assume that clonal hematopoiesis is associated with CVD through inflammatory mechanisms. Limitations include the use of observational data, which could lead to confounding factors, and although we attempt to control for potential confounders, residual confounding cannot be entirely ruled out. We rely on existing clinical and genotyping data repositories which may limit the scope of our analyses, as these datasets may not include all relevant variables or populations. The generalizability of our findings may be limited by the specific populations and patient cohorts included in the studies and results may not apply to all demographic groups or clinical settings. Finally, limited information on clonal hematopoiesis and cardiovascular events in malignancy may affect the interpretation of our findings.

By acknowledging these assumptions and limitations, we hope to provide a balanced perspective on the contributions and potential implications of our research within the broader context of cancer and CVD research.

Chapter 2

Genetic Meta-Analysis of Cancer Diagnosis Following Statin Use Identifies New Associations and Implicates Human Leukocyte Antigen (HLA) in Women.

Objective 1: Investigate the potential of non-specific inhibition of inflammation in reducing the risk of cancer.

Previously several population-based studies have alluded to a potential association between statin use and cancer. Individual post-hoc analyses of clinical trials, designed to assess the efficacy of statin use on cardiovascular endpoints in primary or secondary settings, revealed divergent results. Given that statins harbor anti-inflammatory properties in addition to their lipid-lowering abilities, there is a possibility that targeting inflammation may be conducive to reducing the risk of cancer incidence. The use of a genome wide association studies (GWAS) in statin users for the risk of incident cancer could reveal underlying mechanisms of action that are related to the purported association of statin and cancer. Furthermore, revealing such mechanisms could potentially identify individuals at an increased risk of such associations.

Given the conflicting evidence on the association between statin use and cancer risk based on clinical trials, we sought to focus on potential genetic factors that may be linked to the risk of incident cancer. For that purpose, we conducted a GWAS for the risk of incident cancer among statin users using data from two phase IV randomized-controlled clinicals trials (TNT & IDEAL) in patients with available genotyping information. We also leveraged genotyping data from a large, prospective cohort as an external replication data as necessary (UK Biobank). Age- and sex-stratified analyses were anticipated.

Our analyses identified a genetic variant (rs13210472) that was significantly associated with a higher risk of incident cancer diagnosis in women with CAD at baseline taking statins (development cohort using TNT & IDEAL). Its effect remained significantly associated with incident cancer in an external cohort of women statin users with prevalent CAD. Interaction term

analyses suggest that the association between the variant and incident cancer was sex-specific, but not statin-specific. Functional annotation was conducted to better extrapolate the biological and physiological roles of the involved genes, and further underlined the potential immunological mechanisms and sex-specific effect with rs13210472 (as the effect was undetectable in men). This work represents the first GWAS on statin use and the risk of incident cancer. Its findings point towards a tri-factor relationship between inflammation, cancer, and CVD.

Contributions: Maxine Sun and Marie-Pierre Dubé conceptualized and designed the study approach. A collaboration between the MHI and Pfizer to establish access to clinical data and biological material was established by Jean-Claude Tardif and Marie-Pierre Dubé. Clinical trial data preparation and validation was overseen by Sylvie Provost and Marie-Pierre Dubé. Genotyping was conducted at the Pharmacogenomics Centre by Diane Valois and Ian Mongrain, overseen by Marie-Pierre Dubé. Genetic and clinical data were integrated by Géraldine Asselin, Sylvie Provost, and Yassmin Feroz Zada. GWAS were conducted by Géraldine Asselin. Replication data implementation and optimization on the *sql* server was performed by Marc-André Legault and overseen by Sylvie Provost. Maxine Sun performed (except for GWAS) and interpreted statistical analyses. Audrey Lemaçon performed functional annotation analyses of the identified genetic variant and drafted a scientific interpretation of those findings. Amina Barhdadi and Marie-Pierre Dubé assisted Maxine Sun in the interpretation of the results and provided critical suggestions for re-analyses. Maxine Sun drafted the first version of the manuscript. Subsequent drafts were critically revised by Marie-Pierre Dubé, Hugues Aschard, and Jean-Claude Tardif.

Genetic Meta-Analysis of Cancer Diagnosis Following Statin Use Identifies New Associations and Implicates Human Leukocyte Antigen (HLA) in Women

Maxine Sun [1,2,3], Audrey Lemaçon[1,2], Marc-André Legault [1,2,3], Géraldine Asselin[1,2], Sylvie Provost[1,2], Hugues Aschard[4], Amina Barhdadi[1,2], Yassamin Feroz Zada[2], Diane Valois[1,2], Ian Mongrain[1,2], Jean-Claude Tardif[2,3*], Marie-Pierre Dubé[1,2,3*]

[1] Université de Montréal Beaulieu-Saucier Pharmacogenomics Center, Montreal, Quebec, Canada
[2] Montreal Heart Institute, Montreal, Quebec, Canada
[3] Faculty of Medicine, Université de Montréal, Montreal, Quebec, Canada
[4]Centre de Bioinformatique, Biostatistique et Biologie Intégrative (C3BI), Institut Pasteur, Paris, France
[*] Equal contribution

Pharmacogenomics J. 2021;21(4):446-457. doi:10.1038/s41397-021-00221-z

2.1 Abstract

We sought to perform a genomic evaluation of the risk of incident cancer in statin users, free of cancer at study entry. Patients who previously participated in two phase IV trials (TNT and IDEAL) with genetic data were used (n_{pooled}=11,196). A GWAS meta-analysis using Cox modeling for the prediction of incident cancer was conducted in the pooled cohort and sex-stratified. rs13210472 (near HLA-DOA gene) was associated with higher risk of incident cancer amongst women with prevalent coronary artery disease (CAD) taking statins (hazard ratio [HR]: 2.66, 95% confidence interval [CI]: 1.88 to 3.76, P=3.5×10^{-08}). Using the UK Biobank and focusing exclusively on women statin users with CAD (n_{female} =2,952), rs13210472 remained significantly associated with incident cancer (HR: 1.71, 95% CI: 1.14 to 2.56, P=9.0×10^{-03}). The association was not observed in non-statin users. In this genetic meta-analysis, we have identified a variant in women statin users with prevalent CAD that was associated with incident cancer, possibly implicating the human leukocyte antigen pathway.

2.2 Keywords

GWAS**;** Meta-analysis**;** Statin**;** Coronary artery disease**;** Cancer incidence**;** Human leukocyte antigen**;** Inflammation

2.3 Introduction

Statins, also known as HMG-CoA reductase inhibitors, are a class of lipid-lowering medications that have shown to effectively reduce low density lipoprotein (LDL) cholesterol levels and decrease the incidence of cardiovascular and cerebrovascular events in various settings.[278,279]

Although formal meta-analyses based on data from randomized-controlled clinical trials have consistently refuted that an association exists between statin use and the risk of cancer[280], occasional post-hoc analyses have implicated otherwise.

For example, within the Treating to New Targets (TNT) clinical trial, daily high-dose atorvastatin for women was associated with an increase in cancer death compared to low-dose atorvastatin (n_{female}=1,902, *P*=0.006).[281] Similarly, women randomized to pravastatin in the Cholesterol and Recurrent Events (CARE) saw an increased risk of recurrent/incident breast cancer compared to placebo (n_{female}=576, *P*=0.002).[183] Within the West of Scotland Coronary Prevention Study Group (WOSCOPS) trial, a higher incidence of prostate cancer was observed in men treated with pravastatin compared to placebo (2.7% vs. 1.8%, *P*=0.03).[282] Within the PROspective Study of Pravastatin in the Elderly at Risk (PROSPER) focusing on those with an average age of 75 years old, pravastatin users had 25% higher risk of developing any cancer compared to placebo (n_{female}=3,000, *P*=0.02).[222] Outside of the clinical trial setting, large population-based studies have also shown in various occasions, although inconsistently, that statins may be associated with the risk of cancer.[228,283–285]

Although the use of statins is generally not associated with incident cancer, it is possible that this risk may vary for some segments of the population given a specific genetic profile. Based on such premises, it may be postulated that a higher dose of statin may in fact modulate that risk for some individuals. Against such backdrop, we sought to conduct and report on the first pharmacogenomic evaluation via a meta-analysis of genome-wide-association studies (meta-GWAS) for incident cancer occurrences in cancer-free individuals using two randomized-controlled phase IV trials in patients with coronary artery disease (CAD).[286,287] Furthermore, we relied on the UK Biobank as a mean of replicating our findings. Our hypothesis stated that cancer incidence may vary significantly for certain segments of the population of statin users as determined by genetic variants.

2.4 Materials and Methods

2.4.1 Data Sources

Discovery cohorts (TNT & IDEAL):

We relied on data from two previously conducted randomized-controlled phase IV trials, namely the Treating to New Targets (TNT , n=10,001) and the Incremental Decrease in End Points through Aggressive Lipid Lowering (IDEAL, n=8,888, **Suppl. Materials, A.1.1**), both of which performed a comparative effectiveness study of high-dose (atorvastatin at 80 mg for both trials) vs. lower-dose statin (atorvastatin at 10 mg for TNT or simvastatin at 20/40 mg for IDEAL) for the prevention of cardiovascular outcomes in the secondary setting.[286–289] For both trials, prevalent cancer was

defined using baseline medical history questionnaires and baseline medication documentation, and excluded. This resulted in 9,697 patients for TNT and 8,279 patients for IDEAL (**Suppl. Methods, A.1.2)**.

Of those, we focused only on Caucasians with available genetic information. Specifically, genome-wide genotyping was performed using 200 ng of genomic DNA in GLP-environment at the Beaulieu-Saucier Pharmacogenomics Centre (Montreal, Canada). Following this step, genetic data were available for a pooled sample of 11,196 patients from both TNT and IDEAL, which formed the basis of our discovery cohort. Stratified analyses were conducted amongst men (n_{pooled}=9,139) and women (n_{pooled}=2,057). A full description of genotyping methodologies is available in the **Suppl. Methods, A.1.3**)

Our primary endpoint of interest was an incident cancer diagnosis or cancer-related death, hereafter termed incident cancer. For both trials, these events were captured using adverse event records and from adjudicated cause of death file (**Suppl. Methods, A.1.2**). The date of the first event was recorded and subtracted from the index date (randomization date) for time-to-event analyses. Only the first cancer event and associated date were retained for patients with multiple cancers. Additional covariates such as treatment type (statin dose), age at study entry, sex, baseline body mass index (BMI), hypertension, diabetes mellitus, smoking status, as well as ten principal components (PCs 1-10) to account for genetic ancestry were included in various models.

Replication cohort (UK Biobank):

For our replication cohort, we relied on the UK Biobank, a major international health resource with over 500,000 participants aged between 40-69 years old, recruited between 2006-2010 across the UK, with available genotyping information for all participating individuals. For the purpose of our analyses, we focused on cancer-free Caucasian patients at study entry of the UK Biobank (version 2). Cancer-free status was determined using cancer registry files, including ICD-9 and ICD-10 diagnostic codes and the corresponding diagnosis dates, as well as those who self-reported having had cancer during the initial verbal interview. Using the date of baseline as our study's index date (time 0), a cancer diagnosis that occurred before this time (time<0) was considered a prevalent case, and excluded. Of note, patients with a prevalent non-melanoma skin cancer diagnosis were kept within our analysis cohort, unless another malignant diagnosis was identified prior to the index date. Incident cancer diagnosis or death from cancer was determined whenever any cancer event was recorded after the study's entry date (excluding diagnosis of non-melanoma skin cancer). Sensitivity analyses, elaborated in the Supplemental Materials (**Suppl. Methods, A.1.4**), were performed in an attempt to cross-validate cancer diagnoses obtained from cancer registry files to those captured in hospitalization files (via inpatient files). To better simulate the TNT and IDEAL population cohorts, which consisted of statin users with cardiovascular morbidities, we further extracted information on baseline statin use and CAD.

Baseline statin medication is self-reported and was obtained using a previously designed algorithm which converts medication records into the standardized ATC code.[290] Prevalent and incident CAD was defined using composite codes for ischemic heart disease, angina pectoris, acute myocardial infarction, PCI, and CABG (**Suppl. Methods, A.1.5**).

This allowed the identification of 12,462 (9,510 men and 2,952 women) cancer-free individuals receiving statins (past or present) with a prevalent CAD constituting our main replication cohort. Median follow-up was 7.2 years. Incident cancer diagnosis or cancer death was recorded in 1,048 individuals (8.4%). Additional covariates were extracted, including baseline measure of high-sensitivity C-reactive protein (hs-CRP), age at data entry, BMI, smoking status, baseline hypertension and diabetes, and PCs (1-10) for genetic ancestry. Patients not receiving statins at baseline as well as those without CAD (neither prevalent nor incident) were used for additional exploratory analyses in subgroups. Overall, 385,212 cancer-free individuals were available for exploratory analyses.

2.4.2 Statistical Analyses

First, separate Cox regressions were performed to analyze genome-wide associations with prediction of incident cancer (diagnosis or death related to cancer) in the TNT and IDEAL cohorts separately and stratified according to sex. Subsequently, results were pooled by meta-analysis (n_{total}=11,196). Adjustment was made for age, sex (except in sex-stratified analyses), treatment type (high-dose vs. low/usual-dose), and 10 principal components (PCs) for ancestry. Models including a variant-by-treatment (high/low-dose statin) and variant-by-sex interaction terms were conducted in the event of statistically significant SNPs. All models including an interaction term also included the variables of the interaction as covariates. A GWAS was performed using genipe (version 1.3.1 [291]) for TNT and genetest (version 0.2.2) and meta-analyses with GWAMA (version 2.2.2; [292]) for both genotyped and imputed variants with a MAF≥0.01 with a fixed-effect model. Statistical tests in the discovery phase were two-sided and adjusted to account for the multiple testing of common SNPs using a significance threshold of 5.0×10^{-8}. Heterogeneity was assessed using I^2 inconsistency metric and the *P*-value for Cochran's Q statistic.[293–295] Variants that would emerge as statistically significant would be re-examined in the replication cohort (UK Biobank) in statin users with CAD. In the main replication models, adjustment would be made for age, sex, and PCs (1-10). Statistically significant variants were re-assessed in models considering a variant-by-treatment (i.e., statin vs. no statin use), and a variant-by-sex interaction terms. In additional and sensitivity analyses, the consideration of baseline hypertension and diabetes, smoking status, and hs-CRP were included, as well as non-statin use and those without CAD at baseline. A summary flow diagram characterizing the main cohorts and analyses performed is depicted in **Suppl. Figure A.2.1** The study was performed under the terms of the Declaration of Helsinki and the study protocol was approved by the local review boards or ethics committees and all patients gave written informed consent.

2.5 Results

2.5.1 Discovery Cohort: Multivariable Cox Regression Analyses Predicting the Risk of Incident Cancer Via a Meta-GWAS Approach

Overall, incident cancer diagnosis or cancer-related death was recorded in 1,078 individuals (618 from TNT and 460 from IDEAL; 9.6%). Four SNPs were associated with incident cancer (**Table 2.1**): two intergenic variants located between *IFNA6* and *IFNA13* were identified amongst men only, rs10964972 (hazard ratio [HR]: 2.32, 95% confidence interval [CI]: 1.77 to 3.04, $P=1.1\times10^{-09}$) and rs10964973 (HR: 2.30, 95% CI: 1.76 to 3.01, $P=1.1\times10^{-09}$, **Figure 2.1A, Suppl. Fig.** A.2.2); and the other two variants were identified amongst women only, rs7846694 located on *PEBP4* (HR: 1.74, 95% CI: 1.43 to 2.13, $P=3.9\times10^{-08}$), as well as rs13210472 located between the *HLA-DOA* and *HLA-DPA1* genes (HR: 2.66, 95% CI: 1.88 to 3.76, $P=3.5\times10^{-08}$, **Figure 2.1B,** A.2.2). Given that statistical significance for these four SNPs on incident cancer was observed only in sex-specific sub-cohorts, we also tested a variant-by-sex interaction term within the overall pooled cohort. Only rs7846693-by-sex and rs13210472-by-sex interaction terms were statistically significant ($P=9.0\times10^{-09}$ and 1.8×10^{-06}, respectively, **Table 2.1**). None of the identified SNPs interacted with treatment assignment to higher or lower dose statin in individual studies or within the pooled cohorts (**Table 2.2**). There was no evidence of heterogeneity among studies for any of the loci of interest (**Suppl. Table** A.3.1).

Table 2.1 SNPs With P≤5.0×10^{-08} From the Genetic Meta-Analysis Predicting Time To Incidence of Cancer.

	Cohort†	SNP (Nearest gene)	Chr:pos	EAF	Beta	SE	HR	95% CI	*P*
	Men								
Discovery	**Discovery (n_{total}=9,126)**	rs10964972 (IFNA6, IFNA13)	9:21,362,409	0.0164	**0.84**	**0.14**	**2.32**	**1.77-3.04**	**1.1×10^{-09}**
	TNT (n=4,039)				0.57	0.19	1.76	1.21-2.57	3.3×10^{-03}
	IDEAL (n=5,087)				1.13	0.20	3.11	2.11-4.58	1.0×10^{-08}
	Pooled with variant×sex*				0.12				
Replication	UK Biobank (n=9,488)				-0.02	0.19	0.98	0.67-1.43	0.93
Discovery	**Discovery (n_{total}=9,127)**	rs10964973 (IFNA6, IFNA13)	9: 21,362,772	0.0170	**0.83**	**0.14**	**2.30**	**1.76-3.04**	**1.1×10^{-09}**
	TNT (n=4,040)				0.57	0.19	1.76	1.22-2.55	2.7×10^{-03}
	IDEAL (n=5,087)				1.12	0.20	3.08	2.09-4.54	1.3×10^{-08}
	Pooled with variant×sex*				0.11				
Replication	UK Biobank (n=9,488)				-0.03	0.19	0.95	0.67-1.42	0.89
	Women								
Discovery	**Discovery (n_{total}=2,057)**	rs7846693 (PEBP4)	8: 32,995,081	0.0333	**0.56**	**0.10**	**1.74**	**1.43-2.13**	**3.9×10^{-08}**
	TNT (n=927)				0.44	0.14	1.56	1.19-2.05	1.5×10^{-03}
	IDEAL (n=1,130)				0.68	0.15	1.97	1.48-2.62	3.6×10^{-06}
	Pooled with variant×sex*				9.0×10^{-09}				
Replication	UK Biobank (n=2,865)				-0.08	0.12	0.92	0.73-1.17	0.51
Discovery	**Discovery (n_{total}=2,052)**	rs13210472 (HLA-DOA, HLA-DPA1)	6: 22,617,312	0.2489	**0.98**	**0.18**	**2.66**	**1.88-3.76**	**3.5×10^{-08}**
	TNT (n=925)				0.90	0.22	2.46	1.58-3.81	5.9×10^{-05}
	IDEAL (n=1,127)				1.11	0.29	3.04	1.72-5.38	1.2×10^{-04}
	Pooled with variant×sex*				1.8×10^{-06}				
Replication	UK Biobank (n=2,952)				0.54	0.21	1.71	1.14-2.56	9.0×10^{-03}

In replication analyses, no adjustment was made for treatment as there is no statin dosage available within the database.
†The number of patients available for analyses may differ slightly per SNP depending on whether the genetic information was available for all patients. Cox regression models were based on the pooled cohorts of TNT and IDEAL, and adjusted for age, treatment, sex-specific principal components (SC1-SC10).
**The shown P-value corresponds to the P-value of the snp×sex interaction term, which was obtained via a multivariable Cox regression model that also adjusted for age, treatment, the SNP, and sex based on a meta-analytic approach from the pooled cohorts of TNT and IDEAL. SNP: single nucleotide polymorphism, Chr: chromosome, pos: position, EAF: effect allele frequency, SE: standard error, HR: hazard ratio, CI: confidence interval.*

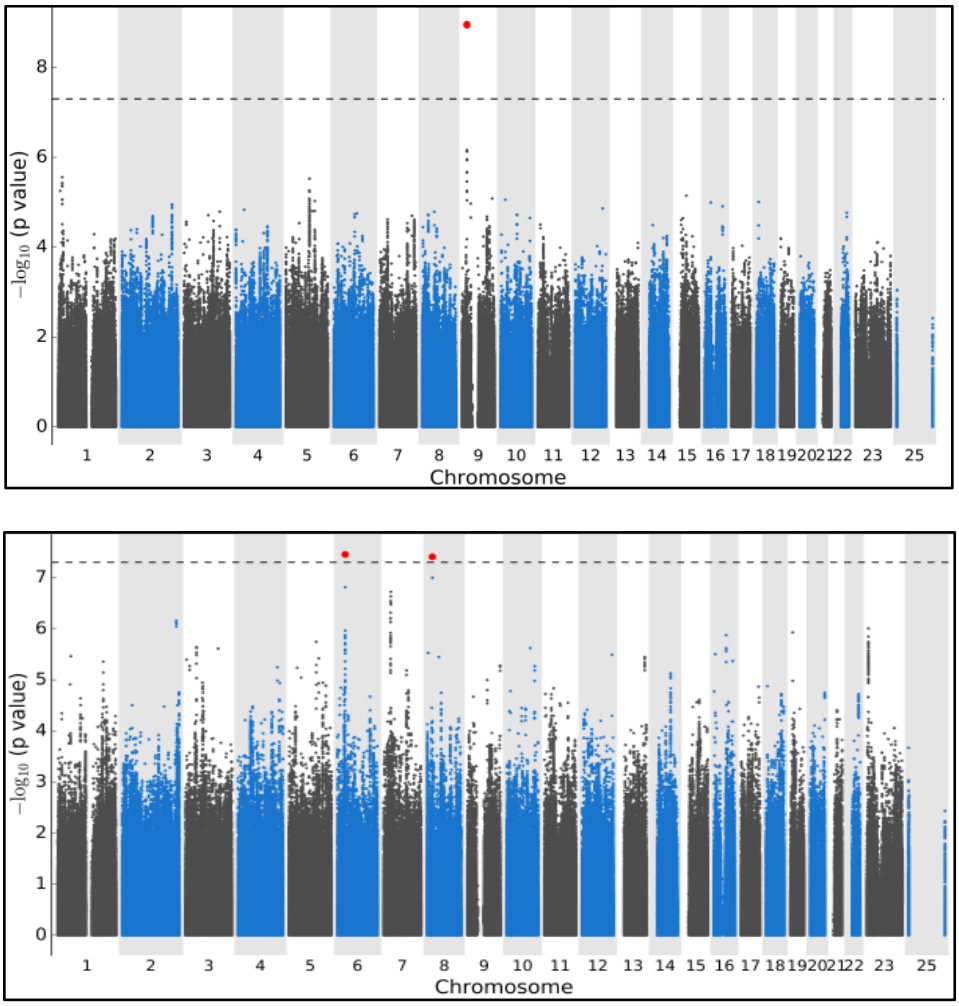


Figure 2.1 Manhattan Plot of the Genome-Wide Association Meta-Analysis Analysis of Genetic Variants of Minor Allele Frequency ≥1% For Time to Occurrence of Cancer Using a Cox Proportional Hazards Regression From the TNT And IDEAL Cohorts.

***A (top)**. Men only analysis controlling for age, treatment, 10 sex-specific principal components (6,133,088 genetic variants and 9,139 samples); **B (bottom)**. Women only analysis controlling for age, treatment, 10 sex-specific principal components (6,087,742 variants and 2,057 samples).*

2.5.2 Replication Cohort: Multivariable Cox Regression Analyses Predicting the Risk of Incident Cancer

Using the UK Biobank, rs10964972 and rs10964973 amongst men statin users with CAD, as well as rs7846693 amongst women statin users with CAD were not associated with incident cancer (**Table 2.2**). In contrast, rs13210472 remained associated with incident cancer amongst women

with prevalent CAD taking statins (HR: 1.71, 95% CI: 1.14 to 2.56, P=9.0×10^{-03}, **Table 2.2, Fig. 2.1**). In the replication cohort, the variant-by-sex interaction in statin users with CAD was indicative of an interaction but did not pass the 0.05 significance threshold ($P_{variant\times sex}$=0.05), and the variant-by-treatment (statin yes/no) interaction was not significant ($P_{variant\times statin}$=0.15). Of note, amongst the 2,952 women who were taking statins with prevalent CAD, incident cancer was observed in 6.5%. Of those, the most common cancer types were breast (23%), gastrointestinal (22%), and respiratory-related (16%, (**Suppl. Fig.** A.2.3). Cancer incidence rates according to rs13210472 genotypes were 6.2% for A/A, 9.8% for A/C, and 25.0% for C/C (χ^2 P=0.029), with more pronounced differences observed according to specific cancers (**Suppl. Table A.3.2**).

Table 2.2 Analyses Assessing the Variant-By-Treatment Type Within the TNT And IDEAL Cohort (Discovery) and the UK Biobank (Replication).

	Cohort	N	Events	SNP	HR (95% CI)	*P*
Men						
Discovery	**High-dose**			**rs10964972**		
	TNT	2014	255		1.62 (0.94-2.79)	0.082
	IDEAL	2509	194		2.00 (1.02-3.93)	0.044
	Low/usual-dose					
	TNT	2017	237		1.87 (1.10-3.17)	0.020
	IDEAL	2500	170		4.18 (2.58-6.79)	7.3×10^{-09}
	Pooled TNT & IDEAL					
	variant×treatment	9040	856		$P_{interaction}$=0.666	
Replication	**UK biobank**					
	Statin	9488	855		1.11 (0.89-1.39)	0.353
	No statin	1690	162		1.19 (0.71-1.99)	0.512
	variant ×statin	11178	1017		$P_{interaction}$=0.879	
Discovery	**High-dose**			**rs10964973**		
	TNT	2015	256		1.68 (0.99-2.84)	0.053
	IDEAL	2508	194		2.00 (1.02-3.92)	0.447
	Low/usual-dose					
	TNT	2017	237		1.80 (1.06-3.06)	0.029
	IDEAL	2501	170		4.11 (2.53-6.68)	1.1×10^{-08}
	Pooled TNT & IDEAL					
	variant ×treatment	9041	857		$P_{interaction}$=0.321	
Replication	**UK biobank**					
	Statin	9488	855		0.98 (0.68-1.43)	0.926
	No statin	1690	162		0.90 (0.33-2.42)	0.828
	variant ×statin	11178	1017		$P_{interaction}$=0.818	
Women						
Discovery	**High-dose**			**rs7846693**		
	TNT	442	61		1.55 (1.04-2.30)	0.031
	IDEAL	538	48		2.23 (1.43-2.48)	3.9×10^{-04}
	Low/usual-dose					
	TNT	483	64		1.62 (1.10-2.40)	0.015
	IDEAL	576	48		1.72 (1.15-2.55)	0.008
	Pooled TNT & IDEAL					
	variant ×treatment	2039	221		$P_{interaction}$=0.473	

Replication	**UK biobank**					
	Statin	2865	192		0.91 (0.72-1.16)	0.443
	No statin	1329	89		1.08 (0.75-1.53)	0.686
	variant ×statin	4194	278		$P_{interaction}$=0.452	
Discovery	**High-dose**			**rs13210472**		
	TNT	442	61		1.65 (0.77-3.54)	0.196
	IDEAL	538	48		2.89 (1.13-7.35)	0.026
	Low/usual-dose					
	TNT	483	64		3.21 (1.93-5.34)	7.1×10^{-06}
	IDEAL	576	48		3.65 (1.70-7.87)	9.3×10^{-04}
	Pooled TNT & IDEAL					
	variant ×treatment	2039	221		$P_{interaction}$=0.638	
Replication	**UK biobank**					
	Statin	2952	193		1.71 (1.14-2.56)	9.1×10^{-03}
	No statin	1370	88		0.85 (0.37-1.95)	0.704
	variant ×statin	4322	281		$P_{interaction}$=0.151	

It is noteworthy that the number of samples may differ from the whole population given that some patients may have missing dosage information per variant. Furthermore the treatment in the interaction of variant×treatment is high-dose vs. low-dose statin, whereas the statin in the interaction of variant×statin is statin yes vs. no. SNP: single nucleotide polymorphism, Chr: chromosome, pos: position, EAF: effect allele frequency, SE: standard error, HR: hazard ratio, CI: confidence interval

Additional exploratory analyses in women with prevalent CAD not taking statins (n=1,370), or women without prevalent CAD taking statins (n=20,358) did not show association between rs13210472 and incident cancer (**Table 2.3**). Interaction terms between rs13210472 and sex, statin, and/or CAD were also not significant. The association between rs13210472 and incident cancer was also not significant in men, regardless of statin use and CAD status.

Table 2.3 Replication Multivariable Cox Regression Analyses Assessing the Effect of rs13210472 For Prediction of Incident Cancer Amongst Cancer-Free Individuals Within the UK Biobank.

Cohort	N	Events	HR	95% CI	*P*
A) With prevalent CAD					
Statin users	12,462	1,048	1.21	0.99-1.47	0.05
Women	2,952	193	1.71	1.14-2.56	**9.1×10^{-03}**
Men	9,510	855	1.11	0.89-1.39	0.36
With a variant-by-sex interaction term	12,462	1,048	1.74	1.16-2.60	**7.1×10^{-03}** ($P_{variant\times sex}$=0.05)
Non-statin users	3,061	250	1.08	0.70-1.67	0.72
Women	1,370	88	1.09	0.71-1.69	0.70
Men	1,691	162	1.19	0.71-1.99	0.51
With a variant-by-sex interaction term	3,061	250	0.85	0.37-1.95	0.70 ($P_{variant\times sex}$=0.48)
Women	4,322	281	1.44	1.00-2.07	0.05
With a variant-by-statin interaction term	4,322	281	0.87	0.38-1.98	0.74 ($P_{variant\times statin}$=0.15)
B) Without prevalent CAD					

Women	199,207	10,364	1.02	0.95-1.09	0.65
Women statin users	20,358	1,394	0.96	0.80-1.17	0.70
Women non-statin users	178,849	8,970	1.02	0.95-1.10	0.54
Women with a variant-by-statin interaction term	199,207	10,364	1.02	0.95-1.10	0.53 ($P_{variant \times statin}$=0.57)
C) With or without prevalent CAD					
Women	203,529	10,645	1.03	0.96-1.10	0.44
Women statin users	23,310	1,587	1.05	0.88-1.24	0.61
Women non-statin users	180,219	9,058	1.02	0.95-1.10	0.56
Women with a variant-by-statin interaction term	203,529	10,645	1.02	0.95-1.10	0.54 ($P_{variant \times statin}$= 0.82)
Women with a variant-statin-CAD interaction term	203,529	10,645	1.03	0.95-1.10	0.51 ($P_{variant \times statin \times CAD}$=0.12)
All patients	385,212	21,783	1.03	0.98-1.08	0.23
With a variant-by-sex interaction term	385,212	21,783	1.03	0.96-1.10	0.45 ($P_{variant \times sex}$ =0.92)
With a variant -by-statin interaction term	385,212	21,783	1.04	0.98-1.09	0.17 ($P_{variant \times statin}$ =0.50)
With a variant -by- CAD interaction term	385,212	21,783	1.11	0.93-1.33	0.24 ($P_{variant \times CAD}$=0.36)
With a variant -sex-statin- CAD interaction term	385,212	21,783	1.32	0.84-2.08	0.23 ($P_{variant \times sex \times statin \times CAD}$=0.41)

All models report the additive genetic effect of rs13210472 on incident cancer adjusted for age, incident or recurrent coronary artery disease depending on whether the analyses were performed in patients with or without prevalent coronary artery disease, principal components for genetic ancestry, sex, and statin use except in analyses that focused on a specific sex or statin utilization sub-group. HR: hazard ratio, CI: confidence interval, CAD: coronary artery disease

2.5.3 Sensitivity Analyses

To exclude the possibility of having indirectly identified a genetic variant associated with CAD through collider bias, we ruled out the association between rs13210472 and CAD (CardiogramPlusC4D P=0.466[296], as well as with recurrent CAD events amongst women taking statins with prevalent CAD in the UK Biobank (P=0.83, **Suppl. Table A.3.3**). We also ruled out possible confounding effects by adjusting for shared risk factors between CAD and cancer[297] in the association path between rs13210472 and incident cancer by including hypertension, diabetes, BMI, and smoking status at baseline as covariates within the discovery cohort (TNT and IDEAL) and again in the replication cohort (UK Biobank). In those analyses, the significant relationship between rs13210472 and incident cancer was maintained across all tested populations TNT (HR: 2.51, 95% CI: 1.61 to 3.91, P=4.5×10^{-05}), IDEAL (HR: 3.18, 95% CI: 1.79 to 5.64, P=7.5×10^{-05}), as well as within the UK Biobank (HR: 1.71, 95% CI: 1.14 to 2.56, P=9.2×10^{-03}, **Suppl. Table A.3.3**).

Since rs13210472 is located within the class II region of the HLA system[298], which has strong implications with inflammation, we further examined the relationship between rs13210472 and baseline C-reactive Protein (CRP) in the UK Biobank. Measured high sensitivity (hs)-CRP levels increased significantly according to rs13210472-C allele for women taking statins with prevalent CAD (**Suppl. Fig.** A.2.4). A multivariable linear regression model for prediction of natural logarithm transformation of CRP levels (ln(CRP)) demonstrated that women carriers of the rs13210472-C allele who were taking statins and had prevalent CAD had higher ln(CRP) (β: 0.240, $P=8.7\times10^{-04}$, **Suppl. Table** A.3.4). Although the association between rs13210472 and ln(CRP) was not significant in women with CAD who were not statin users, the variant-by-statin use interaction term in all women with prevalent CAD for prediction of ln(CRP) ($P=0.103$) was not significant. In multivariable Cox regression analyses for prediction of incident cancer in women taking statins with prevalent CAD, the additional adjustment for ln(CRP) did not alter the effect of rs13210472 (HR: 1.72, 95% CI: 1.14-2.60, $P=9.7\times10^{-03}$). No significant interaction was found between rs13210472 and ln(CRP) with respect to incident cancer ($P=0.371$).

2.5.4 Functional Annotation

The variant rs13210472 is an expression quantitative trait loci (eQTL) for HLA-DOA and HLA-DPA1 in blood tissues, as well as for HLA-DPB1 in the pituitary gland tissue[299,300]. The function of the MHC-II molecules is to present antigens to CD4+ T cells.[301] The eQTL data suggests that an increase in HLA-DOA and HLA-DPB1 expression with increasing number of C alleles (**Suppl. Methods** A.1.6). This overexpression could lead to the presence of more HLA-DP heterodimers at the surface of antigen-presenting cells (macrophages, dendritic cells, and B-lymphocytes). The accumulation of HLA-DP presenting antigens would then induce the activation of additional CD4+ helper T cells. Previous reports have implicated CD4+ T cells in diseases such as rheumatoid arthritis in Asians, where they were shown to play a central role in the pathological immune response leading to joint damage.[302] Patients with rheumatoid arthritis have also been shown to be at an increased risk of certain types of cancer– a likely result of chronic inflammatory state.[303] The action of overstimulated $CD4^+$ T cells could be a possible explanation for the increased risk of cancer that we have observed for rs13210472-C carriers.

The sex-specific effect of rs13210472 on incident cancer is supported by the fact that in the adaptive context, females exhibit higher $CD4^+$ T cell counts associated with an increased $CD4^+/CD8^+$ T-cell ratio compared to males.[304] This has been previously confirmed by the overrepresentation of women with autoimmune diseases.[305,306] The putative implication of statin on the increased risk of cancer may underlie a recently reported mechanism[307], where atorvastatin and simvastatin have shown to repress human dendritic cells maturation induced by oxidized LDL to limit T-cell activation, and promote an anti-inflammatory cell response and induction of T regulatory cells through the downregulation of the microRNAs (miRNA) let-7c, a potential tumor suppressor.[308–311]

2.6 Discussion

In the current study, we sought to identify a segment of the population that may be at risk for incident cancer while using statins through a GWAS meta-analysis of two previously conducted clinical trials. Several findings merit further discussion. First, although we were able to identify four genetic variants associated with a higher risk of incident cancer, the effect of these variants did not appear to be potentiated when given at a higher versus lower statin dosage, as highlighted by the lack of significance in the variant-by-treatment interaction terms. This implies that the effect of higher statin dose on cancer incidence compared to lower statin dose was not different across the genotypes of the genetic variants.

Second, while we could not replicate the association signal of three out of four identified genetic variants in the discovery cohort, the association of one genetic variant, rs13210472, with incident cancer was statistically significant in both the discovery and replication cohorts. Specifically, the rs13210472-C allele portended to a 2.6-fold risk of incident cancer compared to the A allele amongst women participants of the TNT and IDEAL trials. To replicate the clinical profile of women participants of the clinical trials, we exclusively selected women with prevalent CAD who reported taking statins at baseline in the UK Biobank. In this replication cohort, the rs13210472-C allele conveyed a 1.7-fold in risk of incident cancer as compared to the A allele. Interestingly, the effect of rs13210472 on cancer failed to be reproduced in men within both the discovery and replication cohorts, which suggests that the association may be mediated by sex-specific pathways.

Because rs13210472 is located within the HLA system, we further hypothesized that the purported effect of the genetic variant on cancer may be mediated via the known effect of inflammation on cancer[49], which prompted us to evaluate hs-CRP, a conventional biomarker of general inflammation. Interestingly, the levels of hs-CRP increased significantly with increasing copies of the rs13210472-C allele, but only amongst women statin users with prevalent CAD, and not amongst men statin users with prevalent CAD, nor amongst individuals without prevalent CAD with/out statin intakes. The association of rs13210472 on incident cancer persisted despite the additional control for measured hs-CRP in the time-to-event model.

The identified SNP (rs13210472) is located between genes HLA-DOA and HLA-DPA1, which involves the HLA system. The HLA complex is genetically inherited [312] and is composed of a cluster of genes encoding the MHC proteins, a highly polymorphic region in the human genome located on the short arm of chromosome 6 with >200 genes, representing the most important genetic area in the vertebrate genome.[313] The MHC proteins encode antigens located on the cell membrane of leukocytes in humans and are involved in infection and autoimmunity, in adaptive and innate immunity, as well as inflammatory pathways.[314,315]

The MHC is the most gene-dense part of the human genome, exhibiting haplotype-specific linkage disequilibrium patterns, and is a hot spot for disease associations.[316] Over the last decade, genome-wide association studies (GWAS) have unraveled a large number of disease associations with MHC region variants, namely in rheumatoid arthritis[317], type 1 diabetes and celiac disease[318], as well as drug hypersensitivities[319]. Nonetheless, the true extent of the involvement of the MHC region in disease genetics is probably underestimated given that customary GWAS do not have sufficient power to uncover all potential associations of this region.[320]

Our study focused on the study of patients with stable CAD, who are likely to have suffered from chronic low-grade inflammation, which arguably, aside from its involvement in atherosclerotic disease[321], is also known to perpetuate the development of cancer and stimulate tumor cell survival and proliferation.[322] From a clinical perspective, our results potentially indicate the specific implication of the HLA system in modulating the risk of carcinogenesis and tumor progression in women statin users with an established vascular morbidity. Because our results showed that the increase in cancer risk determined by the genetic variant in the HLA system did not appear to be worse in higher-dose statin users, the dose-relation of the effect of statin on cancer incidence in individuals at a higher genetic risk could not be supported here. This tempers the hypothesis that it would be the statin use itself that would interact with the HLA pathway in modulating the risk of incident cancer in women carrying the risk allele in the TNT and IDEAL studies. However, we cannot entirely exclude the role of statins, as we did not observe an effect of the risk allele on incidence of cancer in women with established CAD who were not using statins in the UK Biobank. Because the effect of the genetic variant on incident cancer was not observed in men statin users with prevalent CAD, it is plausible to suspect that somehow through the use of statins for women presenting with CAD, as a result of an underlying systemic inflammation, an undiagnosed cancer progression could be hastened as a manifestation of immunity malfunctions that is intertwined with the HLA system. It would also follow that this potential explanation does not apply to men. That being said, we could not readily confirm this given that the effect of rs13210472 appeared to be independent of hs-CRP, a traditional biomarker of inflammation that was measured at baseline in the UK Biobank. A better understanding of the tri-factor relationship between cardiovascular disease, cancer, and inflammation has crucial implications in the treatment management of patients with high levels of inflammation at high risk of both cardiovascular disease and cancer.

Despite its strengths, our study was not devoid of limitations. First, our primary endpoint of interest (i.e. incident cancer) was partially captured from the two clinical trials' reported adverse events, which is subject to variability and incompleteness. It is possible that some patients had asymptomatic cancer which remained undetected at baseline, and cancer information was at times unspecific and likely variable between study sites. Second, it is unclear to what extent the TNT and IDEAL study participants were screened for previous history of cancers. Although

patients with life-limiting disease were excluded upfront; patients were not specifically screened for cancers. Third, statin use within the UK Biobank was self-reported and only available at baseline. It does not distinguish current and past users, and does not represent a validated claim of statin prescription. Future studies may rely on the UK Biobank's linkage to general practitioner records, where prescriptions of statins are available. However, caution is still warranted in that setting, as a prescription record does not guarantee usage. Fourth, hs-CRP is known to be volatile. However, we were unable to assess other biomarkers of inflammation. Fifth, as neither of the two clinical trials was designed for the purpose of examining cancer incidence as a pre-specified endpoint, factors typically considered in cancer studies have not been collected (e.g. family history). In that regard, given the study designs and indications of TNT and IDEAL, we could not test the statin treated vs. non-treated hypothesis in the discovery cohort, and could only explore that potential using the UK Biobank. Finally, we could not confidently confirm a gene-environment interaction (between rs13210472 and statin use) as in spite of the large genotyping sample size of the UK Biobank, our analyses focused on the specific sub-group of women with CAD taking statins. As such, the study remained underpowered to detect high-order models of risk effects considering that the various interactive risks were of moderate effect sizes. Future consortium-sized samples with over 100,000 genotyped women with and without CAD, and with and without statins use will be necessary to confirm the specificity of the effect of genetic variant rs13210472 on cancer risk to women with CAD using statins.

2.7 Conclusions

The use of statins could interact with the HLA pathway in modulating the risk of incident cancer amongst women with CAD. The association of rs13210472 with incident cancer in women, however, could not be shown to be specific to statin users or to those with a history of CAD due to statistical limitations.

2.8 Funding/Acknowledgements

The TNT and IDEAL trials were sponsored by Pfizer. Pfizer granted access to data but had no role in the design of the current study, the drafting of this report, or the decision to submit these analyses for publication. The current study was funded in part by grants from Genome Canada and Genome Quebec and the Canadian Institutes of Health Research (CIHR). This research has been conducted using the UK Biobank Resource under Application Number 20168. M Sun is supported by a scholarship from FRQS. MA Legault is supported by a scholarship from the CIHR. MP Dubé holds the Canada Research Chair in Precision Medicine Data Analysis, JC Tardif holds the Canada Research Chair in Personalized Medicine.

2.9 Conflicts of Interest

Dr. Dubé has received research support from AstraZeneca, DalCor, Pfizer, GSK, Servier and honoraria from DalCor, Servier and GSK. Dr. Tardif has received research support from Amarin, AstraZeneca, DalCor, Eli-Lilly, Hoffmann-LaRoche, Merck, Pfizer, Sanofi and Servier, and honoraria (to his institution) from Hoffmann-LaRoche, Pfizer, Servier and Valeant. Drs. Tardif and Dubé have an equity interest in DalCor. M Sun has received grant funding from Pfizer, BMS, and Exelixis. All other authors have no conflicts of interest or disclosures to state.

Chapter 3

Somatic Mosaic Chromosomal Alterations and Death of Cardiovascular Disease Causes Among Cancer Survivors: A Cohort Analysis of the UK Biobank

Objective 2: Evaluate clonal hematopoiesis as a biomarker for CVD risk in cancer survivors.

In recent years, several studies have reported on the impact of CH on various health outcomes. CH is a process in which a hematopoietic stem cell that has acquired a somatic mutation passes that mutation onto its daughter cells, eventually forming a clonal subpopulation of cells with a mutation that differs from germline DNA. Epidemiological studies have reported associations of mCA, a type of CH that is characterized by expanded large structural variants, with all-cause mortality and cardiovascular events in the general population. Consequently, we sought to assess the impact of mCA on the risk of death from CVD-related causes and all causes by focusing on patients diagnosed with cancer. We focused exclusively on patients diagnosed with cancer due to an increasing number of cancer survivors worldwide resulting from ageing populations and improvements in early cancer detection and treatment modalities. In fact, it is estimated that over 26 million people in the United States alone will be living with a history of cancer by the year 2040. As we know, cancer patients exhibit an increased predisposition to CVD that is multifactorial. Traditional cardiology markers do not appear to accurately identify individuals at such excess risk of cardiac dysfunction at cancer diagnosis. Therefore, there was a clinical interest to develop a marker that would be better in risk stratifying patients for heightened cardiovascular-related outcomes than current benchmarks.

We assessed and compared the risk of death from CVD causes, coronary artery disease (CAD) causes, and of any cause between carriers of mCA and non-mCA carriers diagnosed with cancer within the UK Biobank. mCA calls were obtained from a dataset return within the UK Biobank (return #3094, application #19808), based on a previously published methodology which relied on DNA genotyping array intensity data and long-range chromosomal phase information inferred from participants.[242,243] Primarily, mCAs were dichotomized as those with $\geq$1 mCA vs. none.

Alternatively, they were also considered as autosomal mCAs, mosaic loss of the Y chromosome (in men), mosaic loss of the X chromosome (in women), and expanded mCAs (cell fraction of ≥10%). Exploratory analyses included sub-analyses according to sex, smoking status, and chemotherapy status. Additional exploratory endpoints included various incident cardiovascular-related endpoints based on a previous publication.[323] Analyses were ultimately assessed across cancer types.

Our analyses revealed interesting results, with a focus on patients diagnosed with cancer within the UK Biobank, we found that carriers of mCA (20.6%) were significantly associated with a higher risk of death from CAD causes, despite adjustment for confounders. In sub-analyses, we found that carriers of mCAs diagnosed with kidney cancer had a 2.0- fold and 3.6-fold higher risk of death from CVD causes and CAD causes, respectively. Women diagnosed with breast cancer with an mCA also demonstrated a 2.5-fold higher risk of death from CAD causes.

Contributions for Article 2: Maxine Sun and Marie-Pierre Dubé conceptualized and designed the study approach. Data curation was conducted by Louis-Philippe Lemieux Perreault and Marie-Pierre Dubé. Johanna Sandoval and Louis-Philippe Lemieux Perreault then optimized data access and data quality assessments. Maxine Sun conducted the analysis, with assistance from Marie-Christyne Cyr. Maxine Sun, Marie-Christyne Cyr and Marie-Pierre Dubé interpretated results and provided critical suggestions for re-analyses. Maxine Sun drafted the first version of the manuscript, where the overall content was supervised, and drafts critically revised by Lambert Busque, Jean-Claude Tardif, and Marie-Pierre Dubé. Administrative and technical support were provided by Johanna Sandoval and Louis-Philippe Lemieux Perreault.

Somatic Mosaic Chromosomal Alterations and Death of Cardiovascular Disease Causes Among Cancer Survivors

Running title: A Cohort Analysis of the UK Biobank

Maxine Sun[1,2,3], Marie-Christyne Cyr[1,2], Johanna Sandoval[1,2], Louis-Philippe Lemieux Perreault[1,2], Lambert Busque[3,4], Jean-Claude Tardif[1,3], Marie-Pierre Dubé[1,2,3]

[1]Montreal Heart Institute, Montreal, Canada
[2]Université de Montréal Beaulieu-Saucier Pharmacogenomics Centre, Montreal, Canada
[3]Department of Medicine, Faculty of Medicine, Université de Montréal, Montreal, Canada
[4]Hôpital Maisonneuve-Rosemont

Cancer Epidemiol Biomarkers Prev. 2023;EPI-22-1290. doi:10.1158/1055-9965.EPI-22-1290

3.1 Abstract

Background: Cancer survivors are at an increased risk of cardiovascular disease (CVD) compared to the general population. We sought to evaluate the impact of mosaic chromosomal alterations (mCAs) on death of CVD causes, coronary artery disease (CAD) causes, and of any cause in patients with a cancer diagnosis.

Methods: The study was a prospective cohort analysis of 48,919 UK Biobank participants with a cancer diagnosis. mCAs were characterized using DNA genotyping array intensity data and long-range chromosomal phase inference. Multivariable Cox regression models were used to ascertain the associations of mCAs. Exploratory endpoints included various incident cardiovascular phenotypes.

Results: Overall, 10,070 individuals (20.6%) carried ≥1 mCA clone. In adjusted analyses, mCA was associated with an increased risk of death of CAD causes (HR: 1.37, 95% CI: 1.09-1.71, *P*=0.006). In sub-analyses, we found that carriers of mCAs diagnosed with kidney cancer had an increased risk of death of CVD causes (HR: 2.03, 95% CI: 1.11-3.72, *P*=0.022) and CAD causes (HR: 3.57, 95% CI: 1.44-8.84, *P*=0.006). Women diagnosed with breast cancer who carried a mCA also had a higher risk of death of CAD causes (HR: 2.46, 95% CI: 1.23-4.92, *P*=0.011).

Conclusions: Among cancer survivors, carriers of any mCA are at an increased risk of CAD death compared to non-carriers. Mechanistic studies should be considered to better ascertain the biological mechanisms underneath the observed associations between mCAs and cardiovascular events for specific cancer types.

Impact: There may be clinical relevance in considering mCAs in patients diagnosed with cancer and undergoing treatment.

3.2 Keywords

clonal hematopoiesis, CHIP, mosaic chromosomal alterations, cardiovascular mortality, UK Biobank

3.3 Introduction

The number of survivors of cancer is growing worldwide due to the aging populations and improvements in early cancer detection and treatment modalities.[324] It is estimated that over 26 million people in the United States alone will be living with a history of cancer by the year 2040.[325,326] Among cancer survivors, a pressing clinical problem is their increased predisposition to cardiovascular disease (CVD)[327,328] and treatment-related cardiac dysfunction[213,329]. Currently, there are no guideline recommendations with respect to CVD screening for patients with cancer, possibly stemming from the lack of CVD-related markers that can better risk-stratify cancer patients beyond existing cardiovascular risk factors for the general population.[330] The identification and development of biomarkers that can eventually be used towards a risk assessment tool for the purpose of discriminating patients diagnosed with cancer who are at a higher risk of CVD may be useful.[330]

Clonal Hematopoiesis (CH) refers to a population of cells derived from a mutated multipotent stem/progenitor cell occurring in the context of aging.[331] CH can be caused by somatic mutation in driver genes[241,332–335] called clonal hematopoiesis of indeterminate potential (CHIP)[336], or by somatic mosaic chromosomal alterations (mCA).[242,243,337–339] Previously, CHIP has been associated with a greater burden of atherosclerotic vessel disease[246], a higher risk of myocardial infarction[340,341], inflammatory response[342], and death of any cause[241,333]. The presence of CHIP has also been associated with treatment-related adverse outcomes in cancer survivors.[343]

Somatic mCAs correspond to large chromosomal gains, loss, or copy neutral losses of heterozygosity which can affect autosomes or sexual chromosomes (bioRxiv 2022.06.24.497515).[344] The most prevalent mCA is the loss of the Y chromosome (mLoY) in aging men[345], which has been associated with all-cause mortality[346,347], Alzheimer's disease[348], autoimmune disease[349], diabetes[346], and cardiovascular events[350].

Given the links between CH and CVD in the normal aging population, and the paucity of data in cancer survivors, we sought to evaluate the impact of mCAs on the risk of death from CVD causes, from cancer, and of any cause among cancer survivors at risk of CVD. Our hypothesis was that carriers of any mCA would be at higher risk of death from CVD causes compared to their mCA non-carriers.

3.4 Materials and Methods

3.4.1 Study Population

The UK Biobank is a large population-based cohort that includes over 500,000 participants aged between 40 and 70 years old, recruited from 2006 to 2010.[351] Baseline interviews regarding their medical history and environmental exposures were conducted in the UK where blood samples for genotyping were obtained and blood analysis performed. Additional health outcome data, including diagnoses of cancer and CVD, have been linked via UK national registries and hospital records managed by the NHS, and genome-wide genotyping of blood-derived DNA was performed by the UK Biobank using 2 genotyping arrays sharing 95% of marker content.[352] The study was approved by the Montreal Heart Institute research ethics committee and complies with the Declaration of Helsinki.

3.4.2 Determination of Mosaic Chromosomal Alterations

As previously described[242,243], allele-specific SNP-array intensity data obtained by genotyping blood-derived DNA from UK Biobank participants were used to call mCAs. Specifically, mCAs were determined from genotype intensities $\log_2$R ratio (LRR) and B-allele frequency (BAF) values, which were used to estimate the total and relative allelic intensities, respectively. Re-phasing was conducted using Eagle2[353] and mCA calling was performed by leveraging long-range phase information searching for allelic imbalances between maternal and paternal allelic fractions across contiguous genomic segments. For the purpose of our study, mCA calls were obtained from dataset Return 3094 from the UK Biobank application 19808.[242,353] From the genetic data of the UK Biobank, 479,435 individuals who passed the sample quality control criteria, including genotypic-phenotypic sex concordance, and without first or second-degree relatives in the dataset were considered (**Figure 3.1**). Of those, we identified 48,919 participants with a diagnosis of cancer before or after the baseline assessment visit based on the cancer register, using ICD-9 and ICD-10 diagnostic codes for bladder, larynx, prostate, corpus uteri, rectum, breast, kidney, non-Hodgkin lymphoma, melanoma of the skin, or lung cancer (**Supplementary Table B.1.1**). These specific cancer types were selected based on a previous publication which identified the top 10 cancer sites with the largest percentage of deaths attributed to CVD.[7]

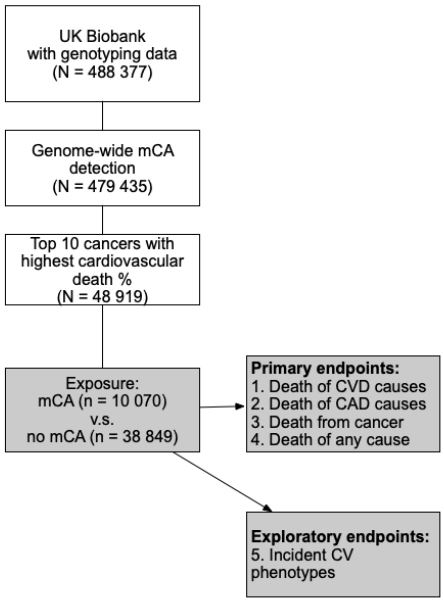


Figure 3.1 Visual Representation of The Study Flow.

mCA: mosaic chromosomal alterations, CVD: cardiovascular disease, CAD: coronary artery disease, CV: cardiovascular

3.4.3 Exposure and Clinical Outcomes

The exposure of interest was the presence of a mCA of any type, which was classified as ≥1 or none. mCAs were also categorized as autosomal mCAs, loss of the X chromosome (mLoX), loss of the Y chromosome (mLoY), and expanded mCAs (defined as mCAs present in at least 10% of peripheral leukocyte DNA indicative of clonal expansion[267]). The study's pre-determined primary endpoints (defined in the **Supplementary Table B.1.2**) consisted of death of CVD causes, death of CAD causes, death from cancer (including recurrent and *de novo* malignancies), and death of any cause based on ICD-10 codes for primary cause of death from the death register records. For each endpoint, the time to death (in years) was calculated from the date of the assessment visit at the time of recruitment into the UK Biobank (baseline) if patients' cancer diagnosis occurred ***before*** baseline, and alternatively, from the date of cancer diagnosis if the cancer diagnosis occurred ***after*** baseline. The median follow-up period was 7.2 years. The number of days between a patient's cancer diagnosis date and baseline was recorded, and it was set as 0 if the cancer diagnosis occurred after baseline. For individuals who were not deceased, the end of follow-up was the last date of death registered based on participant's country of enrolment.

Other exploratory endpoints also considered various *incident* CV phenotypes[354] recorded after patients' cancer diagnosis date (**Supplementary Table B.1.2**). For these exploratory endpoints, the time to the event (i.e., incident CV event of interest or death of CVD causes) was calculated from the date of cancer diagnosis. For those who did not experience any incident CV events,

censor date was set to the date of death of non-CVD causes, or the last hospitalization date known for each patient.

3.4.4 Statistical Analyses

In our primary analyses focusing on our pre-determined primary endpoints, Cox proportional hazards regression models were used to assess the effect of mCA (any vs. none) in the entire cohort of patients with a cancer diagnosis (n=48,919), adjusting for age at baseline, sex, baseline smoking status, treatment with chemotherapy and/or radiotherapy, the number of days between prevalent cancer diagnosis and date of baseline, and principal components for genetic ancestry (PC, 1-10). Sensitivity analyses considered the additional adjustment for baseline measures of alcohol status, use of lipid-lowering medication, type 2 diabetes, hypertension, and body-mass-index status. The influence of mCA on death of CVD and CAD causes was further tested through competing-risks regression models after adjusting for death of other causes.[355]

The association of mCA types on the primary endpoints of interest was also assessed. Exploratory sub-analyses were conducted according to established risk factors of CH, including smoking status, chemotherapy status, and age dichotomized at the 65-years old cut-off. Corresponding mCA-by-smoking, chemotherapy, and age interaction terms were generated for these analyses, respectively. We further explored the effect of mCA with respect to our primary endpoints according to cancer type. Cox regression models were also used for exploratory analyses of incident CV phenotypes. Finally, to determine whether the potential effects of mCA on outcomes were specific to cancer survivors, a mCA-by-cancer status interaction term was assessed with respect to each of the primary endpoints in the overall population including individuals without any history of cancer. The significance threshold for the primary analyses was revised to account for the multiplicity of tests by using a Bonferroni correction (P<0.05/4=0.0125). All other sub-analyses were exploratory in nature and meant to assist in the interpretation of the primary analysis results. All analyses were performed using the survival package in R (version 4.1.2, Bird Hippie). All analytical and summary reports were produced with *gtsummary* (version 1.6.1).[356]

Data availability

The data analyzed in this study were obtained from the UK Biobank and are available at https://biobank.ndph.ox.ac.uk/ukb/. For the purpose of our study, mCA calls were obtained from dataset Return 3094 from the UK Biobank application 19808.

3.5 Results

3.5.1 Baseline Characteristics and mCA Prevalence

Overall, 48,919 participants were diagnosed with bladder, larynx, corpus uteri, prostate, rectal, breast, kidney, non-Hodgkin lymphoma, melanoma, or lung cancer within the UK Biobank (**Table**

3.1). The presence of ≥1 mCA clone was observed in 20.6% (10,070/48,919) of patients **(Supplementary Figure B.2.1)**. Of those, 2,432 were autosomal carriers and 1,910 had mCA in ≥10% of peripheral leukocytes, defined as an expanded mCA clone. Whereas most carriers had only one mCA, 812 individuals carried ≥2 non-overlapping mCAs. In general, those with mCAs were older (median 64 vs. 61 years), where the prevalence of mCA increased with increasing age categories (**Supplementary Figure B.2.1**). Expectedly, 73.6% of patients with mCA were men, likely because the majority of mCA carriers was due to mLoY (n=6,558). In comparison, mCA due to mLoX was observed in 1,641 patients (16.3%). Amongst patients older than 65 years old, the prevalence of mCA according to cancer types was the lowest for breast cancer (14%) and the highest for larynx cancer (43%, **Supplementary Figure B.2.1**).

Table 3.1 Summary Descriptives of Patients With a Cancer Diagnosis Stratified According To Mosaic Chromosomal Alteration Status.

Characteristic	**N**	**Overall**, N = 48,919[1]	**No mCA**, N = 38,849[1]	**Any mCA**, N = 10,070[1]	**p-value**[2]
Age at bsl	48,919				<0.001
N		48,919	38,849	10,070	
N missing		0	0	0	
Mean (SD)		60 (7)	60 (7)	63 (5)	
Median (IQR)		62 (56, 65)	61 (55, 65)	64 (61, 67)	
Range		40, 71	40, 70	40, 71	
Men	48,919	22,403 (45.8%)	14,989 (38.6%)	7,414 (73.6%)	<0.001
Smoking status	48,624				<0.001
Never smoker		23,865 (49.1%)	19,888 (51.5%)	3,977 (39.7%)	
Previous smoker		19,274 (39.6%)	14,755 (38.2%)	4,519 (45.1%)	
Current smoker		5,485 (11.3%)	3,970 (10.3%)	1,515 (15.1%)	
Unknown		295	236	59	
Bladder cancer	48,919	1,743 (3.6%)	1,191 (3.1%)	552 (5.5%)	<0.001
Larynx cancer	48,919	375 (0.8%)	245 (0.6%)	130 (1.3%)	<0.001
Corpus Uteri cancer	48,919	2,331 (4.8%)	2,108 (5.4%)	223 (2.2%)	<0.001
Prostate cancer	48,919	13,274 (27.1%)	8,970 (23.1%)	4,304 (42.7%)	<0.001

Characteristic	N	Overall, N = 48,919[1]	No mCA, N = 38,849[1]	Any mCA, N = 10,070[1]	p-value[2]
Breast cancer	48,919	17,304 (35.4%)	15,678 (40.4%)	1,626 (16.1%)	<0.001
Rectal cancer	48,919	2,268 (4.6%)	1,731 (4.5%)	537 (5.3%)	<0.001
Kidney cancer	48,919	1,910 (3.9%)	1,472 (3.8%)	438 (4.3%)	0.010
Non-Hodgkin lymphoma	48,919	3,038 (6.2%)	2,189 (5.6%)	849 (8.4%)	<0.001
Melanoma	48,919	4,912 (10.0%)	4,031 (10.4%)	881 (8.7%)	<0.001
Lung cancer	48,919	4,215 (8.6%)	3,074 (7.9%)	1,141 (11.3%)	<0.001
Chemotherapy	48,919				<0.001
None		37,512 (76.7%)	29,610 (76.2%)	7,902 (78.5%)	
Yes		11,407 (23.3%)	9,239 (23.8%)	2,168 (21.5%)	
Radiotherapy	48,919				0.036
None		45,846 (93.7%)	36,454 (93.8%)	9,392 (93.3%)	
Yes		3,073 (6.3%)	2,395 (6.2%)	678 (6.7%)	
Prevalent hypertension	48,919	6,077 (12.4%)	4,569 (11.8%)	1,508 (15.0%)	<0.001
Autosomal mCA	48,919				<0.001
Ref.		46,487 (95.0%)	38,849 (100.0%)	7,638 (75.8%)	
Autosomal		2,432 (5.0%)	0 (0.0%)	2,432 (24.2%)	
Loss of X	48,919				<0.001

Characteristic	N	Overall, N = 48,919[1]	No mCA, N = 38,849[1]	Any mCA, N = 10,070[1]	p-value[2]
Ref.		47,278 (96.6%)	38,849 (100.0%)	8,429 (83.7%)	
Loss of X		1,641 (3.4%)	0 (0.0%)	1,641 (16.3%)	
Loss of Y	48,919				<0.001
Ref.		42,361 (86.6%)	38,849 (100.0%)	3,512 (34.9%)	
Loss of Y		6,558 (13.4%)	0 (0.0%)	6,558 (65.1%)	
Expanded mCA	48,919				<0.001
Ref.		47,009 (96.1%)	38,849 (100.0%)	8,160 (81.0%)	
Expanded mCA		1,910 (3.9%)	0 (0.0%)	1,910 (19.0%)	
Death of CVD causes	48,919	813 (1.7%)	550 (1.4%)	263 (2.6%)	<0.001
Death of CAD causes	48,919	368 (0.8%)	226 (0.6%)	142 (1.4%)	<0.001
Death from cancer	48,919	8,433 (17.2%)	6,254 (16.1%)	2,179 (21.6%)	<0.001
Death of any cause	48,919	10,632 (21.7%)	7,785 (20.0%)	2,847 (28.3%)	<0.001

[1] n (%)

[2] Wilcoxon rank sum test; Pearson's Chi-squared test

bsl.: baseline, CAD: coronary artery disease, CVD: cardiovascular disease, IQR: interquartile range, mCA: mosaic chromosomal alteration, Ref.: referent category, SD: standard deviation

3.5.2 mCA and the Risk of Death of CVD Causes, Coronary Artery Disease (CAD) Causes, From Cancer, and Any Cause

In adjusted Cox regression analyses, mCA was associated with an increased risk of death of CAD causes (hazard ratio [HR]: 1.37, 95% confidence interval [CI]: 1.09-1.71, *P*=0.006, **Supplementary Figure B.2.2**) and death of any cause (HR: 1.07, 95% CI: 1.02-1.12, *P*=0.005, **Figure 3.2**). These effects were maintained in sensitivity analyses after the additional adjustment for baseline measures of alcohol status, use of lipid-lowering medication, type 2 diabetes, hypertension, and body-mass-index (**Supplementary Table B.1.3**). Multivariable competing-risks regression analyses replicated the effect of mCA on the risk of death of CAD causes, despite adjustment for death of non-CAD causes and other covariates ($HR_{competing\text{-}risk}$: 1.33, 95% CI: 1.06-1.68, *P*=0.014, **Supplementary Table B.1.4**). The effect of mCA with respect to death of CVD causes ($P_{interaction}$=0.044) and of any cause ($P_{interaction}$=0.011) modulated according to smoking status (**Supplementary Table B.1.5**). In contrast, the impact of mCA was not influenced by use of chemotherapy for any of our primary endpoints (all $P_{interaction} \geq 0.05$, **Supplementary Table B.1.6**).

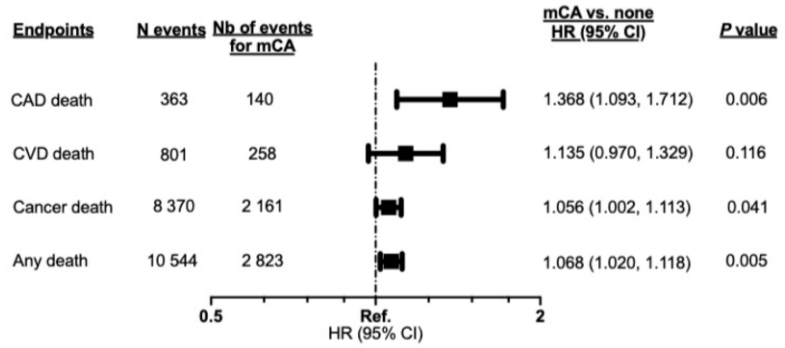


Figure 3.2 Multivariable Cox Regression Models Evaluating the Effect of Mosaic Chromosomal Alterations On Death of Cardiovascular Disease Causes, of Coronary Artery Disease Causes, From Cancer, and Death of Any Cause.

All models are adjusted for age at baseline, sex, smoking status, chemotherapy, radiotherapy, number of days between the date of recruitment and the date of cancer diagnosis, and principal components 1 to 10. CAD: coronary artery disease, CI: confidence interval, CVD: cardiovascular disease, HR: hazard ratio, mCA: mosaic chromosomal alterations, Ref.: referent category (1.0)

Different types of mCAs had different effects on the study endpoints. Notably, autosomal mCAs were associated with death of CVD causes, with death from cancer, as well as with death of any cause (**Supplementary Table** B.1.7). mCA due to mLoX was associated with death of CAD causes (**Supplementary Table** B.1.8), and mCA due to mLoY was associated with death of any cause (**Supplementary Table** B.1.9). Expanded mCAs taken individually had no statistically significant impact on the primary endpoints (**Supplementary Table** B.1.10). When considering the study population of 479,435 participants with, and without a history of any cancer, the interaction terms between cancer and mCA were significant for death from cancer and death of any cause (**Supplementary Table** B.1.11).

3.5.3 mCA and the Risk of Death of CVD Causes, CAD Causes, From Cancer, and Any Cause According to Cancer Site

In multivariable analyses where the effect of mCA was assessed in subgroups according to cancer type, we found that carriers of mCAs diagnosed with kidney cancer had an increased risk of death of CVD causes (HR: 2.03, 95% CI: 1.11-3.72, *P*=0.022) and CAD causes (HR: 3.57, 95% CI: 1.44-8.84, *P*=0.006, **Supplementary Tables** B.1.12 **&** B.1.13) compared to their mCA non-carriers counterparts diagnosed with kidney cancer. Women diagnosed with breast cancer who carried a mCA also had a higher risk of death of CAD causes (HR: 2.46, 95% CI: 1.23-4.92, *P*=0.011). mCA

had no significant impact on deaths from cancer in either cancer type (**Supplementary Table B.1.14**). Women carriers of mCA diagnosed with corpus uteri cancer had an increased risk of death of any cause (HR: 1.35, 95% CI: 1.01-1.80, *P*=0.042, **Supplementary Table B.1.15**).

3.5.4 mCA and the Risk of Incident CV Endpoints (Exploratory Analyses)

In the overall cohort of cancer survivors (n=48,919), carriers of mCA had a higher risk of incident ST-elevation myocardial infarction (STEMI) (HR: 1.19, 95% CI: 1.03-1.37, *P*=0.022) and peripheral vascular disease (HR: 1.17, 95% CI: 1.04-1.31, *P*=0.007, **Figure 3.3**) than mCA non-carriers. Similar findings were obtained when the analyses were restricted to those aged ≥65 years old (**Supplementary Table B.1.16**).

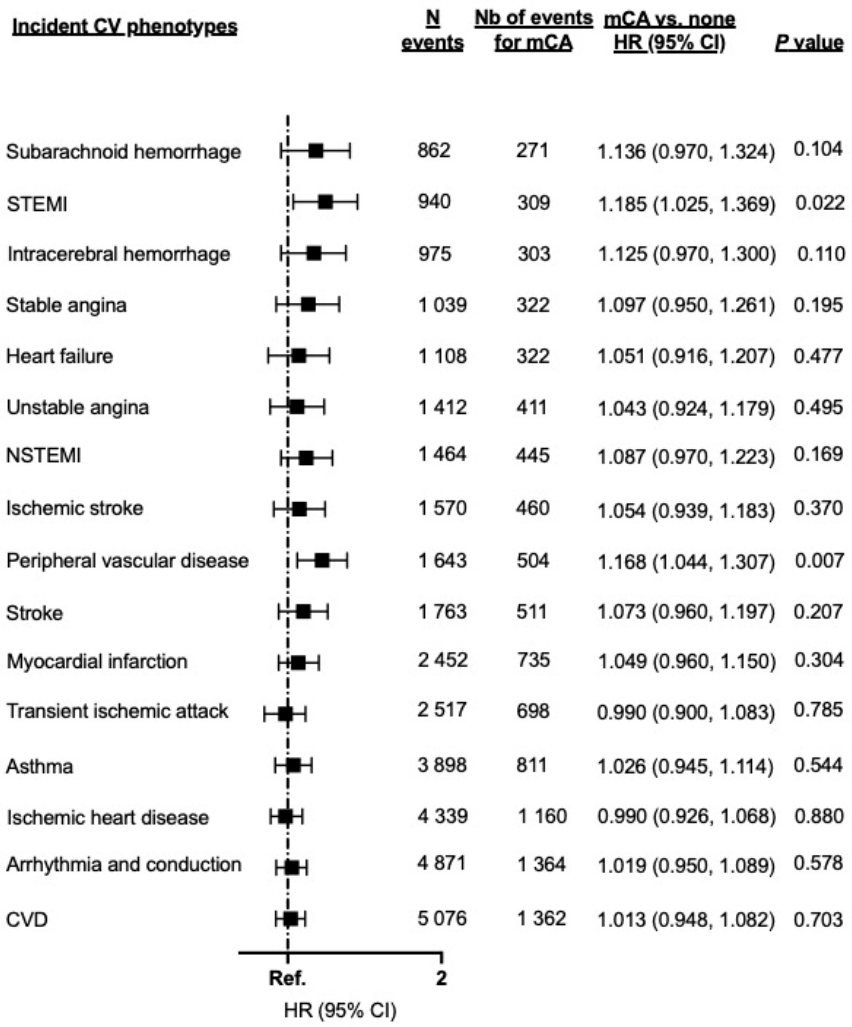


Figure 3.3 Multivariable Cox Regression Models Evaluating the Effect of Mosaic Chromosomal Alterations on Incident Cardiovascular Phenotypes.

All models are adjusted for age at baseline, sex, smoking status, chemotherapy, radiotherapy, number of days between the date of recruitment and the date of cancer diagnosis, and principal components 1 to 10. CVD: cardiovascular disease, CI: confidence interval, HR: hazard ratio, mCA: mosaic chromosomal alteration, MI: myocardial infarction, NSTEMI: non-ST-elevation myocardial infarction, STEMI: ST-elevation myocardial infarction, TIA: transient ischemic attack.

3.6 Discussion

Here, we assessed the impact of somatic mCAs, a subtype of CH, on the risks of death of CVD causes, CAD causes, from cancer, and death of any cause focusing exclusively on individuals diagnosed with one of 10 cancer types known to have the highest reported rates of CVD deaths[7].

Primarily, we found that carriers of mCAs had a notable increased risk of death of CAD causes in all patients diagnosed with cancer. The increased risk of death of CAD causes for carriers of mCAs persisted even after adjusting for death of other causes as a competing event, which included cancer-related deaths. While mCAs were also associated with death of any causes overall, a finding likely driven by death from cancer-related causes, its impact was more subdued. Furthermore, sub-analyses did not reveal a significant association between mCA and cancer-specific deaths for any of the cancer types. On the other hand, mCAs portended to higher risks of both death of CVD and CAD causes in patients diagnosed with kidney cancer and higher risks of death of CAD causes for women diagnosed with breast cancer. Our findings also revealed that some of the associations were more impactful according to mCA subtypes. For example, autosomal mCAs were associated with the risk of death of CVD causes, whereas mCA due to the mLoX was associated with the risk of death of CAD causes, and mCA due to the mLoY was associated with the risk of death of any causes.

In addition, exploratory analyses suggest that carriers of mCAs were at higher risk of incident STEMI and peripheral vascular disease; and for cancer survivors aged >65 years, at a higher risk of incident NSTEMI, stable angina, intracerebral hemorrhage, and subarachnoid hemorrhage. In comparison, a previous mCA study focusing on MI and stroke did not find any significant associations when focusing on carriers of expanded mCAs, regardless of cancer status.[242]

In comparison to previous studies that focused on the general population and found that mCAs were not associated with CAD[243,267], the current study suggests that any mCA carrier status may have prognostic ability of other CV phenotypes in cancer survivors. Past research that has found a relationship between CH and CAD were focused on CHIP mutations. One such study correlated mutations in *DNMT3A* and *TET2* with chronic heart failure in patients with STEMI (n=485).[357] In another study which included approximately 50,000 samples with exome sequencing data, Zekavat et al.[358] found that carriers of CHIP mutations had higher risks of peripheral artery disease.

Second, when considering cancer survivors, it is possible that cancer-directed treatments may have perpetuated the influence of mCA on outcomes.[343] However, our sub-analyses which included a chemotherapy-by-mCA interaction term failed to reach statistical significance, albeit an observed near 2-fold increased risk of death of CAD causes for mCA carriers. Previous work have supported the hypothesis that chemotherapy may lead to clonal expansion of mutations in apoptosis or DNA repair gens (*TP53*, *PPM1D*)[343,359], and that CHIP mutations caused by these genes could worsen the risk of atherosclerosis[358]. It is possible that chemotherapy usage may not be well-captured in the current database. Furthermore, without knowledge on tumor aggressiveness in the current database, it was not possible to exclude the presence of a selection bias where carriers of mCAs may be at a more advanced stage of the disease or have a more aggressive tumor histology, which may make them more likely to undergo additional lines of

chemotherapy/radiotherapy, which could have contributed to their risks of CVD. Because other studies that examined the influence of chemotherapy on CH were concentrated on CHIP mutations, it will be critical to evaluate whether cancer-directed treatments have any additive or synergistic influence on mCAs. Further research will be needed to understand the impact of mCAs on CVD outcomes following modern standards of care, such as treatment with immune-checkpoint inhibitors. In that scenario, mCA carriers may benefit from an in-depth preventive cardiology approach prior to treatment where multi-disciplinary cardio-oncology care could mitigate such risks.

Third, we found that the effect of mCA on the risks of death of CVD and CAD causes, as well as of death of any cause was specific to previous smokers. This upholds the findings of a recent study which established strong causal associations between smoking and mCAs, and postulated that smoking may contribute to the selection of clones bearing somatic mutations.[360] Indeed, we found that cancer survivors who are carriers of mCAs may face added risks if they smoked in the past. Additional work is required to better evaluate how smoking habits can impact mCAs and the downstream consequences of such mutations, especially in patients diagnosed with cancer.

The generalizability of our findings may be limited as the influence of mCA was not consistent across all examined cancer types. This was not entirely unexpected given the heterogeneous biology of cancers included and distinct mechanisms of mCAs with respect to these cancers and treatments that patients received. Nonetheless, the associations observed within this study may inform future studies focusing on the effect of mCA and the risks of death from CVD causes for specific cancer types.

Our study has other limitations. First, our study did not consider CHIP mutations. While the association of CHIP mutations with CVD has been established, it has not been formally tested in patients diagnosed exclusively with non-hematologic cancer. There is reason to believe that carriers of CHIP mutations in addition to being diagnosed with cancer may face higher risks of CVD. In addition, the synergistic effect of both mCA and CHIP mutation carriers amongst cancer survivors will be of interest to explore in future studies. Second, It has been established that detectable mLoY and mLoX chromosomes increase with age, with more age-related pathologies in older individuals than in younger individuals.[361,362] Unexpectedly, the association between the mLoY and death of CVD causes was not significant in the current analysis that focused on cancer survivors. The non-significant association may have been related to the mLoY definition we used, where a larger percentage threshold of blood cells lacking chromosome Y may be more relevant. For example, a recent analysis found that regardless of their cancer status, men with mLoY chromosome in 40% or more of leukocytes displayed a 31% increased risk of dying from any disease related to the circulatory system based on survival data from the UK Biobank.[362] In contrast, the relative risk per 1% increase of detectable mLoY and the risk of death from CVD causes was more subdued (HR: 1.0054, P=0.001).[362] We did also note that mLoY chromosome

was rather impactful for cancer-specific mortality, which may be relevant in considering the significant association found between mLoY and death of any cause. Third, mCAs were detected using allele-specific SNP microarray intensity data obtained by genotyping blood-derived DNA samples from participants, as previously described[242,243], allowing more mCA events than previously developed methodologies[337,338,363]. Alternatively, the use of whole-genome (WGS) or whole-exome sequencing (WES) to detect mosaicism have also been proposed. However, such approaches can be limited in their ability to detect mCAs at low cell fractions.[364] Also, mCA status was evaluated only at study entry. Reasonably, clones carrying a mCA may undergo rapid changes following various environmental exposures (e.g., smoking, cancer-directed therapy). Under this premise, serial measurements of mCA may better help inform the dose-response effect of such factors and mCA. In addition, while adjustment was made for chemotherapy and radiotherapy treatment, some individuals may have undergone such treatments in the outpatient setting, which would not have been properly captured using inpatient procedural codes. In that regard, we also did not have information on the number of cancer-directed treatment cycles, or dosage of therapies in these patients. Furthermore, with regards to variable adjustment, other lifestyle and environmental factors may have also been considered. Future studies focusing specifically on how modifiable lifestyle behaviors (e.g., sleep, diet, physical activity) or the use of medication (e.g., anti-hypertensives, lipid-lowering) may impact the development of clonal hematopoiesis can have considerable clinical relevance to cancer survivors. Finally, although the current study is the first to assess the relationship between mCAs and CVD in cancer survivors, we did not have available independent cohorts to conduct an external replication of the observed associations.

3.7 Conclusions

Our study showed that mCAs are associated with higher risks of death of CAD causes among cancer survivors. Future studies should focus on specific cancer types and their treatments to better ascertain the effect of mCA on the risk of CVD and evaluate if they constitute a useful biomarker in the management of cancer survivors.

3.8 Acknowledgements

We thank the UK Biobank for providing the data under Application Number 20168

3.9 Funding

M. Sun was funded by a doctoral scholarship from the Fonds de Recherche du Québec - Santé (FRQS; #274750). M.-P. Dubé and J.-C. Tardif hold Canada Research Chairs. This project was funded in part by the Health Collaboration Acceleration Fund (FACS) from the Government of

Quebec (J.-C. Tardif and M.-P. Dubé) and by the Canadian Institutes of Health Research (CIHR; #409752) (L. Busque, J.-C. Tardif, M.-P. Dubé).

3.10 Disclosures/Conflicts of interest:

M-P.D. has received personal fees and reports minor equity interest in Dalcor Pharmaceuticals. M-P.D. has a patent Methods for Treating or Preventing Cardiovascular Disorders and Lowering Risk of Cardiovascular Events issued to Dalcor Pharmaceuticals, no royalties received; a patent Genetic Markers for Predicting Responsiveness to Therapy with HDL-Raising or HDL Mimicking Agent issued to Dalcor Pharmaceuticals, no royalties received; and a patent Methods for using low dose colchicine after myocardial infarction, assigned to the Montreal Heart Institute. J-C.T. reports research grants from Amarin, AstraZeneca, Ceapro, DalCor Pharmaceuticals, Esperion, Ionis, Novartis, Pfizer, and RegenXBio; honoraria from AstraZeneca, DalCor Pharmaceuticals, HLS Pharmaceuticals, Pendopharm and Pfizer; minor equity interest from Dalcor Pharmaceuticals; and authorship on a patent Methods for Treating or Preventing Cardiovascular Disorders and Lowering Risk of Cardiovascular Events issued to Dalcor Pharmaceuticals, no royalties received; a patent Genetic Markers for Predicting Responsiveness to Therapy with HDL-Raising or HDL Mimicking Agent issued to Dalcor Pharmaceuticals, no royalties received; a pending patent Early administration of low-dose colchicine after myocardial infarction, and a patent Methods for using low-dose colchicine after myocardial infarction, assigned to the Montreal Heart Institute (J-C.T. has waived his rights in the colchicine patents). Other authors have nothing to declare.

Chapter 4

The Combined Effect of Clonal Hematopoiesis of Indeterminate Potential (CHIP) and Mosaic Chromosomal Alterations (mCA) On Cardiovascular Health Outcomes in Patients Diagnosed With Cancer: An Analysis of the UK Biobank

Objective 2: Evaluate clonal hematopoiesis as a biomarker for CVD risk in cancer survivors.

As a continuation of our analyses on mCAs, recent studies have also shown that CH, characterized by hematologic expansion of clones bearing pathogenic single nucleotide polymorphisms or small insertions/deletions, otherwise known as CHIP (e.g., *DNMT3A*, *TET2*), is associated with a greater burden of atherosclerotic vessel disease[246], higher risk of myocardial infarction[273,341], and death of any cause[243], possibly via mechanisms of inflammation[271].

In a follow-up study, we assessed the risk of death from CVD causes, CAD causes, and of any cause between carriers of CHIP and non-CHIP carriers diagnosed with cancer, excluding prevalent hematological cancers, within the UK Biobank, after gaining access to their whole-exome sequencing data. In addition, we further evaluated the distinct impact of harboring both types of CH, defined as those with CHIP and mCA co-occurrence relative to individuals who had only either type of CH, and to individuals without any CH. An additional outcome was added to this study, which included the risk of cardiovascular-related emergency room admission. Our descriptive statistics corresponded to what others previously observed, where the prevalence of CHIP mutations increased with increasing age. The same trend was also noted for patients with both types of CH. The most common CHIP mutations were *DNMT3A*, *ASXL1*, and *TET2*. CHIP mutations were most prevalent in patients with non-Hodgkin lymphoma and least prevalent in women with corpus uteri cancer. However, the prevalent mCA co-occurrence landscape did not correspond to CHIP mutations prevalence across cancer sites, suggesting a heterogeneous nature of cancer types, that was shown to be independent of age. Overall, carriers of CHIP mutations had higher risk of death from CVD causes compared to non-CHIP carriers, after adjusting for potential confounders, including receipt of chemotherapy and/or radiotherapy. In the evaluation of

contrasts, we also found that carriers of CHIP mutations alone were driving the effect on the risk of death from CVD causes more so than those with mCAs alone. These results somewhat reaffirm what previous studies have shown with regards to CHIP and mCAs, where phenotypes for the two appear to be mostly independent, despite some overlap in the germline genetic predisposition.[365]

Taken together, the clinical implications from both articles 2 and 3 suggest that identifying the CH carrier status, be it mCA or CHIP, in patients diagnosed with cancer can lead to proactive conversations with cardiologists about treatment strategies, thereby facilitating a more preventative approach.

Contributions for Article 3: Maxine Sun and Marie-Pierre Dubé conceptualized and designed the study approach. Data curation was conducted by Johanna Sandoval, Louis-Philippe Lemieux Perreault and Marie-Pierre Dubé. Johanna Sandoval and Louis-Philippe Lemieux Perreault developed and tested a pipeline for CHIP calling using whole-exome sequencing data obtained from the UK Biobank, before optimizing data access and quality control. Maxine Sun and Johanna Sandoval participated in the analysis and interpretation of data. Maxine Sun and Marie-Pierre Dubé interpreted results and provided critical suggestions for re-analyses. Maxine Sun drafted the first version of the manuscript, where the overall content was supervised, and drafts critically revised by Lambert Busque, Jean-Claude Tardif, and Marie-Pierre Dubé. Administrative and technical support were provided by Johanna Sandoval and Louis-Philippe Lemieux Perreault.

The Combined Effect of Clonal Hematopoiesis of Indeterminate Potential (CHIP) and Mosaic Chromosomal Alterations (mCA) On Cardiovascular Health Outcomes in Patients Diagnosed With Cancer: An Analysis of the UK Biobank

Maxine Sun[1,2,3], Johanna Sandoval[1,2], Louis-Philippe Lemieux Perreault[1,2], Amina Barhdadi[1,2,3], Marie-Christyne Cyr[1,2], Lambert Busque[3,4], Jean-Claude Tardif[1,3], Marie-Pierre Dubé[1,2,3]

[1]Montreal Heart Institute, Montreal, Canada
[2]Université de Montréal Beaulieu-Saucier Pharmacogenomics Centre, Montreal, Canada
[3]Department of Medicine, Faculty of Medicine, Université de Montréal, Montreal, Canada
[4]Hôpital Maisonneuve-Rosemont

(To be submitted)

4.1 Abstract

Background: Individuals diagnosed with cancer are at higher risk of CV disease. Recently, clonal hematopoiesis (CH) has been shown to be associated with CV-related endpoints in the general population. In the current study, we sought to assess the effect of clonal hematopoiesis of indeterminate potential (CHIP) on the risk of short- and long-term CV health and survival outcomes in a patient cohort comprising exclusively of individuals diagnosed with cancer. Furthermore, we sought to investigate the clinical significance of the co-occurrence of mosaic chromosomal alterations (mCAs) with CHIP gene mutations on CV-related and survival outcomes.

Methods: Our cohort focused on 19,566 individuals diagnosed with cancer. Whole-exome sequencing data from blood-derived DNA of the UK Biobank participants were used to identify CHIP mutations, and genotyping array data was used to define mCAs. The primary endpoints of interest were time to CV-related emergency admission, time to death from CV disease causes, time to death of coronary artery disease (CAD) causes, and time to death from any causes. Cox regression models were used with adjustment for age, sex, smoking status, chemotherapy, radiotherapy, prevalent CVD, number of days between date of recruitment and date of cancer diagnosis, and inferred genetic ancestry.

Results: A total of 1,166 cancer patients carried at least one CHIP mutation (6.0%). Overall, 302 individuals had both CHIP and mCA (1.2%), whereas 4,512 had either CHIP or mCA alone (CHIP: 864, mCA: 3,648, 23.1%). In multivariable analyses, carriers of CHIP mutations had a higher risk of death from CVD causes compared to CHIP non-carriers (hazard ratio [HR]: 1.66, 95% confidence interval [CI]: 1.14–2.42, *P*=0.008). CHIP alone (HR: 1.74, 95% CI: 1.12–2.87, *P*=0.014) resulted in higher risk of death from CVD causes compared to individuals with neither type of CH.

Conclusions: Our results showed that CHIP mutations are independently associated with a higher risk of death from CVD causes in patients diagnosed with cancer, with or without the co-occurrence of mCAs.

4.2 Keywords

Clonal hematopoiesis, mosaic chromosomal alterations, cardiovascular disease, cancer

4.3 Introduction

There are currently an estimated 17 million individuals alive in the US who have been diagnosed with cancer. Based on current trends, this number is projected to exceed 22 million by the year 2030.[366] Continuous improvements in anti-cancer therapeutics and screening strategies have undoubtedly led to prolonged cancer survivorship. However, the overall success of these achievements is tempered by a simultaneous increase in cardiovascular (CV)- related deaths among cancer survivors. Several reports indicate that adult cancer survivors are at higher risk of heart failure and several other CV disease (CVD) later in life compared to those without cancer.[367] Additionally, they are at higher risk of death from CVD causes[7,368]. While cytotoxic chemotherapy, immunomodulatory agents, and radiation are established contributing factors to the higher risk of cardiac dysfunction among patients diagnosed with cancer, they do not fully explain it.[367,369] Given this backdrop, there is an opportunity to evaluate biomarkers that can better monitor the risk of CVD in cancer patients and improve our understanding of the mechanisms attributed to CV dysfunction in cancer survivors.

Recently, driver gene mutations with a variant allele frequency (VAF) of $\geq$2% in the peripheral blood, even in the absence of clinical criteria of hematological malignancy, have emerged as a major independent risk factor in atherosclerotic CVD, thrombosis, and heart failure in the general population.[241,246,247,271,370] This phenomenon is termed clonal hematopoiesis of indeterminate potential (CHIP). Clonal hematopoiesis (CH) occurs when somatic mutations in leukemogenic genes provide a selective advantage to hematopoietic stem and progenitor cells (HSPCs), resulting in the expansion of mutant blood cells.[371] Commonly mutated CHIP genes have been characterized as mediators of epigenetic modifications (*DMNT3A*, *TET2*, and *ASXL1*), hematopoietic cytokine signaling (*JAK2*), DNA damage repair (*PPM1D*, and *TP53*), as well as messenger RNA splicing (*SF3B1*, and *SRSF2*). On the other hand, CH defined by mosaic chromosomal alterations (mCAs)[242,243], classified as gains, losses, and copy-number neutral loss of heterozygosity (CNN LOH) have also been reported, with notable associations with death from coronary artery disease (CAD) among patients diagnosed with cancer[372].

In the current study, we sought to assess the effect of common CHIP mutations on the risk of short- and long-term CV health and survival outcomes in a patient cohort comprising exclusively of individuals diagnosed with cancer. Furthermore, we sought to investigate the clinical

significance of mCAs in the context of co-occurring CHIP gene mutations and assess its impact on CV-related and survival outcomes.

4.4 Methods

4.4.1 Data Source

The UK Biobank project is a prospective cohort study with genetic and phenotypic data based on approximately 500,000 participants from across the United Kingdom, aged between 40 and 60 years old at study recruitment.[352] Baseline assessments were conducted across assessment centers, which included collection of genome-wide genotyping samples on all participants. Of ~200,633 sequenced exomes (release: October 2020), we focused on participants consenting to genetic analyses after exclusion criteria (prevalent hematologic cancers, individuals without genotypic-phenotypic sex concordance, one of each pair of 1st and 2nd degree relatives at random), with a select cancer diagnosis captured via ICD-10 diagnostic codes based on linked cancer registry data, as previously described[372] (bladder, larynx, prostate, corpus uteri, rectal, breast, kidney, melanoma, non-Hodgkin lymphoma, and lung, **Supplementary Table C.1.1**). These cancer types were selected based on a previous publication which identified top cancer sites with the largest percentage of deaths attributed to CVD.[7] This resulted in 19,566 participants for analyses.

4.4.2 CHIP Calling

Previously, whole-exome sequencing (WES) was captured with the IDT gem Exome Research Panel v1.0 including supplemental probes. Multiplexed samples were sequenced with dual-indexed 75x75 bp paired-end reads on the Illumina NovaSeq 6000 platform using S2 (initial 50k release) and S4 flow cells (all subsequent samples).[373] We implemented the GATK best practices for somatic short variant in tumor only mode. The OQFE CRAM alignment files were used to call for somatic small variants using MuTect2 (v4.2.6.1). Common germline variants and sequencing artifacts were excluded as previously described.[246,258] Sample annotation of any CHIP, performed according to previous approach[246,258], occurred when MuTect2 identified the presence of ≥1 pre-specified list of pathogenic somatic variants, namely *ASXL1*, *CBL*, *DNMT3A*, *GNAS*, *GNB1*, *JAK2*, *PPM1D*, *SF3B1*, *SRSF2*, *TET2* and *TP53*.[342]

4.4.3 mCA Calling

As previously described[242,243], allele-specific SNP-array intensity data obtained by genotyping blood-derived DNA from UK Biobank participants were used to call mCAs. Specifically, mCAs were determined from genotype intensities $\log_2$R ratio (LRR) and B-allele frequency (BAF) values, which were used to estimate the total and relative allelic intensities, respectively. Re-phasing was conducted using Eagle2[353] and mCA calling was performed by leveraging long-range phase information searching for allelic imbalances between maternal and paternal allelic fractions

across contiguous genomic segments. For the purpose of our study, mCA calls were obtained from dataset return #3094 from the UK Biobank application #19808.[242,353]

4.4.4 Phenotypic Definitions

Our primary endpoints consisted of the following: 1) time to CV-related emergency room admission, 2) time to death from CVD causes, 3) time to death from CAD causes, and 4) time to death from any causes. Cause of death determination was based on ICD-10 codes for primary cause of death per death register records. For each endpoint, the time to death (in years) was calculated from the date of the assessment visit at the time of recruitment into the UK Biobank (baseline) if patients had received their cancer diagnosis before baseline. Alternatively, if the cancer diagnosis occurred after baseline, the time to death was calculated from the date of cancer diagnosis. To overcome time bias incurred due to prevalent cancer diagnoses, we calculated the number of days between a patient's prevalent cancer diagnosis date and study recruitment, setting it as 0 if the cancer diagnosis occurred after baseline. For individuals who were not deceased, the end of follow-up was the last date of death registered based on participant's country of enrolment.

Other exploratory endpoints were considered and defined as CV events[323] that occurred after a patients' cancer diagnosis date (**Supplementary Table C.1.2**). For these exploratory endpoints, the time to the event (i.e., incident CV event of interest or death of CVD causes) was calculated from the date of cancer diagnosis. For those who did not experience any incident CV events, censor date was set to the date of death of non-CVD causes, or the last hospitalization date known for each patient.

4.4.5 Statistical Analysis

In our primary analyses focusing on pre-determined endpoints, Cox proportional hazards regression models were used to assess the effect of CHIP (any vs. none) in the entire cohort of patients with a cancer diagnosis (n=19,566), adjusting for age, sex, smoking status, chemotherapy, radiotherapy, prevalent CVD, number of days between date of recruitment and date of cancer diagnosis, and genetic ancestry principal components 1 through 10. Alternatively, the effect of the most prevalent CHIP mutations on endpoints were also considered. In sensitivity analyses, competing-risk regression models were used to further account for death of other causes including cancer-specific mortality, as described by Fine and Gray.[374]

To assess the influence of gene mutations that coincide with mCAs on our primary outcomes, we constructed predefined contrasts of CH status, which was determined using both CHIP and mCA. We achieved this by adjusting the categorical variable's reference level to enable comparisons across different levels using alternative reference groups. Specifically, we set the reference groups to individuals without CHIP or mCA (1) and those with mCA alone (2). Corresponding mCA-

by-CHIP interaction terms were also generated for all of the primary endpoints. Finally, in supporting analyses, Cox regression models were used to assess the impact of CHIP (any vs. none) on the risk of incident CV phenotypes.

All analyses were performed using the Jupyter notebooks (version 5.0) on the DNAnexus Platform of the UK Biobank with the PYTHON_R feature (Python 3.6.5 libraries and R 4.1.3 libraries).[375] All analytical and summary reports were produced with gtsummary (version 1.6.1).[376] Kaplan-Meier survival curves and corresponding log-rank *P* comparisons were generated using ggsurvfit (version 0.3.0).[377]

4.5 Results

4.5.1 Patient Cohort and CHIP Characteristics

Overall, a total of 19,566 individuals diagnosed with bladder, larynx, corpus uteri, prostate, breast, rectal, kidney, non-Hodgkin lymphoma, and lung cancers were included (**Table 4.1**). A total of 1,166 patients carried at least one CHIP mutation (6.0%), where the most prevalent gene mutations were *DNMT3A* (50%), *TET2* (19%), and *ASXL1* (18%, **Figure 4.1**). Carriers of CHIP mutations were significantly older (mean: 63 vs. 60 years) and more frequently men (49.9% vs. 44.7%) than CHIP non-carriers (both *P*<0.001, **Table 4.1**).

Table 4.1 Overall Descriptives of Cancer Survivors Stratified According To Clonal Hematopoiesis of Indeterminate Potential.

Characteristic	**Overall**, N = 19,566[1]	**None**, N = 18,400[1]	**Any CHIP**, N = 1,166[1]	**p-value**[2]
Age at bsl				<0.001
N	19,566	18,400	1,166	
Mean (SD)	60 (7)	60 (7)	63 (5)	
Median (IQR)	62 (56, 65)	62 (56, 65)	64 (60, 67)	
Range	40, 70	40, 70	41, 70	
Sex				<0.001
Female	10,757 (55.0%)	10,173 (55.3%)	584 (50.1%)	
Male	8,809 (45.0%)	8,227 (44.7%)	582 (49.9%)	
Smoking status				0.105
Never smoker	9,725 (50.0%)	9,184 (50.2%)	541 (47.0%)	
Previous smoker	7,720 (39.7%)	7,241 (39.6%)	479 (41.7%)	
Current smoker	2,004 (10.3%)	1,874 (10.2%)	130 (11.3%)	
Unknown	117	101	16	
Prevalent CVD	4,275 (21.8%)	3,984 (21.7%)	291 (25.0%)	0.008
Bladder cancer	653 (3.3%)	604 (3.3%)	49 (4.2%)	0.090
Larynx cancer	135 (0.7%)	122 (0.7%)	13 (1.1%)	0.071

Characteristic	**Overall**, N = 19,566[1]	**None**, N = 18,400[1]	**Any CHIP**, N = 1,166[1]	**p-value**[2]
Corpus Uteri cancer	944 (4.8%)	903 (4.9%)	41 (3.5%)	0.032
Prostate cancer	5,462 (27.9%)	5,112 (27.8%)	350 (30.0%)	0.099
Breast cancer	7,249 (37.0%)	6,865 (37.3%)	384 (32.9%)	0.003
Rectal cancer	901 (4.6%)	862 (4.7%)	39 (3.3%)	0.034
Kidney cancer	767 (3.9%)	721 (3.9%)	46 (3.9%)	0.964
Lung cancer	1,464 (7.5%)	1,336 (7.3%)	128 (11.0%)	<0.001
CV-related ER admission	2,761 (14.1%)	2,566 (13.9%)	195 (16.7%)	0.008
Death from CVD	291 (1.5%)	257 (1.4%)	34 (2.9%)	<0.001
Death from CAD	121 (0.6%)	107 (0.6%)	14 (1.2%)	0.009
Death from cancer	3,157 (16.1%)	2,876 (15.6%)	281 (24.1%)	<0.001
Any death	3,942 (20.1%)	3,580 (19.5%)	362 (31.0%)	<0.001

[1]n (%)

[2]Pearson's Chi-squared test; Wilcoxon rank sum test; Fisher's exact test
CAD: coronary artery disease, CHIP: clonal hematopoiesis of indeterminate potential, CVD: cardiovascular disease, ER: emergency room, IQR: interquartile range, SD: standard deviation

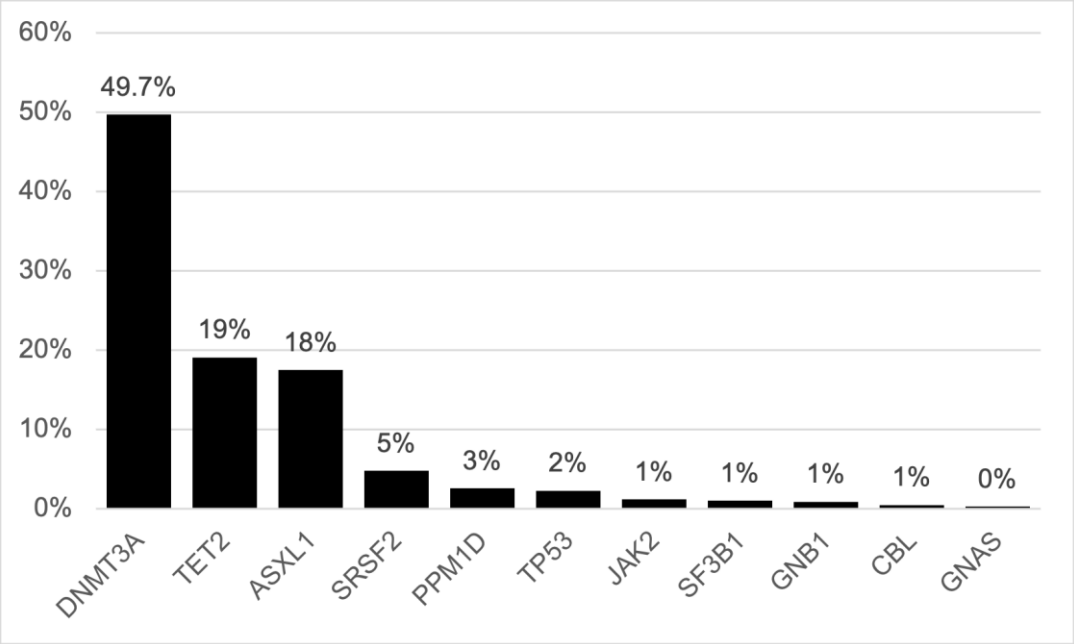


Figure 4.1 Distribution of CHIP Mutations Among Individuals Diagnosed With Cancer.

Numbers at the top of the bars represent n per CHIP mutation type. CHIP: clonal hematopoiesis of indeterminate potential

4.5.2 Co-occurrence of CHIP and mCA

Overall, 302 individuals had both CHIP and mCA (1.2%), whereas 4,512 had either CHIP or mCA alone (CHIP: 864, mCA: 3,648, 23.1%). Both the occurrence of CHIP mutations and co-occurrence of CHIP and mCA increased with increasing age (**Figure 4.2**). Individuals with either CHIP or mCA alone ($P<2.2\text{x}10^{-16}$), or those with both acquired gene mutations and chromosomal alterations ($P=4.0\text{x}10^{-14}$) were significantly older than those without any type of CH (**Suppl. Figure C.2.1**). The co-occurrence of mCA and CHIP was notably higher for the most common CHIP mutations (*DNMT3A*, *ASXL1*, and *TET2* (**Suppl. Figure C.2.2**). The prevalence of CHIP mutations was highest in patients diagnosed with non-Hodgkin lymphoma, and lowest in those diagnosed with corpus uteri cancer (**Suppl. Figure C.2.3**), possibly linked with cancer-specific age distributions (**Suppl. Figure C.2.3**). However, co-occurrence of mCA across cancer sites was not consistently associated with increasing prevalence of CHIP mutations, thereby reflecting the heterogenous nature of cancer types, independent of age.

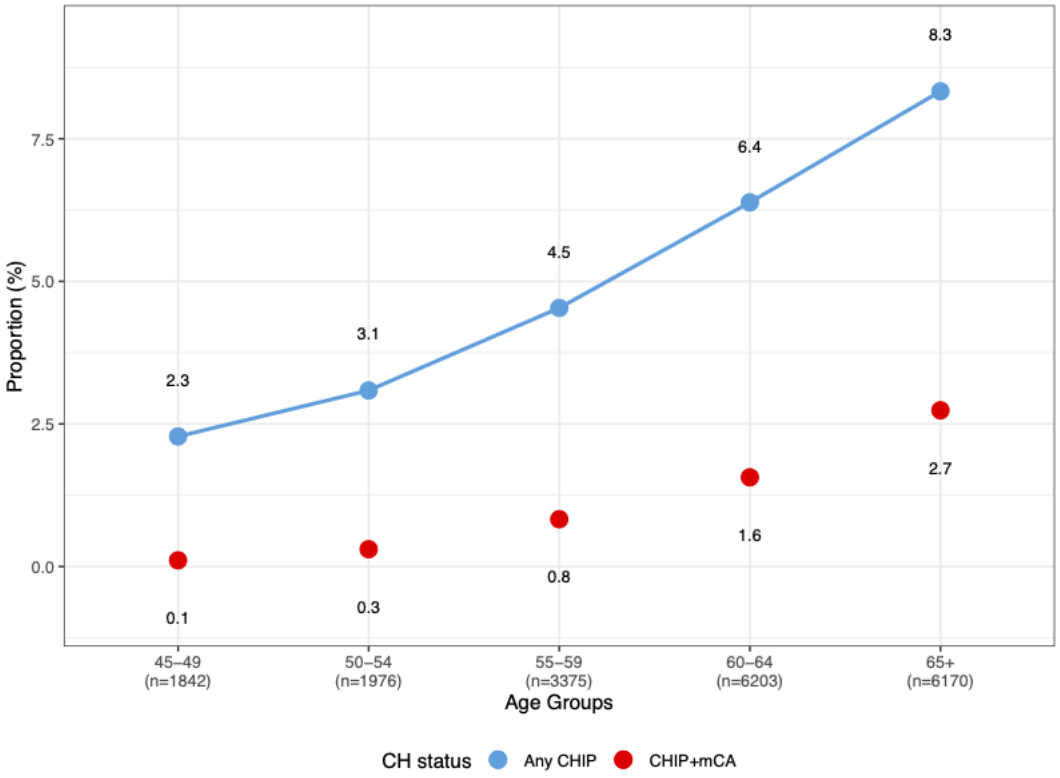


Figure 4.2 Proportion of Patients (%) With Detected CHIP Mutations (Blue), With mCAs (Red) According To Increasing Age Categories.

CH: clonal hematopoiesis, CHIP: clonal hematopoiesis of indeterminate potential, mCA: mosaic chromosomal alterations

4.5.3 CHIP and Death From CVD Causes

Carriers of CHIP mutations had a higher risk of death from CVD causes compared to CHIP non-carriers (hazard ratio [HR]: 1.66, 95% confidence interval [CI]: 1.14–2.42, *P*=0.008), after adjusting for potential confounders (**Suppl. Figure** C.2.4, **Table 4.2**), which held true in competing-risks analyses ($HR_{competing\text{-}risk}$: 1.52, 95% CI: 1.05–2.20, *P*=0.026, **Suppl. Table** C.1.2). The association was notable for carriers of *TET2* (HR: 3.15, 95% CI: 1.40–7.09, *P*=0.005) and *ASXL1* (HR: 2.19, 95% CI: 1.08–4.44, *P*=0.029, **Suppl. Table** C.1.3) but not for *DNMT3A* mutation carriers. In analyses that considered contrasts of CHIP and mCA co-occurrence, CHIP alone (HR: 1.74, 95% CI: 1.12–2.87, *P*=0.014) resulted in higher risk of death from CVD causes compared to non-CH carriers, whereas mCA alone did not (HR: 0.91, 95% CI: 0.68–1.22, *P*=0.531, **Table 4.3**). Moreover, patients with CHIP alone also experienced higher risks of death from CVD causes compared to carriers of mCA alone (HR: 1.91, 95% CI: 1.18–3.10, *P*=0.009, **Table 4.3**).

Table 4.2 Multivariable Cox Regression Analyses On The Influence of CHIP (Any Vs. None) On the Risks of Death From CV-Related Causes and Emergency Room Admission.

Characteristic	N	Event N	HR[1]	95% CI[1]	p-value
Time to CV death					
None	18,299	255	—	—	
Any CHIP	1,150	31	1.662	1.143, 2.416	0.008
Time to CAD death					
None	18,299	107	—	—	
Any CHIP	1,150	11	1.339	0.716, 2.505	0.361
Time to any death					
None	18,299	3550	—	—	
Any CHIP	1,150	354	1.450	1.299, 1.618	<0.001

Models adjusted for age at baseline, sex, smoking status, chemotherapy, radiotherapy, prevalent CVD, number of days between date of recruitment and date of cancer diagnosis, and genetic principal components 1-10. CAD: coronary artery disease, CHIP: clonal hematopoiesis of indeterminate potential, CI: confidence interval, CV: cardiovascular, HR: hazard ratio.

Table 4.3 Multivariable Cox Regression Analyses On the Influence of CHIP and mCA Status Co-Occurrence Contrasts On the Risks of Death.

Characteristic	N	Event N	HR[1]	95% CI[1]	p-value
Time to CV death					
Neither (ref.)	14,671	185	–	–	–
CHIP alone	852	22	1.740	1.117, 2.871	0.014
mCA alone	3,268	70	0.911	0.682, 1.218	0.531
CHIP & mCA	298	9	1.460	0.742, 2.871	0.273
mCA alone (ref.)	3,268	70	–	–	–
CHIP alone	852	22	1.909	1.175, 3.101	0.009
Time to CAD death					
Neither (ref.)	14,671	71	–	–	–
CHIP alone	852	6	1.178	0.511, 2.718	0.700
mCA alone	3,628	36	1.105	0.725, 1.684	0.642
CHIP & mCA	298	5	1.844	0.731, 4.657	0.195
mCA alone (ref.)	3,628	36	–	–	–
CHIP alone	852	6	1.066	0.446, 2.549	0.885
Time to any death					
Neither (ref.)	14.671	2,632	–	–	–
CHIP alone	852	236	1.443	1.262, 1.649	<0.001
mCA alone	3,268	918	1.079	0.996, 1.169	0.063
CHIP & mCA	298	118	1.825	1.515, 2.200	<0.001
mCA alone (ref.)	3,268	918	–	–	–
CHIP alone	852	236	1.337	1.157, 1.545	<0.001

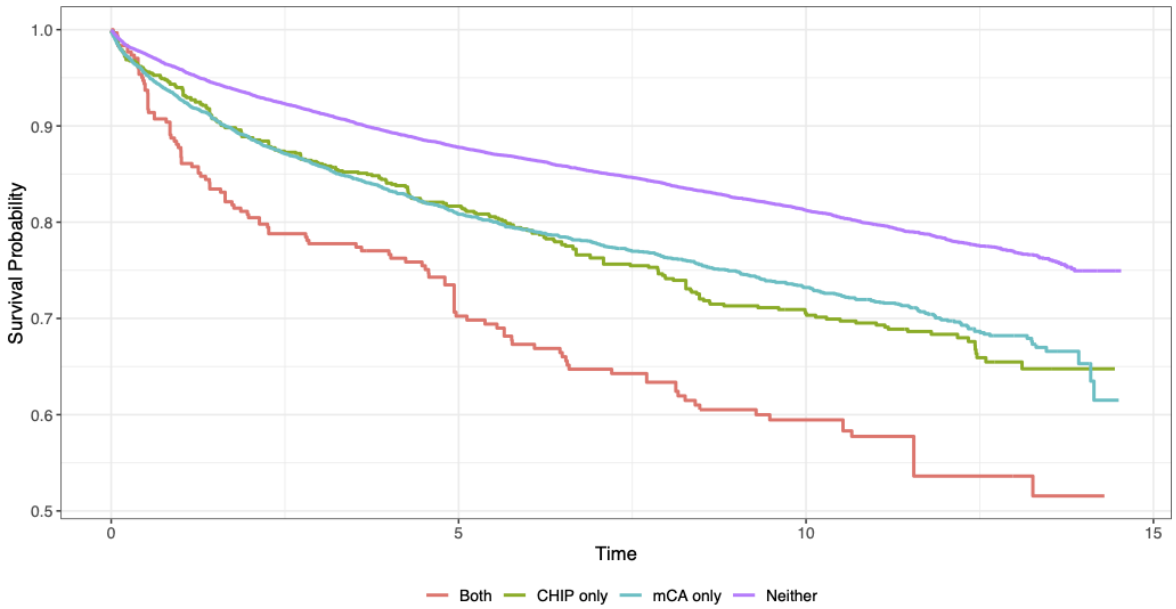
Models adjusted for age at baseline, sex, smoking status, chemotherapy, radiotherapy, prevalent CVD, number of days between date of recruitment and date of cancer diagnosis, and genotyping principal components 1-10. CAD: coronary artery disease, CHIP: clonal hematopoiesis of indeterminate potential, CI: confidence interval, CV: cardiovascular, HR: hazard ratio, mCA: mosaic chromosomal alterations, ref.: referent category.

4.5.4 CHIP and Death From CAD Causes and ER Admission Due to CV-Related Causes

Carriers of CHIP were not significantly associated with the risk of death from CAD causes (HR: 1.34, 95% CI: 0.72–2.51, *P*=0.361, **Suppl. Figure C.2.4**, **Table 4.2**; $HR_{competing\text{-}risk}$: 1.21, 95% CI: 0.67–2.19, *P*=0.530, **Suppl. Table C.1.2**). The lack of significance for the risk of death from CAD causes was sustained across the most common CHIP mutations, except for carriers *ASXL1* who portended to higher risk of death from CAD causes (HR: 3.18, 95% CI: 1.29–7.82, *P*=0.012, **Suppl. Table C.1.3**). Interaction term analyses focusing on death from CVD and CAD causes showed no significant CHIP-by-mCA effect ($P_{interaction}$=0.844 for death of CVD causes; $P_{interaction}$=0.589 for death of CAD causes, data not shown).

4.5.5 CHIP and Overall Mortality

Carriers of CHIP mutations had higher risk of death from any causes compared to non-carriers (HR: 1.45, 95% CI: 1.30–1.62, *P*<0.001, **Figure 4.3**, **Table 4.2**). Both *TET2* and *ASXL1* were significantly associated with a higher risk of death of any cause compared to their non-*TET2* and non-*ASXL1* counterparts (**Suppl. Table C.1.3**). Compared to persons with neither types of CH, those with CHIP alone (HR: 1.44, 95% CI: 1.26–1.65, *P*<0.001) were more likely to succumb to death of any cause (**Table 4.3**), whereas it was not significant for those with mCA alone (HR: 1.08, 95% CI: 1.00–1.17, *P*=0.063). While the combined effect of CHIP and mCA on the risk of any cause of death was the largest relative to those without CH (HR: 1.83, 95% CI: 1.52–2.20, *P*<0.001), it appeared to be principally driven by CHIP (CHIP alone vs. mCA alone, HR: 1.34, 95% CI: 1.16–1.55, *P*<0.001, **Table 4.3**). However, the CHIP-by-mCA interaction term for death of any cause was not statistically significant ($P_{interaction}$ =0.182, data not shown).



At Risk				
Both	302	173	105	0
CHIP only	864	601	357	0
mCA only	3648	2372	1353	0
Neither	14752	10637	6458	0
Events				
Both	0	86	110	120
CHIP only	1	154	224	242
mCA only	8	673	854	924
Neither	15	1733	2387	2656

Figure 4.3 Kaplan-Meier Overall Survival Curves Stratified According To CHIP and mCA Status Where Individuals With Both Types of CH Are Depicted in Red, With CHIP Only In Green, With mCA Only in Blue, and Without Any CH in Purple.

Log-rank P-value comparisons for CHIP+mCA vs. CHIP only (P=0.023), CHIP+mCA vs. mCA only (P=0.002), CHIP+mCA vs. none (P<0.001), CHIP only vs. mCA only (P=0.584), CHIP only vs. none (P<0.001), mCA only vs. none (P<0.001). CH: clonal hematopoiesis, CHIP: clonal hematopoiesis of indeterminate potential, mCA: mosaic chromosomal alterations

4.5.6 CHIP and Exploratory Incident CV Phenotypes

CHIP mutations were significantly associated with several exploratory incident CV phenotypes (**Suppl. Table** C.1.4). Notable differences were also observed for the influence of CHIP mutations on the risk of death from CV-related, non-CV related, and any causes across cancer types (**Suppl. Figure** C.2.5).

4.6 Discussion

Several years ago, researchers discovered an enrichment of monoclonality in atherosclerotic plaques.[244] At the time, they were unaware that individuals with a higher prevalence of peripheral blood lymphocytes had a greater atherosclerotic burden, partly due to inadequate ability of ageing hematopoietic stem cells to prevent and repair DNA replication errors.[239] Clonal mutational expansion has historically been an established hallmark of carcinogenesis. With the advent of large population-based repositories that contain blood DNA, CHIP has recently been shown to be associated with CAD in the general population.[246,271,273]

The primary objective of the current study was to assess whether carriers of common somatic CHIP mutations would have a higher risk of atherosclerosis, even when focusing on patients with a cancer diagnosis. Using a large prospective cohort, our results confirmed the impact of CHIP mutations on death from CVD, with effect sizes comparable to what was previously reported in the general population (current HR: 1.67, previous HR: 1.27 to 1.90[246,273]). We also found that patients with *TET2* mutations had markedly higher risks of death from CVD causes (HR: 3.15) than those without mutations. Previous experimental designs have confirmed that atherosclerotic plaque size increases in hypercholesterolemic mice transplanted with bone marrow containing various CHIP mutations, including loss of *TET2*.[271] In this regard, macrophages with CHIP mutations tend to drive the production of inflammatory cytokines, including interleukin-1β (IL-1β) and IL-6 via NLRP3 inflammasome activation in *TET2* CHIP.[271,378] Indeed, human studies show that elevated circulating cytokines are prevalent amongst those with CHIP, such as plasma IL-6 in *TET2* CHIP carriers, thereby supporting the inflammasome-cytokine production pathway (NLRP3-IL-1β-IL-6) as an underlying pathway from CHIP to CAD.[258] Besides atherosclerotic plaque formation, thrombosis is also a major cause of CAD. Our analysis confirmed that carriers of CHIP mutations had a higher risk of peripheral vascular disease in cancer survivors, as previously observed in the general population for carriers of *TP53*.[269]

We also assessed the combined effect of CHIP and mCAs on the risk of CV dysfunction. When CH was defined as the co-occurrence of CHIP and mCA, its impact on the risk of death from CVD causes was no longer significant. However, significance re-emerged when we compared carriers of CHIP mutations alone vs. carriers of mCAs alone. A similar observation may be made with respect to all-cause mortality. Overall, these results suggest the absence of additive risks when mCAs are considered alongside CHIP, as previously observed in a Japanese cohort.[268] In support of these results, the interaction terms of CHIP-by-mCA also did not reach significance for all primary endpoints examined, suggesting non-synergistic effect of both CH types on CV-related and overall mortality. A recent evaluation of phenotypic overlap between CHIP and types of mCAs found that CHIP was largely distinct from mCA phenotypes, with more than 80% of CHIP carriers having no co-occurring mCA, even in age-adjusted analyses (odds ratio 1.02, P=0.27).[365] However,

a nominal shared genetic predisposition between CHIP and mosaic loss of the Y chromosome was observed (*P*=0.014), despite distinctive CH phenotypes.[365]

Previous studies that characterized the presentation of mCAs in the context of CHIP gene mutations have suggested that genomic imbalances gain a selection advantage through complex patterns of co-occurrence,[379] including recurring composite genotypes that indicate co-operative or epistatic interactions, as well as loss of gatekeeper function (e.g., *TP53*) accompanied by multiple chromosomal aneuploidies.[379] To gain a better understanding of how both types of CH evolve and co-occur over time, it will be necessary to accurately identify gene mutations and chromosomal alterations simultaneously at the single-cell level.

The current study has important clinical implications, given the growing number of cancer survivors in the US and worldwide.[366] Adult cancer survivors have up to 40% higher risk of CVD than individuals without a history of cancer, and common CV risk factors do not entirely account for this excess risk.[7,367] Cancer-related therapies such as chemotherapy and radiotherapy are likely to contribute to the incidence of cardiotoxicity in cancer patients. However, the current study has adjusted for receipt of both treatment types, and CHIP remained independently associated with an increased risk of CVD mortality, reaffirming an underlying shared biological mechanism between CVD and cancer, irrespective of treatment type. It is also worth mentioning that according to previous studies, CV risk varies widely across cancer types at baseline (e.g., CVD HR for breast cancer: 1.32 vs. 2.37 for lung cancer)[367], which could then be exacerbated by the presence of clonal somatic mutations, as evidenced in the variable effect of CHIP on such outcomes in the current study, across cancer sites.

Ultimately, as intent-to-treat cancer therapies leading to CVD-related complications will continue to augment in upcoming years, it is imperative to encourage multi-disciplinary care for patients who require both cardiology and oncology treatments. At the health care systems level, primary prevention of CVD in cancer patients at diagnosis will be crucial for the purpose of reducing the burden of multi-morbidities. The use of CH as a marker of cardiac dysfunction in patients with cancer can prompt discussions with cardiologists about potential intervention and management strategies at the outset.

The present study has some limitations that are worth mentioning. First, the UK Biobank is predominantly composed of individuals of European ancestry and limited to individuals living in the UK, so replication of our findings in diverse populations is necessary to ensure the validity of the results in other populations. Second, our analyses may have been underpowered for some endpoints, notably the risk of death from CAD causes and for the exploratory analysis with the CV endpoints. Another limitation related to our exploratory endpoints is potential misclassification of case status, which can occur in administrative linked data. Misclassification can erroneously result in the loss of statistical power for discovery and bias results toward the null. Fourth,

selection bias from participation to the UK Biobank, differential loss of follow-up status, missingness in covariates may be present given the large-scale nature of biobank repositories. Fifth, while we adjusted for patients who received chemotherapy and radiotherapy, individuals who received such interventions in the outpatient setting would not be captured in the inpatient data used. Furthermore, we could not adjust for treatment cycles or dosage-response relationships. Finally, DNA genotyping was obtained at recruitment centers at study baseline. It is likely that clones appear over time, and may also grow or shrink depending on varying exposures. Longitudinal studies characterizing the mutational fitness and clone evolution over time will be critical in better understanding the opportunities for risk stratification and potential therapeutics.

4.7 Conclusions

In summary, our results showed that CHIP mutations are independently associated with a higher risk of death from CVD causes in patients diagnosed with cancer, with or without the co-occurrence of mCAs.

Chapter 5

Other Scientific Contributions

5.1 High-Sensitivity CRP Is Associated With Clonal Hematopoiesis of Indeterminate Potential

At the beginning of my doctoral studies, my director (Dre. Marie-Pierre Dubé) and co-director (Dr. Jean-Claude Tardif) invited me to work on a project that had just begun taking place in collaboration with Dr. Lambert Busque, a hematologist at the *Maisonneuve-Rosemont* hospital focusing on the association between hs-CRP levels and CHIP mutations using data from the Montreal Heart Institute (MHI) Biobank. At that time, there had been little confirmation of CHIP and inflammation association, if any, in human studies. We used data derived from plasma and DNA samples from participants that had been recruited at the MHI Biobank (n=1,887), aged 70 years old and above. We investigated the link between 11 genes previously implicated in CHIP, and hs-CRP levels. Our findings showed that carriers of any CHIP mutations had 21% higher hs-CRP levels compared to non-CHIP carriers. The correlation between CHIP and hs-CRP was notable in patients with prevalent CAD (n=1,359). That study was among the first to confirm the CHIP and hs-CRP relationship using human data.

Author contributions: Lambert Busque, Jean-Claude Tardif, Marie-Pierre Dubé conceived the study. Manuel Buscarlet performed and analyzed the sequencing data with the support of Sami Ayachi and Vincent Bourgoin. Manuel Buscarlet generated figures. Maxine Sun performed statistical analyses with the support of Yassamin Feroz Zada and Sylvie Provost under the supervision of Marie-Pierre Dubé. Maxine Sun drafted the first version of the manuscript. Lambert Busque, Jean-Claude Tardif, and Marie-Pierre Dubé provided critical suggestions for subsequent revisions. Luigina Mollica, Marlies Meisel, Reinhard Hinterleitner and Bana Jabri revised the manuscript and provided insightful suggestions.

High-Sensitivity CRP Is Associated With Clonal Hematopoiesis of Indeterminate Potential

Lambert Busque[1,2,3,*¶], Maxine Sun[3,4], Manuel Buscarlet[1], Sami Ayachi[1,3],Yassamin Feroz Zada[4], Sylvie Provost[4],Vincent Bourgoin[1], Luigina Mollica[1,2,3], Marlies Meisel[5], Reinhard Hinterleitner[5], Bana Jabri[6], Marie-Pierre Dubé[3,4*] and Jean-Claude Tardif[3,4*¶]

[1]Research Center, [2]Hematology Division, Hôpital Maisonneuve-Rosemont, [3]Université de Montréal, Montréal, Canada

[4]Beaulieu-Saucier Pharmacogenomics Center, Montreal Heart Institute, Montréal, Canada.

[5]Department of Immunology, University of Pittsburgh School of Medicine, Pittsburgh, PA 15261, USA

[6]Department of Medicine, University of Chicago, Chicago, Illinois 60637, USA University of Pittsburgh School of Medicine

Blood Adv. 2020;4(11):2430-2438. doi:10.1182/bloodadvances.2019000770

5.1.1 Abstract

Clonal Hematopoiesis of Indeterminate Potential (CHIP) is predictive of hematological cancers and cardiovascular diseases, but the etiology of CHIP initiation and clonal expansion is unknown. Several lines of evidence suggest that proinflammatory cytokines may favor mutated Hematopoietic Stem Cell (HSC) expansion. To investigate the potential link between inflammation and CHIP, we performed targeted deep sequencing of 11 genes previously implicated in CHIP in 1887 subjects aged over 70 years-old from the Montreal Heart Institute (MHI) Biobank, of which 1359 had prior coronary artery disease (CAD) and 528 controls did not. We assessed association of CHIP with log (ln) transformed high-sensitivity C-reactive protein (hs-CRP), a validated biomarker of inflammation. CHIP was identified in 427 of the 1887 subjects (22.6%). CHIP mutations were more frequently identified in *DNMT3A* (11.6%) and *TET2* (6.1%), with a higher proportion of *TET2* mutations occurring in controls than in patients with CAD (9.0% vs 4.9%, $P<0.001$). CHIP carriers had 21% higher hs-CRP compared to their non-carrier counterparts ($e^{\beta}=1.21$, 95% confidence interval [CI]: 1.08-1.36, $P=0.001$). A similar effect was observed in the subgroup of patients with known CAD ($e^{\beta}=1.22$, 95% CI: 1.06-1.41, $P=0.005$). These findings confirm the association between inflammation and CHIP. This may open investigational avenues aimed at documenting mechanisms linking inflammation to clonal progression and ultimately support prevention interventions to attenuate CHIP's impact on cardiovascular disease and cancer.

5.1.2 Introduction

Clonal Hematopoiesis (CH) occurring in normal aging subjects, initially suggested by X-chromosome inactivation (XCI) studies,[380,381] is caused by acquired mutations in genes recurrently mutated in hematological cancers[241,332,334,382] and in non-driver candidates.[235,382] CH prevalence increases significantly in patients aged above 60 years and confers an increased risk of progression to hematological cancers and cardiovascular diseases.[235,241,246,382] The precise risk associated with the presence of CH in healthy individuals is uncertain, hence the creation of a clinical entity named Clonal Hematopoiesis of Indeterminate Potential (CHIP).[336]

Little is known about the etiology of clone initiation and clonal expansion. Genetic predisposition is controversial. Zink *et al.* showed an association between a small germline deletion in intron 3 of the telomerase reverse transcription (*TERT*) gene[235] and CH. Studying a cohort of sib-ships, we identified a significant 2.7-fold increase in the familial risk for mutation in *TET2* but not in *DNMT3A.*[383] However, a strong genetic contribution to CH was recently refuted when the concordance of CH was studied in monozygotic and dizygotic twin pairs[384,385].

The association between CHIP and both cardiovascular[241,246,382] and chronic pulmonary disease[235,383] raised the possibility that age-associated chronic inflammation[386] may be a key common denominator between these medical conditions.[387] Studies in mice have supported the role of inflammation in clonal expansion of mutated Hematopoietic Stem Cells (HSC). Abegunde *et al*. demonstrated that a proinflammatory environment supported by TNF-α promotes the expansion of *Tet2* mutant clones in mice,[388] and we found that inflammation was a key driver of preleukemic myeloproliferation in *Tet2* deficient mice ($Tet2^{-/-}$).[389] In humans, Jaiswal *et al*. demonstrated higher IL-8 levels in patients with CHIP and cardiovascular disease in a subset of 12 with *TET2* mutations. [246] A trend towards increased levels of IL-6 was also observed in patients with CHIP.[390]

We report here a statistically significant correlation between high-sensitivity C-Reactive Protein (hs-CRP), a validated and routinely available biomarker of inflammation[391], and CHIP.

5.1.3 Methods

5.1.3.1 Study Population

We selected 1940 study subjects from participants of the Montreal Heart Institute (MHI) biobank, an ongoing prospective cohort including 23,000 individuals for the purpose of clinical and genetic research.[392] Participants are recruited on a voluntary basis during any hospital visit, regardless of the presence or stage of heart disease. All subjects undergo a medical questionnaire by a research nurse and their electronic chart is reviewed. DNA, plasma and serum are collected at baseline.

For the purpose of the current study, patients with coronary artery disease (CAD) were defined as those with a prior history of myocardial infarction (MI), percutaneous coronary intervention (PCI), and/or coronary artery bypass graft (CABG) surgery. Subjects also needed to be aged 70 years and above. The protocol was approved by MHI's ethics committee and performed in accordance to the Helsinki declaration.

5.1.3.2 hs-CRP

hs-CRP concentration was measured by quantitative immuno-nephelometric analysis on a Dimension Vista 500 Intelligent Lab System (Siemens Healthineers).

5.1.3.3 CHIP Determination By Next Generation Sequencing (NGS)

Subject's DNA (n=1940) was sequenced at high coverage (95% >500x) on an Ion Proton sequencer using a custom Ampliseq "CHIP" panel (Thermo Fisher Scientific) designed to target the top 11 genes reported in CHIP (*ASXL1, CBL*, *DNMT3A*, *GNAS*, *GNB1*, *JAK2* (chr9:5073674- 5073808), *PPM1D*, *SF3B1* (exons 14-16), *SRSF2*, *TET2* and *TP53*)[241,334,382,383] with 202 amplicons covering 38.49 kb. The panel coverage, specificity and sensitivity were validated (see **Supplemental Figure D.1.1 & D.1.2**). Mutations were considered present if the Variant Allele Frequency (VAF) was ≥ 2% as defined by Steensma *et al.*[336] Base calling, alignment (hg19) and variant calling were performed in instrument by TorrentServer v5.8.0 (Thermo Fisher Scientific). Subsequently, mutations were annotated and filtered using IonReporter v5.8 (Thermo Fisher Scientific) and only exonic and splice site mutations with a minor allele frequency (MAF) ≤0.001 were kept for further annotation. Frameshift, nonsense, in-frame deletions or insertions, splice sites and predicted consequential missense mutations (based on ClinVar, FATHMM or PolyPhen) were considered significant. In absence of non-hematological tissue for germline status confirmation, mutations with a VAF of 50% or 100% (± 4%) were considered potentially germline and excluded (n=9) except if they co-occurred with other somatic mutations (n=9) (see **Supplemental Table D.1.1**).

5.1.3.4 Statistical Analyses

Bivariate associations were evaluated using the Fisher's Exact test and the Kruskal-Wallis test for categorical and continuously coded variables, respectively. hs-CRP was modelled as ln(hs-CRP) due to its non-normal distribution. The geometric mean was calculated by taking the antilog of the mean of the log-transformed hs-CRP data, as previously described.[393] Normality of residuals from a generalized linear regression (GLM) model adjusted for age, sex, body mass index (BMI), and previous history of CAD with ln(CRP) was confirmed. GLMs tested the associations of CHIP-associated mutations and levels of ln(hs-CRP) after adjusting for age, sex, body mass index (BMI), and CAD status at baseline. Stratified analyses were performed according to CAD status. Back transformation of the regression beta term for ln(CRP) was derived as e^{β}, and the 95% confidence interval (CI) as $e^{\beta \pm 1.96 \times \text{standard error}}$. The percent difference in hs-CRP between CHIP+ and CHIP- was

derived as $\left(e^{\beta}\text{-}1\right)\times 100$. Additional analyses were performed according to the CHIP VAF, and categorized as VAF ≥0.10 vs. no mutation. Analyses were performed using SAS v.9.4 and R version 3.5.1. In sensitivity analyses, we adjusted for additional confounding factors, such as diabetes mellitus at baseline, as well as statin, aspirin, and beta-blocker use at baseline. In sub-group analyses, we focused on those without a history of cancer at study entry.

5.1.4 Results

5.1.4.1 Study Population

The population was comprised of 1940 subjects, of which 9 were excluded because of potential germline mutations (see below) and another 44 were excluded because they had cardiovascular disease without CAD at baseline. The remaining 1887 participants included 1359 patients with at least one previous account of MI, PCI, or CABG, and 528 patients without a previous history of CAD.

5.1.4.2 CHIP-Associated Mutations

We identified 427 CHIP carriers among the 1887 participants (22.6%, **Figure 5.1.1**). Of these, 331 (17.5%) had a single mutation and 96 (5.1%) had more than one mutation. The prevalence and the relative proportion between the 11 candidate genes were similar to previous reports.[235,241,334,382,383] Mutations in *DNMT3A*, *TET2* and *ASXL1* accounted for the majority of mutations (82.9%). The mean VAF was 13.3%. The prevalence of CHIP carriers was slightly higher in the non-CAD than in the CAD cohort (25.3% vs 21.5%, *P*=0.075). The relative prevalence between the different genes was similar between the two cohorts, except for *TET2* which prevalence was higher in the non-CAD cohort (9.0% vs 4.9%, *P*<0.001). All mutations are described in **Supplemental Table D.1.1**.

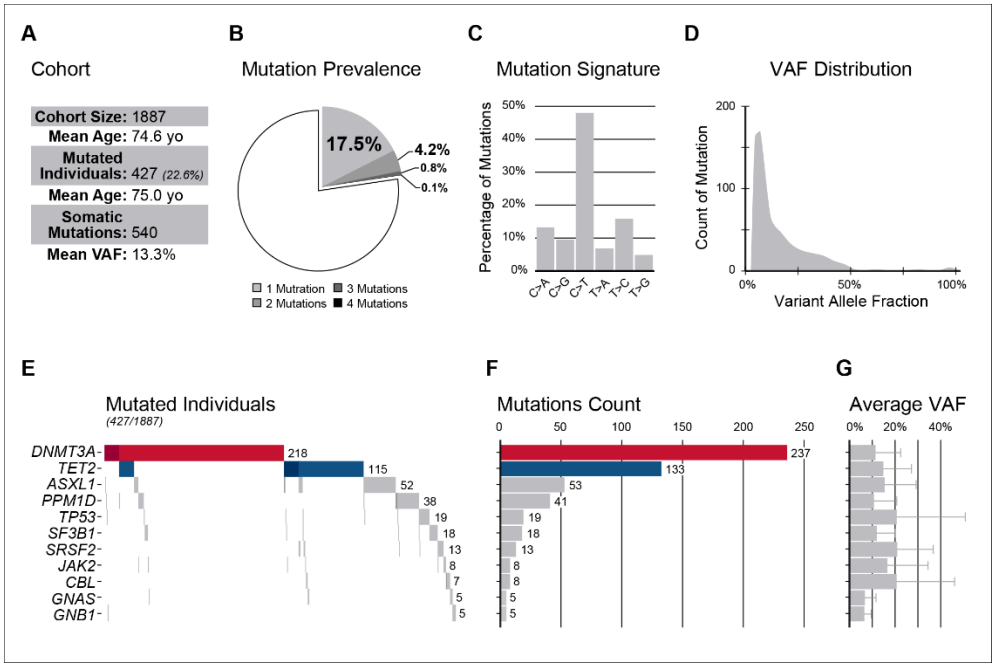


Figure 5.1 Prevalence and Distribution of Somatic Mutations.

(A) Description of the cohort (n=1887). (B) Prevalence of the 540 somatic mutations identified in 427 individuals of the cohort. (C) Mutation signature of all single nucleotide substitutions (see Supplemental Figure 2). (D) Variant allele fraction (VAF) distribution of all somatic mutations. (E) Co-occurrence of the somatic mutations in the 427 mutated individuals. Darker shades represent double mutation in the same gene. (F) Contribution of individual genes to the total number of observed somatic mutations. (G) Average VAF of somatic mutation for each gene.

5.1.4.3 Population Characteristics According To CHIP Carrier Status

Subjects were segregated according to their CHIP carrier status (CHIP- or CHIP+) in three groups: (i) all subjects, (ii) subjects with CAD, and (iii) subjects without CAD. **Table 5.1.1** describes the baseline characteristics of the three groups. Univariable analyses showed that CHIP carriers were significantly older (mean: 75.0 vs. 74.4 years, *P*=0.002), were less affected by dyslipidemia (85% vs. 88%, *P*=0.041) had fewer previous PCI (30% vs. 38%, *P*=0.006) than CHIP non-carriers. hs-CRP was significantly higher in CHIP carriers compared to non-carriers (median: 1.60 vs. 1.41 mg/L, *P*=0.009). Smoking, past or current, was not associated with CHIP.

Table 5.1 Baseline Characteristics of the Study Population, According To Clonal Hematopoiesis of Indeterminate Potential (CHIP) Status (Univariate Analysis).

	All patients		**Without baseline coronary artery disease**		**With baseline coronary artery disease**	
	CHIP+	CHIP-	CHIP+	CHIP-	CHIP+	CHIP-
N (%)	427(22.6)	1460 (77.4)	134(25.3)	394(74.7)	293(21.5)	1066(78.5)
Age, yr Mean (SD)	75.0 (3.4)*	74.4 (3.3)	74.9 (3.2)*	74.1 (3.3)	75.0 (3.5)*	74.583 (3.3)
Weight, kg Mean (SD)	77.5 (15.5)	78.2 (14.4)	77.0 (16.5)	75.0 (14.9)	77.8 (15.0)	79.4 (14.1)
Height, m Mean (SD)	1.65 (0.09)	1.66 (0.08)	1.64 (0.10)	1.63 (0.09)	1.66 (0.08)	1.67 (0.08)
BMI, Mean (SD)	28.197 (5.041)	28.245 (4.423)	28.390 (5.265)	27.915 (4.446)	28.108 (4.942)	28.367 (4.410)
Male, *n* (%)	292 (68.4%)	1044 (71.5%)	67 (50.0%)	183 (46.4%)	225 (76.8%)	861 (80.8%)
Cardiovascular risk factors – *n* (%)						
Dyslipidemia	361 (84.5%)*	1287 (88.3%)	84 (62.7%)	267 (67.9%)	277 (94.5%)	1020 (95.8%)
Hypertension	327 (76.8%)	1095 (75.3%)	82 (61.7%)	239 (60.7%)	245 (83.6%)	856 (80.7%)
Diabetes	120 (28.1%)	408 (28.0%)	26 (19.4%)	73 (18.5%)	94 (32.1%)	335 (31.5%)
Ever smoker	301 (70.7%)	988 (67.7%)	79 (59.4%)	218 (55.3%)	222 (75.8%)	770 (72.3%)
Current smoker	21 (4.9%)	71 (4.9%)	1 (0.8%)*	18 (4.6%)	20 (6.8%)	53 (5.0%)
CVD history – *n* (%)						
Coronary heart disease	293 (68.6%)	1066 (73.0%)	-	-	-	-
Previous MI	193 (45.3%)	649 (44.5%)	0 (0.0%)	0 (0.0%)	193 (66.1%)	649 (60.9%)
Previous PCI	130 (30.4%)*	550 (37.7%)	0 (0.0%)	0 (0.0%)	130 (44.4%)*	550 (51.6%)
Previous CABG	163 (38.2%)	601 (41.2%)	0 (0.0%)	0 (0.0%)	163 (55.6%)	601 (56.4%)
Stroke	42 (9.9%)	154 (10.6%)	0 (0.0%)	0 (0.0%)	42 (14.4%)	154 (14.5%)
Angina	250 (58.7%)	924 (63.5%)	0 (0.0%)	0 (0.0%)	250 (85.6%)	924 (87.1%)
PVD	85 (19.9%)	306 (21.0%)	0 (0.0%)	0 (0.0%)	85 (29.0%)	306 (28.8%)
Angiography	279 (65.3%)	1018 (69.7%)	15 (11.2%)	52 (13.2%)	264 (90.1%)	966 (90.6%)
CHF	72 (16.9%)	195 (13.4%)	1 (0.7%)	2 (0.5%)	71 (24.3%)*	193 (18.2%)
Medication – n (%)						
Aspirin	310 (72.6%)	1123 (76.9%)	66 (49.3%)	193 (49.0%)	244 (83.3%)	930 (87.2%)
Antiplatelet	329 (77.0%)	1164 (79.7%)	67 (50.0%)	199 (50.5%)	262 (89.4%)	965 (90.5%)
Statin	360 (84.3%)	1276 (87.4%)	83 (61.9%)	259 (65.7%)	277 (94.5%)	1017 (95.4%)
ACE	144 (33.7%)	538 (36.8%)	24 (17.9%)	77 (19.5%)	120 (41.0%)	461 (43.2%)
ARB	157 (36.8%)	496 (34.0%)	43 (32.1%)	130 (33.0%)	114 (38.9%)	366 (34.3%)
Beta-blockers	287 (67.2%)	970 (66.4%)	55 (41.0%)	143 (36.3%)	232 (79.2%)	827 (77.6%)
Lipids, mmol/L – mean (SD)						
Total cholesterol	3.89 (1.00)	3.86 (0.98)	4.44 (1.02)	4.59 (1.11)	3.63 (0.88)	3.59 (0.76)
Triglycerides	1.92 (1.03)	1.99 (1.02)	1.92 (1.08)	2.04 (1.03)	1.93 (1.00)	1.97 (1.01)
HDL	1.23 (0.37)	1.22 (0.36)	1.32 (0.36)	1.34 (0.42)	1.19 (0.36)	1.18 (0.33)

LDL	1.80 (0.84)	1.75 (0.83)	2.28 (0.89)	2.36 (0.99)	1.58 (0.71)	1.52 (0.62)
hs-CRP mg/L – median (IQR), min-max						
hs-CRP	1.60 (2.96)*, 0.09-115.0	1.41 (2.29), 0.08-80.70	1.55 (2.99), 0.09-92.80	1.45 (2.23), 0.08-61.30	1.60 (2.88)*, 0.09-115.0	1.39 (2.30), 0.08-80.70
Cancer history – *n* (%)						
Any cancer	89 (20.8%)	257 (17.6%)	26 (19.4%)	67 (17.0%)	63 (21.5%)	190 (17.8%)
Hematological cancer	6 (1.4%)	17 (1.1%)	3 (2.2%)*	0 (0.0%)	3 (1.0%)	17 (1.5%)
Incident/recurrent cancer – *n* (%)						
Any cancer	34 (7.9%)	111 (7.6%)	11 (8.2%)	25 (6.3%)	23 (7.8%)	86 (8.0%)
Hematological cancer	7 (1.6)	21 (1.4%)	0 (0.0%)	2 (0.5%)	2 (0.6%)	4 (0.3%)

*ACE: angiotensin-converting enzyme; ARB: angiotensin receptor blockers; BMI: body mass index; CABG: coronary artery bypass graft; CAD: coronary artery disease; CRP: C-reactive protein; CVD: cardiovascular disease; HDL: high density lipoprotein; HF: heart failure; IQR: interquartile range; LDL: low density lipoprotein; MI: myocardial infarction; PCI: percutaneous coronary revascularization intervention; STD: standard deviation. * When P<0.05 per Fisher's Exact test or Kruskal-Wallis two-sided significance test.*

5.1.4.4 Primary Analysis (Entire Cohort)

Our primary analysis focused on all patients, regardless of CAD status, and consisted in performing a multivariable generalized linear regression analysis for prediction of hs-CRP (n=1887). CHIP carriers of any gene mutations had significantly higher hs-CRP than non-CHIP carriers in adjusted analyses (e^{β}=1.21, 95% CI: 1.08-1.36, *P*=0.001, **Table 5.1.2**). Focusing on mutations in individual genes, CHIP carriers of mutations in *DNMT3A* had higher hs-CRP than those without mutations in that gene (e^{β}=1.17, 95% CI: 1.01-1.36, *P*=0.04). No other significant effect was observed for any of the remaining individual genes.

Table 5.2 Association Between CHIP and hs-CRP At Study Entry (Multivariable Analysis).

	All patients (*n* = 1887)		**Without baseline coronary artery disease (*n* = 528)**		**With baseline coronary artery disease (*n* = 1359)**	
	CHIP+	CHIP-	CHIP+	CHIP-	CHIP+	CHIP-
Any mutations						
n	427	1460	134	394	293	1066
hs-CRP GM, mg/L (IQR)	1.85 (0.81-3.77)	1.5 (0.71-3.00)	1.86 (0.83-3.82)	1.51 (0.77-3.02)	1.84 (0.81-3.72)	1.49 (0.69-3.01)
Back-transformed β (95% CI)*	1.21 (1.08-1.36)		1.15 (0.94-1.40)		1.22 (1.06-1.41)	
% difference	21%		15%		22%	
P value*	0.001		0.178		0.005	
DNMT3A						
n	218	1669	66	462	152	1207
hs-CRP GM, mg/L (IQR)	1.82 (0.80-3.79)	1.54 (0.73-1.43)	1.91 (0.81-4.72)	1.55 (0.77-3.10)	1.79 (0.78-3.66)	1.54 (0.71-3.13)
Back-transformed β (95% CI)*	1.17 (1.01-1.36)		1.23 (0.95-1.59)		1.15 (0.95-1.38)	
% difference	17%		23%		15%	
P value*	0.04		0.125		0.148	
TET2						
n	115	1772	48	480	67	1292
hs-CRP GM, mg/L (IQR)	1.78 (0.81-3.46)	1.56 (0.73-3.19)	1.77 (0.82-2.77)	1.58 (0.77-3.27)	1.79 (0.76-4.01)	1.55 (0.71-3.17)
Back-transformed β (95% CI)*	1.13 (0.92-1.38)		0.92 (0.68-1.24)		1.24 (0.94-1.62)	
% difference	13%		-8%		24%	
P value*	0.253		0.585		0.128	
ASXL1						
n	52	1835	14	514	38	1321
hs-CRP GM, mg/L (IQR)	1.93 (0.88-3.64)	1.56 (0.73-3.19)	1.77 (0.96-2.45)	1.59 (0.77-3.27)	1.99 (0.83-4.25)	1.55 (0.71-3.17)
Back-transformed β (95% CI)*	1.26 (0.94-1.69)		1.09 (0.64-1.86)		1.31 (0.92-1.87)	
% difference	26%		9%		31%	
P value	0.128		0.751		0.138	
PPM1D						

n	38	1849	9	519	29	1330
hs-CRP GM, mg/L (IQR)	1.64 (0.67-3.43)	1.57 (0.74-3.19)	1.84 (0.60-5.09)	1.59 (0.78-3.23)	1.59 (0.74-3.28)	1.56 (0.71-3.18)
Back-transformed β (95% CI)*	0.98 (0.69-1.39)		1.14 (0.59-2.20)		0.95 (0.63-1.43)	
% difference	-2%		14%		-5%	
P value	0.909		0.705		0.805	

CHIP: clonal hematopoiesis of indeterminate potential, CHIP+: carrier of mutation (s). CHIP-: Non-carrier of mutation, CAD: coronary artery disease, GM: geometric mean, IQR: interquartile range, Std. Error: Standard error
*N.B. The sample size related to carriers of mutations in TP53, SF3B1, SRSF2, CBL, and JAK2 were insufficient to perform analyses in these sub-groups. The presented P values for mutations yes vs. no status differences in log-transformed changes in hs-CRP. * Generalized linear regression model-derived ln(hs-CRP) with adjustment for age, sex, and coronary artery disease status at baseline-adjusted where CHIP- was the referent category.*

5.1.4.5 Secondary Analyses (CAD and Non-CAD Subgroups)

Among the patients with CAD at baseline (n=1359), CHIP carriers were older (mean: 75.1 vs. 74.5 years), had fewer previous PCI (44% vs. 52%), more congestive heart failure (24% vs 18%), and higher hs-CRP (median: 1.60 vs. 1.39 mg/L) than non-CHIP carriers (all *P*<0.05, **Table 5.1.1**). In the non-CAD cohort (n=528), CHIP carriers were also significantly older than non-carriers (74.9 vs. 74.1 years) but had lower rates of current smoking (0.8% vs. 4.6%), and hs-CRP levels were not significantly different between those two groups. In multivariable generalized linear regression analyses focusing on patients with a history of CAD at baseline, CHIP carriers of any mutations had higher hs-CRP than non-CHIP carriers after adjusting for covariates (e^{β}=1.22, 95% CI: 1.06-1.41, *P*=0.005, **Table 5.1.2**). This association was not significant in the smaller subgroup of patients without a history of CAD at baseline (e^{β}=1.15, 95% CI: 0.94-1.40, *P*=0.178).

5.1.4.6 Sensitivity, Sub-Group and Additional Analyses

When limiting analyses to individuals with VAF ≥10%, as compared to those without CHIP, carriers of any gene mutation also had higher ln(CRP) (e^{β}=1.23, 95% CI: 1.05-1.45, *P*=0.038) in adjusted analyses (data not shown). Due to limited sample size in patients with VAF ≥10% (n=191), a clear correlation between ln(CRP) and CHIP carriers could not be identified. In sensitivity analyses, we adjusted for additional confounders (statin, aspirin, beta-blocker use at baseline, as well as diabetes mellitus) that could have influenced the results of our analyses. The additional adjustment of these confounding factors in our multivariable analyses did not affect the magnitude of our findings. Finally, because patients with a cancer diagnosis at study entry are likely to have received chemotherapy or radiation therapy, which could result in clonal expansion, we also conducted a sub-group analysis in patients without cancer at baseline (n=1541). The relationship between CHIP and ln(CRP) in this cohort remained unchanged.

5.1.5 Discussion

Young *et al.* documented mutations in *TET2* or *DNMT3A* at very low frequencies in 95% of individuals aged 50 to 60 years, suggesting that these genes are almost ubiquitously mutated after the age of 50. [394] However, only a fraction of mutated individuals will have clonal expansion of a magnitude corresponding to CHIP,[336] and associated increased risk of hematological cancers and cardiovascular disease. It is therefore of prime importance to identify which are the factors associated with clonal progression. Therefore, we evaluated the hypothesis that inflammation is associated with clonal development of CHIP-related mutations in a cohort of 1887 individuals aged over 70, comprising individuals with (1359) and without CAD (528) at baseline. We showed that hs-CRP is significantly higher in subjects with CHIP than those without CHIP, supporting the hypothesis that there is a link between inflammation and CHIP's pathogenesis in aging individuals (**Figure 5.1.2**).

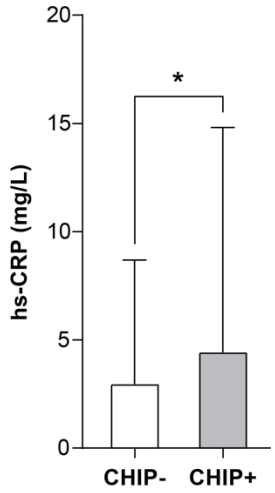


Figure 5.2 High Sensitivity C-Reactive Protein Concentration (Mg/L) in CHIP- and CHIP+ Individuals.

*Mean hs-CRP concentration (mg/L) in CHIP- and CHIP+ individuals with standard deviation. * When P<0.05 per Kruskal-Wallis two-sided significance test.*

CRP was discovered in 1930 in the serum of patients with pneumococcal pneumonia.[395] It is a pentraxin produced by the liver and an acute phase reactant[396]. IL-6, IL-1 and TNFα are the main inducers of CRP.[397,398] CRP measurement is commonly used to evaluate tissue injury, infection and inflammatory diseases[399]. The development of high-sensitivity assays[400] has been invaluable for the investigation of low grade chronic inflammation in different human diseases[399]. The value of hs-CRP as a predictor of cardiovascular events was demonstrated 20 years ago.[401] For example, in a cohort of 28,263 normal women, hs-CRP measured only once at baseline was the strongest predictor of coronary heart disease death, nonfatal myocardial infarction, stroke, or the need for

coronary revascularization procedures amongst other biomarkers including IL-6. We adopted a similar approach using baseline hs-CRP measurement in a cohort of patients with and without CAD.

The primary objective of this study was to document an association between inflammation and CHIP carrier status. This was demonstrated in both univariable and multivariable analyses for the entire study population and for the cohort with CAD. This is the first clear demonstration of the potential role of inflammation in clonal expansion of CHIP-related mutations. Further, this demonstration was made using an adequately large cohort of older individuals and a commonly available biomarker of inflammation. It is noteworthy to state, that despite this demonstration, we could not readily test the cause-and-effect hypothesis and cannot entirely omit that CHIP can aggravate inflammation.

Despite similar differences in hs-CRP levels between CHIP and non-CHIP carriers within the CAD and non-CAD cohorts, statistical significance was not reached in the latter subgroup. This might be related to the fact that our non-CAD cohort was smaller than the CAD counterpart (528 vs 1359). Bick et al. analyzed exome sequences from 35,416 individuals from the UK biobank without prevalent cardiovascular disease (similar to our non-CAD cohort)[273], and reported that hs-CRP level was not significantly different between CHIP and non-CHIP carriers in a multivariable analysis. It is possible that the association between inflammation and CHIP may be more important in CAD than in non-CAD patients. Granted, there is also the possibility that our results are not generalizable beyond the current study cohort. Notably, a recent study, failed to identify a significant association between CRP and CHIP, although CRP was aggregated with IL-6 (P=0.004) and IL-1b (P=0.026), both of which showed a positive correlation with CHIP.[258]

The relative difference of hs-CRP levels between CHIP carriers and non-CHIP carriers was relatively modest (±20%). The underlying difference in inflammation may nevertheless be sufficient to impact the marrow microenvironment and provide growth advantage to mutated stem cells. Myers SanMiguel *et al*. demonstrated that transplantation of *Dnmt3a*$^{R8878H+}$ bone marrow cells into young versus old congenic recipient mice, lead to an accelerated expansion of the *Dnmt3a*$^{R8878H+}$ cells in the older mice [402]. TNF-α and M-CSF were identified by RNA-seq to have higher expression levels in older mice. Furthermore, in a recent meta-analysis of inflammatory cytokine profile of 697 individuals with myelodysplastic syndromes versus controls, Shi et al. demonstrated a small but significant increase levels of TNF, IL-8 and IL-6[403]. Relatively modest quantitative difference in inflammatory cytokines may have a lever effect if it is linked with increase receptor expression. Such a model has been demonstrated for chronic myelogenous leukemia (CML) in which stem cells have increased IL-1 receptor activity,[404] and

increased proliferation and survival in a proinflammatory environment.[405] Interestingly, treatment with the anti-IL1b antibody canakinumab led to significant reductions of not only cardiovascular disease[190] but also incident lung cancer in patients with CAD.[209] Arguably, the relationship between clonal expansion and cardiovascular events in consideration of inflammation remains complex. For example, Bick et al. demonstrated that a relatively common coding mutation in the IL6-Receptor (pAsp358Ala) which significantly reduces signalization of IL-6, attenuated cardiovascular risk of CHIP carriers, but not in non-CHIP carriers.[273] The CHIP-by-IL-6 interaction term emerged as significant, further supporting the three-way association. The recent demonstration that utilization of a low dose of the anti-inflammatory drug colchicine reduces cardiovascular events after MI may further justify the prospective evaluation of such approaches in CHIP carriers.[201]

We selected patients aged above 70 years to maximize the prevalence of CHIP, which was indeed 22.6% in this cohort. Surprisingly, patients without CAD at baseline had slightly higher prevalence of CHIP than those with CAD at baseline (25.3% vs 21.5%, P<0.075). Since CHIP is associated with an increased risk of CAD, we expected a high prevalence of CHIP in this subgroup. We speculate that selecting patients aged >70 years may have systematically introduced a survival bias, where CHIP carriers with CAD may have been left out due to earlier death. It is also intriguing that the prevalence of *TET2* mutations was significantly lower in the CAD cohort that the non-CAD subgroup. It is possible that specific CHIP-associated genes have a different survival impact. Expanding the cohort to include younger patients may ultimately allow us to answer this question.

This study also allowed us to make additional observations. CHIP regroups alterations occurring in several different genes, but clinical outcomes have been usually estimated according to the CHIP-carrier status[241,382] rather than gene-specific estimation. Recently, some studies have demonstrated gene specificities in regards to risk of progression to acute myeloid leukemia,[406] genetic predisposition[383] or lineage restriction[407]. We wanted to study the relationship between inflammation and specific CHIP-associated genes. We were able to document a positive association between *DNMT3A* and hsCRP in the entire cohort. No other association was statistically significant, although quantitative differences were similar for *DNMT3A*, *TET2* and *ASXL1*. A significantly larger cohort would be necessary to address this question. We documented a significant association between heart failure and CHIP carrier status. This is in line with several previous observations in mice[408] and humans.[247,409] In contrast, we did not find an association with smoking in this cohort where prevalence was very low.

5.1.6 Conclusions

In conclusion, this study highlights the role of inflammation in CHIP. The etiology of CHIP is probably multifactorial and several other factors need to be identified. Clinical trials should test whether anti-inflammatory therapy can reduce CHIP progression and related diseases.

5.1.7 Acknowledgements

This work was supported by a grant from the Canadian Institutes of Health Research (CIHR) and by the Leukemia Lymphoma Society of Canada (LLSC).

5.1.8 Conflict-of-interest disclosure

The authors declare no competing financial interests.

5.2 Low-Dose Colchicine and High-Sensitivity C-Reactive Protein After Myocardial Infarction: A Combined Analysis Using Individual Patient Data From the COLCOT and Lodoco-MI Studies

In the original landmark trial publication of COLCOT, investigators performed sub-analyses focusing on the association between colchicine use and hs-CRP levels pre- and post-intervention.[201] However, not all patients consented to provide biomarker data, and hence, sample size of this sub-analysis was low which rendered the interpretation of results difficult. In an attempt to augment the number of patients in the evaluation of colchicine and hs-CRP, Dr. Jean-Claude Tardif (Principal Investigator [PI] of COLCOT and co-director of my doctoral studies) and Dre. Marie-Pierre Dubé coordinated access to data from the LoDoCo-MI[410] trial by reaching out to investigators (and colleagues) of that trial. Thus, in collaboration with Dr. Thomas Hennessy, Dr. Carl J Schultz, Dr. Graham S Hillis (Department of Cardiology at *Royal Perth Hospital*), we combined colchicine and hs-CRP pre- and post-intervention data from both the COLCOT and LoDoCo-MI studies. Our primary findings showed that treatment with colchicine significantly increased the likelihood of achieving post-treatment hs-CRP values $\leq$1 mg/L compared to placebo, which is considered the "low risk" cutoff according to the American Heart Association and US Centers for Disease Control and Prevention guidelines.[411] The results of this pooled analysis supports the hypothesis that targeting inflammation with colchicine through the NLRP3 inflammasome may be relevant in the treatment of atherosclerosis.

Author contributions: Maxine Sun, Marie-Pierre Dubé, and Jean-Claude Tardif conceptualized the study design and approach. Jean-Claude Tardif and Graham S Hillis provided access to data, which were previously populated by Thomas Hennessy and David Rhainds. Maxine Sun, Thomas Hennessy, and Amina Barhdadi optimized combination of data and methodological approaches, which were revised and approved by Marie-Pierre Dubé, Graham S Hillis, and Jean-Claude Tardif. Maxine Sun performed the analyses which were overseen by Amina Barhdadi. Interpretation of data and results were conducted by Maxine Sun, Marie-Pierre Dubé, Thomas Hennessy, Amina Barhdadi, Graham S Hillis and Jean-Claude Tardif. Maxine Sun wrote the first draft of the manuscript, which was revised by Marie-Pierre Dubé, Thomas Hennessy, Carl J Schultz, Graham S

Hillis and Jean-Claude Tardif. Additional administrative and technical support were provided by David Rhainds.

Low-Dose Colchicine and High-Sensitivity C-Reactive Protein After Myocardial Infarction: A Combined Analysis Using Individual Patient Data From the COLCOT and Lodoco-MI Studies

Maxine Sun[1,2,3*‡], Marie-Pierre Dubé[1,2,3*‡], Thomas Hennessy[4‡], Carl J. Schultz[4‡], Amina Barhdadi[1,2,3‡], David Rhainds[1‡], Graham S. Hillis[4†‡], Jean-Claude Tardif[1,3†‡]

[1]Montreal Heart Institute, Montreal, Canada
[2]Université de Montréal Beaulieu-Saucier Pharmacogenomics Centre, Montreal, Canada
[3]Department of Medicine, Faculty of Medicine, Université de Montréal, Montreal, Canada
[4]Department of Cardiology, Royal Perth Hospital, Perth, Australia and Medical School, University of Western Australia.

* Contributed equally
† Co-corresponding authors
‡ This author takes responsibility for all aspects of the reliability and freedom from bias of the data presented and their discussed interpretation.

Int J Cardiol. 2022;363:20-22. doi:10.1016/j.ijcard.2022.06.028

5.2.1 Introduction

Inflammation is a central contributor to the development and progression of atherosclerosis. Inhibition of the interleukin-1β-interleukin-6 signaling pathway, a process initiated at the level of the NLRP3 inflammasome, is associated with reduced cardiovascular events, but anti-inflammatory agents that fail to alter levels of interleukin-1β, interleukin-6, or C-reactive protein (CRP) appear to be ineffective in reducing such events.[200,412] Colchicine is a repurposed anti-inflammatory medication that binds tubulin and affects neutrophil functions and possibly the NLRP3 (NLR family pyrin domain containing 3) inflammasome. Low-dose colchicine (0.5 mg once daily) reduced the risks of recurrent cardiovascular events following an acute myocardial infarction (MI) in the COLCOT trial.[201] However, the influence of colchicine on inflammation, as determined using high-sensitivity CRP (hs-CRP), a conventional marker of chronic inflammation, remains inconclusive[410]. In the current study, we conducted a combined individual patient data analysis from the COLCOT[201] and LoDoCo-MI[410] studies to assess the effect of low-dose colchicine on hs-CRP in patients with acute MI.

5.2.2 Methods

Adopting the approach of a patient-level meta-analysis (as opposed to conventional summary statistics meta-analyses), we performed a combined analysis using individual patient data from two clinical trials. The study designs and populations have previously been described.[201,410] Paired

pre-treatment (baseline) and post-treatment hs-CRP (mg/L) were available in 222 out of 237 patients for LoDoCo-MI, and 207 out of 4745 patients for COLCOT. Although this represented a sample subgroup for COLCOT, baseline characteristics of these patients remained similar to those of the overall population, as previously reported[201]. The pooled sample size resulted in 429 subjects available for analysis. Continuously coded study outcomes included post-treatment hs-CRP, absolute change hs-CRP (calculated as post-treatment hs-CRP minus pre-treatment hs-CRP), and percent change (% change) (calculated as absolute change hs-CRP divided by pre-treatment hs-CRP and multiplied by 100). hs-CRP was also dichotomized at the cutoff of 1.0 mg/L, which is considered the "low-risk" cutoff according to the American Heart Association and U.S. Centers for Disease Control and Prevention[411]. The primary variable of interest was study treatment (colchicine versus placebo). To account for study differences in the time delay from index MI to randomization and follow-up duration, a mixed-model multi-level regression analysis[413] was used to obtain estimates of the effect of colchicine versus placebo on post-treatment hs-CRP. Adjustment was made for age, sex, body-mass index, and baseline hs-CRP as fixed effects. Study (COLCOT or LoDoCo-MI) and baseline hs-CRP were modeled as random effects. hs-CRP % change and absolute change were inverse transformed to a normal distribution of mean 0 and standard deviation of 1.[414]

5.2.3 Results

The two cohorts differed with respect to baseline hs-CRP (median: 4.28 vs. 7.40 mg/L for COLCOT and LoDoCo-MI, respectively, *P*<0.001, **Table 5.2.1**), likely due to different duration between patients' index MI and baseline hs-CRP measurements (mean of 16.4 vs. 1.4 days for COLCOT and LoDoCo-MI, respectively). Overall median post-treatment hs-CRP was 1.26 mg/L for COLCOT and 1.70 mg/L for LoDoCo-MI (*P*=0.056) and the proportion of patients reaching hs-CRP values $\leq$1 mg/L post-treatment was comparable (38% vs. 34% for COLCOT and LoDoCo-MI, respectively, *P*=0.48).

The median (interquartile range [IQR]) post-treatment hs-CRP values for the colchicine and placebo groups were 1.12 mg/L (0.77-2.10) and 1.39 mg/L (0.90-2.65) respectively in COLCOT and 1.60 mg/L (0.70-3.40) and 2.00 mg/L (0.90-3.95) in LoDoCo-MI **(Table 5.2.1)**. The median % change in hs-CRP for the colchicine and placebo groups were -65% and -67% respectively in COLCOT and -78% and -64% in LoDoCo-MI. In the combined analysis of COLCOT and LoDoCo-MI using a mixed-model multi-level regression analysis, treatment with colchicine was not significantly associated with post-treatment hs-CRP (beta: -0.41, *P*=0.429), absolute change in hs-CRP (beta: -0.03, *P*=0.687), or with % change in hs-CRP (beta: -0.11, *P*=0.173, **Table 5.2.2**). In contrast, treatment with colchicine significantly increased the likelihood of achieving post-treatment hs-CRP values $\leq$1 mg/L compared to placebo (OR: 1.64, 95% CI: 1.07 to 2.51, *P*=0.024, **Table 5.2.2**).

Table 5.3 Summary Statistics of Pre-Treatment and Post-Treatment Hs-CRP For COLCOT and Lodoco-MI.

Characteristic	COLCOT				LoDoCo-MI			
	Overall, N = 207[1]	**Placebo**, N = 108[1]	**Colchicine**, N = 99[1]	**p-value**[2]	**Overall**, N = 222[1]	**Placebo**, N = 111[1]	**Colchicine**, N = 111[1]	**p-value**[2]
Pre-treatment hs-CRP, mg/L (continuous)				0.170				0.891
Mean (SD)	8.43 (12.09)	8.95 (11.28)	7.87 (12.95)		20.91 (38.20)	20.55 (37.79)	21.26 (38.77)	
Median (IQR)	4.28 (2.28, 8.78)	4.83 (2.47, 11.83)	4.00 (2.13, 6.98)		7.40 (3.00, 17.58)	7.70 (3.00, 18.90)	6.90 (3.10, 15.60)	
Range	0.47, 93.33	0.47, 80.18	0.65, 93.33		0.20, 225.00	0.20, 225.00	0.50, 202.90	
Post-treatment hs-CRP, mg/L (continuous)				0.091				0.115
Mean (SD)	2.45 (3.96)	2.58 (3.72)	2.31 (4.23)		3.83 (6.64)	4.09 (6.29)	3.58 (6.99)	
Median (IQR)	1.26 (0.83, 2.39)	1.39 (0.90, 2.65)	1.12 (0.77, 2.10)		1.70 (0.80, 3.60)	2.00 (0.90, 3.95)	1.60 (0.70, 3.40)	
Range	0.20, 36.84	0.20, 29.48	0.37, 36.84		0.05, 50.20	0.20, 37.60	0.05, 50.20	
Absolute change				0.353				0.439
Mean (SD)	-5.99 (12.37)	-6.37 (11.22)	-5.57 (13.55)		-17.07 (37.82)	-16.47 (38.05)	-17.68 (37.75)	
Median (IQR)	-2.26 (-6.34, -0.80)	-2.67 (-9.42, -0.82)	-2.10 (-5.31, -0.80)		-4.00 (-13.93, -1.00)	-3.30 (-12.50, -0.90)	-4.30 (-13.75, -1.10)	

Characteristic	COLCOT				LoDoCo-MI			
	Overall, N = 207[1]	**Placebo**, N = 108[1]	**Colchicine**, N = 99[1]	**p-value**[2]	**Overall**, N = 222[1]	**Placebo**, N = 111[1]	**Colchicine**, N = 111[1]	**p-value**[2]
Range	-90.42, 31.69	-76.90, 13.72	-90.42, 31.69		-224.00, 33.50	-224.00, 20.30	-201.80, 33.50	
% change in hs-CRP (%, continuous)[4]				0.640				0.088
Mean (SD)	-38.33 (123.30)	-45.00 (70.57)	-31.05 (162.54)		-38.61 (115.62)	-34.01 (132.07)	-43.20 (96.82)	
Median (IQR)	-65.94 (-81.75, -39.78)	-66.79 (-81.18, -34.88)	-64.76 (-83.17, -40.69)		-73.53 (-89.32, -35.70)	-64.29 (-87.61, -31.06)	-77.85 (-91.56, -46.91)	
Range	-98.29, 1,349.67	-97.74, 312.64	-98.29, 1,349.67		-99.62, 1,133.33	-99.56, 1,133.33	-99.62, 497.62	
Post-treatment hs-CRP (≤1 mg/L)	78 (38%)	34 (31%)	44 (44%)	0.055	76 (34%)	33 (30%)	43 (39%)	0.157

Absolute change =post-treatment hs-CRP – pre-treatment hs-CRP
Percent change was calculated by substracting post-treatment hs-CRP to pre-treatment hs-CRP, divided by pre-treatment hs-CRP, multipled by 100
[1]n (%)
[2] Wilcoxon rank sum test; Pearson's Chi-squared test
hs-CRP: high-sensitivity C-reactive protein, SD: standard deviation, IQR: interquartile range

Table 5.4 Effects of Colchicine Versus Placebo On Post-Treatment hs-CRP in the Combined COLCOT and Lodoco-MI Randomized Clinical Trials

Endpoints	N COLCOT	N LoDoCo-MI	Combined effect of colchicine vs. placebo on post-treatment hs-CRP	
			Estimates[a] (95% CI)	*P*
Post-treatment hs-CRP, mg/L	207	222	-0.41 (-1.42, 0.61)	0.429
Absolute change in hs-CRP, mg/L [b, c]			-0.03 (-0.20, 0.14)	0.687
% change in hs-CRP, %[c]			-0.11 (-0.27, 0.05)	0.173
Post-treatment hs-CRP, dichotomized at ≤1 mg/L			1.64 (1.07, 2.51)	0.024

CI, confidence interval; hs-CRP, high-sensitivity C-reactive protein

[a] *Estimates are regression coefficients for post-treatment hs-CRP values, as well as % change in hs-CRP as continuous variables, and odds ratio for post-treatment hs-CRP dichotomized at ≤1 mg/L. All analyses were performed using a mixed-model multi-level regression with adjustment for age, sex, body-mass index and baseline hs-CRP as fixed effects, and trial and baseline hs-CRP as random effects.*

[b] *Absolute change was calculated as post-treatment hs-CRP minus pre-treatment hs-CRP. In this model, adjustment was not made for baseline hs-CRP as a fixed effect*

[c] *% change and absolute change in hs-CRP were inverse normal transformed (INT) with mean of 0 and standard deviation of 1.0.*

5.2.4 Discussion

This combined analysis of individual patient data from the COLCOT and LoDoCo-MI trials has shown that low-dose colchicine can lower hs-CRP beyond the low-risk threshold of $\leq$1 mg/L when administered following an acute MI. However, post-treatment hs-CRP and % change were not significantly different between treatment groups. This may be related to the inherent differences with respect to the time from index MI to randomization, as well as in the follow-up duration between the two studies. Although the mixed modeling approach can mitigate data heterogeneity, there may remain unaccounted study-specific variability in hs-CRP measurements that can limit statistical power. Another limitation is that although the current analysis is the largest ever-reported assessment of the effect of colchicine on post-treatment hs-CRP (n=429), sample size remains small. Our study supports the hypothesis that targeting inflammation with colchicine through the NLPR3 inflammasome may be important in the treatment of atherosclerosis. In conclusion, the current findings imply that a reduction in inflammation was a component in the clinical efficacy of low-dose colchicine observed in the COLCOT trial and supports the consideration of hs-CRP as a potential predictor of such efficacy.

Chapter 6

Discussion, Conclusions & Perspectives

6.1 Discussion

The overarching objective of this doctoral thesis was to explore how inflammation is implicated in the pathogenic occurrence of both cancer and CVD. Collectively, the findings from the conducted work provide strong supporting evidence implicating a shared common role of inflammatory mechanisms in both disease phenotypes.

Objective 1: Investigate the potential of non-specific inhibition of inflammation in reducing the risk of cancer.

Summary of findings: In an effort to reconcile the inconsistent findings regarding statin use and cancer risk in patients with CAD as reported by clinical trials and population-based cohorts, we performed the first genetic meta-analysis using data from two randomized-controlled phase IV trials to evaluate the impact of higher dose statin on the risk of incident cancer (see **Article 1**). We identified a genetic variant (rs13210472:C) that was associated with a higher risk of incident cancer in women only, and successfully replicated its effect using the UK Biobank. Interestingly, the variant was specific to women statin users, but not statin use itself. The functional annotation revealed additional insights on the identified variant. Specifically, it was shown to be associated with increased expression of *HLA-DOA*, *HLA-DPA1*, and *HLA-DPB1* genes. These genes encode for MHC-II molecules (major histocompatibility complex, class II), which are responsible for presenting antigens to $CD4^+$ T cells. Expression quantitative trait loci data suggest that an increase in HLA-DP heterodimers at the surface of antigen-presenting cells could lead to the activation of additional $CD4^+$ helper T cells. Furthermore, the effect of the genetic variant on cancer risk appeared to exhibit heterogeneity across different cancer types, as evidenced in sub-analyses; however, caution is warranted in interpreting these results due to very low sample sizes and limited numbers of events.

Clinical interpretation: In women, the heightened risk of cancer associated with carrying the rs13210472 variant and taking statins may be rooted in a complex, sex-specific interplay between

immune regulation and drug response. Carriers of rs13210472 are predisposed to differential expression of HLA-DOA and HLA-DPA1 genes, pivotal players in the HLA system responsible for the regulation of our immune responses. Women inherently exhibit higher $CD4^+$ T cell counts and an increased $CD4^+/CD8^+$ T-cell ratio compared to men, characteristics often linked to autoimmune susceptibility. When these immune profiles interact with the rs13210472 variant, it may be speculated that it produces a uniquely sensitized environment conducive to cancer initiation or progression. The addition of statins into this equation may further complicate the risk profile. Statins are functionally known to suppress dendritic cell maturation and promote an anti-inflammatory response, potentially weakening the immune system's ability to ward off cancer. Therefore, the cumulative effects of the rs13210472 variant, inherent sex-based immune differences, and statin-induced immune modulation might converge to amplify cancer risk in women. Additional analyses from observational data may be undertaken to explore the association between rs13210472 and CD4+/CD8+ T-cell ratios, notably whether the effect of the variant on incident cancer is modulated by such ratios according to sex. *In vitro* experiments can be eventually designed to utilize immune cells to assess the impact of rs13210472 on the expression of the HLA-DOA and HLA-DPA1 genes. These studies can help determine how such genetic variation influences CD4+/CD8+ T-cell ratios in men and in women. Subsequent cellular assays can be conducted to investigate the modulatory effects of statin medications on these immune cell functions. This will allow for a more nuanced understanding of how genetics and pharmacological interventions collectively influence immune response according to men and women.

Altogether, these findings highlight the intricate interplay between genetics, immune function, and disease risk. The lack of a significant interaction term between the genetic variant and statin use with regard to any cancer risk further implies that the identified SNP simply reflects complex underlying processes that regulate the immune gradient balance, which can shift towards either a weakened immune system that fails to exert cancer immunosurveillance or an overactive immune system that leads to sustained chronic inflammation. As a result of, by selecting individuals with CAD at baseline who were on statins, we inadvertently also selected individuals that were prone to such immune dysregulation. It may then be concluded that a general anti-

inflammatory inhibitor, such as statins, may ultimately be insufficient for persons suffering from some autoimmunity that has been ongoing for an extended period. This indicates that there may be an opportunity to develop therapeutics or tools that can regulate a patient's immune gradient effectively to prevent the onset of health conditions that may arise from an overactive immune response. Furthermore, these findings also serve as a reminder that immune (dys)regulation is still a highly complex and a heterogenous process.

Objective 2: Evaluate clonal hematopoiesis as a biomarker for CVD risk in cancer survivors.

Summary of findings: At the beginning of my doctoral studies, the link between CH and atherosclerosis was newly uncovered. Naturally, investigators also sought to investigate the potential influence of mCAs, another type of CH, on CVD and CV-related outcomes. In those studies, mCAs were not significantly associated with CV endpoints.[242,243] However, there was interest in assessing their impact in patients diagnosed with cancer. Cancer survivors are at a higher risk of CVD-related death compared to the general population. Traditional CV factors do not adequately risk stratify patients at this excess risk of CV morbidity. A reliable cardiac biomarker that can better identify patients at higher risk of developing CV complications at cancer diagnosis can have tremendous impact on treatment management and cardiology surveillance.

Under this premise, we relied on the UK Biobank to identify patients diagnosed with select cancers at risk of CVD-related death (see **Article 2**). After characterizing those with and without mCAs, using a returned data, we systematically evaluated the influence of any mCA (e.g., ≥1 alteration), and mCA types (autosomal, mosaic loss of X, mosaic loss of Y, expanded cell fraction) on the risk of death from CVD causes, CAD causes and any cause. Several sub-analyses were proposed including those who received chemotherapy, radiotherapy, by sex, age groups, and across cancer sites. Our findings suggest that for individuals diagnosed with cancer, those with at least one type of mCA had a higher risk of death from CAD causes. The effect was not particularly modulated by receipt of chemotherapy or radiotherapy. The effect of mCAs was also found to be highly variable across different types of cancer, with some exhibiting a particularly strong association (e.g., kidney cancer, breast cancer).

As a follow-up to that study, we obtained whole-exome data for all participants of the UK Biobank and sought to examine the influence of CHIP on death from CVD causes, CAD causes and any cause among those same cancer patients (see **Article 3**). Our analyses revealed significant associations between CHIP and death from CVD causes. Furthermore, our findings also indicated that co-occurrence of CHIP and mCA did not result in worse outcomes than those with CHIP mutations alone with respect to death from CVD causes. However, those with both types of CH did have a higher risk of all-cause mortality. Taken together, these findings appear to support the possibility of considering CH (either CHIP or mCA) as a marker of CV-profile risk stratification in patients diagnosed with cancer.

Clinical interpretation: Increasing evidence now suggests that formation of CH is dependent on an inflammatory environment (see section 1.5.2.3) Early on, a study showed that circulating IL-1β levels were higher in *TET2*-mutant human CHIP.[258] In an initial study, we have also assessed the correlation between CHIP mutations and hs-CRP in patients recruited at the Montreal Heart Institute (MHI) Biobank (see **Article 4**). Our findings showed that CHIP carriers had 21% higher hs-CRP levels compared to their non-CHIP counterparts.[342] A recent study has shown that genetic *Il-1r1* loss could rescue mutant *TET2*-associated hematopoietic stem and progenitor cells (HSPC) abnormalities and immune perturbations.[415] The ability to overcome chronic inflammation may also contribute to the fitness advantage conferred by other genetic drivers of CH. Age-related signaling molecules such as TNFα and interferon gamma promote the selective expansion of murine *DNMT3A*-mutant hematopoietic stem and progenitor cells, with potential therapeutic targets including TNF-receptor 1[416] and Tnxip[417]. Similarly, nr4a1 has been shown to enhance the fitness of *ASXL1*-mutant zebrafish by mediating the resistance to inflammation.[418] Thus, development and progression of CH-expansion initiates as a result of an enhanced fitness imparted by increased self-renewal. In turn, with ageing, this leads to increased chronic expression of pro-inflammatory cytokines and other mediators that alter the hematopoietic milieu. This environment goes to suppress normal hematopoietic stem and progenitor cells, CH-mutant cells on the other hand adapt and gain fitness advantage because of the inflammation. A recent study showed that by targeting inflammatory cytokines, associated receptors, signaling,

and adaptive mechanisms can break up the CH-inflammation cycle, thereby decreasing the risks of subsequent health sequalae.[419]

Indeed, a study involving three macaque non-human primates with notable *TET2*, *DNMT3A*, and *ASXL1* clones that were treated with an antibody that blocked IL-6 signaling found a deceleration of *TET2* CH clones.[420] Similarly, a previous post-hoc analysis showed that in patients with previous MI and *TET2*-mutant CHIP mutations, they experienced less adverse cardiac events following treatment with canakinumab[277], an anti-IL-1β inhibitor[190] (see section 1.4.3.1).

Thus, an important question is whether effective anti-inflammatory therapy for atherosclerosis could impact CH clonal dynamics in humans, resulting in a slowing down of CH expansion, thereby limiting events related to CVD. Specifically, if CH is associated with inflammation, would an anti-inflammatory treatment, such as colchicine that can effectively reduce inflammation (e.g., hs-CRP) be able to break up the CH-inflammation cycle? There may be a basis for this rationale, as we have previously conducted a combined-analysis of patients with available pre- and post-intervention hs-CRP levels in two clinical trials comparing colchicine to placebo (COLCOT[201] & LoDoCo-MI[410], see section 1.4.3.2). Our findings revealed that colchicine was significantly associated with post-treatment hs-CRP values of ≤1.0 mg/L compared to placebo (see **Article 5**).[421]

6.2 Strengths and limitations

In this thesis, several strengths can be highlighted, including the utilization of large, well-characterized cohorts such as two randomized clinical trial datasets and the UK Biobank, which provided a robust sample size and comprehensive data for investigation. The exploration of the complex interplay between immune function and the co-occurrence of CVD and cancer contributed to a more comprehensive understanding of the underlying disease processes. Additionally, the examination of the influence of CHIP and mCA on CVD outcomes in cancer survivors shed light on its potential prognostic value as a biomarker for CV-profile risk stratification during routine clinical practice.

Specific limitations per article are worth mentioning. First, for **Article 1**, the fact that the development cohort, which comprised of phase IV trial datasets from TNT and IDEAL, assessed statin use as high versus lower dose groups did not allow us to properly assess the impact of statin and non-statin use on the risk of incident cancer. In addition, statin use in the validation cohort (UK Biobank) further digressed from its original definition and was obtained using self-reported data. Limited information regarding dosage over time and utilization duration further aggravates confounding in our analyses.

Second, for **Article 2**, although abstraction of mCAs was improved relative to more historical studies, they were nonetheless inferred. Similarly, for **Article 3**, CHIP was defined based on previously developed pipelines, which arbitrarily defined mutations with a VAF of $\geq$2% as CHIP mutations. This cut-off although widely used across CHIP studies is nevertheless an operative definition more than a biological one. Furthermore, both studies were not able to assess temporal evolution of clones. Clonal expansion may be further influenced by exposure to chemo/radiotherapy treatments, which may have been underestimated in the UK Biobank. Specifically, chemotherapy and/or radiotherapy were identified based on in-hospital procedure codes. In reality, some of such treatments may have been prescribed in the outpatient setting, or taken at home (e.g., oral agents) which would have been missed. With respect to treatment-induced clonal expansion, the UK Biobank recruited patients during the pre-targeted therapy era. The influence of tyrosine kinase inhibitors (e.g., sunitinib) and modern immunotherapy (e.g., ipilimumab) on clonal evolution may be more clinically relevant to assess since contemporary patients are more likely to be treated with such modalities.

An important limitation in **Article 4** relates to patient selection bias, where the significant effect of CHIP was only observed in patients with CAD. In contrast, those without CAD were not subjected to a significant effect for CHIP on hs-CRP levels. A possible explanation may be related to the smaller sample size of patients without CAD. However, some recent studies also failed to observe a significant relationship between CHIP and hs-CRP.[277] The difference observed in our study compared to other studies may be related to the time-point of when hs-CRP was measured. In **Article 5**, our analyses revealed that hs-CRP levels can fluctuate significantly depending on the

time that the blood sample was taken. Because of that variability, we had to explore measures of imputation in order to render hs-CRP levels between COLCOT and LoDoCo comparable.

In addition to these aforementioned limitations, the current thesis also has some overall limitations that are worth stating. First, the use of observational data could lead to confounding factors. Although we attempted to control for potential confounders, residual confounding cannot entirely be ruled out. The presence of unknown or unmeasured confounders may influence the observed associations, and causal inferences should be made cautiously. Second, we relied on existing clinical and genotyping data repositories, which may naturally limit the scope of our analyses, given its post-hoc setting. These datasets may not include all relevant variables or populations, and the generalizability of our findings are rendered limited. The specific patient cohorts included in the studies may not be representative of all clinical settings, further limiting the applicability of our results to diverse populations. Third, the analyses may not capture the full complexity of immune (dys)regulation, as it is still a highly complex and heterogeneous process. The factors contributing to defective immunity may be multifaceted, and our understanding of these processes is still evolving. The nature of the study designs without a fundamental sub-aim implies that our findings, although interesting, remain predominantly hypothesis-generating. Further research will be necessary to elucidate the precise mechanisms underlying immune dysregulation and its role in the co-occurrence of CVD and cancer. Fourth, the effect of CH on CVD outcomes in cancer patients was found to be variable across different types of cancer, which may limit the applicability of the findings to specific cancer types. This variability suggests that the associations between CH and outcomes may be context-dependent, and additional research is needed to understand the factors that contribute to this variability.

6.3 Future Perspectives

The current setting of CH research has turned towards the targeting of inflammatory cytokines in order to break the CH-inflammatory cycle to lower the risks of atherosclerosis. What will help achieve this goal is to better understand what shapes the fitness landscape of CH. Whether targeting inflammatory pathways can lead to improved outcomes for both malignancy and atherosclerosis will also involve a better characterization of how clones evolve as a result

exposure to cancer therapy.[359] In that regard, serial blood sampling of CH clones prior to intervention, and periodically as they are exposed to anti-inflammatory therapy would be ideal. Preliminary data of pediatric cancers suggest distinct dynamics of CH as a result of treatment exposure to chemotherapy.[422] Repeated time series evaluation of larger clone size was a significant predictor of age-related CH-expansion in the long-term, while therapy-related CH remained stable decades thereafter. Alternatively, clonal evolution may now also be inferred using bioinformatic tools developed for the purpose of understanding clonal expansion with only a single measure.[423] There is also the question as to whether clones' growth rate is permanently thwarted following an effective anti-inflammatory target, or could eventually rebound under treatment discontinuation, as previously seen in macaques.[420]

As a specific endeavor, our work exploring the relationship between CH and CV health in cancer survivors revealed variable results for different cancer types. Notably, we observed significant effects of CH on CV outcomes in patients diagnosed with kidney cancer. Going forward, I felt that this finding merits further investigation into the mechanisms linking CH and kidney cancer.

In my submitted post-doctoral grant applications, I have proposed several approaches to advance our understanding of this relationship. First, I intend to estimate the genetic correlations of kidney cancer and previously identified hematopoietic phenotypes. This analysis will help identify potential shared genetic factors contributing to both conditions. It could be conducted by estimating the genetic correlations of renal cell carcinoma and hematopoietic phenotypes using GWAS summary statistics. Second, I wish to attempt to assess the impact of CH on CV outcomes in patients diagnosed with kidney cancer by leveraging the Dana-Farber/Harvard Cancer Center biobank. CH may be defined using CHIP and mCA. This analysis will provide valuable insights into the role of CH in the CV outcomes of kidney cancer patients. Moreover, considering the use of tyrosine kinase inhibitors (e.g., sunitinib, sorafenib) and immune checkpoint inhibitors (e.g., nivolumab, ipilimumab) in the current treatment landscape for kidney cancer, will be highly relevant given the current treatment management culture of patients diagnosed with kidney cancer. Finally, I plan to consider comparing the inflammatory gene expression profiles between kidney cancer patients carrying common CHIP mutations (e.g. in DNMT3A and TET2) and those without CHIP mutations by analyzing bulk RNA-seq data from peripheral mononuclear cells.

Additional analyses should focus on patients who experienced a CV event to further explore the relationship between CH and CV outcomes in kidney cancer patients.

6.4 Conclusions

In conclusion, findings from our work on the potential of non-specific inhibition of inflammation in reducing the risk of cancer (**Article 1**) emphasize the complex interplay between genetics, immune function, and disease risk in the context of statin use and cancer risk among patients with CAD. The identification of a sex-specific genetic variant highlights the need for further investigation into the intricate mechanisms underlying immune dysregulation and its role in disease development. Our results suggest that conventional anti-inflammatory treatments, such as statins, may be insufficient for addressing long-term autoimmunity and underline the potential for developing targeted therapeutics or tools to effectively regulate patients' immune gradients. As the immune system remains a highly complex and heterogenous process, future research should continue to explore the multifaceted relationships between genetics, immune function, and various health outcomes to improve patient care and inform personalized treatment strategies.

In our work on the influence of CH on the risk of cardiovascular-related health in cancer survivors **(Article 2),** we utilized the UK Biobank resource to identify patients diagnosed with select cancers at risk of CVD-related death and systematically evaluated the influence of mCAs on the risk of death from CVD causes, CAD causes, cancer causes, and any cause. Our findings suggested that individuals diagnosed with cancer and at least one type of mCA had a higher risk of death from CAD causes, with the effect being highly variable across different types of cancer. This effect did not appear to be significantly modulated by the receipt of chemotherapy or radiotherapy. In a follow-up study **(Article 3)**, we examined the influence of CHIP on death from CVD causes, CAD causes, and any cause among cancer patients with available WES and genotyping information. Our analyses revealed significant associations between CHIP and death from CVD causes. Furthermore, the co-occurrence of CHIP and mCA did not result in worse outcomes than those with CHIP mutations alone with respect to death from CVD causes, but those with both types of CH had a higher risk of all-cause mortality. In our other scientific contributions **(Article 4)**, we

assessed the correlation between CHIP mutations and hs-CRP concentrations in patients recruited at the MHI Biobank. Our findings showed that CHIP carriers had 21% higher hs-CRP levels compared to their non-CHIP counterparts. Furthermore, we demonstrated that colchicine, an anti-inflammatory treatment, was significantly associated with post-treatment hs-CRP values of $\leqslant$1.0 mg/L compared to placebo **(Article 5)**.

Overall, our findings support the possibility of considering CH (either CHIP or mCA) as a marker of CV-profile risk stratification in patients diagnosed with cancer, potentially impacting treatment management and cardiology surveillance.

Bibliography

1. GBD 2019 Diseases and Injuries Collaborators. Global burden of 369 diseases and injuries in 204 countries and territories, 1990-2019: a systematic analysis for the Global Burden of Disease Study 2019. *Lancet* **396**, 1204–1222 (2020).

2. Miller, K. D. *et al.* Cancer treatment and survivorship statistics, 2022. *CA Cancer J. Clin.* **72**, 409–436 (2022).

3. Tsao, C. W. *et al.* Heart disease and stroke statistics-2022 update: A report from the American Heart Association. *Circulation* **145**, e153–e639 (2022).

4. Armenian, S. H. *et al.* Cardiovascular Disease Among Survivors of Adult-Onset Cancer: A Community-Based Retrospective Cohort Study. *J. Clin. Oncol.* **34**, 1122–1130 (2016).

5. Mulrooney, D. A. *et al.* Cardiac outcomes in a cohort of adult survivors of childhood and adolescent cancer: retrospective analysis of the Childhood Cancer Survivor Study cohort. *BMJ* **339**, b4606 (2009).

6. Haugnes, H. S. *et al.* Cardiovascular risk factors and morbidity in long-term survivors of testicular cancer: a 20-year follow-up study. *J. Clin. Oncol.* **28**, 4649–4657 (2010).

7. Sturgeon, K. M. *et al.* A population-based study of cardiovascular disease mortality risk in US cancer patients. *Eur. Heart J.* **40**, 3889–3897 (2019).

8. Hasin, T. *et al.* Heart failure after myocardial infarction is associated with increased risk of cancer. *J. Am. Coll. Cardiol.* **68**, 265–271 (2016).

9. Meijers, W. C. *et al.* Heart failure stimulates tumor growth by circulating factors. *Circulation* **138**, 678–691 (2018).

10. Simes, J. *et al.* D-dimer predicts long-term cause-specific mortality, cardiovascular events, and cancer in patients with stable coronary heart disease: LIPID study. *Circulation* **138**, 712–723 (2018).

11. Sundbøll, J., Veres, K., Horváth-Puhó, E., Adelborg, K. & Sørensen, H. T. Risk and prognosis of cancer after lower limb arterial thrombosis. *Circulation* **138**, 669–677 (2018).

12. Velders, M. A., Hagström, E. & James, S. K. Temporal trends in the prevalence of cancer and its impact on outcome in patients with first myocardial infarction: A nationwide study. *J. Am. Heart Assoc.* **9**, e014383 (2020).

13. Morris, P. B. *et al.* Cardiovascular effects of exposure to cigarette smoke and electronic cigarettes. *J. Am. Coll. Cardiol.* **66**, 1378–1391 (2015).

14. Gandini, S. *et al.* Tobacco smoking and cancer: a meta-analysis. *Int. J. Cancer* **122**, 155–164 (2008).

15. Steppan, C. M. *et al.* The hormone resistin links obesity to diabetes. *Nature* **409**, 307–312 (2001).

16. Giovannucci, E. *et al.* Diabetes and cancer: a consensus report. *CA Cancer J. Clin.* **60**, 207–221 (2010).

17. Renehan, A. G. *et al.* Insulin-like growth factor (IGF)-I, IGF binding protein-3, and cancer risk: systematic review and meta-regression analysis. *Lancet* **363**, 1346–1353 (2004).

18. Chen, W. *et al.* Phenotypes and genotypes of insulin-like growth factor 1, IGF-binding protein-3 and cancer risk: evidence from 96 studies. *Eur. J. Hum. Genet.* **17**, 1668–1675 (2009).

19. Nöthlings, U., Ford, E. S., Kröger, J. & Boeing, H. Lifestyle factors and mortality among adults with diabetes: findings from the European Prospective Investigation into Cancer and Nutrition-Potsdam study. *J. Diabetes* **2**, 112–117 (2010).

20. Moslehi, J. J. Cardiovascular Toxic Effects of Targeted Cancer Therapies. *N. Engl. J. Med.* **375**, 1457–1467 (2016).

21. Lenihan, D. J. & Cardinale, D. M. Late cardiac effects of cancer treatment. *J. Clin. Oncol.* **30**, 3657–3664 (2012).

22. Rothe, D., Paterson, I., Cox-Kennett, N., Gyenes, G. & Pituskin, E. Prevention of cardiovascular disease among cancer survivors: The role of pre-existing risk factors and cancer treatments. *Curr. Epidemiol. Rep.* **4**, 239–247 (2017).

23. Strongman, H. *et al.* Medium and long-term risks of specific cardiovascular diseases in survivors of 20 adult cancers: a population-based cohort study using multiple linked UK electronic health records databases. *Lancet* **394**, 1041–1054 (2019).

24. Maayah, Z. H. *et al.* Breast cancer diagnosis is associated with relative left ventricular hypertrophy and elevated endothelin-1 signaling. *BMC Cancer* **20**, 751 (2020).

25. Tadic, M. *et al.* Left ventricular strain in chemotherapy-naive and radiotherapy-naive patients with cancer. *Can. J. Cardiol.* **34**, 281–287 (2017).

26. Próchnicki, T. & Latz, E. Inflammasomes on the Crossroads of innate immune recognition and metabolic control. *Cell Metab.* **26**, 71–93 (2017).

27. Grivennikov, S. I., Greten, F. R. & Karin, M. Immunity, inflammation, and cancer. *Cell* **140**, 883–899 (2010).

28. Furman, D. *et al.* Chronic inflammation in the etiology of disease across the life span. *Nat. Med.* **25**, 1822–1832 (2019).

29. Medzhitov, R. Origin and physiological roles of inflammation. *Nature* **454**, 428–435 (2008).

30. Medzhitov, R. The spectrum of inflammatory responses. *Science* **374**, 1070–1075 (2021).

31. Besedovsky, L., Lange, T. & Haack, M. The sleep-immune crosstalk in health and disease. *Physiol. Rev.* **99**, 1325–1380 (2019).

32. Franceschi, C., Garagnani, P., Parini, P., Giuliani, C. & Santoro, A. Inflammaging: a new immune–metabolic viewpoint for age-related diseases. *Nat. Rev. Endocrinol.* **14**, 576–590 (2018).

33. López-Otín, C., Blasco, M. A., Partridge, L., Serrano, M. & Kroemer, G. The hallmarks of aging. *Cell* **153**, 1194–1217 (2013).

34. Colaço, H. G. & Moita, L. F. Initiation of innate immune responses by surveillance of homeostasis perturbations. *FEBS J.* **283**, 2448–2457 (2016).

35. Kotas, M. E. & Medzhitov, R. Homeostasis, inflammation, and disease susceptibility. *Cell* **160**, 816–827 (2015).

36. Medzhitov, R., Schneider, D. S. & Soares, M. P. Disease tolerance as a defense strategy. *Science* **335**, 936–941 (2012).

37. Ayres, J. S. & Schneider, D. S. Tolerance of infections. *Annu. Rev. Immunol.* **30**, 271–294 (2012).

38. Chaudhry, A. & Rudensky, A. Y. Control of inflammation by integration of environmental cues by regulatory T cells. *J. Clin. Invest.* **123**, 939–944 (2013).

39. Serhan, C. N. Pro-resolving lipid mediators are leads for resolution physiology. *Nature* **510**, 92–101 (2014).

40. Buckley, C. D., Gilroy, D. W., Serhan, C. N., Stockinger, B. & Tak, P. P. The resolution of inflammation. *Nat. Rev. Immunol.* **13**, 59–66 (2013).

41. Kennedy, B. K. *et al.* Geroscience: linking aging to chronic disease. *Cell* **159**, 709–713 (2014).

42. Berg, A. H. & Scherer, P. E. Adipose tissue, inflammation, and cardiovascular disease. *Circ. Res.* **96**, 939–949 (2005).

43. Frayn, K., Bernard, S., Spalding, K. & Arner, P. Adipocyte triglyceride turnover is independently associated with atherogenic dyslipidemia. *J. Am. Heart Assoc.* **1**, e003467 (2012).

44. Rodríguez-Iturbe, B., Pons, H., Quiroz, Y. & Johnson, R. J. The immunological basis of hypertension. *Am. J. Hypertens.* **27**, 1327–1337 (2014).

45. Esposito, K. *et al.* Inflammatory cytokine concentrations are acutely increased by hyperglycemia in humans: role of oxidative stress. *Circulation* **106**, 2067–2072 (2002).

46. Balkwill, F. & Mantovani, A. Inflammation and cancer: back to Virchow? *Lancet* **357**, 539–545 (2001).

47. Binnewies, M. *et al.* Understanding the tumor immune microenvironment (TIME) for effective therapy. *Nat. Med.* **24**, 541–550 (2018).

48. Okabe, Y. & Medzhitov, R. Tissue biology perspective on macrophages. *Nat. Immunol.* **17**, 9–17 (2016).

49. Coussens, L. M. & Werb, Z. Inflammation and cancer. *Nature* **420**, 860–867 (2002).

50. Zhou, X. *et al.* Circuit design features of a stable two-cell system. *Cell* **172**, 744-757.e17 (2018).

51. Koebel, C. M. *et al.* Adaptive immunity maintains occult cancer in an equilibrium state. *Nature* **450**, 903–907 (2007).

52. McGranahan, N. & Swanton, C. Cancer evolution constrained by the immune microenvironment. *Cell* vol. 170 825–827 (2017).

53. Zilionis, R. *et al.* Single-cell transcriptomics of human and mouse lung cancers reveals conserved myeloid populations across individuals and species. *Immunity* **50**, 1317-1334.e10 (2019).

54. Kuper, H., Adami, H.-O. & Trichopoulos, D. Infections as a major preventable cause of human cancer. *J. Intern. Med.* **249**, 61–74 (2001).

55. Wahl, L. M. & Kleinman, H. K. Tumor-associated macrophages as targets for cancer therapy. *Journal of the National Cancer Institute* vol. 90 1583–1584 (1998).

56. Brigati, C., Noonan, D. M., Albini, A. & Benelli, R. Tumors and inflammatory infiltrates: friends or foes? *Clin. Exp. Metastasis* **19**, 247–258 (2002).

57. Tsung, K., Dolan, J. P., Tsung, Y. L. & Norton, J. A. Macrophages as effector cells in interleukin 12-induced T cell-dependent tumor rejection. *Cancer Res.* **62**, 5069–5075 (2002).

58. Schoppmann, S. F. *et al.* Tumor-associated macrophages express lymphatic endothelial growth factors and are related to peritumoral lymphangiogenesis. *Am. J. Pathol.* **161**, 947–956 (2002).

59. DeNardo, D. G. *et al.* CD4(+) T cells regulate pulmonary metastasis of mammary carcinomas by enhancing protumor properties of macrophages. *Cancer Cell* **16**, 91–102 (2009).

60. Langowski, J. L., Kastelein, R. A. & Oft, M. Swords into plowshares: IL-23 repurposes tumor immune surveillance. *Trends Immunol.* **28**, 207–212 (2007).

61. Smyth, M. J., Dunn, G. P. & Schreiber, R. D. Cancer immunosurveillance and immunoediting: the roles of immunity in suppressing tumor development and shaping tumor immunogenicity. *Adv. Immunol.* **90**, 1–50 (2006).

62. Galon, J. *et al.* Type, density, and location of immune cells within human colorectal tumors predict clinical outcome. *Science* **313**, 1960–1964 (2006).

63. Swann, J. B. & Smyth, M. J. Immune surveillance of tumors. *J. Clin. Invest.* **117**, 1137–1146 (2007).

64. Roberts, S. J. *et al.* Characterizing tumor-promoting T cells in chemically induced cutaneous carcinogenesis. *Proc. Natl. Acad. Sci. U. S. A.* **104**, 6770–6775 (2007).

65. Aspord, C. *et al.* Breast cancer instructs dendritic cells to prime interleukin 13-secreting CD4+ T cells that facilitate tumor development. *J. Exp. Med.* **204**, 1037–1047 (2007).

66. Langowski, J. L. *et al.* IL-23 promotes tumour incidence and growth. *Nature* **442**, 461–465 (2006).

67. Lin, W.-W. & Karin, M. A cytokine-mediated link between innate immunity, inflammation, and cancer. *J. Clin. Invest.* **117**, 1175–1183 (2007).

68. Erdman, S. E. *et al.* CD4+CD25+ regulatory lymphocytes induce regression of intestinal tumors in ApcMin/+ mice. *Cancer Res.* **65**, 3998–4004 (2005).

69. Mantovani, A., Allavena, P., Sica, A. & Balkwill, F. Cancer-related inflammation. *Nature* **454**, 436–444 (2008).

70. de Visser, K. E., Eichten, A. & Coussens, L. M. Paradoxical roles of the immune system during cancer development. *Nat. Rev. Cancer* **6**, 24–37 (2006).

71. Acha-Sagredo, A., Ganguli, P. & Ciccarelli, F. D. Somatic variation in normal tissues: friend or foe of cancer early detection? *Ann. Oncol.* **33**, 1239–1249 (2022).

72. Hill, W. *et al.* Lung adenocarcinoma promotion by air pollutants. *Nature* **616**, 159–167 (2023).

73. Rodriguez-Meira, A. *et al.* Single-cell multi-omics identifies chronic inflammation as a driver of TP53-mutant leukemic evolution. *Nat. Genet.* **55**, 1531–1541 (2023).

74. Palomero, L. *et al.* Immune cell associations with cancer risk. *iScience* **23**, 101296 (2020).

75. Sayaman, R. W. *et al.* Germline genetic contribution to the immune landscape of cancer. *Immunity* **54**, 367-386.e8 (2021).

76. Fagny, M., Platig, J., Kuijjer, M. L., Lin, X. & Quackenbush, J. Nongenic cancer-risk SNPs affect oncogenes, tumour-suppressor genes, and immune function. *Br. J. Cancer* **122**, 569–577 (2020).

77. Chin, I. S. *et al.* Germline genetic variation and predicting immune checkpoint inhibitor induced toxicity. *NPJ Genom. Med.* **7**, 73 (2022).

78. Han, Y., Chen, W., Li, P. & Ye, J. Association between coeliac disease and risk of any malignancy and gastrointestinal malignancy: A meta-analysis. *Medicine (Baltimore)* **94**, e1612 (2015).

79. Simon, T. A., Thompson, A., Gandhi, K. K., Hochberg, M. C. & Suissa, S. Incidence of malignancy in adult patients with rheumatoid arthritis: a meta-analysis. *Arthritis Res. Ther.* **17**, 212 (2015).

80. Song, L. *et al.* The risks of cancer development in systemic lupus erythematosus (SLE) patients: a systematic review and meta-analysis. *Arthritis Res. Ther.* **20**, 270 (2018).

81. Vaengebjerg, S., Skov, L., Egeberg, A. & Loft, N. D. Prevalence, incidence, and risk of cancer in patients with psoriasis and psoriatic arthritis: A systematic review and Meta-analysis. *JAMA Dermatol.* **156**, 421–429 (2020).

82. Wan, Q. *et al.* Inflammatory bowel disease and risk of gastric, small bowel and colorectal cancer: a meta-analysis of 26 observational studies. *J. Cancer Res. Clin. Oncol.* **147**, 1077–1087 (2021).

83. Rothwell, P. M. *et al.* Effect of daily aspirin on long-term risk of death due to cancer: analysis of individual patient data from randomised trials. *Lancet* **377**, 31–41 (2011).

84. Rothwell, P. M. *et al.* Effect of daily aspirin on risk of cancer metastasis: a study of incident cancers during randomised controlled trials. *Lancet* **379**, 1591–1601 (2012).

85. Ridker, P. M. *et al.* Rosuvastatin to prevent vascular events in men and women with elevated C-reactive protein. *N. Engl. J. Med.* **359**, 2195–2207 (2008).

86. Puram, S. V. *et al.* Single-cell transcriptomic analysis of primary and metastatic tumor ecosystems in head and neck cancer. *Cell* **171**, 1611-1624.e24 (2017).

87. Grivennikov, S. I. *et al.* Adenoma-linked barrier defects and microbial products drive IL-23/IL-17-mediated tumour growth. *Nature* **491**, 254–258 (2012).

88. Karin, M. Nuclear factor-kappaB in cancer development and progression. *Nature* **441**, 431–436 (2006).

89. Wu, S. *et al.* A human colonic commensal promotes colon tumorigenesis via activation of T helper type 17 T cell responses. *Nat. Med.* **15**, 1016–1022 (2009).

90. Waldner, M. J. & Neurath, M. F. Colitis-associated cancer: the role of T cells in tumor development. *Semin. Immunopathol.* **31**, 249–256 (2009).

91. Greten, F. R. & Grivennikov, S. I. Inflammation and Cancer: Triggers, Mechanisms, and Consequences. *Immunity* **51**, 27–41 (2019).

92. Punturieri, A., Szabo, E., Croxton, T. L., Shapiro, S. D. & Dubinett, S. M. Lung cancer and chronic obstructive pulmonary disease: needs and opportunities for integrated research. *J. Natl. Cancer Inst.* **101**, 554–559 (2009).

93. Tuncman, G. *et al.* Functional in vivo interactions between JNK1 and JNK2 isoforms in obesity and insulin resistance. *Proc. Natl. Acad. Sci. U. S. A.* **103**, 10741–10746 (2006).

94. Calle, E. E. Obesity and cancer. *BMJ (Clinical research ed.)* vol. 335 1107–1108 (2007).

95. Quail, D. F. *et al.* Obesity alters the lung myeloid cell landscape to enhance breast cancer metastasis through IL5 and GM-CSF. *Nat. Cell Biol.* **19**, 974–987 (2017).

96. Quail, D. F. & Dannenberg, A. J. The obese adipose tissue microenvironment in cancer development and progression. *Nat. Rev. Endocrinol.* **15**, 139–154 (2019).

97. Fearon, E. R. & Vogelstein, B. A genetic model for colorectal tumorigenesis. *Cell* **61**, 759–767 (1990).

98. Canli, Ö. *et al.* Myeloid cell-derived reactive oxygen species induce epithelial Mutagenesis. *Cancer Cell* **32**, 869-883.e5 (2017).

99. Chang, W.-C. L. *et al.* Loss of p53 enhances the induction of colitis-associated neoplasia by dextran sulfate sodium. *Carcinogenesis* **28**, 2375–2381 (2007).

100. Meira, L. B. *et al.* DNA damage induced by chronic inflammation contributes to colon carcinogenesis in mice. *J. Clin. Invest.* **118**, 2516–2525 (2008).

101. Gronke, K. *et al.* Interleukin-22 protects intestinal stem cells against genotoxic stress. *Nature* **566**, 249–253 (2019).

102. Grivennikov, S. I. Inflammation and colorectal cancer: colitis-associated neoplasia. *Semin. Immunopathol.* **35**, 229–244 (2013).

103. Schwitalla, S. *et al.* Intestinal tumorigenesis initiated by dedifferentiation and acquisition of stem-cell-like properties. *Cell* **152**, 25–38 (2013).

104. Mehlen, P. & Puisieux, A. Metastasis: a question of life or death. *Nat. Rev. Cancer* **6**, 449–458 (2006).

105. Peitzsch, C., Tyutyunnykova, A., Pantel, K. & Dubrovska, A. Cancer stem cells: The root of tumor recurrence and metastases. *Semin. Cancer Biol.* **44**, 10–24 (2017).

106. Dominguez, C., David, J. M. & Palena, C. Epithelial-mesenchymal transition and inflammation at the site of the primary tumor. *Semin. Cancer Biol.* **47**, 177–184 (2017).

107. Panigrahy, D. *et al.* Preoperative stimulation of resolution and inflammation blockade eradicates micrometastases. *J. Clin. Invest.* **129**, 2964–2979 (2019).

108. Francart, M.-E. *et al.* Epithelial-mesenchymal plasticity and circulating tumor cells: Travel companions to metastases. *Dev. Dyn.* **247**, 432–450 (2018).

109. Akkari, L. *et al.* Distinct functions of macrophage-derived and cancer cell-derived cathepsin Z combine to promote tumor malignancy via interactions with the extracellular matrix. *Genes Dev.* **28**, 2134–2150 (2014).

110. Veglia, F., Perego, M. & Gabrilovich, D. Myeloid-derived suppressor cells coming of age. *Nat. Immunol.* **19**, 108–119 (2018).

111. Yang, L. *et al.* Abrogation of TGF beta signaling in mammary carcinomas recruits Gr-1+CD11b+ myeloid cells that promote metastasis. *Cancer Cell* **13**, 23–35 (2008).

112. Bindea, G. *et al.* Spatiotemporal dynamics of intratumoral immune cells reveal the immune landscape in human cancer. *Immunity* **39**, 782–795 (2013).

113. Sharma, P. & Allison, J. P. The future of immune checkpoint therapy. *Science* **348**, 56–61 (2015).

114. Topalian, S. L., Drake, C. G. & Pardoll, D. M. Immune checkpoint blockade: a common denominator approach to cancer therapy. *Cancer Cell* **27**, 450–461 (2015).

115. Tsoupras, A., Lordan, R. & Zabetakis, I. Inflammation, not cholesterol, is a cause of chronic disease. *Nutrients* **10**, 604 (2018).

116. Ross, R. & Glomset, J. A. Atherosclerosis and the Arterial Smooth Muscle Cell. *Science* **180**, 1332–1339 (1973).

117. Ross, R. Atherosclerosis--an inflammatory disease. *N. Engl. J. Med.* **340**, 115–126 (1999).

118. Liuzzo, G. *et al.* The prognostic value of C-reactive protein and serum amyloid a protein in severe unstable angina. *N. Engl. J. Med.* **331**, 417–424 (1994).

119. Biasucci, L. M. *et al.* Elevated levels of interleukin-6 in unstable angina. *Circulation* **94**, 874–877 (1996).

120. Tsiantoulas, D. *et al.* APRIL limits atherosclerosis by binding to heparan sulfate proteoglycans. *Nature* **597**, 92–96 (2021).

121. Hansson, G. K. & Hermansson, A. The immune system in atherosclerosis. *Nat. Immunol.* **12**, 204–212 (2011).

122. Roy, P., Orecchioni, M. & Ley, K. How the immune system shapes atherosclerosis: roles of innate and adaptive immunity. *Nat. Rev. Immunol.* **22**, 251–265 (2022).

123. Camici, P. G., Rimoldi, O. E., Gaemperli, O. & Libby, P. Non-invasive anatomic and functional imaging of vascular inflammation and unstable plaque. *Eur. Heart J.* **33**, 1309–1317 (2012).

124. Nagenborg, J., Goossens, P., Biessen, E. A. L. & Donners, M. M. P. C. Heterogeneity of atherosclerotic plaque macrophage origin, phenotype and functions: Implications for treatment. *Eur. J. Pharmacol.* **816**, 14–24 (2017).

125. Niessner, A. *et al.* Pathogen-sensing plasmacytoid dendritic cells stimulate cytotoxic T-cell function in the atherosclerotic plaque through interferon-alpha. *Circulation* **114**, 2482–2489 (2006).

126. Combadière, C. *et al.* Combined inhibition of CCL2, CX3CR1, and CCR5 abrogates Ly6C(hi) and Ly6C(lo) monocytosis and almost abolishes atherosclerosis in hypercholesterolemic mice. *Circulation* **117**, 1649–1657 (2008).

127. Hamers, A. A. J. *et al.* Human monocyte heterogeneity as revealed by high-dimensional mass cytometry. *Arterioscler. Thromb. Vasc. Biol.* **39**, 25–36 (2019).

128. Robbins, C. S. *et al.* Extramedullary hematopoiesis generates Ly-6C(high) monocytes that infiltrate atherosclerotic lesions. *Circulation* **125**, 364–374 (2012).

129. Tolani, S. *et al.* Hypercholesterolemia and reduced HDL-C promote hematopoietic stem cell proliferation and monocytosis: studies in mice and FH children. *Atherosclerosis* **229**, 79–85 (2013).

130. Williams, J. W. *et al.* Limited proliferation capacity of aortic intima resident macrophages requires monocyte recruitment for atherosclerotic plaque progression. *Nat. Immunol.* **21**, 1194–1204 (2020).

131. Robbins, C. S. *et al.* Local proliferation dominates lesional macrophage accumulation in atherosclerosis. *Nat. Med.* **19**, 1166–1172 (2013).

132. Depuydt, M. A. C. *et al.* Microanatomy of the human atherosclerotic plaque by single-cell transcriptomics. *Circ. Res.* **127**, 1437–1455 (2020).

133. Zernecke, A. *et al.* Meta-analysis of leukocyte diversity in atherosclerotic mouse aortas. *Circ. Res.* **127**, 402–426 (2020).

134. Cochain, C. *et al.* Single-cell RNA-seq reveals the transcriptional landscape and heterogeneity of aortic macrophages in Murine atherosclerosis. *Circ. Res.* **122**, 1661–1674 (2018).

135. Murray, P. J. Macrophage polarization. *Annu. Rev. Physiol.* **79**, 541–566 (2017).

136. Fernandez, D. M. *et al.* Single-cell immune landscape of human atherosclerotic plaques. *Nat. Med.* **25**, 1576–1588 (2019).

137. Kim, K. *et al.* Transcriptome analysis reveals nonfoamy rather than foamy plaque macrophages are proinflammatory in atherosclerotic Murine models. *Circ. Res.* **123**, 1127–1142 (2018).

138. Deguchi, J.-O. *et al.* Inflammation in atherosclerosis. *Circulation* **114**, 55–62 (2006).

139. Goossens, P. *et al.* Myeloid type I interferon signaling promotes atherosclerosis by stimulating macrophage recruitment to lesions. *Cell Metab.* **12**, 142–153 (2010).

140. Merad, M., Sathe, P., Helft, J., Miller, J. & Mortha, A. The dendritic cell lineage: ontogeny and function of dendritic cells and their subsets in the steady state and the inflamed setting. *Annu. Rev. Immunol.* **31**, 563–604 (2013).

141. Subramanian, M. & Tabas, I. Dendritic cells in atherosclerosis. *Semin. Immunopathol.* **36**, 93–102 (2014).

142. Yilmaz, A. *et al.* Emergence of dendritic cells in rupture-prone regions of vulnerable carotid plaques. *Atherosclerosis* **176**, 101–110 (2004).

143. Maganto-García, E. *et al.* Foxp3+-inducible regulatory T cells suppress endothelial activation and leukocyte recruitment. *J. Immunol.* **187**, 3521–3529 (2011).

144. Robertson, A.-K. L. *et al.* Disruption of TGF-beta signaling in T cells accelerates atherosclerosis. *J. Clin. Invest.* **112**, 1342–1350 (2003).

145. Subramanian, M., Thorp, E., Hansson, G. K. & Tabas, I. Treg-mediated suppression of atherosclerosis requires MYD88 signaling in DCs. *J. Clin. Invest.* **123**, 179–188 (2013).

146. Weber, C. *et al.* CCL17-expressing dendritic cells drive atherosclerosis by restraining regulatory T cell homeostasis in mice. *J. Clin. Invest.* **121**, 2898–2910 (2011).

147. Soehnlein, O. Multiple roles for neutrophils in atherosclerosis. *Circ. Res.* **110**, 875–888 (2012).

148. Naruko, T. *et al.* Neutrophil infiltration of culprit lesions in acute coronary syndromes. *Circulation* **106**, 2894–2900 (2002).

149. Warnatsch, A., Ioannou, M., Wang, Q. & Papayannopoulos, V. Inflammation. Neutrophil extracellular traps license macrophages for cytokine production in atherosclerosis. *Science* **349**, 316–320 (2015).

150. Moschonas, I. C. & Tselepis, A. D. The pathway of neutrophil extracellular traps towards atherosclerosis and thrombosis. *Atherosclerosis* **288**, 9–16 (2019).

151. Zernecke, A. *et al.* Protective role of CXC receptor 4/CXC ligand 12 unveils the importance of neutrophils in atherosclerosis. *Circ. Res.* **102**, 209–217 (2008).

152. Silvestre-Roig, C. *et al.* Externalized histone H4 orchestrates chronic inflammation by inducing lytic cell death. *Nature* **569**, 236–240 (2019).

153. Fuchs, T. A. *et al.* Extracellular DNA traps promote thrombosis. *Proc. Natl. Acad. Sci. U. S. A.* **107**, 15880–15885 (2010).

154. Saigusa, R., Winkels, H. & Ley, K. T cell subsets and functions in atherosclerosis. *Nat. Rev. Cardiol.* **17**, 387–401 (2020).

155. Ketelhuth, D. F. J. & Hansson, G. K. Adaptive response of T and B cells in atherosclerosis. *Circ. Res.* **118**, 668–678 (2016).

156. Wolf, D. *et al.* Pathogenic autoimmunity in atherosclerosis evolves from initially protective apolipoprotein B100-reactive CD4+ T-regulatory cells. *Circulation* **142**, 1279–1293 (2020).

157. Kyaw, T. *et al.* Cytotoxic and proinflammatory CD8+ T lymphocytes promote development of vulnerable atherosclerotic plaques in apoE-deficient mice. *Circulation* **127**, 1028–1039 (2013).

158. van Duijn, J., Kuiper, J. & Slütter, B. The many faces of CD8+ T cells in atherosclerosis. *Curr. Opin. Lipidol.* **29**, 411–416 (2018).

159. Hwang, Y. *et al.* Expansion of CD8+ T cells lacking the IL-6 receptor α chain in patients with coronary artery diseases (CAD). *Atherosclerosis* **249**, 44–51 (2016).

160. Bergström, I., Backteman, K., Lundberg, A., Ernerudh, J. & Jonasson, L. Persistent accumulation of interferon-γ-producing CD8+CD56+ T cells in blood from patients with coronary artery disease. *Atherosclerosis* **224**, 515–520 (2012).

161. Brennan, P. J., Brigl, M. & Brenner, M. B. Invariant natural killer T cells: an innate activation scheme linked to diverse effector functions. *Nat. Rev. Immunol.* **13**, 101–117 (2013).

162. Getz, G. S. & Reardon, C. A. Natural killer T cells in atherosclerosis. *Nat. Rev. Cardiol.* **14**, 304–314 (2017).

163. Bobryshev, Y. V. & Lord, R. S. A. Co-accumulation of dendritic cells and natural killer T cells within rupture-prone regions in human atherosclerotic plaques. *J. Histochem. Cytochem.* **53**, 781–785 (2005).

164. Binder, C. J., Papac-Milicevic, N. & Witztum, J. L. Innate sensing of oxidation-specific epitopes in health and disease. *Nat. Rev. Immunol.* **16**, 485–497 (2016).

165. Kyaw, T. *et al.* B1a B lymphocytes are atheroprotective by secreting natural IgM that increases IgM deposits and reduces necrotic cores in atherosclerotic lesions. *Circ. Res.* **109**, 830–840 (2011).

166. Sage, A. P., Tsiantoulas, D., Binder, C. J. & Mallat, Z. The role of B cells in atherosclerosis. *Nat. Rev. Cardiol.* **16**, 180–196 (2019).

167. Nus, M. *et al.* Marginal zone B cells control the response of follicular helper T cells to a high-cholesterol diet. *Nat. Med.* **23**, 601–610 (2017).

168. Doran, A. C., Yurdagul, A., Jr & Tabas, I. Efferocytosis in health and disease. *Nat. Rev. Immunol.* **20**, 254–267 (2020).

169. Kojima, Y., Weissman, I. L. & Leeper, N. J. The role of efferocytosis in atherosclerosis. *Circulation* **135**, 476–489 (2017).

170. Cai, B. *et al.* MerTK receptor cleavage promotes plaque necrosis and defective resolution in atherosclerosis. *J. Clin. Invest.* **127**, 564–568 (2017).

171. Doran, A. C. *et al.* CAMKIIγ suppresses an efferocytosis pathway in macrophages and promotes atherosclerotic plaque necrosis. *J. Clin. Invest.* **127**, 4075–4089 (2017).

172. Thorp, E. *et al.* Shedding of the Mer tyrosine kinase receptor is mediated by ADAM17 protein through a pathway involving reactive oxygen species, protein kinase Cδ, and p38 mitogen-activated protein kinase (MAPK). *J. Biol. Chem.* **286**, 33335–33344 (2011).

173. Kojima, Y. *et al.* Cyclin-dependent kinase inhibitor 2B regulates efferocytosis and atherosclerosis. *J. Clin. Invest.* **124**, 1083–1097 (2014).

174. Kojima, Y. *et al.* CD47-blocking antibodies restore phagocytosis and prevent atherosclerosis. *Nature* **536**, 86–90 (2016).

175. Overton, C. D., Yancey, P. G., Major, A. S., Linton, M. F. & Fazio, S. Deletion of macrophage LDL receptor-related protein increases atherogenesis in the mouse. *Circ. Res.* **100**, 670–677 (2007).

176. Advani, R. *et al.* CD47 blockade by Hu5F9-G4 and rituximab in non-Hodgkin's lymphoma. *N. Engl. J. Med.* **379**, 1711–1721 (2018).

177. Ansell, S. M. *et al.* Phase I study of the CD47 blocker TTI-621 in patients with relapsed or refractory hematologic malignancies. *Clin. Cancer Res.* **27**, 2190–2199 (2021).

178. Westlake, S. L. *et al.* Tumour necrosis factor antagonists and the risk of cardiovascular disease in patients with rheumatoid arthritis: a systematic literature review. *Rheumatology (Oxford)* **50**, 518–531 (2011).

179. Jarr, K.-U. *et al.* The pleiotropic benefits of statins include the ability to reduce CD47 and amplify the effect of pro-efferocytic therapies in atherosclerosis. *Nat Cardiovasc Res* **1**, 253–262 (2022).

180. Ridker, P. M., Cushman, M., Stampfer, M. J., Tracy, R. P. & Hennekens, C. H. Inflammation, aspirin, and the risk of cardiovascular disease in apparently healthy men. *N. Engl. J. Med.* **336**, 973–979 (1997).

181. Ridker, P. M., Rifai, N., Rose, L., Buring, J. E. & Cook, N. R. Comparison of C-reactive protein and low-density lipoprotein cholesterol levels in the prediction of first cardiovascular events. *N. Engl. J. Med.* **347**, 1557–1565 (2002).

182. Pai, J. K. *et al.* Inflammatory markers and the risk of coronary heart disease in men and women. *N. Engl. J. Med.* **351**, 2599–2610 (2004).

183. Sacks, F. M. *et al.* The effect of pravastatin on coronary events after myocardial infarction in patients with average cholesterol levels. Cholesterol and Recurrent Events Trial investigators. *N. Engl. J. Med.* **335**, 1001–1009 (1996).

184. Albert, M. A., Danielson, E., Rifai, N., Ridker, P. M. & PRINCE Investigators. Effect of statin therapy on C-reactive protein levels: the pravastatin inflammation/CRP evaluation (PRINCE): a randomized trial and cohort study. *JAMA* **286**, 64–70 (2001).

185. Cannon, C. P. *et al.* Intensive versus Moderate Lipid Lowering with Statins after Acute Coronary Syndromes. *N. Engl. J. Med.* **350**, 1495–1504 (2004).

186. Ridker, P. M. *et al.* Relative efficacy of atorvastatin 80 mg and pravastatin 40 mg in achieving the dual goals of low-density lipoprotein cholesterol <70 mg/dl and C-reactive protein <2 mg/l: an analysis of the PROVE-IT TIMI-22 trial. *J. Am. Coll. Cardiol.* **45**, 1644–1648 (2005).

187. Ridker, P. M. *et al.* C-Reactive Protein Levels and Outcomes after Statin Therapy. *N. Engl. J. Med.* **352**, 20–28 (2005).

188. Engelen, S. E., Robinson, A. J. B., Zurke, Y.-X. & Monaco, C. Therapeutic strategies targeting inflammation and immunity in atherosclerosis: how to proceed? *Nat. Rev. Cardiol.* (2022) doi:10.1038/s41569-021-00668-4.

189. Ridker, P. M. *et al.* Effects of interleukin-1β inhibition with canakinumab on hemoglobin A1c, lipids, C-reactive protein, interleukin-6, and fibrinogen: a phase IIb randomized, placebo-controlled trial. *Circulation* **126**, 2739–2748 (2012).

190. Ridker, P. M. *et al.* Antiinflammatory Therapy with Canakinumab for Atherosclerotic Disease. *N. Engl. J. Med.* **377**, 1119–1131 (2017).

191. Ridker, P. M. *et al.* Relationship of C-reactive protein reduction to cardiovascular event reduction following treatment with canakinumab: a secondary analysis from the CANTOS randomised controlled trial. *Lancet* **391**, 319–328 (2018).

192. Ridker, P. M. *et al.* Modulation of the interleukin-6 signalling pathway and incidence rates of atherosclerotic events and all-cause mortality: analyses from the Canakinumab Anti-Inflammatory Thrombosis Outcomes Study (CANTOS). *Eur. Heart J.* **39**, 3499–3507 (2018).

193. Hui, M. *et al.* The British society for rheumatology guideline for the management of gout. *Rheumatology (Oxford)* **56**, e1–e20 (2017).

194. Sprigings, D. & B. Chambers, J. Acute pericarditis. in *Acute Medicine - A Practical Guide to the Management of Medical Emergencies, 5th Edition* 340–343 (John Wiley & Sons, Ltd, 2017).

195. Paschke, S. *et al.* Technical advance: Inhibition of neutrophil chemotaxis by colchicine is modulated through viscoelastic properties of subcellular compartments. *J. Leukoc. Biol.* **94**, 1091–1096 (2013).

196. Leung, Y. Y., Yao Hui, L. L. & Kraus, V. B. Colchicine--Update on mechanisms of action and therapeutic uses. *Semin. Arthritis Rheum.* **45**, 341–350 (2015).

197. Dalbeth, N., Lauterio, T. J. & Wolfe, H. R. Mechanism of action of colchicine in the treatment of gout. *Clin. Ther.* **36**, 1465–1479 (2014).

198. Cronstein, B. N. *et al.* Colchicine alters the quantitative and qualitative display of selectins on endothelial cells and neutrophils. *J. Clin. Invest.* **96**, 994–1002 (1995).

199. Nidorf, S. M., Eikelboom, J. W., Budgeon, C. A. & Thompson, P. L. Low-dose colchicine for secondary prevention of cardiovascular disease. *J. Am. Coll. Cardiol.* **61**, 404–410 (2013).

200. Ridker, P. M. *et al.* Low-Dose Methotrexate for the Prevention of Atherosclerotic Events. *N. Engl. J. Med.* **380**, 752–762 (2019).

201. Tardif, J.-C. *et al.* Efficacy and Safety of Low-Dose Colchicine after Myocardial Infarction. *N. Engl. J. Med.* (2019) doi:10.1056/NEJMoa1912388.

202. Nidorf, S. M. *et al.* Colchicine in Patients with Chronic Coronary Disease. *N. Engl. J. Med.* **383**, 1838–1847 (2020).

203. Ridker, P. M. *et al.* Inflammation and cholesterol as predictors of cardiovascular events among patients receiving statin therapy: a collaborative analysis of three randomised trials. *Lancet* **401**, 1293–1301 (2023).

204. Denise Martin, E., De Nicola, G. F. & Marber, M. S. New therapeutic targets in cardiology: p38 alpha mitogen-activated protein kinase for ischemic heart disease. *Circulation* **126**, 357–368 (2012).

205. Dean, J. L. E., Sarsfield, S. J., Tsounakou, E. & Saklatvala, J. p38 Mitogen-activated protein kinase stabilizes mRNAs that contain cyclooxygenase-2 and tumor necrosis factor AU-rich elements by inhibiting deadenylation. *J. Biol. Chem.* **278**, 39470–39476 (2003).

206. Elkhawad, M. *et al.* Effects of p38 mitogen-activated protein kinase inhibition on vascular and systemic inflammation in patients with atherosclerosis. *JACC Cardiovasc. Imaging* **5**, 911–922 (2012).

207. O'Donoghue, M. L. *et al.* Effect of losmapimod on cardiovascular outcomes in patients hospitalized with acute myocardial infarction: A randomized clinical trial. *JAMA* **315**, 1591–1599 (2016).

208. Newby, L. K. *et al.* Losmapimod, a novel p38 mitogen-activated protein kinase inhibitor, in non-ST-segment elevation myocardial infarction: a randomised phase 2 trial. *Lancet* **384**, 1187–1195 (2014).

209. Ridker, P. M. *et al.* Effect of interleukin-1β inhibition with canakinumab on incident lung cancer in patients with atherosclerosis: exploratory results from a randomised, double-blind, placebo-controlled trial. *Lancet* **390**, 1833–1842 (2017).

210. Narayan, V. *et al.* Mechanistic Biomarkers Informative of Both Cancer and Cardiovascular Disease: JACC State-of-the-Art Review. *J. Am. Coll. Cardiol.* **75**, 2726–2737 (2020).

211. *Novartis top-line results for CANOPY-1 Phase III study support further evaluation of canakinumab in lung cancer |*.

212. Paz-Ares, L. *et al.* 1194MO Canakinumab (CAN) + docetaxel (DTX) for the second- or third-line (2/3L) treatment of advanced non-small cell lung cancer (NSCLC): CANOPY-2 phase III results. *Ann. Oncol.* **32**, S953–S954 (2021).

213. Moslehi, J. J., Salem, J.-E., Sosman, J. A., Lebrun-Vignes, B. & Johnson, D. B. Increased reporting of fatal immune checkpoint inhibitor-associated myocarditis. *Lancet* **391**, 933 (2018).

214. Drobni, Z. D. *et al.* Association between immune checkpoint inhibitors with cardiovascular events and atherosclerotic plaque. *Circulation* **142**, 2299–2311 (2020).

215. Poels, K. *et al.* Immune checkpoint inhibitor therapy aggravates T cell-driven plaque inflammation in atherosclerosis. *JACC CardioOncol* **2**, 599–610 (2020).

216. Axelrod, M. L. *et al.* T cells specific for α-myosin drive immunotherapy-related myocarditis. *Nature* **611**, 818–826 (2022).

217. Won, T. *et al.* Cardiac myosin-specific autoimmune T cells contribute to immune-checkpoint-inhibitor-associated myocarditis. *Cell Rep.* **41**, 111611 (2022).

218. Dale, K. M., Coleman, C. I., Henyan, N. N., Kluger, J. & White, C. M. Statins and cancer risk: a meta-analysis. *JAMA* **295**, 74–80 (2006).

219. Egen, J. G., Ouyang, W. & Wu, L. C. Human anti-tumor immunity: Insights from immunotherapy clinical trials. *Immunity* **52**, 36–54 (2020).

220. Pedersen, T. R. *et al.* Randomised trial of cholesterol lowering in 4444 patients with coronary heart disease: the Scandinavian Simvastatin Survival Study (4S). 1994. *Atheroscler. Suppl.* **5**, 81–87 (2004).

221. Heart Protection Study Collaborative Group. Effects on 11-year mortality and morbidity of lowering LDL cholesterol with simvastatin for about 5 years in 20,536 high-risk individuals: a randomised controlled trial. *Lancet* **378**, 2013–2020 (2011).

222. Shepherd, J. *et al.* Pravastatin in elderly individuals at risk of vascular disease (PROSPER): a randomised controlled trial. *Lancet* **360**, 1623–1630 (2002).

223. Nielsen, S. F., Nordestgaard, B. G. & Bojesen, S. E. Statin use and reduced cancer-related mortality. *The New England journal of medicine* vol. 368 576–577 (2013).

224. Blais, L., Desgagné, A. & LeLorier, J. 3-Hydroxy-3-methylglutaryl coenzyme A reductase inhibitors and the risk of cancer: a nested case-control study. *Arch. Intern. Med.* **160**, 2363–2368 (2000).

225. Graaf, M. R., Beiderbeck, A. B., Egberts, A. C. G., Richel, D. J. & Guchelaar, H.-J. The risk of cancer in users of statins. *J. Clin. Oncol.* **22**, 2388–2394 (2004).

226. Friis, S. *et al.* Cancer risk among statin users: a population-based cohort study. *Int. J. Cancer* **114**, 643–647 (2005).

227. Kaye, J. A. & Jick, H. Statin use and cancer risk in the General Practice Research Database. *Br. J. Cancer* **90**, 635–637 (2004).

228. Poynter, J. N. *et al.* Statins and the risk of colorectal cancer. *N. Engl. J. Med.* **352**, 2184–2192 (2005).

229. Bonovas, S., Filioussi, K., Flordellis, C. S. & Sitaras, N. M. Statins and the risk of colorectal cancer: a meta-analysis of 18 studies involving more than 1.5 million patients. *J. Clin. Oncol.* **25**, 3462–3468 (2007).

230. Beck, P., Wysowski, D. K., Downey, W. & Butler-Jones, D. Statin use and the risk of breast cancer. *J. Clin. Epidemiol.* **56**, 280–285 (2003).

231. Cauley, J. A. *et al.* Lipid-lowering drug use and breast cancer in older women: a prospective study. *J. Womens. Health (Larchmt)* **12**, 749–756 (2003).

232. Coogan, P. F. *et al.* Statin use and the risk of breast and prostate cancer. *Epidemiology* **13**, 262–267 (2002).

233. Boudreau, D. M. *et al.* The association between 3-hydroxy-3-methylglutaryl conenzyme A inhibitor use and breast carcinoma risk among postmenopausal women: a case-control study. *Cancer* **100**, 2308–2316 (2004).

234. Gibson, C. J. & Steensma, D. P. New Insights from Studies of Clonal Hematopoiesis. *Clin. Cancer Res.* **24**, 4633–4642 (2018).

235. Zink, F. *et al.* Clonal hematopoiesis, with and without candidate driver mutations, is common in the elderly. *Blood* **130**, 742–752 (2017).

236. Franco, I. *et al.* Whole genome DNA sequencing provides an atlas of somatic mutagenesis in healthy human cells and identifies a tumor-prone cell type. *Genome Biol.* **20**, 285 (2019).

237. Park, S. *et al.* Clonal dynamics in early human embryogenesis inferred from somatic mutation. *Nature* **597**, 393–397 (2021).

238. Abascal, F. *et al.* Somatic mutation landscapes at single-molecule resolution. *Nature* **593**, 405–410 (2021).

239. Geiger, H., de Haan, G. & Florian, M. C. The ageing haematopoietic stem cell compartment. *Nat. Rev. Immunol.* **13**, 376–389 (2013).

240. Duncan, B. K. & Miller, J. H. Mutagenic deamination of cytosine residues in DNA. *Nature* **287**, 560–561 (1980).

241. Jaiswal, S. *et al.* Age-related clonal hematopoiesis associated with adverse outcomes. *N. Engl. J. Med.* **371**, 2488–2498 (2014).

242. Loh, P.-R., Genovese, G. & McCarroll, S. A. Monogenic and polygenic inheritance become instruments for clonal selection. *Nature* **584**, 136–141 (2020).

243. Loh, P.-R. *et al.* Insights into clonal haematopoiesis from 8,342 mosaic chromosomal alterations. *Nature* **559**, 350–355 (2018).

244. Benditt, E. P. & Benditt, J. M. Evidence for a monoclonal origin of human atherosclerotic plaques. *Proc. Natl. Acad. Sci. U. S. A.* **70**, 1753–1756 (1973).

245. Nakao, T. & Natarajan, P. Clonal hematopoiesis, multi-omics and coronary artery disease. *Nature Cardiovascular Research* 1–3 (2022).

246. Jaiswal, S. *et al.* Clonal Hematopoiesis and Risk of Atherosclerotic Cardiovascular Disease. *N. Engl. J. Med.* **377**, 111–121 (2017).

247. Dorsheimer, L. *et al.* Association of Mutations Contributing to Clonal Hematopoiesis With Prognosis in Chronic Ischemic Heart Failure. *JAMA Cardiol* **4**, 25–33 (2019).

248. Clonal Hematopoiesis and Atherosclerosis. *N. Engl. J. Med.* **377**, 1400–1402 (2017).

249. Moran-Crusio, K. *et al.* Tet2 loss leads to increased hematopoietic stem cell self-renewal and myeloid transformation. *Cancer Cell* **20**, 11–24 (2011).

250. Bondar, T. & Medzhitov, R. P53-mediated hematopoietic stem and progenitor cell competition. *Cell Stem Cell* **6**, 309–322 (2010).

251. Dudgeon, C. *et al.* Genetic variants and mutations of PPM1D control the response to DNA damage. *Cell Cycle* **12**, 2656–2664 (2013).

252. Silver, A. J., Bick, A. G. & Savona, M. R. Germline risk of clonal haematopoiesis. *Nat. Rev. Genet.* 1–15 (2021).

253. Kilpivaara, O. *et al.* A germline JAK2 SNP is associated with predisposition to the development of JAK2(V617F)-positive myeloproliferative neoplasms. *Nat. Genet.* **41**, 455–459 (2009).

254. Trifa, A. P. *et al.* The G allele of the JAK2 rs10974944 SNP, part of JAK2 46/1 haplotype, is strongly associated with JAK2 V617F-positive myeloproliferative neoplasms. *Ann. Hematol.* **89**, 979–983 (2010).

255. Jones, A. V. *et al.* JAK2 haplotype is a major risk factor for the development of myeloproliferative neoplasms. *Nat. Genet.* **41**, 446–449 (2009).

256. Olcaydu, D. *et al.* A common JAK2 haplotype confers susceptibility to myeloproliferative neoplasms. *Nat. Genet.* **41**, 450–454 (2009).

257. McKerrell, T. *et al.* JAK2 V617F hematopoietic clones are present several years prior to MPN diagnosis and follow different expansion kinetics. *Blood Adv.* **1**, 968–971 (2017).

258. Bick, A. G. *et al.* Inherited causes of clonal haematopoiesis in 97,691 whole genomes. *Nature* **586**, 763–768 (2020).

259. Thompson, D. J. *et al.* Genetic predisposition to mosaic Y chromosome loss in blood. *Nature* **575**, 652–657 (2019).

260. Wright, D. J. *et al.* Genetic variants associated with mosaic Y chromosome loss highlight cell cycle genes and overlap with cancer susceptibility. *Nat. Genet.* **49**, 674–679 (2017).

261. Terao, C. *et al.* GWAS of mosaic loss of chromosome Y highlights genetic effects on blood cell differentiation. *Nat. Commun.* **10**, 4719 (2019).

262. Zhou, W. *et al.* Mosaic loss of chromosome Y is associated with common variation near TCL1A. *Nat. Genet.* **48**, 563–568 (2016).

263. Potus, F. *et al.* Novel mutations and decreased expression of the epigenetic regulator TET2 in pulmonary arterial hypertension. *Circulation* **141**, 1986–2000 (2020).

264. Newton-Cheh, C. *et al.* Genome-wide association study identifies eight loci associated with blood pressure. *Nat. Genet.* **41**, 666–676 (2009).

265. Safarova, M. S. *et al.* Targeted sequencing study to uncover shared genetic susceptibility between peripheral artery disease and coronary heart disease-brief report. *Arterioscler. Thromb. Vasc. Biol.* **39**, 1227–1233 (2019).

266. Dichgans, M. *et al.* Shared genetic susceptibility to ischemic stroke and coronary artery disease: a genome-wide analysis of common variants. *Stroke* **45**, 24–36 (2014).

267. Zekavat, S. M. *et al.* Hematopoietic mosaic chromosomal alterations increase the risk for diverse types of infection. *Nat. Med.* (2021) doi:10.1038/s41591-021-01371-0.

268. Saiki, R. *et al.* Combined landscape of single-nucleotide variants and copy number alterations in clonal hematopoiesis. *Nat. Med.* **27**, 1239–1249 (2021).

269. Zekavat, S. M. *et al.* TP53-mediated clonal hematopoiesis confers increased risk for incident atherosclerotic disease. *Nature Cardiovascular Research* **2**, 144–158 (2023).

270. Wolach, O. *et al.* Increased neutrophil extracellular trap formation promotes thrombosis in myeloproliferative neoplasms. *Sci. Transl. Med.* **10**, eaan8292 (2018).

271. Fuster, J. J. *et al.* Clonal hematopoiesis associated with TET2 deficiency accelerates atherosclerosis development in mice. *Science* **355**, 842–847 (2017).

272. Zhang, Q. *et al.* Tet2 is required to resolve inflammation by recruiting Hdac2 to specifically repress IL-6. *Nature* **525**, 389–393 (2015).

273. Bick, A. G. *et al.* Genetic IL-6 Signaling Deficiency Attenuates Cardiovascular Risk in Clonal Hematopoiesis. *Circulation* (2019) doi:10.1161/CIRCULATIONAHA.119.044362.

274. Wainstein, M. V. *et al.* Elevated serum interleukin-6 is predictive of coronary artery disease in intermediate risk overweight patients referred for coronary angiography. *Diabetol. Metab. Syndr.* **9**, 67 (2017).

275. Fidler, T. P. *et al.* The AIM2 inflammasome exacerbates atherosclerosis in clonal haematopoiesis. *Nature* **592**, 296–301 (2021).

276. Wang, W. *et al.* Macrophage inflammation, erythrophagocytosis, and accelerated atherosclerosis in *Jak2* ^{V617F} mice. *Circ. Res.* **123**, (2018).

277. Svensson, E. C. *et al.* TET2-Driven Clonal Hematopoiesis and Response to Canakinumab: An Exploratory Analysis of the CANTOS Randomized Clinical Trial. *JAMA Cardiol* **7**, 521–528 (2022).

278. Taylor, F. *et al.* Statins for the primary prevention of cardiovascular disease. *Cochrane Database Syst Rev.* **31**, CD004816 (2013).

279. Silverman, M. G. *et al.* Association Between Lowering LDL-C and Cardiovascular Risk Reduction Among Different Therapeutic Interventions A Systematic Review and Meta-analysis. *JAMA* **316**, 1289–1297 (2016).

280. Cholesterol Treatment Trialists (CTT) Collaborators. Lack of effect of lowering LDL cholesterol on cancer: meta-analysis of individual data from 175,000 people in 27 randomised trials of statin therapy. *PLoS ONE* **7**, e29849 (2012).

281. Wenger, N. K. *et al.* Beneficial effects of aggressive low-density liprotein cholesterol lowering in women with stable coronary heart disease in the Treating to New Targets (TNT) study. *BMJ* **94**, 434–439 (2007).

282. Ford, I. *et al.* Long-Term Follow-up of the West of Scotland Coronary Prevention Study. *New England Journal of Medicine* **357**, 1477–1486 (2007).

283. Blais, L., Desgagné, A. & LeLorier, J. 3-hydroxy-3-methylglutaryl coenzyme A reductase inhibitors and the risk of cancer. *JAMA intern med* **160**, 2363–2368 (2000).

284. Kuoppala, J., Lamminpaa, A. & Pukkala, E. Statins and cancer: a systematic review and meta-analysis. *EJC* **44**, 2122–2132 (2008).

285. Tran, K. T. *et al.* Statin use and risk of liver cancer: evidence from two population-based studies. *International Journal of Cancer* **0**, (2019).

286. LaRosa, J. C. *et al.* Intensive lipid lowering with atorvastatin in patients with stable coronary disease. *N. Engl. J. Med.* **352**, 1425–1435 (2005).

287. Pedersen, T. R. *et al.* High-dose atorvastatin vs usual-dose simvastatin for secondary prevention after myocardial infarction: the IDEAL study: a randomized controlled trial. *JAMA* **294**, 2437–2445 (2005).

288. Pedersen, T. R. *et al.* Design and baseline characteristics of the Incremental Decrease in End Points through Aggressive Lipid Lowering study. *The American Journal of Cardiology* **94**, 720–724 (2004).

289. Waters, D. D. *et al.* Treating to New Targets (TNT) Study: does lowering low-density lipoprotein cholesterol levels below currently recommended guidelines yield incremental clinical benefit? *Am. J. Cardiol.* **93**, 154–158 (2004).

290. Wu, Y. *et al.* Genome-wide association study of medication-use and associated disease in the UK Biobank. *Nature Comm* **10**, (2019).

291. Lemieux-Perreault, L. P., Legault, M. A., Asselin, G. & Dubé, M. P. Genipe: an automated genome-wide imputation pipeline with automatic reporting and statistical tools. *Bioinformatics* **32**, 3661–3663 (2016).

292. Mägi, R. & Morris, A. P. GWAMA: software for genome-wide association meta-analysis. *BMC Bioinformatics* **11**, 288 (2010).

293. Higgins, J. P., Thompson, S. G., Deeks, J. J. & Altman, D. G. Measuring inconsistency in meta-analyses. *BMJ* **327**, 557–560 (2003).

294. Higgins, J. P. & Thompson, S. G. Quantifying heterogeneity in a meta-analysis. *StatMed* **21**, 1539–1558 (2002).

295. Cochran, W. G. The combination of estimates from different experiements. *Biometrics* **10**, 101–129 (1954).

296. Nelson, C. P. *et al.* Association analyses based on false discovery rate implicate new loci for coronary artery disease.

297. Meijers, W. C. & Boer, R. A. Common risk factors for heart failure and cancer. *Cardiovasc Res* **115**, 844–853 (2019).

298. Shiina, T., Hosomichi, K., Inoko, H. & Kulski, J. K. The HLA genomic loci map: expression, interaction, diversity and disease. *Journal of Human Genetics* **54**, 15–39 (2009).

299. Staley, J. R. *et al.* PhenoScanner: a database of human genotype–phenotype associations. *Bioinformatics* **32**, 3207–3209 (2016).

300. Kamat, M. A. *et al.* PhenoScanner V2: an expanded tool for searching human genotype–phenotype associations. *Bioinformatics* **35**, 4851–4853 (2019).

301. Couture, A. *et al.* HLA-Class II Artificial Antigen Presenting Cells in CD4+ T Cell-Based Immunotherapy. *Front. Immunol.* **10**, 1081 (2019).

302. Chemin, K., Gerstner, C. & Malmström, V. Effector Functions of CD4+ T Cells at the Site of Local Autoimmune Inflammation-Lessons From Rheumatoid Arthritis. *Front. Immunol.* **10**, 353 (2019).

303. Smitten, A. L., Simon, T. A., Hochberg, M. C. & Suissa, S. A meta-analysis of the incidence of malignancy in adult patients with rheumatoid arthritis. *Arthritis Res. Ther.* **10**, R45 (2008).

304. Capone, I., Marchetti, P., Ascierto, P. A., Malorni, W. & Gabriele, L. Sexual Dimorphism of Immune Responses: A New Perspective in Cancer Immunotherapy. *Front. Immunol.* **9**, 552 (2018).

305. van Vollenhoven, R. F. Sex differences in rheumatoid arthritis: more than meets the eye. *BMC Med.* **7**, 12 (2009).

306. Nair, B., Taylor-Gjevre, R., Wu, L., Jin, S. & Quail, J. M. Incidence and prevalence of rheumatoid arthritis in Saskatchewan, Canada: 2001-2014. *BMC Rheumatol* **3**, 28 (2019).

307. Frostegård, J., Zhang, Y., Sun, J., Yan, K. & Liu, A. Oxidized Low‐Density Lipoprotein (OxLDL)‐Treated Dendritic Cells Promote Activation of T Cells in Human Atherosclerotic Plaque and Blood, Which Is Repressed by Statins: microRNA let‐7c Is Integral to the Effect. *J. Am. Heart Assoc.* **5**, e003976.

308. Nadiminty, N. *et al.* MicroRNA let-7c is downregulated in prostate cancer and suppresses prostate cancer growth. *PLoS One* **7**, e32832 (2012).

309. Mizuno, R., Kawada, K. & Sakai, Y. The Molecular Basis and Therapeutic Potential of Let-7 MicroRNAs against Colorectal Cancer. *Can J Gastroenterol Hepatol* **2018**, 5769591 (2018).

310. Biamonte, F. *et al.* MicroRNA let-7g acts as tumor suppressor and predictive biomarker for chemoresistance in human epithelial ovarian cancer. *Sci. Rep.* **9**, 5668 (2019).

311. Zhao, B. *et al.* MicroRNA let-7c inhibits migration and invasion of human non-small cell lung cancer by targeting ITGB3 and MAP4K3. *Cancer Lett.* **342**, 43–51 (2014).

312. Rhodes, D. A. & Trowsdale, J. Genetics and molecular genetics of the MHC. *Rev. Immunogenet.* **1**, 21–31 (1999).

313. Nangalia, J. & Campbell, P. J. Genome Sequencing during a Patient's Journey through Cancer. *New England Journal of Medicine* **381**, 2145–2156 (2019).

314. Trowsdale, J. Genomic structure and function in the MHC. *Trends Genet.* **9**, 117–122 (1993).

315. Horton, R. *et al.* Gene map of the extended human MHC. *Nat. Rev. Genet.* **5**, 889–899 (2004).

316. Howell, W. M. HLA and disease: guilt by association. *International Journal of Immunogenetics* **41**, 1–12 (2014).

317. Klareskog, L., Catrina, A. I. & Paget, S. Rheumatoid arthritis. *The Lancet* **373**, 659–672 (2009).

318. Busch, R. *et al.* On the perils of poor editing: regulation of peptide loading by HLA-DQ and H2-A molecules associated with celiac disease and type 1 diabetes. *Expert reviews in molecular medicine* **14**, e15–e15 (2012).

319. Illing, P. T., Vivian, J. P., Purcell, A. W., Rossjohn, J. & McCluskey, J. Human leukocyte antigen-associated drug hypersensitivity. *Current opinion in immunology* **25**, 81–89 (2013).

320. Kennedy, A. E., Ozbek, U. & Dorak, M. T. What has GWAS done for HLA and disease associations? *Int. J. Immunogenet.* **44**, 195–211 (2017).

321. Hansson, G. K. Inflammation, atherosclerosis, and coronary artery disease. *N Engl J Med* **352**, 1685–1695 (2005).

322. Schreiber, R. D., Old, L. J. & Smyth, M. J. Cancer immunoediting: integrating immunity's roles in cancer suppression and promotion. *Science* **331**, 1565–1570 (2011).

323. Wan, E. Y. F. *et al.* Blood pressure and risk of cardiovascular disease in UK Biobank: A Mendelian randomization study. *Hypertension* **77**, 367–375 (2021).

324. Coleman, M. P. *et al.* EUROCARE-3 summary: cancer survival in Europe at the end of the 20th century. *Ann Oncol* **14 Suppl 5**, v128-49 (2003).

325. Bluethmann, S. M., Mariotto, A. B. & Rowland, J. H. Anticipating the "Silver Tsunami": Prevalence Trajectories and Comorbidity Burden among Older Cancer Survivors in the United States. *Cancer Epidemiol. Biomarkers Prev.* **25**, 1029–1036 (2016).

326. Siegel, R. L., Miller, K. D. & Jemal, A. Cancer statistics, 2016. *CA: A Cancer Journal for Clinicians* **66**, 7–30 (2016).

327. Armenian, S. H. *et al.* Cardiovascular Disease Among Survivors of Adult-Onset Cancer: A Community-Based Retrospective Cohort Study. *J Clin Oncol* **34**, 1122–1130 (2016).

328. Haugnes, H. S. *et al.* Cardiovascular risk factors and morbidity in long-term survivors of testicular cancer: a 20-year follow-up study. *J Clin Oncol* **28**, 4649–4657 (2010).

329. Emery, J. *et al.* Management of common clinical problems experienced by survivors of cancer. *Lancet* **399**, 1537–1550 (2022).

330. Lamberg, M. *et al.* Next Generation Risk Markers in Preventive Cardio-oncology. *Curr. Atheroscler. Rep.* **24**, 443–456 (2022).

331. Khoury, J. D. *et al.* The 5th edition of the World Health Organization Classification of Haematolymphoid Tumours: Myeloid and Histiocytic/Dendritic Neoplasms. *Leukemia* **36**, 1703–1719 (2022).

332. Busque, L. *et al.* Recurrent somatic TET2 mutations in normal elderly individuals with clonal hematopoiesis. *Nat Genet* **44**, 1179–1181 (2012).

333. Genovese, G. *et al.* Clonal hematopoiesis and blood-cancer risk inferred from blood DNA sequence. *N. Engl. J. Med.* **371**, 2477–2487 (2014).

334. Xie, M. *et al.* Age-related mutations associated with clonal hematopoietic expansion and malignancies. *Nat. Med.* **20**, 1472–1478 (2014).

335. McKerrell, T. *et al.* Leukemia-associated somatic mutations drive distinct patterns of age-related clonal hemopoiesis. *Cell Rep* **10**, 1239–1245 (2015).

336. Steensma, D. P. *et al.* Clonal hematopoiesis of indeterminate potential and its distinction from myelodysplastic syndromes. *Blood* **126**, 9–16 (2015).

337. Laurie, C. C. *et al.* Detectable clonal mosaicism from birth to old age and its relationship to cancer. *Nat Genet* **44**, 642–650 (2012).

338. Jacobs, K. B. *et al.* Detectable clonal mosaicism and its relationship to aging and cancer. *Nat. Genet.* **44**, 651–658 (2012).

339. Forsberg, L. A. *et al.* Age-related somatic structural changes in the nuclear genome of human blood cells. *Am J Hum Genet* **90**, 217–228 (2012).

340. Bick, A. G. *et al.* Genetic Interleukin 6 Signaling Deficiency Attenuates Cardiovascular Risk in Clonal Hematopoiesis. *Circulation* **141**, 124–131 (2020).

341. Honigberg, M. C. *et al.* Premature menopause, clonal hematopoiesis, and coronary artery disease in postmenopausal women. *Circulation* **143**, 410–423 (2021).

342. Busque, L. *et al.* High-sensitivity C-reactive protein is associated with clonal hematopoiesis of indeterminate potential. *Blood Adv* **4**, 2430–2438 (2020).

343. Coombs, C. C. *et al.* Therapy-Related Clonal Hematopoiesis in Patients with Non-hematologic Cancers Is Common and Associated with Adverse Clinical Outcomes. *Cell Stem Cell* **21**, 374-382.e4 (2017).

344. Lin, S. H. *et al.* Incident disease associations with mosaic chromosomal alterations on autosomes, X and Y chromosomes: insights from a phenome-wide association study in the UK Biobank. *Cell Biosci* **11**, 143 (2021).

345. Forsberg, L. A. *et al.* Mosaic loss of chromosome Y in leukocytes matters. *Nat Genet* **51**, 4–7 (2019).

346. Loftfield, E. *et al.* Predictors of mosaic chromosome Y loss and associations with mortality in the UK Biobank. *Sci. Rep.* **8**, 12316 (2018).

347. Forsberg, L. A. *et al.* Mosaic loss of chromosome Y in peripheral blood is associated with shorter survival and higher risk of cancer. *Nat. Genet.* **46**, 624–628 (2014).

348. Dumanski, J. P. *et al.* Mosaic Loss of Chromosome Y in Blood Is Associated with Alzheimer Disease. *Am J Hum Genet* **98**, 1208–1219 (2016).

349. Persani, L. *et al.* Increased loss of the Y chromosome in peripheral blood cells in male patients with autoimmune thyroiditis. *J Autoimmun* **38**, J193-6 (2012).

350. Haitjema, S. *et al.* Loss of Y Chromosome in Blood Is Associated With Major Cardiovascular Events During Follow-Up in Men After Carotid Endarterectomy. *Circ Cardiovasc Genet* **10**, e001544 (2017).

351. Sudlow, C. *et al.* UK biobank: an open access resource for identifying the causes of a wide range of complex diseases of middle and old age. *PLoS Med* **12**, e1001779 (2015).

352. Bycroft, C. *et al.* The UK Biobank resource with deep phenotyping and genomic data. *Nature* **562**, 203–209 (2018).

353. Loh, P.-R. *et al.* Reference-based phasing using the Haplotype Reference Consortium panel. *Nat. Genet.* **48**, 1443–1448 (2016).

354. Wan, E. Y. F. *et al.* Blood Pressure and Risk of Cardiovascular Disease in UK Biobank. *Hypertension* **77**, 367–375 (2021).

355. Nolan, E. K. & Chen, H. Y. A comparison of the Cox model to the Fine-Gray model for survival analyses of re-fracture rates. *Arch Osteoporos* **15**, 86 (2020).

356. Sjoberg DD, Whiting K, Curry M, Lavery JA & J, Larmarange. Reproducible summary tables with the gtsummary package. *The R Journal* **13**, 570–580 (2021).

357. Wang, S. *et al.* Prevalence and prognostic significance of DNMT3A- and TET2- clonal haematopoiesis-driver mutations in patients presenting with ST-segment elevation myocardial infarction. *EBioMedicine* **78**, 103964 (2022).

358. Zekavat, S. M. *et al.* TP53-mediated clonal hematopoiesis confers increased risk for incident peripheral artery disease. *bioRxiv* (2021) doi:10.1101/2021.08.22.21262430.

359. Bolton, K. L. *et al.* Cancer therapy shapes the fitness landscape of clonal hematopoiesis. *Nature Genetics* **52**, 1219–1226 (2020).

360. Levin, M. G. *et al.* Genetics of smoking and risk of clonal hematopoiesis. *Sci. Rep.* **12**, 7248 (2022).

361. Machiela, M. J. *et al.* Female chromosome X mosaicism is age-related and preferentially affects the inactivated X chromosome. *Nat. Commun.* **7**, 11843 (2016).

362. Sano, S. *et al.* Hematopoietic loss of Y chromosome leads to cardiac fibrosis and heart failure mortality. *Science* **377**, 292–297 (2022).

363. Machiela, M. J. & Chanock, S. J. Detectable clonal mosaicism in the human genome. *Semin. Hematol.* **50**, 348–359 (2013).

364. MoChA. MOsaic CHromosomal Alterations (MoChA) caller. https://github.com/freeseek/mocha (2022).

365. Kessler, M. D. *et al.* Common and rare variant associations with clonal haematopoiesis phenotypes. *Nature* **612**, 301–309 (2022).

366. Emery, J. *et al.* Management of common clinical problems experienced by survivors of cancer. *Lancet* **399**, 1537–1550 (2022).

367. Florido Roberta *et al.* Cardiovascular Disease Risk Among Cancer Survivors. *J. Am. Coll. Cardiol.* **80**, 22–32 (2022).

368. Stoltzfus, K. C. *et al.* Fatal heart disease among cancer patients. *Nat. Commun.* **11**, 2011 (2020).

369. Koene, R. J., Prizment, A. E., Blaes, A. & Konety, S. H. Shared Risk Factors in Cardiovascular Disease and Cancer. *Circulation* **133**, 1104–1114 (2016).

370. Pascual-Figal, D. A. *et al.* Clonal hematopoiesis and risk of progression of heart failure with reduced left ventricular ejection fraction. *J. Am. Coll. Cardiol.* **77**, 1747–1759 (2021).

371. Marnell, C. S., Bick, A. & Natarajan, P. Clonal hematopoiesis of indeterminate potential (CHIP): Linking somatic mutations, hematopoiesis, chronic inflammation and cardiovascular disease. *J. Mol. Cell. Cardiol.* **161**, 98–105 (2021).

372. Sun, M. *et al.* Somatic mosaic chromosomal alterations and death of cardiovascular disease causes among cancer survivors: an analysis of the UK Biobank. *bioRxiv* (2022) doi:10.1101/2022.08.20.22279019.

373. Szustakowski, J. D. *et al.* Advancing human genetics research and drug discovery through exome sequencing of the UK Biobank. *Nat. Genet.* **53**, 942–948 (2021).

374. Fine, J. & Gray, R. A Proportional Hazards Model for the Subdistribution of a Competing Risk. *J. Am. Stat. Assoc.* **94**, 496–509 (1999).

375. Using DXJupyterLab. *DNAnexus Documentation* https://documentation.dnanexus.com/user/jupyter-notebooks.

376. Sjoberg, D., Whiting, K., Curry, M., Lavery, J. & Larmarange, J. Reproducible Summary Tables with the gtsummary Package. *R J.* **13**, 570 (2021).

377. Sjoberg, DD Baillie, M Haesendonckx, S Treis, T. ggsurvfit: Flexible time-to-event figures. *ggsurvfit* https://github.com/ddsjoberg/ggsurvfit, http://www.danieldsjoberg.com/ggsurvfit/. (2023).

378. Rauch, P. J. *et al.* Loss-of-function mutations in Dnmt3a and Tet2 lead to accelerated atherosclerosis and convergent macrophage phenotypes in mice. *Blood* **132**, 745–745 (2018).

379. Gao, T. *et al.* Interplay between chromosomal alterations and gene mutations shapes the evolutionary trajectory of clonal hematopoiesis. *Nat. Commun.* **12**, 338 (2021).

380. Fey, M. F. *et al.* Clonality and X-inactivation patterns in hematopoietic cell populations detected by the highly informative M27 beta DNA probe. *Blood* **83**, 931–938 (1994).

381. Busque, L. *et al.* Nonrandom X-inactivation patterns in normal females: lyonization ratios vary with age. *Blood* **88**, 59–65 (1996).

382. Genovese, G. *et al.* Clonal hematopoiesis and blood-cancer risk inferred from blood DNA sequence. *N Engl J Med* **371**, 2477–2487 (2014).

383. Buscarlet, M. *et al.* DNMT3A and TET2 dominate clonal hematopoiesis and demonstrate benign phenotypes and different genetic predispositions. *Blood* **130**, 753–762 (2017).

384. Hansen, J. W. *et al.* Clonal hematopoiesis in elderly twins: concordance, discordance and mortality. *Blood* (2019) doi:10.1182/blood.2019001793.

385. Fabre, M. A. *et al.* Concordance for clonal hematopoiesis is limited in elderly twins. *Blood* (2019) doi:10.1182/blood.2019001807.

386. Franceschi, C., Garagnani, P., Vitale, G., Capri, M. & Salvioli, S. Inflammaging and "Garb-aging." *Trends Endocrinol Metab* **28**, 199–212 (2017).

387. Busque, L., Buscarlet, M., Mollica, L. & Levine, R. L. Concise Review: Age-Related Clonal Hematopoiesis: Stem Cells Tempting the Devil. *Stem Cells* **36**, 1287–1294 (2018).

388. Abegunde, S. O., Buckstein, R., Wells, R. A. & Rauh, M. J. An inflammatory environment containing TNFalpha favors Tet2-mutant clonal hematopoiesis. *Exp Hematol* **59**, 60–65 (2018).

389. Meisel, M. *et al.* Microbial signals drive pre-leukaemic myeloproliferation in a Tet2-deficient host. *Nature* **557**, 580–584 (2018).

390. Cook, E. K. *et al.* Comorbid and inflammatory characteristics of genetic subtypes of clonal hematopoiesis. *Blood Adv* **3**, 2482–2486 (2019).

391. Ridker, P. M. From C-Reactive Protein to Interleukin-6 to Interleukin-1: Moving Upstream To Identify Novel Targets for Atheroprotection. *Circ Res* **118**, 145–156 (2016).

392. Montreal Heart Institute Hospital Biobank. https://www.icm-mhi.org/en/research/infrastructures-services/mhis-hospital-biobank.

393. Tardif, J. C. *et al.* Genotype-Dependent Effects of Dalcetrapib on Cholesterol Efflux and Inflammation: Concordance With Clinical Outcomes. *Circ Cardiovasc Genet* **9**, 340–348 (2016).

394. Young, A. L., Challen, G. A., Birmann, B. M. & Druley, T. E. Clonal haematopoiesis harbouring AML-associated mutations is ubiquitous in healthy adults. *Nat Commun* **7**, 12484 (2016).

395. Tillett, W. S. & Francis, T. Serological Reactions in Pneumonia with a Non-Protein Somatic Fraction of Pneumococcus. *J Exp Med* **52**, 561–571 (1930).

396. Oliveira, E. B., Gotschlich, C. & Liu, T. Y. Primary structure of human C-reactive protein. *J Biol Chem* **254**, 489–502 (1979).

397. Moshage, H. J. *et al.* The effect of interleukin-1, interleukin-6 and its interrelationship on the synthesis of serum amyloid A and C-reactive protein in primary cultures of adult human hepatocytes. *Biochem Biophys Res Commun* **155**, 112–117 (1988).

398. Yoshizaki, K. Pathogenic role of IL-6 combined with TNF-alpha or IL-1 in the induction of acute phase proteins SAA and CRP in chronic inflammatory diseases. *Adv Exp Med Biol* **691**, 141–150 (2011).

399. Pepys, M. B. & Berger, A. The renaissance of C reactive protein. *BMJ* **322**, 4–5 (2001).

400. Rifai, N. & Ridker, P. M. High-sensitivity C-reactive protein: a novel and promising marker of coronary heart disease. *Clin Chem* **47**, 403–411 (2001).

401. Ridker, P. M., Hennekens, C. H., Buring, J. E. & Rifai, N. C-Reactive Protein and Other Markers of Inflammation in the Prediction of Cardiovascular Disease in Women. *N. Engl. J. Med.* **342**, 836–843 (2000).

402. SanMiguel, J. M. *et al.* Cell-Extrinsic Stressors from the Aging Bone Marrow (BM) Microenvironment Promote Dnmt3a-Mutant Clonal Hematopoiesis. *Blood* **134**, 5 (2019).

403. Shi, X. *et al.* The inflammatory cytokine profile of myelodysplastic syndromes: A meta-analysis. *Medicine (Baltimore)* **98**, e15844 (2019).

404. Zhang, B. *et al.* Microenvironmental protection of CML stem and progenitor cells from tyrosine kinase inhibitors through N-cadherin and Wnt-beta-catenin signaling. *Blood* **121**, 1824–1838 (2013).

405. Zhang, B. *et al.* Inhibition of interleukin-1 signaling enhances elimination of tyrosine kinase inhibitor-treated CML stem cells. *Blood* **128**, 2671–2682 (2016).

406. Abelson, S. *et al.* Prediction of acute myeloid leukaemia risk in healthy individuals. *Nature* **559**, 400–404 (2018).

407. Buscarlet, M. *et al.* Lineage restriction analyses in CHIP indicate myeloid bias for TET2 and multipotent stem cell origin for DNMT3A. *Blood* **132**, 277–280 (2018).

408. Sano, S. *et al.* Tet2-Mediated Clonal Hematopoiesis Accelerates Heart Failure Through a Mechanism Involving the IL-1beta/NLRP3 Inflammasome. *J Am Coll Cardiol* **71**, 875–886 (2018).

409. Dorsheimer, L. *et al.* Hematopoietic alterations in chronic heart failure patients by somatic mutations leading to clonal hematopoiesis. *Haematologica* (2019) doi:10.3324/haematol.2019.224402.

410. Hennessy, T. *et al.* The Low Dose Colchicine after Myocardial Infarction (LoDoCo-MI) study: A pilot randomized placebo controlled trial of colchicine following acute myocardial infarction. *Am. Heart J.* **215**, 62–69 (2019).

411. Pearson, T. A. *et al.* Markers of inflammation and cardiovascular disease: application to clinical and public health practice: A statement for healthcare professionals from the Centers for

Disease Control and Prevention and the American Heart Association. *Circulation* **107**, 499–511 (2003).

412. STABILITY Investigators. Darapladib for Preventing Ischemic Events in Stable Coronary Heart Disease. *New England Journal of Medicine* **370**, 1702–1711 (2014).

413. Gelman, A. & Hill, J. *Data analysis using regression and multilevel/hierarchical models*. 648 (Cambridge University Press, 2006).

414. McCaw, Z. R., Lane, J. M., Saxena, R., Redline, S. & Lin, X. Operating characteristics of the rank-based inverse normal transformation for quantitative trait analysis in genome-wide association studies. *Biometrics* **76**, 1262–1272 (2020).

415. Burns, S. S. *et al.* Il-1r1 drives leukemogenesis induced by Tet2 loss. *Leukemia* **36**, 2531–2534 (2022).

416. SanMiguel, J. M. *et al.* Distinct tumor necrosis factor alpha receptors dictate stem cell fitness versus lineage output in Dnmt3a-mutant clonal hematopoiesis. *Cancer Discov.* **12**, 2763–2773 (2022).

417. Zhang, C. R. *et al.* Txnip enhances fitness of Dnmt3a-mutant hematopoietic stem cells via p21. *Blood cancer discovery* vol. 3 220–239 (2022).

418. Avagyan, S. *et al.* Resistance to inflammation underlies enhanced fitness in clonal hematopoiesis. *Science* **374**, 768–772 (2021).

419. Caiado, F. *et al.* Aging drives Tet2+/- clonal hematopoiesis via IL-1 signaling. *Blood* **141**, 886–903 (2023).

420. Shin, T.-H. *et al.* A macaque clonal hematopoiesis model demonstrates expansion of TET2-disrupted clones and utility for testing interventions. *Blood* **140**, 1774–1789 (2022).

421. Sun, M. *et al.* Low-dose colchicine and high-sensitivity C-reactive protein after myocardial infarction: A combined analysis using individual patient data from the COLCOT and LoDoCo-MI studies. *Int. J. Cardiol.* **363**, 20–22 (2022).

422. Hagiwara, K. *et al.* Dynamics of age- versus therapy-related clonal hematopoiesis in long-term survivors of pediatric cancer. *Cancer Discov.* (2023) doi:10.1158/2159-8290.CD-22-0956.

423. Weinstock, J. S. *et al.* Aberrant activation of TCL1A promotes stem cell expansion in clonal haematopoiesis. *Nature* (2023) doi:10.1038/s41586-023-05806-1.

Supplementary Files for Chapter 2

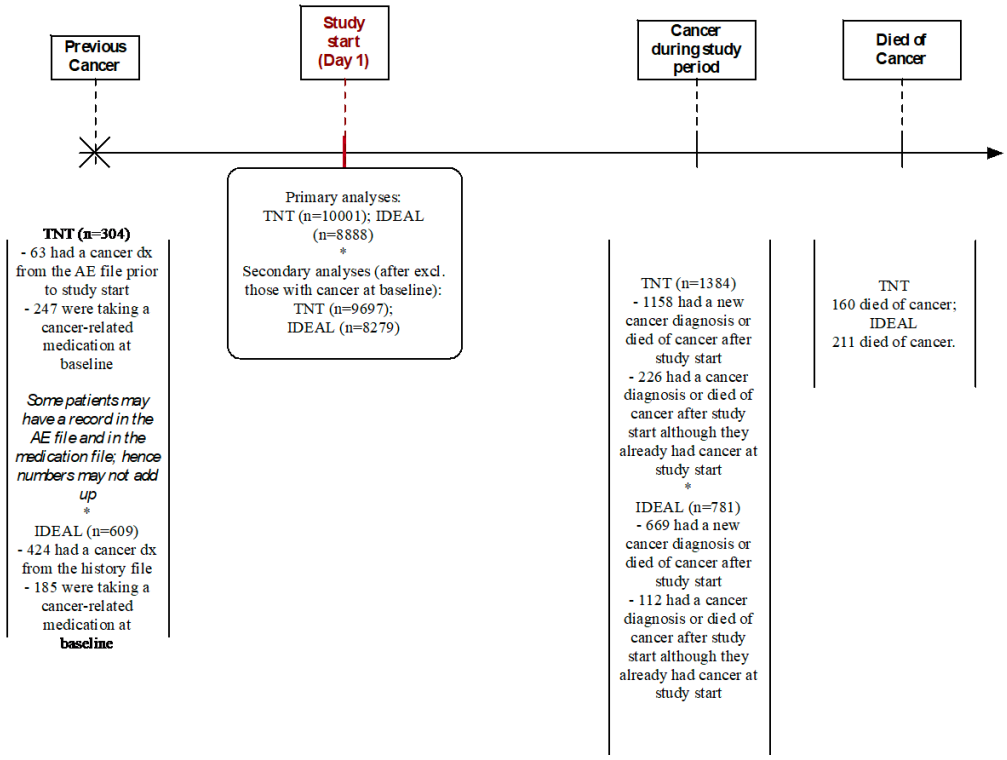
A.1 Supplementary Methods

A.1.1 Description of the TNT and IDEAL Clinical Trials

Both trials evaluated the efficacy and safety of high-dose statin therapy (atorvastatin at 80 mg) vs. lower-dose statin therapy (atorvastatin 10 mg in the TNT study and simvastatin 20 or 40 mg in IDEAL) for the prevention of cardiovascular events in patients who previously experienced coronary artery disease (CAD). In the TNT study, 10001 patients with a history of myocardial infarction, angina with objective evidence of atherosclerotic coronary arterial disease, or a history of coronary revascularization were enrolled. After a median follow-up of 4.9 years, the risk for the occurrence of a major cardiovascular event (i.e., fatal or nonfatal myocardial infarction, fatal or nonfatal stroke, fatal coronary heart disease, or resuscitation after cardiac arrest) was significantly lower in favor of high-dose atorvastatin compared to low-dose atorvastatin (hazard ratio [HR]: 0.78, 95% confidence interval [CI]: 0.69 to 0.89, P<0.001). In the IDEAL trial, 8888 patients with a history of myocardial infarction were recruited. After a median follow-up of 4.8 years, the risk for the occurrence of a major cardiovascular event was significantly lower in favour of high-dose atorvastatin compared to simvastatin (HR: 0.87, 95% CI: 0.78 to 0.98, *P*=0.02).

A.1.2 Definition of Cancer Status for All Patients Enrolled in the TNT and IDEAL Clinical Trials

Cancer diagnosis during the study period was determined from the adverse event records based on reports of cancer that occurred after randomization and from an adjudicated cause of death of cancer. Previous history of cancer was determined using the baseline medical history questionnaires and using baseline medication documentation.



Extended Data Fig. 1 Timeline Defining Cancer Status Per TNT and IDEAL.

Sixty-three patients in TNT and 424 in IDEAL were identified as having had malignant tumors from the baseline questionnaires, same as 247 and 185 additional patients who were using anti-cancer medication for a cancer indication at baseline in TNT and IDEAL, respectively. Overall, 1,384 (13.8%) patients in TNT and 781 (8.8%) patients in IDEAL had a cancer diagnosis during the study period. Of those, 226 in TNT and 112 in IDEAL also had cancer prior to study start. As such, cancer incidence (i.e., new onset of cancer) was recorded in 1158/10001 (11.6%) patients in TNT and 669/8888 (7.5%) in IDEAL.

In the TNT study, both a new onset of cancer and a history of cancer information was abstracted from the 'Adverse Effect' file. To distinguish between a new onset of cancer from those with a history of cancer, we relied on the date of the reported adverse effect. If the date of a cancer event was recorded before randomization, then the patient was assumed to have been diagnosed with cancer prior to study start. If the date of a cancer event was recorded after study start date, then the patient was assumed to have been diagnosed with cancer after the study start. In the

IDEAL study, a baseline diagnosis of cancer was captured from the patients' history or medication related to cancer used prior to study start. Cancer status following study start was determined using the adverse event file and its associated date.

Power Analysis:

We estimated that a genome-wide association study with 11,196 subjects (1078 with incident cancer and 10,118 without), at a two-sided 5x10-8 significance level using a logistic regression model with a likelihood ratio test, for an additive genetic model, that we had 80% power to detect a genetic variant with effect size of OR=1.35 and 1.53 for a genetic variant of allele frequency 0.30 and 0.10 respectively. In the subgroup of 9139 males, we had 80% power to detect a genetic variant with effect size of OR= 1.37 and 1.57 for a genetic variant of allele frequency 0.30 and 0.10 respectively, and in the subgroup of 2057 females of OR= 1.92 and OR=2.39 for a genetic variant of allele frequency 0.30 and 0.10 respectively.

For genetic variant rs13210472 which has a minor allele frequency of 0.033, we estimated 80% power to detect a SNP-by-sex interaction effect with ORGxE=2.09, assuming 20% females, no main SNP or sex effects, at a two-sided 0.05 significance level, using a logistic regression with a likelihood ratio test. Similarly, in our replication study of 4322 women with CAD, we estimated 80% power to detect a SNP-by-statin interaction effect with ORGxE=3.5 or greater, assuming 68% statin users, no main SNP or statin effects, at a two-sided 0.05 significance level, using a logistic regression with a likelihood ratio test.

A.1.3 Genotyping

Genotyping

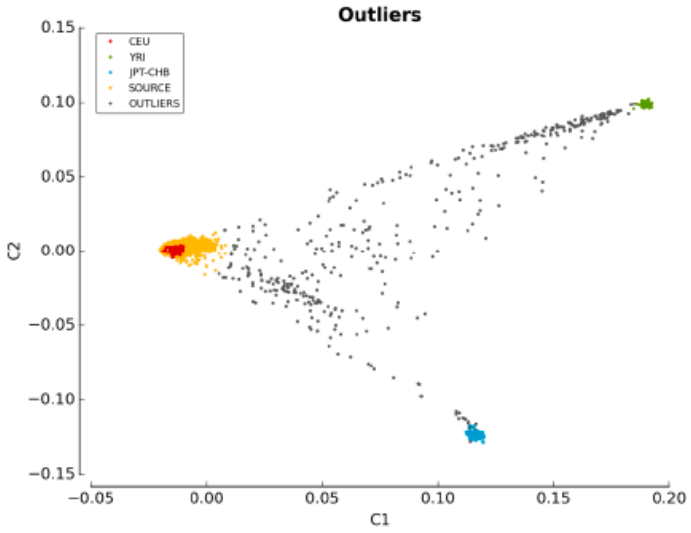
TNT: Genome-wide genotyping was performed using 200ng of genomic DNA in GLP-environment at the Beaulieu-Saucier Pharmacogenomics Centre (Montreal, Canada). The Illumina Infinium Multi-Ethnic Global Array (MEGA) Consortium v1 BeadChip (Illumina, San Diego, CA) including 1,705,969 genomic markers were used and processed according to the manufacturer's specifications. Each BeadChip was scanned and analyzed using the Illumina iScan Reader. Scanned images were analyzed using Illumina's Beeline v1.0.37.0 with the data manifest MEGA_Consortium_15063755_B2, without manual cluster adjustment and using the manufacturer's cluster file HapMap_MEGA_2015. Genotype data files were produced in three instalments of comparable size as data became available.

IDEAL: Genome-wide genotyping was performed using 200ng of genomic DNA in a GLP-environment at the Beaulieu-Saucier Pharmacogenomics Centre (Montreal, Canada). The Illumina Infinium Multi-Ethnic Global Array (MEGA) Consortium v2 BeadChip (Illumina, San Diego, CA) including 2,036,060 genomic markers was used and processed according to the manufacturer's specifications. BeadChips were scanned using the Illumina iScan Reader and analyzed using Illumina's Beeline v2 with the data manifest MEGA_Consortium_v2_15070954_A2.bpm, with minor manual cluster adjustment for ADME genes and using a custom cluster file.

Data cleanup for Genome-wide association study (GWAS)

TNT: The Beeline final report files were used to generate gender plots, LRR and BAF graphics. PyGenClean (1) version 1.7.1 and PLINK (2) version 1.07 were used for the quality checks (QC) and genetic data cleanup process. The genotyping experiment consisted of 67 plates of DNA samples. There was one control per hybridization run (corresponding to 2 plates), selected from NA19119, NA18980, NA19147 and NA12878. The pairwise concordance of Coriell samples ranged from 0.99989 to 0.999993. The comparison of Coriell genotypes to expectation from the 1000 Genomes data provided concordance ranging from 0.9940 to 0.9958.

Detailed genetic data cleanup results are presented below. Duplicated SNPs were evaluated for concordance, completion rate, allele call and MAF. SNPs with different allele calls or different MAF were retained. Identical and concordant SNPs were merged. Genotyping completion rate for samples and SNPs was set to 98%. SNPs with genotyping plate bias (based on the 96 well plates used to dilute DNA samples) were flagged but not removed as the effect of genetic ancestry could not be excluded. Pairwise identity-by-state (IBS) was used to conduct close familial relationship checks. We flagged and removed all but one pair-member of related pairs and sample duplicates (IBS2*ratio > 0.80) based on a selection of uncorrelated SNPs ($r2 < 0.1$). The pairwise IBS matrix was used as a distance metric to identify cryptic relatedness among samples and sample outliers by multi-dimensional scaling (MDS). The first two MDS components of each subject were plotted including the genotypes of HapMap CEU, JPT-CHB, and YRI data (keeping only founder individuals). Outliers from the main Caucasian cluster were flagged and removed by k-nearest neighbor (k=3) with a threshold of 1.9 σ in PyGenClean (v1.7.1):



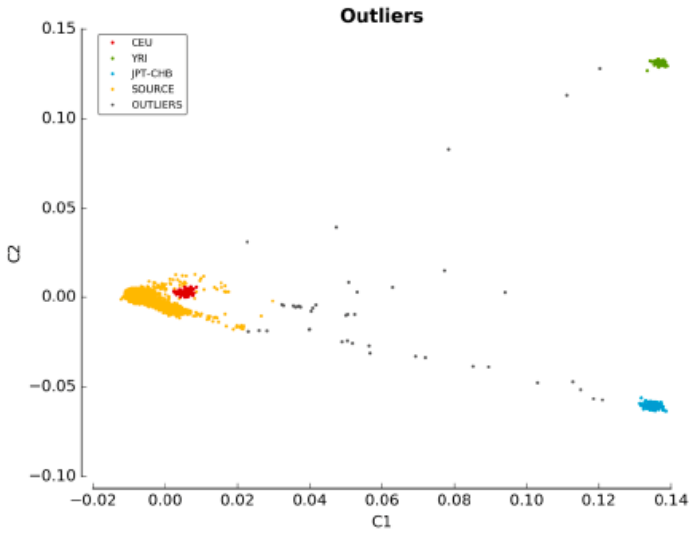
Extended Data Fig 2. For TNT: Multi-dimensional scaling (MDS) plots showing the first two principal components of the source dataset with the reference panels. The outliers of the Northern Europeans from Utah (CEU) population are shown in grey, while samples of the source dataset that resemble the CEU population is shown in orange. A multiplier of 1.9 was used to find the 411 outliers.

Principal components were next generated on the selected study samples only, and the scree plot and the cumulative explained variance were used to select the principal components to control for confounding by ancestry.(3)

IDEAL: The Beeline final report files were used to generate gender plots, LRR and BAF graphics. PyGenClean (1) version 1.8.3 and PLINK (2) version 1.07 were used for the quality checks (QC) and genetic data cleanup process. The genotyping experiment consisted of 69 plates of DNA samples. There was one control per hybridization run (corresponding to 2 plates), selected from NA19119, NA18980, NA19147 and NA12878. The pairwise concordance of Coriell samples ranged from 0.9997338 to 0.9999979. The comparison of Coriell genotypes to expectation from the 1000 Genomes data provided concordance ranging from 0.9961 to 0.9980.

Detailed genetic data cleanup results are presented below. Duplicated SNPs were evaluated for concordance, completion rate, allele call and MAF. SNPs with different allele calls or different MAF were retained. Identical and concordant SNPs were merged. The completion rate threshold for genotypes and samples was set to 98%. SNPs with genotyping plate bias (based on the 96 well

plates used to dilute DNA samples) were flagged but not removed as the effect of genetic ancestry could not be excluded. Pairwise IBS was used to conduct close familial relationship checks. We flagged and removed all but one member of related samples (IBS2*ratio > 0.80) based on a selection of uncorrelated SNPs (r2 < 0.1). The pairwise IBS matrix was used as a distance metric to identify cryptic relatedness among samples and sample outliers by MDS. The first two MDS components of each subject were plotted including the genotypes of HapMap CEU, JPT-CHB, and YRI data (keeping only founder individuals). Outliers from the main Caucasian cluster were flagged and removed by k-nearest neighbor (k=3) with a threshold of 1.9 σ in PyGenClean (v1.8.3):



Extended Data Fig 3. For IDEAL: Multidimensional scaling plots showing the first two principal components of the source dataset with the reference panels. The outliers of the CEU population are shown in grey, while samples of the source dataset that resemble the CEU population is shown in orange. A multiplier of 1.9 was used to find the 42 outliers.

Principal components were next generated on the selected study samples only, and the scree plot and the cumulative explained variance were used to select the principal components to control for confounding by ancestry.(3)

(a) Clean-up step	N	Procedure	
		SNP	IDs
Number of SNPs in genotyping file	1,705,969		
Number of samples in genotyping file	6,399		
SNPs without physical position, Insertion/deletion variants and tri allelic variants removed	39,960	-39,960	
Genotyping controls			-47
Duplicate samples excluded	181		-181
Failed markers	149,888	-149,888	
Replicated SNPs excluded	23,567	-23,567	
Samples with >10% missing genotypes	182		-182
SNPs with >2% missing genotypes	58,738	-58,738	
Samples with >2% missing genotypes	254		-254
SNPs with plate-bias $P < 4 \cdot 10^{-8}$	1,304	flagged	
SNPs used for IBS analysis	74,134		
SNPs used for MDS analysis	49,024		
Gender problem	56		-56
Discordant duplicated samples	42		-42
Twins with discordant ids	30		-30
Discordant age for parent-child relationship	3		-3
Related	20		-20
Internal QC controls	8		-8
Individuals from specific problematic sites	26		-26
Ethnicity other than Caucasian by MDS cluster	403		-403
Haploid genotypes (after gender issues processed)	2,451,835	Set to missing	
SNPs with low completion rate on X chromosome	460	-460	
SNPs with MAF=0	327,578	-327,578	
HWE test $4.5 \cdot 10^{-8} < P < 10^{-4}$	35,755	flagged	
HWE test $P < 4.5 \cdot 10^{-8}$ (0.05/ 1,106,238)	25,645	-25,645	
SNPs discordant between Megav1 and 1000G	1951	-1951	
Final number of SNPs for analysis	**1,078,182**		
Final number of samples for analysis	**5,147**		

(b) Clean-up step	N	Procedure	
		SNP	IDs
Number of SNPs in genotyping file	2,036,060		
Number of samples in genotyping file	6,526		
SNPs without physical position, Insertion/deletion variants and tri allelic variants removed	58,318	-58,318	
Genotyping controls	97		-97
Failed markers	279,977	-279,977	
Replicated SNPs excluded	20,778	-20,778	
Samples with >10% missing genotypes	21		-21
SNPs with >2% missing genotypes	16,924	-16,924	
Samples with >2% missing genotypes	47		-47
SNPs with plate-bias $P < 1 \cdot 10^{-7}$	116	flagged	
SNPs used for IBS analysis	67,822		

SNPs used for MDS analysis	65,381		
Gender problem	18		-18
Discordant duplicated samples	27		-27
Related	67		-67
Ethnicity other than Caucasian by MDS cluster	42		-42
Haploid genotypes (after gender issues processed)	1,362,960	Set to missing	
SNPs with low completion rate on X chromosome	426	-426	
SNPs with MAF=0	543,833	-543,833	
HWE test $4.4 \cdot 10^{-8} < P < 10^{-4}$	3,603	flagged	
HWE test $P < 4.4 \cdot 10^{-8}$ (0.05/ 1,123,529)	1,723	-1,723	
Final number of SNPs for analysis	**1,114,081**		
Final number of samples for analysis	**6,207**		

Extended Data Table 1a/b. Summary table of genetic data clean-up procedures performed prior to statistical analysis for TNT (a) and IDEAL (b).

Imputation of Chromosome X

TNT & IDEAL: Imputation was performed using IMPUTE2 (v2.3.2) (4,5,6) and phasing was performed using the SHAPEIT2 algorithm (v.2r790)(7). Strand alignment was solved by flipping non ambiguous SNPs and 133,946 ambiguous A/T and C/G SNPs were considered missing and were imputed. Imputation was performed based on 1,058,606 genetic variants for TNT and 1,049,714 genetic variants for IDEAL using the phased 1000 Genomes reference data with singleton sites removed released on June 16 2014 and which include samples from all populations (distributed through the IMPUTE2 website). Using an all-inclusive reference panel is known to improve imputation accuracy.(8) The pseudo-autosomal regions on the X chromosome were imputed separately from the rest of the chromosome. Internal cross-validation was performed with IMPUTE2 and provided a mean genotype concordance of 98.05%. Any missing genotypes at the genotyped SNPs were also imputed. For TNT, A total of 11,693,523 genetic variants with imputation probability of 0.90 or greater and completion rate of 98% or greater were obtained. For the genome-wide analysis, there were a total of 4,973,077 genetic variants with a minor allele frequency (MAF) greater than 5% and 6,579,943 with a MAF greater than 1%. For IDEAL, a total of 12,159,506 genetic variants with imputation probability of 0.90 or greater and completion rate of 98% or greater were obtained. For the genome-wide analysis, there were a total of 5,261,127 genetic variants with a MAF greater than 5% and 6,804,917 with a MAF greater than 1%.

References: 1.Lemieux Perreault, L.P., et al., pyGenClean: efficient tool for genetic data clean-up before association testing. Bioinformatics, 2013. 29(13): p. 1704-5. 2.Purcell, S., et al., PLINK: a tool set for whole-genome association and population-based linkage analyses. Am J Hum Genet, 2007. 81(3): p. 559-75. 3.Price, A.L., et al., Principal components analysis corrects for stratification in genome-wide association studies. Nat Genet, 2006. 38(8): p. 904-9. 4.Howie BN, Donnelly P, Marchini J: A flexible and accurate genotype imputation method for the next generation of genome-wide association studies. PLoS genetics 2009, 5(6):e1000529.[DOI:10.1371/journal.pgen.1000529]. 5.Howie B, Marchini J, Stephens M: Genotype imputation with thousands of genomes. G3: Genes, Genomes, Genetics 2011, 1(6):457–470. [DOI:10.1534/g3.111.001198]. 6.Howie B, Fuchsberger C, Stephens M, Marchini J, Abecasis GR: Fast and accurate genotype imputa-tion in genome-wide association studies through pre-phasing. Nature Genetics 2012, 44(8):955–959. [DOI:10.1038/ng.2354]. 7.Delaneau O, Zagury JF, Marchini J: Improved whole-chromosome phasing for disease and population genetic studies. Nature methods 2013, 10:5–6. [DOI:10.1038/nmeth.2307]. 8. Huang, L., Li, Y., Singleton, A. B., Hardy, J. A., Abecasis, G., Rosenberg, N. A., et al. (2009). Genotype-imputation accuracy across worldwide human populations. Am. J. Hum. Genet. 84, 235–250. doi: 10.1016/j.ajhg.2009.01.013

A.1.4 Validation of Cancer Diagnosis Per Hospitalization Records and Cancer Registry Files

For the purpose of understanding the concordance of cancer occurrence that was captured per hospitalization files versus cancer registry files, we performed a cross-validation of cancer diagnoses obtained from both data sources. The rationale for this exercise stemmed from differing censoring dates per data sources used. Specifically, data from hospitalization (inpatient) data provided by the NHS Digital, the Information and Statistics Division (ISD), and the Secure Anonymized Information Linkage (SAIL) have a censoring date of March 2017, October 2016, and February 2016, respectively. In contrast, cancer data originating from the NHS Digital and the National Records of Scotland have censoring dates of March 2016 and October 2015, respectively. The lag in time suggests a potential time frame where a patient's cancer diagnosis may not be yet captured per cancer registry files, despite having been hospitalized for a cancer-related issue. Furthermore, one may wonder whether a certain discrepancy occurs between the two data files, where cancer is observed in the cancer registry files, but not in hospitalizations, and vice versa. In consequence, we focused on assessing the extent to which cancer diagnoses do not overlap between the two data sources. Whereas a certain error rate is expected, we postulate that the majority of cancer occurrences are captured in both data sources (>80%).

Methods and data

The data files and their variable identification used to perform this cross-validation are listed below:

- **Hospitalization files**:
  - ICD-9 diagnoses (all) (variable ID#41203)
  - ICD-10 diagnoses (all) (variable ID#41202)
  - Date of first in-patient diagnosis (all ICD10) (variable ID#41262)
  - Date of first in-patient diagnosis (all ICD9) (variable ID#41263)

- **Cancer Registry files**:

o ICD-9 diagnoses (variable ID#40013)

o ICD-10 diagnoses (variable ID#40006)

o Cancer diagnosis dates (variable ID#40005)

Cancer diagnoses were abstracted per ICD-9 and ICD-19 diagnostic codes (types of cancer) along with corresponding dates. For our objective, we focused on the earliest cancer diagnosis date per patient's type of cancer. As such, if a patient had various dates associated with the same type of cancer, only the earliest date was retained. However, if a patient had multiple different cancer diagnoses, the earliest date associated with each type of cancer that was diagnosed was retained. This way, we retain all cancers that a patient ever had until last follow-up. Using the date of baseline as our study's index date (time 0), a cancer diagnosis that occurred before this time (time<0) was considered a prevalent case, whereas a cancer diagnosis that occurred after this time (time$\geq$0) was considered an incident case.

First, a crosstab of prevalent cancers across hospitalization files and cancer registry files was tabulated. Similarly, we repeated this step for incident cancers. Subsequently, we examined the overlap of cancer diagnosis dates obtained in the hospitalization files and cancer registry files by subtracting them. The expected result is that minimal deviation should be expected for the majority of patients following this computation. In addition, we also examined the concordance of prevalent cancers across hospitalization files, registry files, and patients' self-reported questionnaires in which they are asked if they were ever diagnosed with cancer prior to study entry.

Finally, it is possible that an initial malignant cancer diagnosis is immediately followed by a benign histological confirmation after biopsy results. This is relevant and needs to be accounted for, as the cancer could be captured in hospitalization files, but not in registry files. Without the following benign histology confirmation, one may erroneously assume that the patient had a cancer occurrence at the inpatient level. In this step, we abstracted all patients with a benign histology within both cancer registry and hospitalization files and merged it with our cancer data. We presumed that if a histological confirmation occurred within an arbitrary cutoff of $\leq$6 months following a first malignant diagnosis, then the diagnosis was no longer malignant.

Results:

Prevalent cancer

Our cross-validation of prevalent cancer across hospitalization and cancer registry files shows that 5.0% of patients with a cancer diagnosis within cancer registry files did not have a prevalent cancer diagnosis per hospitalization files (24997/498023) vs. 0.4% of patients for those with a

cancer diagnosis within hospitalization files but not within registry files (2134/475160). The remainder of cancer occurrences was concordant.

Registry files		**Hospitalization files**		
		No cancer	Cancer	**Total**
	No cancer	473026	2134	475160
	Cancer	24997	18291	43288
Total		498023	20425	518448

Extended Data Table 2. Cross-validation of prevalent cancer between hospitalization and cancer registry files within the UK Biobank

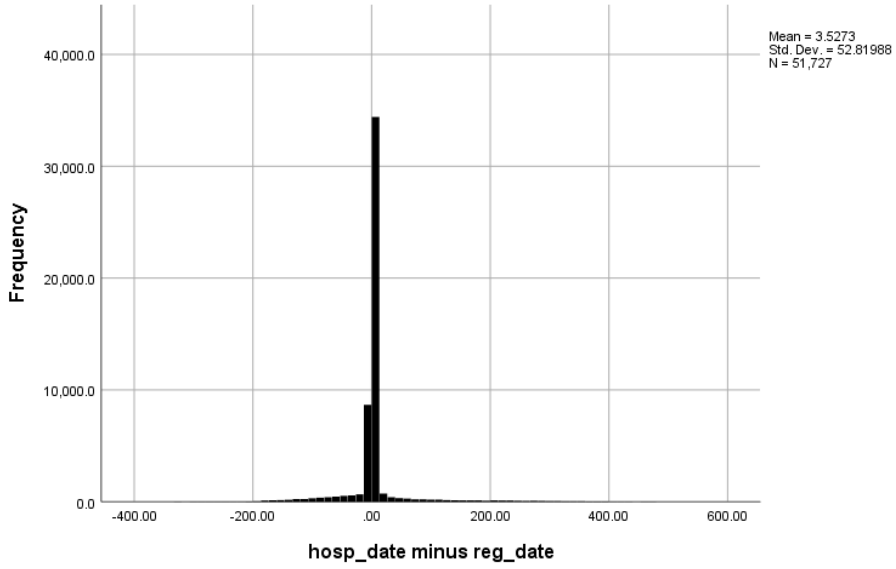
Incident cancer

Our cross-validation of incident cancer across hospitalization and cancer registry files showed that 2.2% of patients with a cancer diagnosis per cancer registry file did not have an incident cancer diagnosis within cancer registry (11164/492042) vs. 1.1% of patients with a cancer diagnosis within cancer registry file but not within hospitalization files. The remainder of cancer occurrences was concordant.

Registry files		**Hospitalization files**		
		No cancer	Cancer	**Total**
	No cancer	480878	5300	486178
	Cancer	11164	21106	32270
Total		492042	26406	518448

Extended Data Table 3. Cross-validation of incident cancer between hospitalization and cancer registry files within the UK Biobank.

Below is a histogram depicting the subtraction of hospitalization date (if available) minus cancer registry date. As illustrated, most cancer diagnosis dates from the two data sources overlap, suggesting even a strong concordance between dates obtained from the two different data files. It is noteworthy that for patients diagnosed with cancer beyond 2015-2016, they are naturally uncaptured within cancer registry files.



Extended Data Fig. 4. Histogram depicting the difference between hospitalization event date and cancer registry event date for the same individual.

In the examination of self-reported cancers at baseline relative to prevalent cancers as recorded per hospitalization and cancer registry files, our analyses revealed that 84% of cancers that appeared in both data files were confirmed by patients at study entry (15287/18291). In contrast, 16% of cancers detected in both files were refuted by patients at study entry (2883/18291).

In this regard, it is noteworthy that many patients whose self-reported answer was 'No' when asked if they had ever been diagnosed with cancer previously at study entry had non-melanoma skin cancer, where the 5-year survival is 100% and rarely spreads to other parts of the body. As such, it may be reasonable to think that such patients did not consider themselves to having cancer.

We examined the number of patients with a first malignant diagnosis followed by a benign histological confirmation within 6 months. For prevalent cases, 33/60 cancers were deemed to be ultimately benign using cancer registry files and 725/2088 cancers using hospitalization files. For incident cases, 8/35 cancers were deemed to be ultimately benign using cancer registry files and 906/3914 using hospitalization files.

Finally, we examined the number of patients who had a benign histology to see if they eventually had a real malignant diagnosis later on. Below is a table which summarizes the breakdown of this analysis, where only patients with a first benign histological confirmation are included:

	# of patients with a subsequent malignant cancer/ # of patients with a benign histology as their first 'cancer' diagnosis.	
	Prevalent	Incident
Hospital files	21/725	1/33
Registry files	75/3008	1/27

Extended Data Table 4. Malignant confirmation in individuals with a benign histology per hospitalization and cancer registry files.

A.1.5 Definition of CAD Based on Diagnostic Codes Within the UK Biobank

Cardiovascular endpoints were dichotomized as prevalent (before the date of baseline; $\leqslant$time 0) or incident (after the date of baseline; >time 0). Coronary artery disease (CAD) was determined using various cardiovascular composite endpoints. Both primary and secondary diagnostic codes (ICD-9, ICD-10) and procedure codes (OPCS) were used to abstract cardiovascular-related information per patient. Patients were considered to have prevalent CAD if the following codes occurred prior to the date of baseline: OPCS codes K40, K41, K42, K43, K44, K45, K46, K49, K50, and K75 for percutaneous coronary intervention (PCI) or coronary artery bypass grafting (CABG), ICD-9 codes 410-414 (except 414.1 for aneurysm), or ICD-10 codes I20-I25. Incident CAD was determined if the following codes occurred after the date of baseline: OPCS codes K40, K41, K42, K43, K44, K45, K46, K49, K50, and K75 for PCI or CABG, ICD-9 codes 410-414 (except 414.1 for aneurysm), or ICD-10 codes I20-I25, as well as those who died of CAD-related death. For patients with multiple CAD-related codes, only the first occurrence (and the corresponding date) was retained for the purpose of our analyses. Below is a description of what each code represents:

Code		**Description**
OPCS		
	K40	Saphenous vein graft replacement of coronary artery
	K41	Other autograft replacement of coronary artery
	K42	Allograft replacement of coronary artery
	K43	Prosthetic replacement of coronary artery

		K44	Other replacement of coronary artery
		K45	Connection of thoracic artery to coronary artery
		K46	Other bypass of coronary artery
		K49	Transluminal balloon angioplasty of coronary artery
		K50	Other therapeutic transluminal operations on coronary artery
		K75	Percutaneous transluminal balloon angioplasty and insertion of stent into coronary artery
ICD-9			
	410-414 (except 414.1)		Ischemic heart disease (except aneurysm)
ICD-10			
		I20	Angina pectoris
		I21	Acute myocardial infarction
		I22	Subsequent myocardial infarction
		I23	Certain current complications following acute myocardial infarction
		I24	Other acute ischemic heart diseases
		I25	Chronic ischemic heart disease

Extended Data Table 5. Overview of CAD-related diagnostic codes within the UK Biobank.

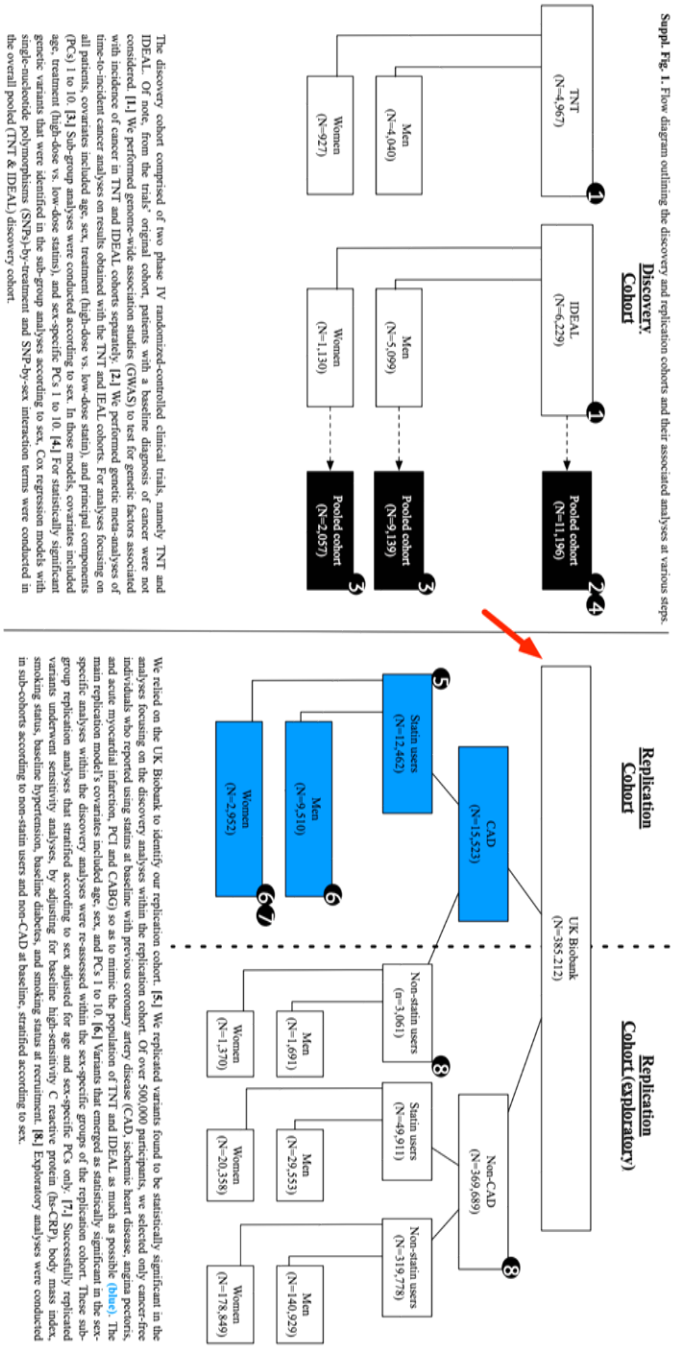
A.1.6 Functional Annotation

The functional annotation was performed using online resources data such as a PhenoScanner(1, 2) and GTEx V8 portals(3) and publicly available data such as chromatin state predicted by the Software ChrommHmm(4). We also used Regulome DB (beta version)(5) and Ensembl Variant Effect Predictor web interface (Ensembl GRCh37 release 99)(6) to assess the predicted effects of our variants. The genes pathways were studied using GeneMANIA website(7). These websites were visited between February and March 2020.

References: 1.Staley JR, Blackshaw J, Kamat MA, Ellis S, Surendran P, Sun BB, et al. PhenoScanner: a database of human genotype-phenotype associations. (1367-4811 (Electronic)). 2.Kamat MA, Blackshaw JA, Young R, Surendran P, Burgess S, Danesh J, et al. PhenoScanner V2: an expanded tool for searching human genotype-phenotype associations.(1367-4811 (Electronic)). 3.Carithers LJ, Ardlie K, Barcus M, Branton PA, Britton A, Buia SA, et al. A Novel Approach to High-Quality Postmortem Tissue Procurement: The GTEx Project. (1947-5543 (Electronic)). 4.Ernst J, Kellis M. ChromHMM: automating chromatin-state discovery and characterization. Nature Methods. 2012;9(3):215-6. 5.Boyle AP, Hong El Fau - Hariharan M, Hariharan M Fau - Cheng Y, Cheng Y Fau -Schaub MA, Schaub Ma Fau - Kasowski M, Kasowski M Fau - Karczewski KJ, et al. Annotation of functional variation in personal genomes using RegulomeDB. (1549-5469 (Electronic)). 6.McLaren W, Gil L, Hunt SE, Riat HS, Ritchie GRS, Thormann A, et al. The Ensembl Variant Effect Predictor. Genome Biology. 2016;17(1):122. 7.Warde-Farley D, Donaldson Sl Fau - Comes O, Comes O Fau - Zuberi K, Zuberi K Fau -Badrawi R, Badrawi R Fau - Chao P, Chao P Fau - Franz M, et al. The GeneMANIA prediction server: biological network integration for gene prioritization and predicting gene function. (1362-4962 (Electronic)).

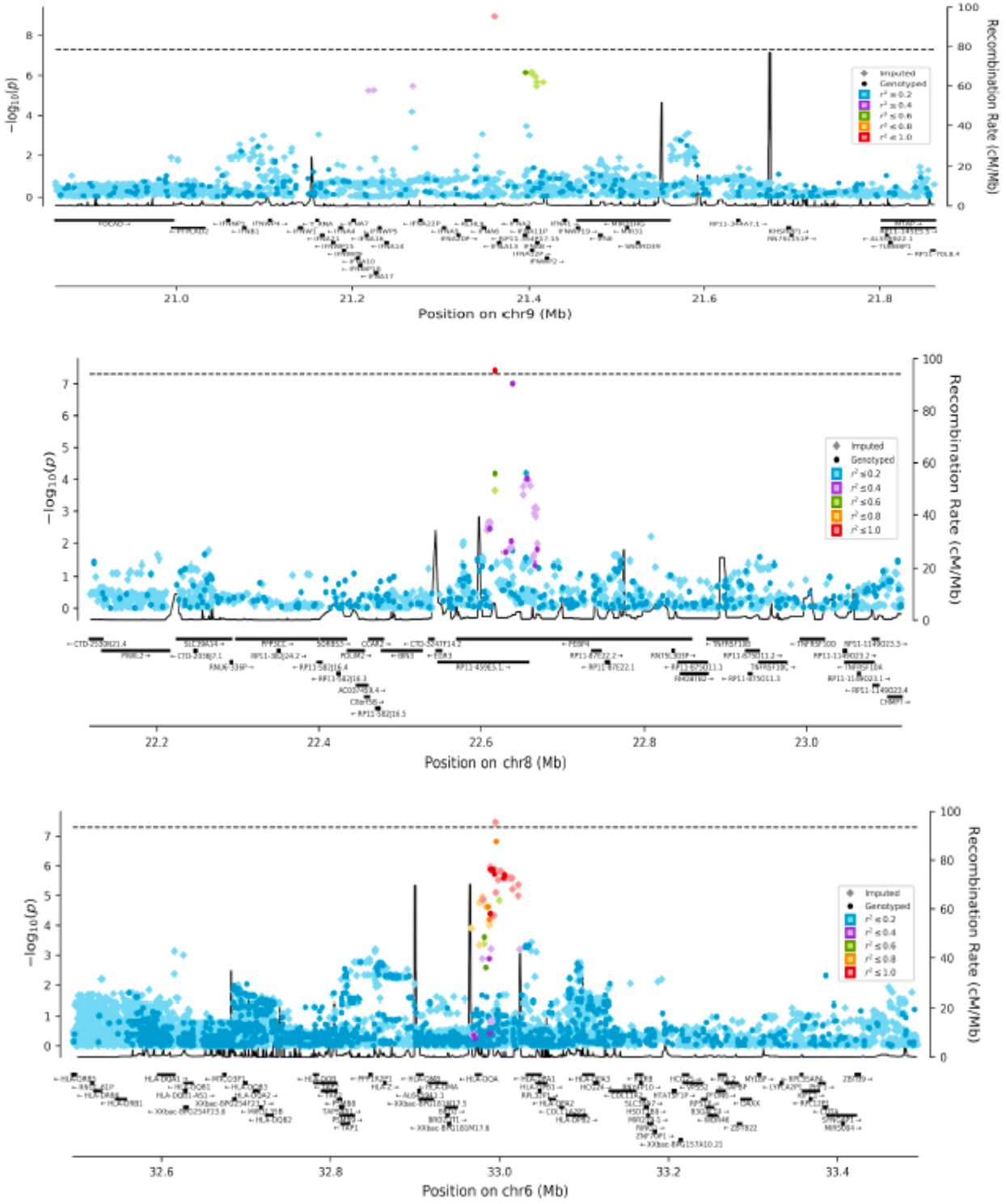
A.2 Supplementary Figures

A.2.1 Supplementary Figure A.2.1



Supplementary Figure A.2.1 Flow Diagram Outlining the Discovery and Replication Cohorts and Their Associated Analyses at Various Stages

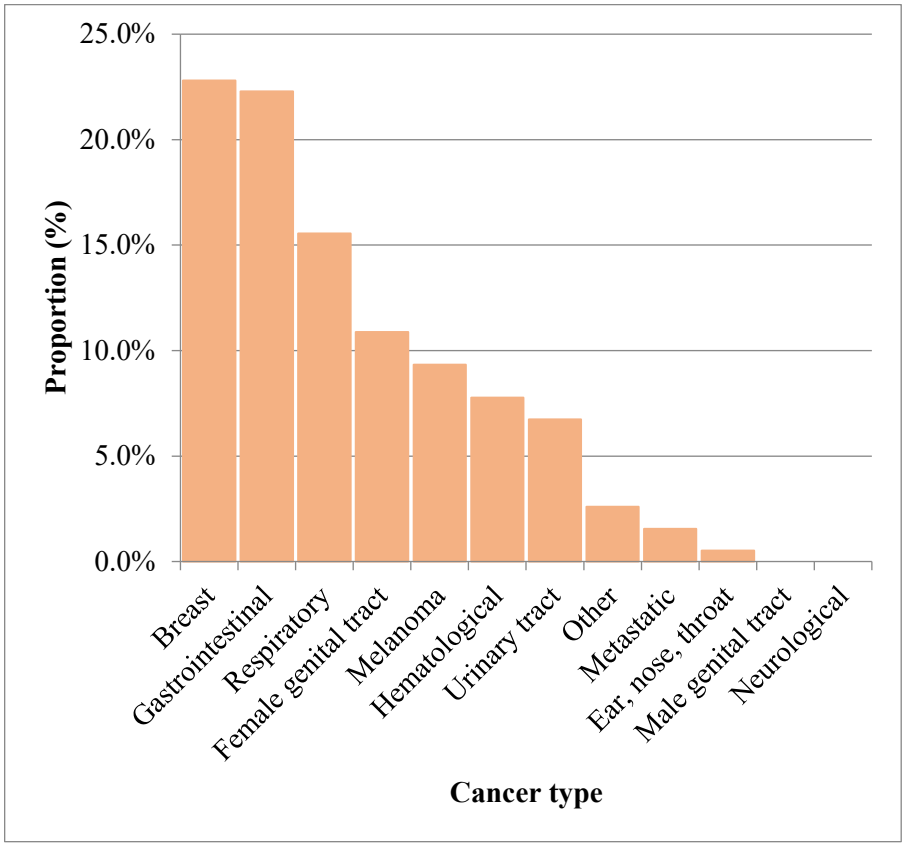
A.2.2 Supplementary Figure A.2.2



Supplementary Figure A.2.2 Plot of Genetic Variants From the Genome-Wide Meta-Analysis of the TNT and IDEAL Cohorts.

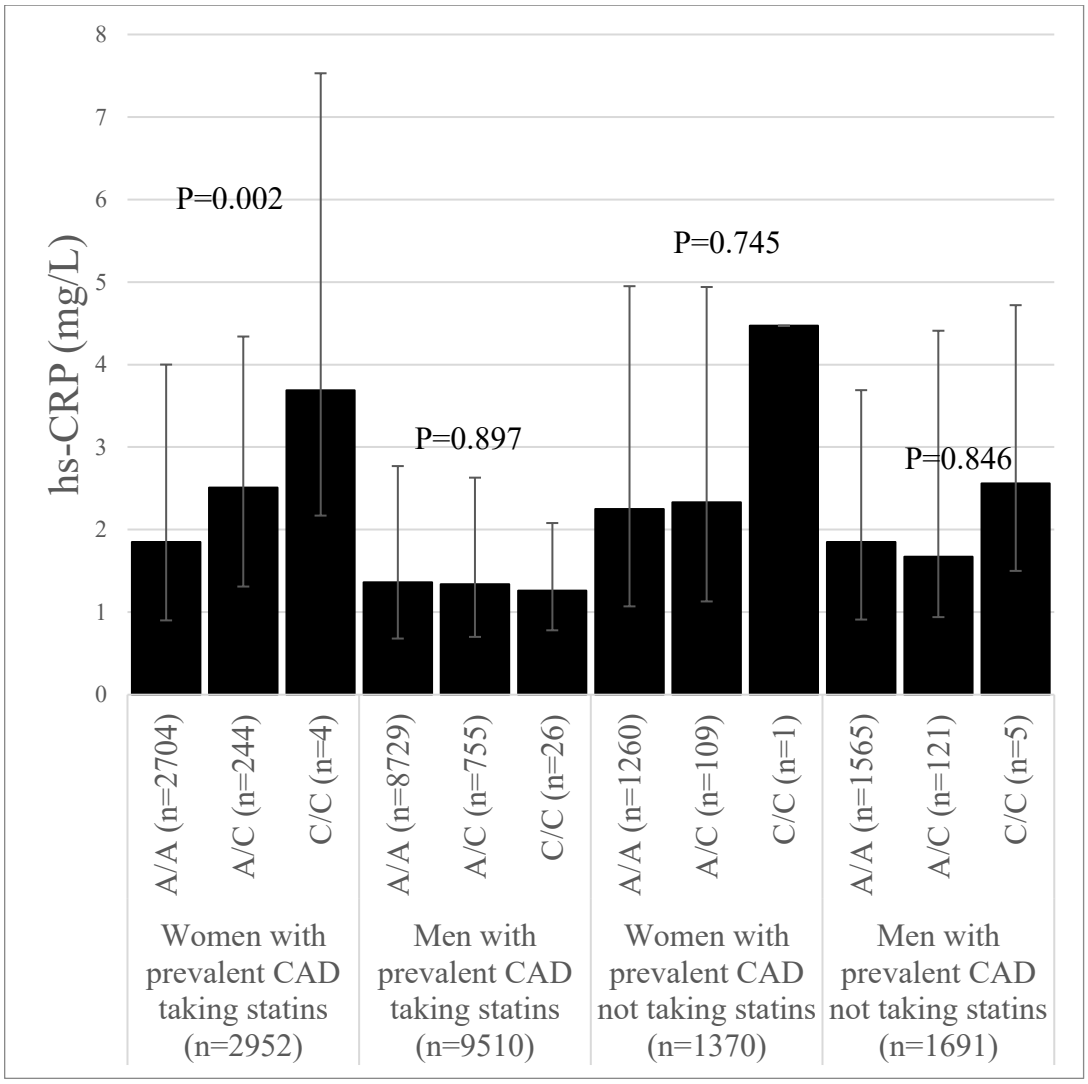
The first y-axis shows the negative log10 of P values for genotyped (circles and darker colors) and imputed (lozenges and paler colors) genetic variants (imputed in the IDEAL genetic dataset), the second y-axis shows the recombination rate from HapMap reference samples (black line). Genes are displayed below the SNPs, base-pair positions are given according to hg19, the degree of linkage disequilibrium (r2) of each genetic variant estimated from the study population is displayed as blue for [0, 0.2], purple for [0.2, 0.4], green for [0.4, 0.6], orange for [0.6, 0.8], and red for [0.8, 1.0]. All gene information comes from Ensembl (build37). ***A (top).*** *Chromosome 9 region (from 20,862,409 to 21,862,409 bp);* ***B (middle).*** *Chromosome 6 (from 32,495,081 to 33,495,081 bp);* ***C (bottom).*** *Chromosome 8 region (from 22,117,312 to 23,117,312 bp)*

A.2.3 Supplementary Figure A.2.3



Supplementary Figure A.2.3 Most Common Cancers Amongst Women Statin Users With Prevalent Coronary Artery Disease With an Incident Cancer (N=2,592)

A.2.4 Supplementary Figure A.2.4



Supplementary Figure A.2.4 Median and Interquartile Ranges of Hs-CRP (mg/L) At Baseline According To rs13210472 Genotype (A/A, A/C, C/C)

Stratified across patients with prevalent coronary artery disease (CAD), stratified according to statin use and sex within the UK Biobank.

A.3 Supplementary Tables

A.3.1 Supplementary Table A.3.1

Supplementary Table A.3.1 Meta-Analysis Statistics For TNT and IDEAL Stratified According To Men (Top) and Women (Bottom).

snp on chr9	pos	gene	hwe	imputed	Reference allele	Other allele	eaf	beta	se	Beta 95L	Beta 95U	p	-log10 pvalue	N samples
rs10964972	**21,362,409**	**IFNA6(dist=11523), IFNA13(dist=4962)**	**na**	**yes**	**A**	**C**	**0.0164**	**0.8418**	**0.138**	**0.5713**	**1.1122**	**1.10E-09**	**8.9639**	**9,126**
rs10964973	**21,362,772**	**IFNA6(dist=11886), IFNA13(dist=4599)**	**na**	**yes**	**A**	**G**	**0.017**	**0.8334**	**0.1368**	**0.5654**	**1.1015**	**1.10E-09**	**8.9467**	**9,127**
rs77051438	21,396,157	IFNA2(dist=10761), IFNA8(dist=12989)	1	no	A	C	0.0213	0.6516	0.1314	0.3941	0.9091	7.20E-07	6.1417	9,139
rs79129546	21,396,392	IFNA2(dist=10996), IFNA8(dist=12754)	na	yes	C	T	0.0213	0.6519	0.1314	0.3943	0.9095	7.20E-07	6.1442	9,137
rs76671143	21,401,709	IFNA2(dist=16313),IFNA8(dist=7437)	na	yes	G	T	0.0213	0.6513	0.1315	0.3937	0.909	7.40E-07	6.1322	9,128
rs12551405	21,402,393	IFNA2(dist=16997), IFNA8(dist=6753)	na	yes	G	A	0.0212	0.6532	0.1315	0.3956	0.9109	6.90E-07	6.1646	9,127
rs12554005	21,402,655	IFNA2(dist=17259), IFNA8(dist=6491)	na	yes	C	G	0.0212	0.6529	0.1315	0.3953	0.9106	6.90E-07	6.1587	9,126
rs117581840	21,403,380	IFNA2(dist=17984), IFNA8(dist=5766)	na	yes	G	T	0.0212	0.653	0.1315	0.3953	0.9106	6.90E-07	6.1594	9,125

snp	region	gene	hwe	imputed	Reference allele	Other allele	eaf	beta	se	Beta 95L	Beta 95U	p	-log10 pvalue	N samples
rs16825596	intronic	PID1	na	yes	C	T	0.4072	0.4738	0.0959	0.2858	0.6618	7.90E-07	6.1	2,045

rs2536214	intronic	PID1	0.2593	no	T	C	0.4071	0.4742	0.0955	0.2871	0.6613	6.90E-07	6.1604	2,057
rs28368744	intronic	PID1	na	yes	A	G	0.4073	0.4722	0.0955	0.2851	0.6593	7.70E-07	6.1154	2,053
rs113132798	intronic	PID1	na	yes	T	TTATC	0.4078	0.4713	0.0959	0.2834	0.6593	9.00E-07	6.0464	2,038
rs13210472	**intergenic**	**HLA-DOA(dist=17692), HLA-DPA1(dist=37265)**	**na**	**yes**	**C**	**A**	**0.0333**	**0.9781**	**0.1772**	**0.6308**	**1.3253**	**3.50E-08**	**7.459**	**2,052**
rs6933319	intergenic	HLA-DOA(dist=18474), HLA-DPA1(dist=36483)	1	no	T	C	0.0386	0.9163	0.1744	0.5744	1.2582	1.50E-07	6.815	2,057
rs62457129	intergenic	CHN2(dist=24455), PRR15(dist=25021)	na	yes	C	G	0.0762	0.7015	0.1393	0.4285	0.9744	4.80E-07	6.3168	2,048
rs62457155	intergenic	PRR15(dist=5086), LOC646762(dist=73541)	na	yes	A	C	0.074	0.7076	0.1419	0.4295	0.9858	6.30E-07	6.2019	2,055
rs17810895	intergenic	PRR15(dist=7580), LOC646762(dist=71047)	na	yes	C	G	0.074	0.7078	0.1419	0.4297	0.986	6.20E-07	6.2068	2,055
rs62457158	intergenic	PRR15(dist=7983), LOC646762(dist=70644)	na	yes	A	G	0.074	0.7078	0.1419	0.4297	0.986	6.20E-07	6.207	2,055
rs62459160	intergenic	PRR15(dist=9195), LOC646762(dist=69432)	na	yes	C	G	0.074	0.7079	0.1419	0.4298	0.986	6.20E-07	6.208	2,055
rs4323392	intergenic	PRR15(dist=9519), LOC646762(dist=69108)	na	yes	C	G	0.074	0.7079	0.1419	0.4298	0.986	6.20E-07	6.2081	2,055
rs185893353	intergenic	PRR15(dist=13437), LOC646762(dist=65190)	na	yes	A	G	0.0742	0.7037	0.142	0.4254	0.982	7.30E-07	6.1338	2,056
rs62459195	intergenic	PRR15(dist=26934), LOC646762(dist=51693)	na	yes	A	G	0.0741	0.7227	0.1407	0.447	0.9984	2.80E-07	6.5473	2,029
rs58619555	ncRNA_intronic	INMT-FAM188B	na	yes	C	T	0.0329	0.9984	0.1915	0.6231	1.3737	1.90E-07	6.7248	2,055
rs79714917	ncRNA_intronic	INMT-FAM188B	na	yes	T	C	0.0331	0.9802	0.1916	0.6046	1.3558	3.20E-07	6.4954	2,056
rs76655111	ncRNA_intronic	INMT-FAM188B	na	yes	T	G	0.0329	0.991	0.1915	0.6158	1.3663	2.30E-07	6.6367	2,056
rs144162786	ncRNA_intronic	INMT-FAM188B	na	yes	AC	A	0.0331	0.981	0.1917	0.6054	1.3567	3.10E-07	6.5029	2,056
rs369084345	ncRNA_intronic	INMT-FAM188B	na	yes	A	G	0.0331	0.9787	0.1916	0.6031	1.3542	3.30E-07	6.4793	2,055

rs7846693	**intronic**	**PEBP4**	**0.2064**	**no**	**T**	**C**	**0.2489**	**0.5561**	**0.1011**	**0.358**	**0.7543**	**3.90E-08**	**7.412**	**2,057**
rs73212892	intronic	PEBP4	na	no	A	G	0.1542	0.6115	0.1147	0.3866	0.8363	1.00E-07	6.9982	2,046
rs113669726	intergenic	CLCN4(dist=190021), MID1(dist=17630)	na	yes	G	A	0.0196	1.1097	0.2265	0.6657	1.5536	9.80E-07	6.0091	2,041
rs17321391	intergenic	CLCN4(dist=191558),MID1(dist=16093)	na	yes	G	A	0.0196	1.1099	0.2265	0.6659	1.5539	9.80E-07	6.0095	2,041

*Common SNPs with p value ≤1.00×10^{-06} and gene information for **time to occurrence of cancer** using a coxph regression with male (or women) only in the TNT and IDEAL cohorts, controlling for age, treatment, sex-specific principal components (C1-C10) with a threshold of 5.00×10-08. **snp** single nucleotide polymorphism; **chr** chromosome; **hwe** hardy weinberg equilibrium; **eaf** effect allele frequency; **se** standard error; **95L** 95% lower limit; **95U** 95% upper limit*

A.3.2 Supplementary Table A.3.2

Supplementary Table A.3.2 Distribution of rs13210472 Genotypes (A/A, A/C, C/C) Across Specific Cancers Amongst Women Participants of the UK Biobank With Prevalent Coronary Artery Disease Taking Statins At Baseline.

Cancer	**A/A**, N = 2704[1]	**A/C**, N = 244[1]	**C/C**, N = 4[1]	**p-value**[2]
Breast				0.12
No	2661 (92%)	243 (8.4%)	4 (0.1%)	
Yes	43 (98%)	1 (2.3%)	0 (0%)	
Ear, nose, & throat				>0.9
No	2703 (92%)	244 (8.3%)	4 (0.1%)	
Yes	1 (100%)	0 (0%)	0 (0%)	
Female genital tract				>0.9
No	2684 (92%)	243 (8.3%)	4 (0.1%)	
Yes	20 (95%)	1 (4.8%)	0 (0%)	
Gastrointestinal				0.086
No	2668 (92%)	238 (8.2%)	3 (0.1%)	
Yes	36 (84%)	6 (14%)	1 (2.3%)	
Hematological				0.032
No	2693 (92%)	240 (8.2%)	4 (0.1%)	
Yes	11 (73%)	4 (27%)	0 (0%)	
Melanoma				0.058
No	2690 (92%)	240 (8.2%)	4 (0.1%)	
Yes	14 (78%)	4 (22%)	0 (0%)	
Metastatic				>0.9
No	2701 (92%)	244 (8.3%)	4 (0.1%)	
Yes	3 (100%)	0 (0%)	0 (0%)	
Other				0.038

Cancer	**A/A**, N = 2704[1]	**A/C**, N = 244[1]	**C/C**, N = 4[1]	**p-value**[2]
No	2702 (92%)	242 (8.2%)	4 (0.1%)	
Yes	2 (50%)	2 (50%)	0 (0%)	
Neurological				
No	2704 (92%)	244 (8.3%)	4 (0.1%)	
Respiratory				0.74
No	2677 (92%)	241 (8.2%)	4 (0.1%)	
Yes	27 (90%)	3 (10%)	0 (0%)	
Urinary				0.089
No	2694 (92%)	241 (8.2%)	4 (0.1%)	
Yes	10 (77%)	3 (23%)	0 (0%)	

[1] Statistics presented: n (%)
[2] Statistical tests performed: Fisher's exact test for comparisons between A/A vs. A/C+C/C genotypes
[*] Percentages presented are in rows

A.3.3 Supplementary Table A.3.3

Supplementary Table A.3.3 Sensitivity Analyses Exploring the Effect of rs13210472 (Additive Model) in Different Contexts For Women Statin Users With Prevalent Coronary Artery Disease (CAD) Within the TNT, IDEAL, And UK Biobank Cohorts.

Cohort	Model	Endpoint	N	Events	rs13210472 HR (95% CI)	*P*
UK Biobank	Age, PCs (1-10)	**Recurrent CAD**	2952	1752	0.98 (0.83-1.16)	0.829
TNT	Age, intervention group, PCs (1-10), diabetes, hypertension, BMI, smoking status	**Incident cancer**	925	125	2.51 (1.61-3.90)	4.5×10^{-05}
IDEAL	Age, intervention group, PCs (1-10), diabetes, hypertension, BMI, smoking status	**Incident cancer**	1105	95	3.18 (1.79-5.64)	7.5×10^{-05}
UK Biobank	Age, PCs (1-10), diabetes, hypertension, BMI, smoking status	**Incident cancer**	2935	192	1.71 (1.14-2.56)	9.2×10^{-03}

PCs: principal components, CAD: coronary artery disease, BMI: body mass index, HR: hazard ratio, CI: confidence interval

A.3.4 Supplementary Table A.3.4

Supplementary Table A.3.4 Assessment of rs13210472 (Additive Model) and High-Sensitivity C Reactive Protein (Hs-CRP) Within the UK Biobank.

Patients	Model	Variables	Endpoint	N	Events	rs13210472 Beta (SE)	*P*
Women statin users with prevalent CAD	Generalized linear regression	Age, PCs (1-10)	ln(hs-CRP)	2805	n/a	0.24 (0.07)	8.7×10^{-05}
Women statin users with prevalent CAD	Generalized linear regression	Age, PCs (1-10), statin use, statin*rs13210472	ln(hs-CRP)	2952	n/a	$P_{interaction}$=0.103	
Men statin users with prevalent CAD	Generalized linear regression	Age, PCs (1-10)	ln(hs-CRP)	9025	n/a	-0.01 (0.04)	0.838
Patients	**Model**	**Variables**	**Endpoint**	**N**	**Events**	**rs13210472 HR (95% CI)**	***P***
Women statin users with prevalent CAD	Cox regression	Age, PCs (1-10), recurrent CAD, ln(hs-CRP)	Incident cancer	2805	184	1.72 (1.14-2.60)	9.7×10^{-03}
Women statin users with prevalent CAD	Cox regression	Age, PCs (1-10), recurrent CAD, ln(hs-CRP) ln(hs-CRP)*rs13210472	Incident cancer	2805	184	$P_{interaction}$=0.371	
Men statin users with prevalent CAD	Cox regression	Age, PCs (1-10), recurrent CAD, ln(hs-CRP)	Incident cancer	9025	810	1.13 (0.90-1.42)	0.311

PCs: principal components, hs-CRP: high-sensitivity C reactive protein, n/a: not available, SE: standard error, HR: hazard ratio, CI: confidence interval, CAD: coronary artery disease

Supplementary files for Chapter 3

A.4 Supplementary Tables

A.4.1 Supplementary Table B.1.1

Supplementary Table A.4.1 Primary Diagnostic Codes For Selected Cancers*

	ICD-9	**ICD-10**
Urinary bladder	188x	C67x
Larynx	161.0, 161.1, 161.3, 161.9	C32x
Prostate	185	C61
Corpus uteri	182	C54x
Rectal	154.1	C20
Breast	174x	C50x
Kidney	189	C64
Non-Hodgkin Lymphoma	200.0 to 200.7, 202.0, 202.1, 202.2, 202.7	C82x, C83x, C84x, C85x
Melanoma of the skin	172.x	C43x
Lung and bronchus	162.x	C34x

*Based on Sturgeon et al.[7] All cancer cases were identified using the cancer register ICD-9 and ICD-10 diagnostic codes (data-fields 40013 and 40006)

A.4.2 Supplementary Table B.1.2

Supplementary Table A.4.2 ICD-9 and ICD-10 Diagnostic and Procedure Codes For Inpatient Cardiovascular-Related Endpoints* and For the Primary Endpoints

Description	**ICD-9**	**ICD-10**
Death for cardiovascular disease causes		I00_I99
Death for coronary artery disease causes		I20, I21, I22, I23, I24 or I25
Death from cancer		C01x, C02x, C03x, C04x, C05x, C06x, C07x, C08x, C09x, C10x, C11x, C12, C13x, C14x, C15x, C16x, C17x, C18x, C19, C20, C21x, C22x, C23x, C25x, C26x, C30x, C31x, C32x, C33, C34x, C37, C38x, C39x, C40x, C41x, C43x, C45x, C46x, C47x, C48x, C49x, C50x, C51x, C52, C53x, C54x, C55, C56, C57x, C48, C60x, C61, C62x, C64, C65, C66, C67x, C68x, C69x, C70x, C71x, C73, C74x, C75x, C76x, C78x, C79x, C80x, C81x, C82x, C83x, C85x, C85x, C86x, C88x, C90x, C91x, C92x, C93x, C94x, C95x, C96x, C97x, D45, D473, D752, D474, D758, D471, D479

Cardiovascular disease (including angina pectoris, STEMI, NSTEMI, other acute ischemic heart diseases, chronic ischemic heart disease, cardiac arrest, nontraumatic subarachnoid hemorrhage, nontraumatic intracerebral hemorrhage, cerebral infarction)	410.x, 411.x, 412.x, 413.x, 414.x, 429.79, 430.x, 431.x, 432.x, 433.x, 434.x, 435.x, 436.x, 437.x, 438.x	I20.x, I21.x, I22.x, I23.x, I24.1, I25.x, I46, I60.x, I61.x, I63.x
Ischemic heart disease	410.x - 414.x	I20.x - I25.x
Myocardial infarction	410.x, 411.x, 412.x, 429.79	I21.x, I22.x, I23.x, I24.1, I25.2
STEMI	410.0 - 410.6, 410.8 - 410.9	I21.0 - I21.3, I22.0, I22.1, I22.8
NSTEMI	410.7	I21.4, I21.9, I22.9
Stable angina	413.1, 413.9	I20.1, I20.8, I20.9
Unstable angina	411.1, 411.81, 411.89	I20.0, I24.0, I24.8, I24.9
Stroke	430.x, 431.x, 432.x, 433.x, 434.x, 435.x, 436.x, 437.x, 438.x	I60.x, I61.x, I63.x, I64.x
Ischemic stroke	434.x,436.x	I63.x, I64.x
Intracerebral hemorrhage	431.x	I61.x
Subarachnoid hemorrhage	430.x	I60.x
Transient ischemic attack	435.x	G45.0, G45.1, G45.2, G45.8, G45.9
Heart Failure	428.x, 425.x	I50.x, I42.0
Peripheral vascular disease	250.6x, 440.2x, 443.1, 443.8, 443.9	I73.x, I74.3, I74.4, I74.5
Arrhythmia and conduction disorders (incl. atrial fibrillation)	426.x, 427.x	I44.x, I48.x
Asthma	493.x	J45.x, J46.x

A.4.3 Supplementary Table B.1.3

Supplementary Table A.4.3 Sensitivity Analyses By Cox Regression Considering Additional Covariates.

Characteristic	N	Event N	HR[1]	95% CI[1]	p-value
Time to CVD death					
mCA					
No mCA	38,612	543	—	—	
Any mCA	10,011	258	1.154	0.985, 1.351	0.076
Time to CAD death					
mCA					
No mCA	38,612	223	—	—	
Any mCA	10,011	140	1.379	1.102, 1.726	0.005
Time to cancer death					
mCA					
No mCA	38,612	6209	—	—	
Any mCA	10,011	2161	1.069	1.015, 1.127	0.012
Time to any death					
mCA					
No mCA	38,612	7721	—	—	
Any mCA	10,011	2823	1.081	1.032, 1.132	<0.001

Models adjusted for age at baseline, sex, smoking status, alcohol status, chemotherapy, radiotherapy, prevalent hypertension, prevalent high cholesterol, prevalent use of lipid-lowering medication, body mass index, number of days between date of recruitment and date of cancer diagnosis, and genotyping principal components 1-10. [1] *CAD: coronary artery disease, CI: confidence interval, CVD: cardiovascular disease, HR: hazard ratio, mCA: mosaic chromosomal alterations*

A.4.4 Supplementary Table B.1.4

Supplementary Table A.4.4 Multivariable Competing-Risks Regression Model For the Prediction of Death Of CVD Causes and CAD Causes.

Cardiovascular disease (CVD)			
	HR	**95% CI**	***P***
mCA vs. none	1.109	0.946, 1.300	0.200
Coronary artery disease (CAD)			
	HR	**95% CI**	***P***
mCA vs. none	1.333	1.060, 1.677	0.014

All models adjusted for age at baseline, sex, chemotherapy, days elapsed between prevalent cancer diagnosis date and study entry date, smoking status, receipt of chemotherapy, principal components 1 thru 10, and death from non-CVD or CAD causes. HR: hazard ratio, CI: confidence interval

A.4.5 Supplementary Table B.1.5

Supplementary Table A.4.5 Cox Regression Analyses Evaluating the Effect of Mosaic Chromosomal Alterations On the Risk of Death of Cardiovascular Disease Causes, Coronary Artery Disease Causes, From Cancer, and Any Cause of Death Stratified By Smoking Status Groups.

	Time to CVD death				Time to CAD death				Time to cancer death				Time to any death			
Characteristic	**Event N**	**HR[1]**	**95% CI[1]**	**p-value**	**Event N**	**HR[1]**	**95% CI[1]**	**p-value**	**Event N**	**HR[1]**	**95% CI[1]**	**p-value**	**Event N**	**HR[1]**	**95% CI[1]**	**p-value**
Never smokers																
No mCA (n=19,888)	198	—	—		70	—	—		2427	—	—		2982	—	—	
Any mCA (n=3,977)	65	1.075	0.801, 1.442	0.632	36	1.479	0.97, 2.259	0.070	592	1.054	0.96, 1.159	0.275	759	1.053	0.97, 1.146	0.225
Previous smokers																
No mCA (n=14,755)	241	—	—		117	—	—		2551	—	—		3239	—	—	
Any mCA (n=4,519)	141	1.259	1.012, 1.566	0.039	79	1.372	1.017, 1.849	0.038	1003	1.080	1.000, 1.167	0.051	1362	1.111	1.039, 1.188	0.002
Current smokers																
No mCA (n=3,970)	104	—	—		36	—	—		1231	—	—		1500	—	—	
Any mCA (n=1,515)	52	0.98	0.682, 1.416	0.926	25	1.278	0.726, 2.250	0.395	566	0.990	0.885, 1.102	0.822	702	0.980	0.892, 1.087	0.761

Models adjusted for age at baseline, sex, chemotherapy, radiotherapy, number of days between date of cancer diagnosis and date of baseline, genotyping, principal components 1 thru 10. [1] *CAD: coronary artery disease, CI: confidence interval, CVD: cardiovascular disease, HR: hazard ratio, mCA: mosaic chromosomal alterations*

A.4.6 Supplementary Table B.1.6

Supplementary Table A.4.6 Cox Regression Analyses Evaluating the Effect Of Mosaic Chromosomal Alterations On the Risk of Death of Cardiovascular Disease Causes, Coronary Artery Disease Causes, From Cancer, and Any Cause of Death By Chemotherapy Status.

	Time to CVD death				Time to CAD death				Time to cancer death				Time to any death			
Characteristic	**Event N**	**HR[1]**	**95% CI[1]**	**p-value**	**Event N**	**HR[1]**	**95% CI[1]**	**p-value**	**Event N**	**HR[1]**	**95% CI[1]**	**p-value**	**Event N**	**HR[1]**	**95% CI[1]**	**p-value**
Never treated with chemotherapy																
No mCA (n=29,425)	437	—	—		188	—	—		3080	—	—		4260	—	—	
Any mCA (n=7,853)	203	1.082	0.907, 1.291	0.384	109	1.242	0.970, 1.593	0.088	1170	1.069	0.990, 1.148	0.071	1682	1.065	1.002, 1.131	0.042
Previously treated with chemotherapy																
No mCA (n=9,188)	106	—	—		35	—	—		3129	—	—		3461	—	—	
Any mCA (n=2,158)	55	1.381	0.970, 1.964	0.073	31	2.151	1.270, 3.641	0.004	991	1.033	0.960, 1.115	0.415	1141	1.058	0.980, 1.137	0.128

Models adjusted for age at baseline, sex, radiotherapy, number of days between date of cancer diagnosis and date of baseline, genotyping principal components 1 thru 10. [1] *CAD: coronary artery disease, CI: confidence interval, CVD: cardiovascular disease, HR: hazard ratio, mCA: mosaic chromosomal alterations*

A.4.7 Supplementary Table B.1.7

Supplementary Table A.4.7 The Effect of Autosomal Mosaic Chromosomal Alterations On Death of Cardiovascular Disease Causes, Coronary Artery Disease Causes, From Cancer and Any Cause of Death.

Characteristic	N	Event N	HR[1]	95% CI[1]	p-value
Time to CVD death					
Autosomal mCA					
Ref.	46,203	739	—	—	
Autosomal	2,421	62	1.353	1.043, 1.754	0.023
Time to CAD death					
Autosomal mCA					
Ref.	46,203	333	—	—	
Autosomal	2,421	30	1.444	0.990, 2.100	0.055
Time to cancer death					
Autosomal mCA					
Ref.	46,203	7841	—	—	
Autosomal	2,421	529	1.137	1.041, 1.242	0.004
Time to any death					
Autosomal mCA					
Ref.	46,203	9870	—	—	
Autosomal	2,421	674	1.143	1.057, 1.236	<0.001

Models adjusted for age at baseline, sex, smoking status, chemotherapy, radiotherapy, number of days between date of cancer diagnosis and date of study recruitment, genotyping principal components 1 thru 10. [1] *CAD: coronary artery disease, CI: confidence interval, CVD: cardiovascular disease, HR: hazard ratio, mCA: mosaic chromosomal alterations, Ref.: referent category includes no mCA or mCAs that were not autosomal*

A.4.8 Supplementary Table B.1.8

Supplementary Table A.4.8 The Effect of Mosaic Loss of X Chromosome On Death of Cardiovascular Disease Causes, Coronary Artery Disease Causes, From Cancer, and Any Cause of Death.

Characteristic	N	Event N	HR[1]	95% CI[1]	p-value
Time to CVD death					
chrX					
Ref.	24,745	258	—	—	
Loss of X	1,632	27	1.194	0.801, 1.780	0.384
Time to CAD death					
chrX					
Ref.	24,745	78	—	—	
Loss of X	1,632	15	2.022	1.157, 3.535	0.013
Time to cancer death					
chrX					
Ref.	24,745	3715	—	—	
Loss of X	1,632	277	1.005	0.889, 1.137	0.932
Time to any death					
chrX					
Ref.	24,745	4497	—	—	
Loss of X	1,632	343	1.000	0.896, 1.118	0.994

Models adjusted for age at baseline, smoking status, chemotherapy, radiotherapy, number of days between date of recruitment and date of cancer diagnosis, and genotyping principal components 1-10. [1] *CAD: coronary artery disease, CI: confidence interval,*

CVD: cardiovascular disease, HR: hazard ratio, mCA: mosaic chromosomal alterations, Ref.: referent category includes no mCA or mCAs that were not loss of X chromosome

A.4.9 Supplementary Table B.1.9

Supplementary Table A.4.9 The Effect of Mosaic Loss of Y Chromosome On Death of Cardiovascular Disease Causes, Coronary Artery Disease Causes, From Cancer and Any Cause of Death.

Characteristic	N	Event N	HR[1]	95% CI[1]	p-value
Time to CVD death					
chrY					
Ref.	15,733	333	—	—	
Loss of Y	6,514	183	1.03	0.856, 1.238	0.757
Time to CAD death					
chrY					
Ref.	15,733	170	—	—	
Loss of Y	6,514	100	1.147	0.890, 1.478	0.289
Time to cancer death					
chrY					
Ref.	15,733	2882	—	—	
Loss of Y	6,514	1496	1.065	0.998, 1.136	0.056
Time to any death					
chrY					
Ref.	15,733	3717	—	—	
Loss of Y	6,514	1987	1.077	1.019, 1.140	0.009

Models adjusted for age at baseline, smoking status, chemotherapy, radiotherapy, number of days between date of recruitment and date of cancer diagnosis, and genotyping principal components 1-10. HR = Hazard Ratio, CI = Confidence Interval, CAD:

coronary artery disease, CI: confidence interval, CVD: cardiovascular disease, HR: hazard ratio, mCA: mosaic chromosomal alterations, Ref.: referent category includes no mCA or mCAs that were not loss of Y chromosome

A.4.10 Supplementary Table B.1.10

Supplementary Table A.4.10 The Effect of Expanded Mosaic Chromosomal Alterations On Death of Cardiovascular Disease Causes, Coronary Artery Disease Causes, From Cancer, and Any Cause of Death.

Characteristic	N	Event N	HR[1]	95% CI[1]	p-value
Time to CVD death					
Expanded mCA					
Ref.	46,722	768	—	—	
Expanded mCA	1,902	33	1.170	0.819, 1.672	0.388
Time to CAD death					
Expanded mCA					
Ref.	46,722	347	—	—	
Expanded mCA	1,902	16	1.624	0.97, 2.724	0.066
Time to cancer death					
Expanded mCA					
Ref.	46,722	8035	—	—	
Expanded mCA	1,902	335	0.997	0.892, 1.114	0.958
Time to any death					
Expanded mCA					
Ref.	46,722	10115	—	—	
Expanded mCA	1,902	429	1.025	0.929, 1.131	0.626

Models adjusted for age at baseline, sex, smoking status, chemotherapy, radiotherapy, number of days between date of recruitment and date of cancer diagnosis, and genotyping principal components 1-10. HR = Hazard Ratio, CI = Confidence

Interval, CAD: coronary artery disease, CI: confidence interval, CVD: cardiovascular disease, HR: hazard ratio, mCA: mosaic chromosomal alterations, Ref.: referent category includes no mCA or not expanded mCAs

A.4.11 Supplementary Table B.1.11

Supplementary Table A.4.11 The Effect of Mosaic Chromosomal Alterations (mCA) On Death of Cardiovascular Disease Causes, Coronary Artery Disease Causes, From Cancer, and Any-Cause Death With a mCA-By-Cancer Status Interaction Term in All Patients.

Characteristic	N	Event N	HR[1]	95% CI[1]	p-value
Time to CVD death					
Cancer status * mCA	479,435	7315	1.016	0.896, 1.154	0.800
Time to CAD death					
Cancer status * mCA	479,435	3793	0.995	0.831, 1.193	0.961
Time to any cancer death					
Cancer status * mCA	479,435	17680	0.62	0.541, 0.711	<0.001
Time to any death					
Cancer status * mCA	479,435	35527	0.718	0.685, 0.754	<0.001

Models adjusted for age at baseline, sex, smoking status, mCA, any cancer, principal components 1-10, and a mCA-by-cancer status interaction term. Of note, any cancer here includes all cancer diagnoses identified within the cancer register. CAD: coronary artery disease, CI: confidence interval, CVD: cardiovascular disease, HR: hazard ratio, mCA: mosaic chromosomal alterations

A.4.12 Supplementary Table B.1.12

Supplementary Table A.4.12 Cox Regression Analyses of the Effect of Mosaic Chromosomal Alterations On the Risk of Death of Cardiovascular Disease Causes Per Cancer Type.

Characteristic	N	Event N	HR[1]	95% CI[1]	p-value
Bladder cancer					
mCA					
No mCA	1,184	30	—	—	
Any mCA	550	25	1.286	0.744, 2.223	0.367
Larynx cancer					
mCA					
No mCA	243	9	—	—	
Any mCA	129	6	1.015	0.347, 2.967	0.979
Corpus uteri					
mCA					
No mCA	2,092	28	—	—	
Any mCA	221	2	0.655	0.155, 2.779	0.567
Prostate cancer					
mCA					
No mCA	8,906	165	—	—	
Any mCA	4,277	104	0.980	0.761, 1.256	0.860
Rectal cancer					
mCA					
No mCA	1,723	25	—	—	
Any mCA	535	13	1.199	0.586, 2.456	0.619
Breast cancer					
mCA					

No mCA	15,606	129	—	—	
Any mCA	1,618	24	1.341	0.863, 2.084	0.192
Kidney cancer					
mCA					
No mCA	1,463	29	—	—	
Any mCA	437	22	2.029	1.106, 3.720	0.022
Non-Hodgkin lymphoma					
mCA					
No mCA	2,173	41	—	—	
Any mCA	845	23	1.206	0.704, 2.063	0.495
Melanoma					
mCA					
No mCA	4,018	41	—	—	
Any mCA	876	16	1.059	0.578, 1.940	0.853
Lung cancer					
mCA					
No mCA	3,039	68	—	—	
Any mCA	1,129	40	1.211	0.785, 1.867	0.387

Models adjusted for age at baseline, sex (except for breast and prostate cancer), smoking status, chemotherapy, radiotherapy, number of days elapsed between date of recruitment and date of cancer diagnosis, and genotyping principal components 1 thru 10. HR: hazard ratio, CI: confidence interval, mCA: mosaic chromosomal alterations.

A.4.13 Supplementary Table B.1.13

Supplementary Table A.4.13 Cox Regression Analyses For the Effect of Mosaic Chromosomal Alterations On the Risk of Death of Coronary Artery Disease Causes By Cancer Type.

Characteristic	N	Event N	HR[1]	95% CI[1]	p-value
Bladder cancer					
mCA					
No mCA	1,184	13	—	—	
Any mCA	550	13	1.587	0.713, 3.532	0.258
Larynx cancer					
mCA					
No mCA	243	6	—	—	
Any mCA	129	5	1.238	0.356, 4.305	0.737
Corpus uteri					
mCA					
No mCA	2,092	7	—	—	
Any mCA	221	0	n/a	n/a	n/a
Prostate cancer					
mCA					
No mCA	8,906	88	—	—	
Any mCA	4,277	55	1.010	0.716, 1.426	0.953
Rectal cancer					
mCA					
No mCA	571	7	—	—	
Any mCA	184	4	1.024	0.279, 3.761	0.972
Breast cancer					
mCA					
No mCA	15,606	32	—	—	

Any mCA	1,618	11	2.458	1.229, 4.916	0.011
Kidney cancer					
mCA					
No mCA	1,463	10	—	—	
Any mCA	437	13	3.571	1.443, 8.838	0.006
Non-Hodgkin lymphoma					
mCA					
No mCA	2,173	16	—	—	
Any mCA	845	11	2.093	0.904, 4.844	0.084
Melanoma					
mCA					
No mCA	4,018	17	—	—	
Any mCA	876	11	1.571	0.702, 3.516	0.272
Lung cancer					
mCA					
No mCA	3,039	31	—	—	
Any mCA	1,129	23	1.418	0.782, 2.572	0.250

Models adjusted for age at baseline, sex (except for breast and prostate cancer), smoking status, chemotherapy, radiotherapy, number of days elapsed between date of recruitment and date of cancer diagnosis, and genotyping principal components 1 thru 10. HR: hazard ratio, CI: confidence interval, mCA: mosaic chromosomal alterations, n/a: not available

A.4.14 Supplementary Table B.1.14

Supplementary Table A.4.14 Cox Regression Analyses of the Effect of Mosaic Chromosomal Alterations On the Risk Of Death From Cancer According To Cancer Type.

Characteristic	N	Event N	HR[1]	95% CI[1]	p-value
Bladder cancer					
mCA					
No mCA	1,184	319	—	—	
Any mCA	550	169	1.079	0.887, 1.313	0.445
Larynx cancer					
mCA					
No mCA	243	57	—	—	
Any mCA	129	23	0.684	0.407, 1.149	0.151
Corpus uteri					
mCA					
No mCA	2,092	291	—	—	
Any mCA	221	45	1.353	0.980, 1.859	0.063
Prostate cancer					
mCA					
No mCA	8,906	958	—	—	
Any mCA	4,277	571	1.025	0.922, 1.140	0.644
Rectal cancer					
mCA					
No mCA	1,723	378	—	—	
Any mCA	535	144	1.168	0.950, 1.435	0.138
Breast cancer					
mCA					
No mCA	15,606	1680	—	—	

Any mCA	1,618	189	1.059	0.909, 1.234	0.464
Kidney cancer					
mCA					
No mCA	1,463	368	—	—	
Any mCA	437	137	0.99	0.805, 1.218	0.923
Non-Hodgkin lymphoma					
mCA					
No mCA	2,173	419	—	—	
Any mCA	845	227	1.15	0.970, 1.362	0.106
Melanoma					
mCA					
No mCA	4,018	325	—	—	
Any mCA	876	109	0.99	0.778, 1.249	0.907
Lung cancer					
mCA					
No mCA	3,039	2000	—	—	
Any mCA	1,129	778	0.98	0.894, 1.071	0.643

Models adjusted for age at baseline, sex (except for breast and prostate cancer), smoking status, chemotherapy, radiotherapy, number of days elapsed between date of recruitment and date of cancer diagnosis, and genotyping principal components 1 thru 10. HR: hazard ratio, CI: confidence interval, mCA: mosaic chromosomal alterations

A.4.15 Supplementary Table B.1.15

Supplementary Table A.4.15 Cox Regression Analyses Evaluating the Effect of Mosaic Chromosomal Alterations On the Risk of Any Cause of Death By Cancer Type.

Characteristic	N	Event N	HR[1]	95% CI[1]	p-value
Bladder cancer					
mCA					
No mCA	1,184	399	—	—	
Any mCA	550	227	1.107	0.933, 1.312	0.244
Larynx cancer					
mCA					
No mCA	243	77	—	—	
Any mCA	129	39	0.837	0.555, 1.261	0.394
Corpus uteri					
mCA					
No mCA	2,092	358	—	—	
Any mCA	221	55	1.348	1.011, 1.797	0.042
Prostate cancer					
mCA					
No mCA	8,906	1337	—	—	
Any mCA	4,277	817	1.032	0.944, 1.129	0.483
Rectal cancer					
mCA					
No mCA	1,723	457	—	—	
Any mCA	535	174	1.097	0.910, 1.321	0.331
Breast cancer					
mCA					

No mCA	15,606	2103	—	—	
Any mCA	1,618	257	1.089	0.95, 1.242	0.207
Kidney cancer					
mCA					
No mCA	1,463	449	—	—	
Any mCA	437	179	1.040	0.865, 1.249	0.678
Non-Hodgkin lymphoma					
mCA					
No mCA	2,173	587	—	—	
Any mCA	845	315	1.136	0.980, 1.311	0.083
Melanoma					
mCA					
No mCA	4,018	429	—	—	
Any mCA	876	148	0.998	0.816, 1.220	0.982
Lung cancer					
mCA					
No mCA	3,039	2200	—	—	
Any mCA	1,129	883	0.990	0.906, 1.075	0.762

Models adjusted for age at baseline, sex (except for breast and prostate cancer), smoking status, chemotherapy, radiotherapy, number of days elapsed between date of recruitment and date of cancer diagnosis, and genotyping principal components 1 thru 10. HR: hazard ratio, CI: confidence interval, mCA: mosaic chromosomal alterations

A.4.16 Supplementary Table B.1.16

Supplementary Table A.4.16 Cox Regression Analyses For the Effect of Mosaic Chromosomal Alterations On the Risk of Incident Cardiovascular Endpoints Aged ≥65 Years Old (N=15,273).

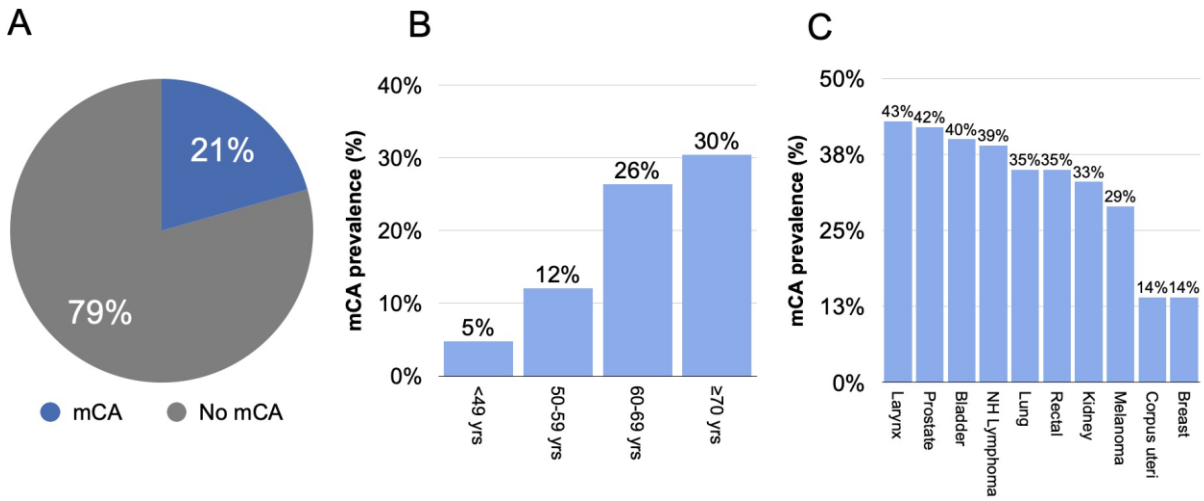
Characteristic	N	Event N	HR[1]	95% CI[1]	p-value
CVD					
mCA	15,273	2365	1.043	0.950, 1.141	0.356
Ischemic heart disease					
mCA	15,273	2032	1.034	0.939, 1.139	0.498
MI					
mCA	15,273	1230	1.114	0.990, 1.259	0.081
STEMI					
mCA	15,273	487	1.229	1.015, 1.488	0.034
NSTEMI					
mCA	15,273	757	1.203	1.031, 1.404	0.019
Stable angina					
mCA	15,273	538	1.204	1.003, 1.445	0.046
Unstable angina					
mCA	15,273	741	1.124	0.960, 1.316	0.145
Stroke					
mCA	15,273	908	1.074	0.931, 1.240	0.328
Ischemic stroke					
mCA	15,273	829	1.085	0.934, 1.260	0.287
Intracerebral hemorrhage					
mCA	15,273	506	1.223	1.013, 1.477	0.036
Subarachnoid hemorrhage					
mCA	15,273	462	1.234	1.013, 1.503	0.037
Heart failure					

mCA	15,273	581	1.121	0.938, 1.340	0.209
TIA					
mCA	15,273	1317	1.033	0.916, 1.164	0.599
Peripheral vascular disease					
mCA	15,273	838	1.197	1.033, 1.388	0.017
Arrhythmia and conduction					
mCA	15,273	2420	1.033	0.945, 1.128	0.475
Asthma					
mCA	15,273	1426	1.004	0.890, 1.132	0.953

Models adjusted for age at baseline, sex, smoking status, chemotherapy, radiotherapy, number of days between cancer diagnosis and date of study recruitment, and genotyping principal components 1 thru 10. HR: hazard ratio, CI: confidence interval, mCA: mosaic chromosomal alterations, MI: myocardial infarction, STEMI: ST-elevation myocardial infarction, NSTEMI: non-ST-elevation myocardial infarction, CVD: cardiovascular disease, TIA: transient ischemic attack

A.5 B.2 Supplementary Figures

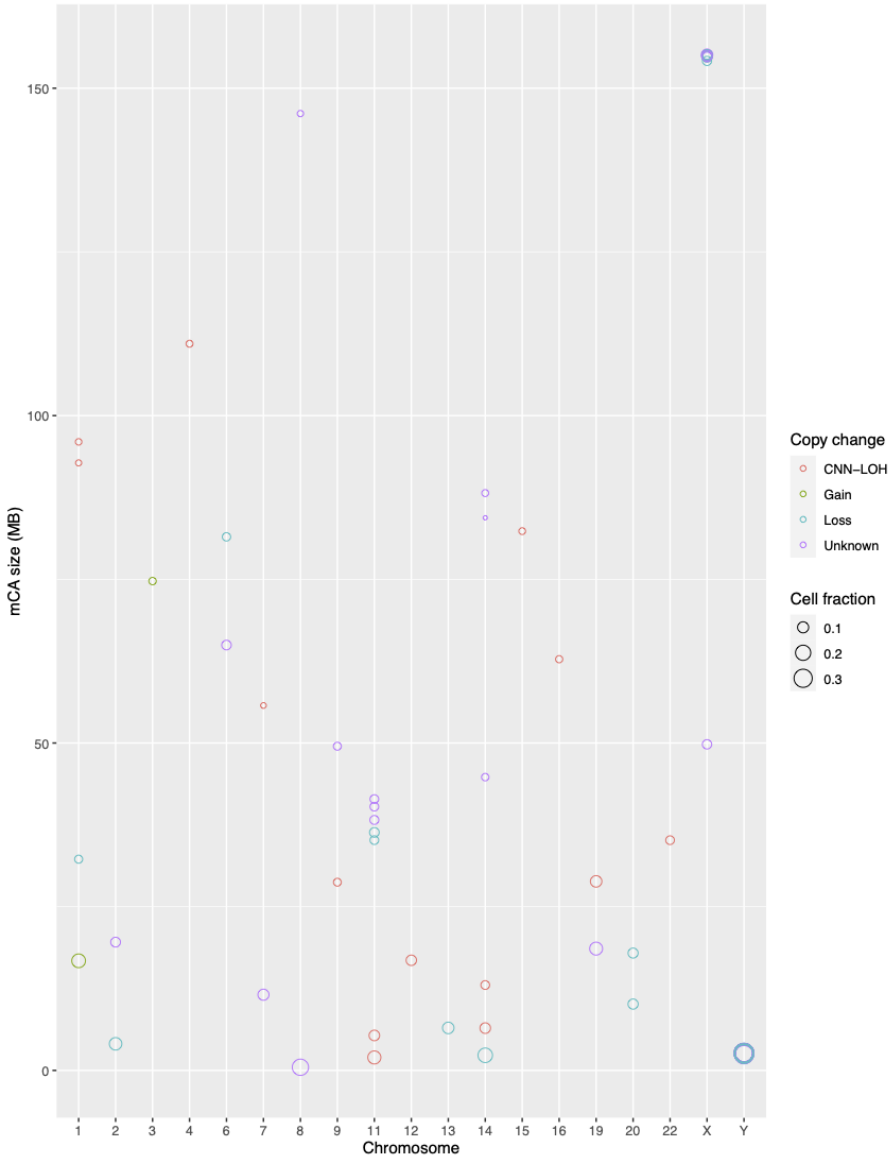
A.5.1 Supplementary Figure B.2.1



Supplementary Figure A.5.1 Visual Depiction of the Proportion of Patients With At Least 1 Mosaic Chromosomal Alteration.

In the overall cohort (A), of the proportion of patients with at least one mosaic chromosomal alteration stratified according to age groups (B), and of the proportion of patients with at least one mosaic chromosomal alteration stratified according to cancer types (C). mCA: mosaic chromosomal alteration

A.5.2 Supplementary Figure B.2.2



Supplementary Figure A.5.2 Visual Mapping of Mosaic Chromosomal Alterations (mCAs) Detected Across the Genome Among Individuals Who Died of Coronary Artery Disease (CAD) Causes.

A total of 155 mCA events, including 5 and 102 events on the X and Y chromosome, respectively, were detected across 140 unique participants. Each point represents one mCA event, with the x-axis as the chromosome, y-axis as the mCA size in mega-bases of DNA (MB), color coded by copy number (loss, gain, copy-number neutral loss of heterozygosity [CNN-LOH], unknown), and the size of the point as the cell fraction of that mCA.

Supplementary Files For Chapter 4

A.6 C.1 Supplementary Tables

A.6.1 Supplementary Table C.1.1

Supplementary Table A.6.1 Primary Diagnostic Codes For Selected Cancers*

	ICD-9	**ICD-10**
Urinary bladder	188x	C67x
Larynx	161.0, 161.1, 161.3, 161.9	C32x
Prostate	185	C61
Corpus uteri	182	C54x
Rectal	154.1	C20
Breast	174x	C50x
Kidney	189	C64
Non-Hodgkin Lymphoma	200.0 to 200.7, 202.0, 202.1, 202.2, 202.7	C82x, C83x, C84x, C85x
Melanoma of the skin	172.x	C43x
Lung and bronchus	162.x	C34x

**Based on Sturgeon et al.*[7] *All cancer cases were identified using the cancer register ICD-9 and ICD-10 diagnostic codes (data-fields 40013 and 40006)*

A.6.2 Supplementary Table C.1.2

Supplementary Table A.6.2 Multivariable Competing-Risk Regression Analyses On the Influence of CHIP (Any Vs. None) On the Risks of Death From Cardiovascular-Related Causes, Accounting For Other-Cause Mortality.

Characteristic	N	HR[1]	95% CI[1]	p-value
Time to CV death				
None	18,299	—	—	
Any CHIP	1,150	1.521	1.051, 2.202	0.026
Time to CAD death				
None	18,299	—	—	
Any CHIP	1,150	1.208	0.667, 2.189	0.530

Models adjusted for age at baseline, sex, smoking status, chemotherapy, radiotherapy, prevalent CVD, number of days between date of recruitment and date of cancer diagnosis, genotyping principal components 1-10, and account for the competing event of other-cause mortality. [1]*HR = Hazard Ratio, CI = Confidence Interval, CAD: coronary artery disease, CVD: cardiovascular disease, CHIP: clonal hematopoiesis of indeterminate potential*

A.6.3 Supplementary Table C.1.3

Supplementary Table A.6.3 Multivariable Cox Regression Analyses On the Influence of DNMT3A, TET2, and ASXL1 On the Risks of Death From CV-Related Causes and Emergency Room Admission.

Characteristic	N	Event N	HR[1]	95% CI[1]	p-value
DNMT3A					
CVD death	576	9	0.883	0.453, 1.719	0.714
CAD death	576	3	0.669	0.210, 2.132	0.496
Any death	576	141	1.138	0.962, 1.347	0.132
CV-related ER admission	576	94	1.085	0.883, 1.333	0.438
TET2					
CVD death	128	6	3.154	1.403, 7.090	0.005
CAD death	128	1	1.256	0.175, 9.014	0.820
Any death	128	42	1.666	1.228, 2.258	0.001
CV-related ER admission	128	23	1.174	0.779, 1.771	0.443
ASXL1					
CVD death	200	8	2.194	1.084, 4.441	0.029
CAD death	200	5	3.177	1.291, 7.818	0.012
Any death	200	71	1.576	1.245, 1.994	<0.001
CV-related ER admission	200	36	1.012	0.728, 1.407	0.945

Models adjusted for age at baseline, sex, smoking status, chemotherapy, radiotherapy, prevalent CVD, number of days between date of recruitment and date of cancer diagnosis, and genotyping principal components 1-10. CAD: coronary artery disease, CHIP: clonal hematopoiesis of indeterminate potential, CI: confidence interval, CV: cardiovascular, ER: emergency room, HR: hazard ratio

A.6.4 Supplementary Table C.1.4

Supplementary Table A.6.4 Multivariable Cox Regression Analyses On the Influence of CHIP (Any Vs. None) On the Risks of Incident CV Phenotypes.

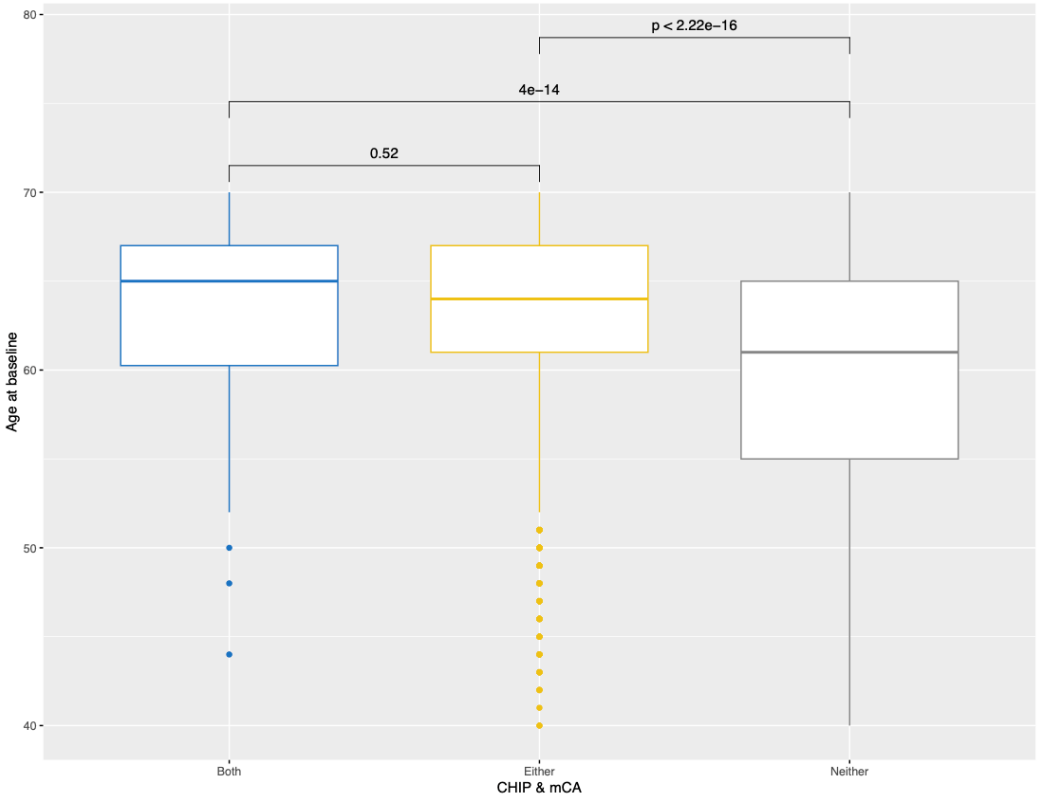
Characteristic	N	Event N	HR[1]	95% CI[1]	p-value
CVD					
None	18,299	1762	—	—	
Any CHIP	1,150	138	1.152	0.968, 1.371	0.111
Ischemic heart disease					
None	18,299	1506	—	—	
Any CHIP	1,150	120	1.167	0.968, 1.406	0.105
MI					
None	18,299	800	—	—	
Any CHIP	1,150	68	1.198	0.934, 1.536	0.155
STEMI					
None	18,299	308	—	—	
Any CHIP	1,150	35	1.614	1.136, 2.292	0.008
NSTEMI					
None	18,299	463	—	—	
Any CHIP	1,150	51	1.550	1.160, 2.072	0.003
Stable angina					
None	18,299	349	—	—	
Any CHIP	1,150	38	1.523	1.088, 2.130	0.014
Unstable angina					
None	18,299	479	—	—	
Any CHIP	1,150	44	1.311	0.962, 1.787	0.087
Stroke					
None	18,299	595	—	—	
Any CHIP	1,150	59	1.421	1.086, 1.858	0.010
Ischemic stroke					
None	18,299	514	—	—	
Any CHIP	1,150	56	1.555	1.179, 2.051	0.002
Intracerebral hemorrhage					
None	18,299	331	—	—	
Any CHIP	1,150	32	1.372	0.953, 1.974	0.089
Subarachnoid hemorrhage					
None	18,299	272	—	—	
Any CHIP	1,150	34	1.753	1.226, 2.508	0.002
Heart failure					
None	18,299	823	—	—	
Any CHIP	1,150	72	1.225	0.962, 1.560	0.099
TIA					

Characteristic	N	Event N	HR[1]	95% CI[1]	p-value
None	18,299	358	—	—	
Any CHIP	1,150	36	1.417	1.005, 1.999	0.047
Peripheral vascular disease					
None	18,299	525	—	—	
Any CHIP	1,150	60	1.654	1.265, 2.163	<0.001
Arrhythmia and conduction					
None	18,299	1678	—	—	
Any CHIP	1,150	131	1.088	0.910, 1.301	0.353
Hypertension					
None	18,299	4664	—	—	
Any CHIP	1,150	322	1.042	0.930, 1.167	0.477
Asthma					
None	18,299	1421	—	—	
Any CHIP	1,150	91	0.993	0.803, 1.228	0.949

Models adjusted for age at baseline, sex, smoking status, chemotherapy, radiotherapy, number of days between cancer diagnosis and date of baseline, and genotyping principal components 1-10. CHIP: clonal hematopoiesis of indeterminate potential, CI: confidence interval, CV: cardiovascular, CVD: cardiovascular disease, HR: hazard ratio, MI: myocardial infarction, STEMI: ST-elevation myocardial infarction, NSTEMI: non-ST-elevation myocardial infarction, TIA: transient ischemic attack

A.7 C.2 Supplementary Figures

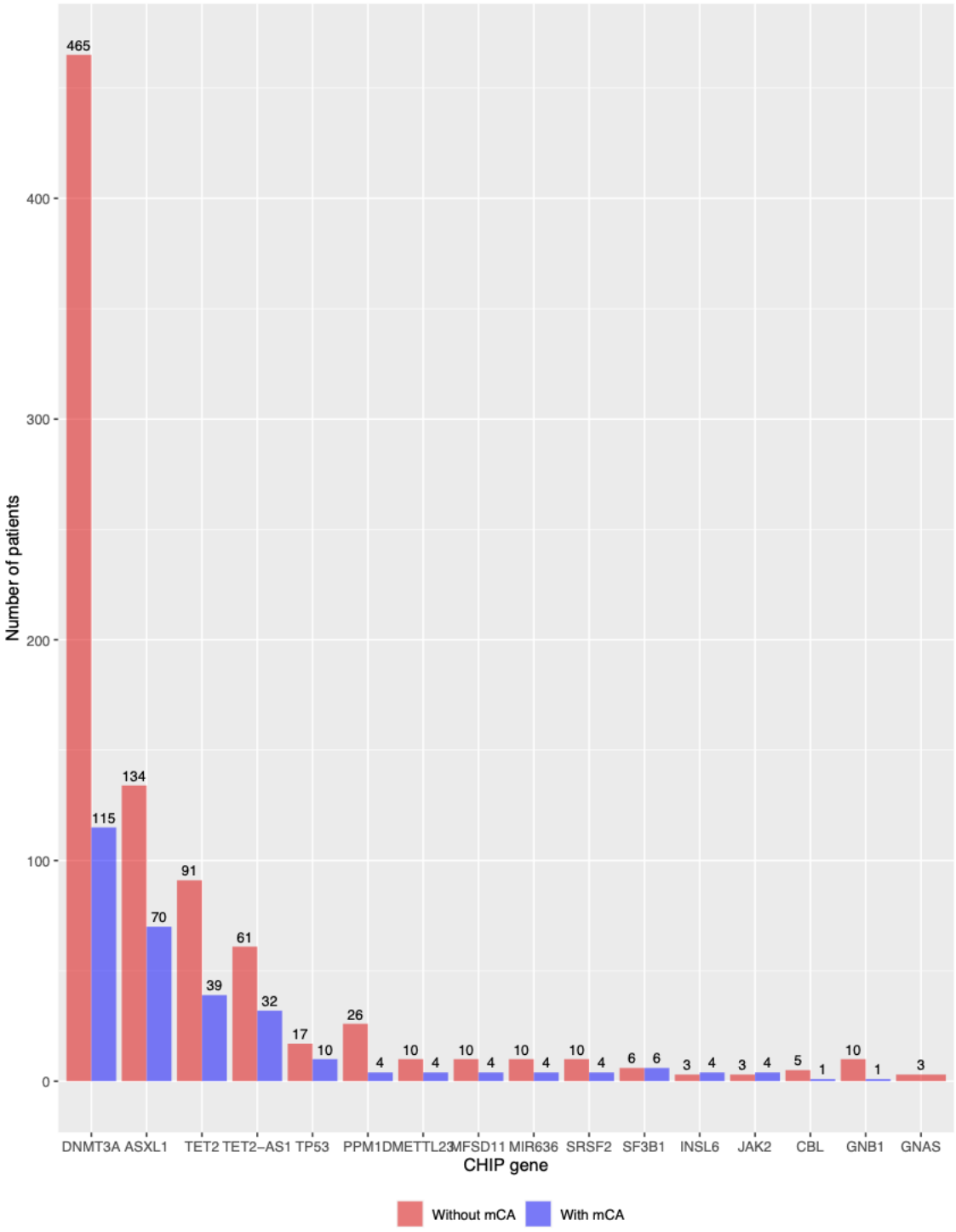
A.7.1 Supplementary Figure C.2.1



Supplementary Figure A.7.1 Age Distribution of Individuals With Both CHIP and Mca (Blue), CHIP or mCA Alone (Yellow), and Neither Type of CH (Grey).

Unadjusted P-values derived from a two-sided Student's t-test are shown. The lower and upper bounds of boxes denote 25th (Q1) and 75th (Q3) percentiles of observed ages, respectively. The lower and upper whiskers indicate the and maximal values. Extreme outliers are shown in dots. CH: clonal hematopoiesis, CHIP: clonal hematopoiesis of indeterminate potential, mCA: mosaic chromosomal alterations

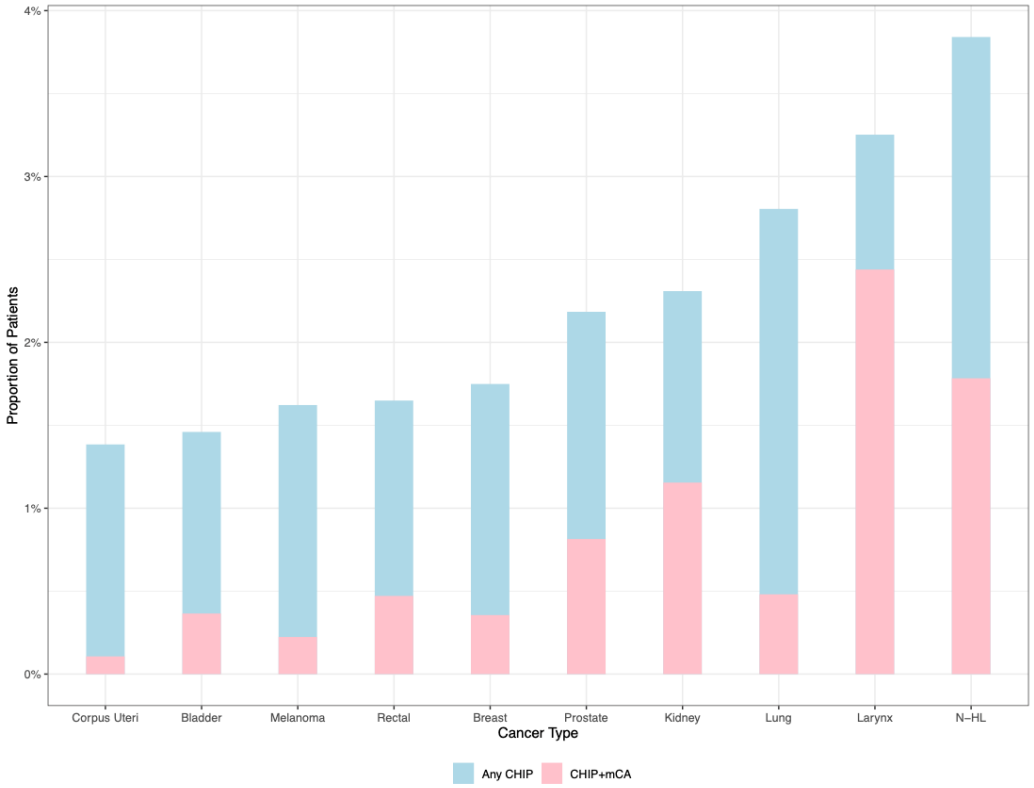
A.7.2 Supplementary Figure C.2.2

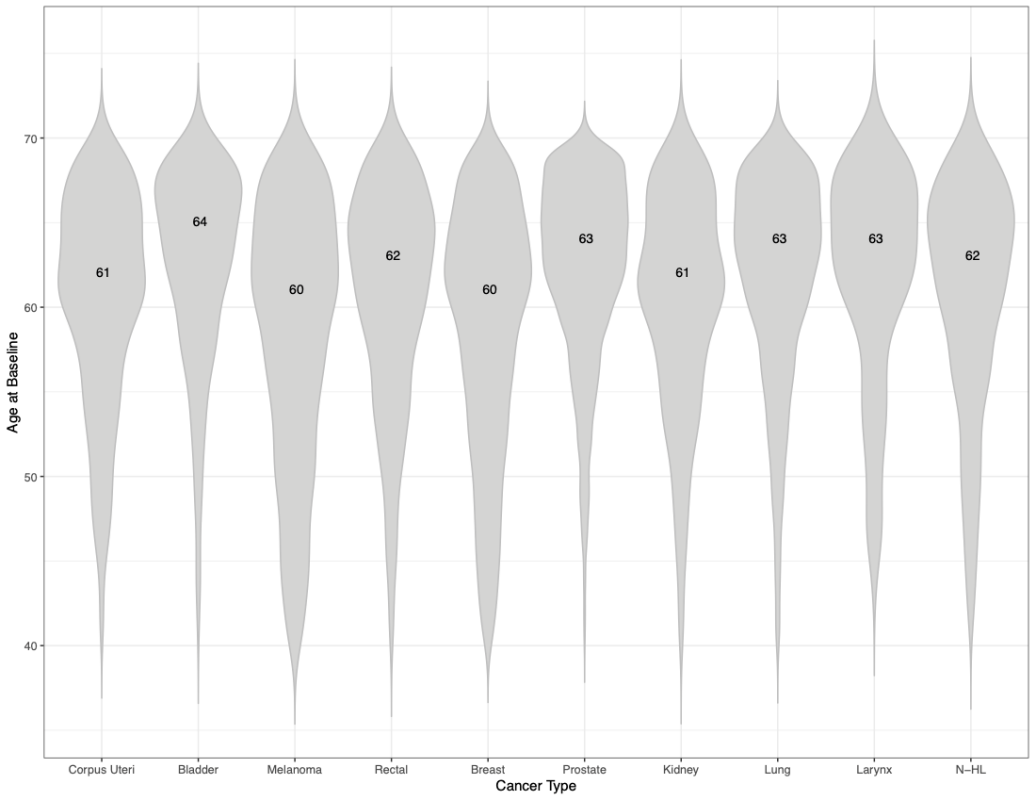


Supplementary Figure A.7.2 Distribution of Somatic CHIP Mutations Without (Blue) and With mCAs (Blue).

CHIP: clonal hematopoiesis of indeterminate potential, mCA: mosaic chromosomal alterations

A.7.3 Supplementary Figure C.2.3

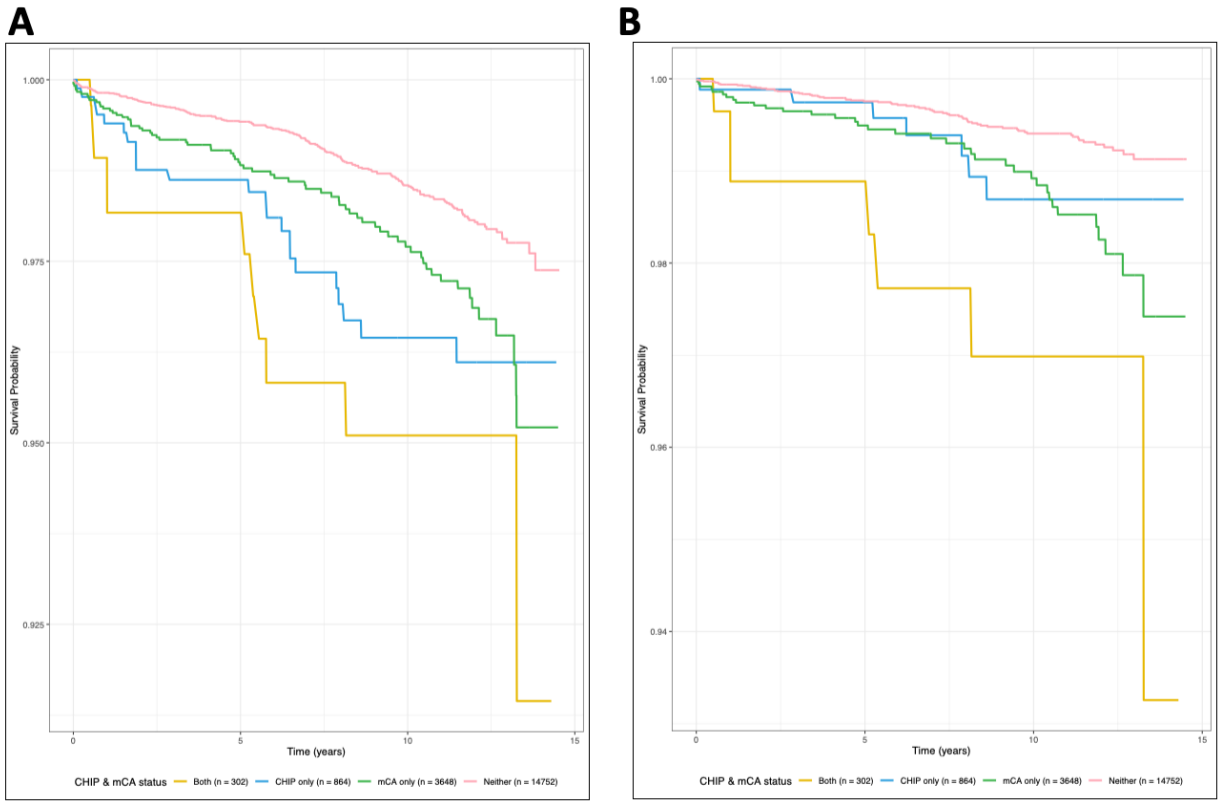




Supplementary Figure A.7.3 Proportion of Individuals With CHIP Mutations (Blue) Enriched With mCAs (Pink) According To Cancer Types (A) With Corresponding Age Distribution Violin Plots (B).

Age-adjusted P-values for carriers of both CHIP and mCA was $P=9.9\times10^{-01}$ for corpus uteri, $P=2.1\times10^{-01}$ for rectal cancer, $P=2.7\times10^{-03}$ for melanoma, $P=2.9\times10^{-07}$ for breast cancer, $P=3.0\times10^{-02}$ for kidney cancer, $P=1.5\times10^{-08}$ for prostate cancer, $P=1.1\times10^{-01}$ for bladder cancer, $P=7.9\times10^{-04}$ for lung cancer, $P=5.8\times10^{-01}$ for non-Hodgkin lymphoma, and $P=4.3\times10^{-01}$ for larynx cancer. CHIP: clonal hematopoiesis of indeterminate potential, mCA: mosaic chromosomal alterations, N-HL: non-Hodgkin lymphoma

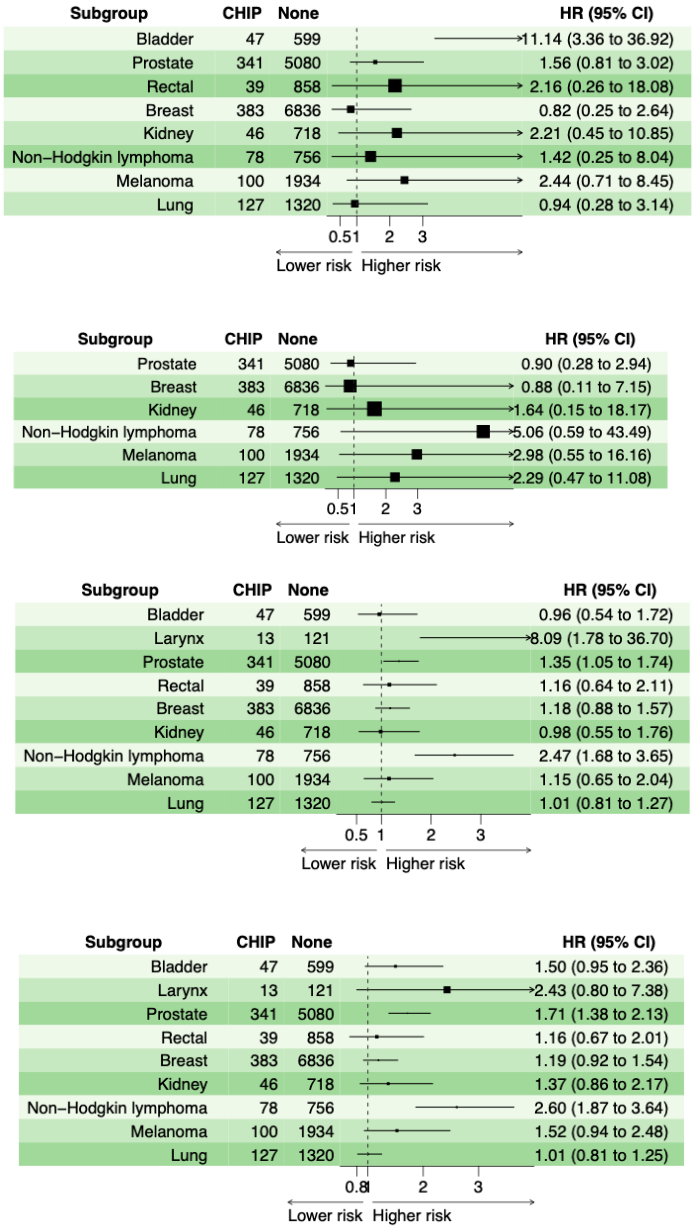
A.7.4 Supplementary Figure C.2.4



Supplementary Figure A.7.4 Kaplan-Meier CVD Mortality-Free (A) and CAD Mortality-Free (B) Survival Curves Stratified According To CHIP and mCA Status Where Individuals With Both Types of CH Are Depicted in Yellow, With CHIP Only in Blue, With mCA Only in Green, And Without Any CH in Pink.

Log-rank P-value comparisons for CVD mortality-free survival, CHIP+mCA vs. CHIP only (P=0.825), CHIP+mCA vs. mCA only (P=0.293), CHIP+mCA vs. none (P=0.023), CHIP only vs. mCA only (P=0.227), CHIP only vs. none (P=0.002), mCA only vs. none (P<0.001). Log-rank P-value comparisons for CAD mortality-free survival, CHIP+mCA vs. CHIP only (P=0.131), CHIP+mCA vs. mCA only (P=0.019), CHIP+mCA vs. none (P<0.001), CHIP only vs. mCA only (P=0.655), CHIP only vs. none (P=0.035), mCA only vs. none (P<0.001). CAD: coronary artery disease, CH: clonal hematopoiesis, CHIP: clonal hematopoiesis of indeterminate potential, CVD: cardiovascular disease, mCA: mosaic chromosomal alterations

A.7.5 Supplementary Figure C.2.5



Supplementary Figure A.7.5 Adjusted Hazard Ratios of CHIP Vs. No CHIP For the Risk of Death From CVD Causes (top), the Risk of Death From CAD Causes (2nd), the Risk of Death From Cancer Causes (3rd), and the Risk of Death From Any Cause (bottom) According To Cancer Types.

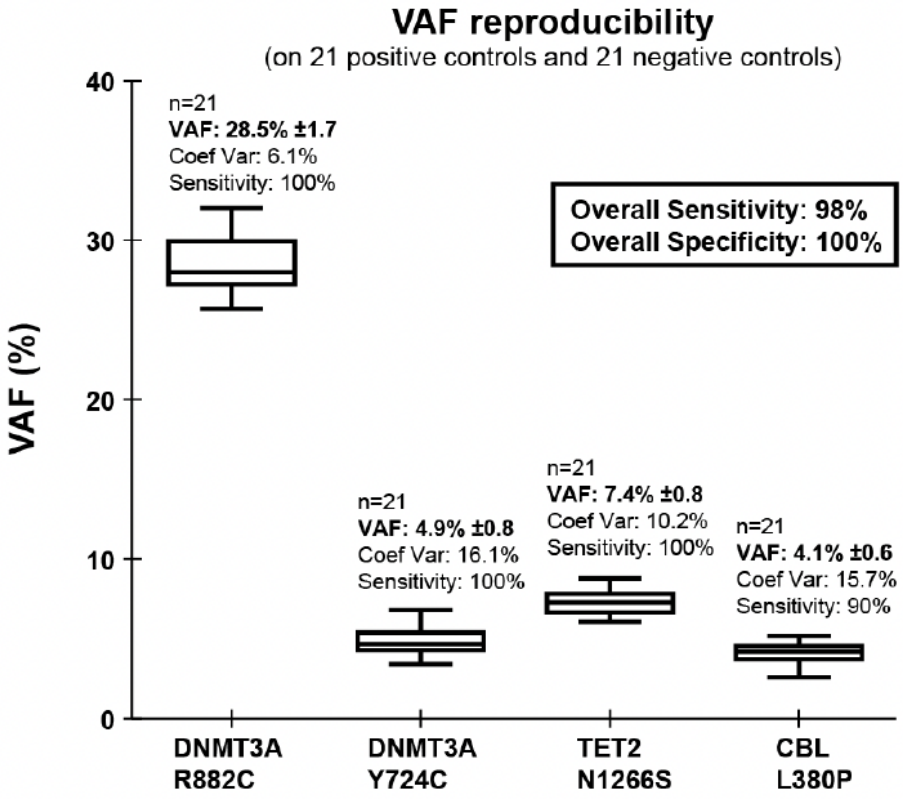
All models adjusted for age at baseline, sex, smoking status, receipt of chemotherapy, receipt of radiotherapy, prevalent CVD, number of days between date of recruitment and date of cancer diagnosis, and genetic principal components 1-10. For cancers of the corpus uteri, the breast, and of the prostate, no adjustment was made for sex. CAD:

coronary artery disease, CH: clonal hematopoiesis, CHIP: clonal hematopoiesis of indeterminate potential, CVD: cardiovascular disease

Supplementary Files For Chapter 5

A.8 Supplementary Figures

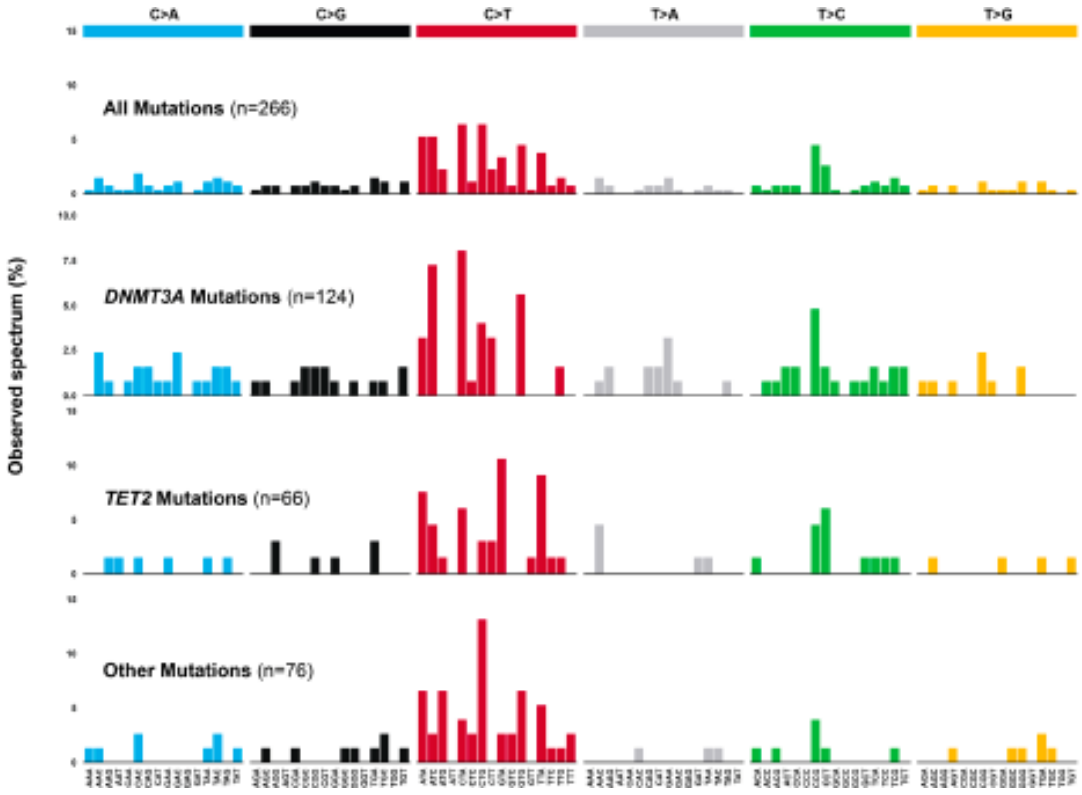
A.8.1 Supplementary Figure D.1.1



Supplementary Figure A.8.1 Ampliseq-Panel Validation On Ion Proton.

Validation of the CHIP-panel was performed by sequencing independently 21 positive control libraries (Individual harboring 4 somatic mutations with VAF of 4.1, 4.9, 7.4 and 28.5%) and 21 negative control libraries (fetal cord sample with no somatic mutation in the tested genes) and showed no false positive (100% specificity) and 2/84 false negatives (98% sensitivity). The sequencing was performed on an Ion Proton instrument (Thermo Fisher Scientific) at >6000x coverage (95% >500x). Variant calling, annotation and filtering was performed using the Ion Reporter software (Thermo Fisher Scientific). VAF; Variant Allele Fraction. Coef Var; Coefficient of Variation.

A.8.2 Supplementary Figure D.1.2



Supplementary Figure A.8.2 CHIP-associated mutation signature.

Trinucleotide context of single nucleotide substitutions present in all, DNMT3A, TET2 or others genes. (Mutalisk, http://mutalisk.org/)

A.9 Supplementary Tables

A.9.1 Supplementary Table D.2.1

Supplementary Table A.9.1 Custom Ampliseq CHIP-Panel Coverage

Type	Name	Chromosome	Chr_Start	Chr_End	Num_Amplicons	Total_Bases	Covered_Bases	Missed_Bases	Overall_Coverage	aa missed	poor coverage (<100x)
GENE	ASXL1	chr20	.	.	33	4762	4695	67	99%	1-19	
GENE	CBL	chr11	.	.	22	2881	2762	119	96%	1-8; 621-647	
GENE	DNMT3A	chr2	.	.	32	3104	3005	99	97%	60-90	646-661
GENE	GNAS	chr20	.	.	29	4161	3641	520	88%	347-464	540-606
GENE	GNB1	chr1	.	.	11	1113	1113	0	100%		
GENOME_REGION	JAK2	chr9	5073674	5073808	1	134	134	0	100%		
GENE	PPM1D	chr17	.	.	13	1878	1682	196	90%	36-101	
GENOME_REGION	SF3B1_14	chr2	198267280	198267550	2	270	270	0	100%		
GENOME_REGION	SF3B1_15	chr2	198266709	198266854	2	145	145	0	100%		
GENOME_REGION	SF3B1_16	chr2	198266466	198266612	2	146	146	0	100%		
GENE	SRSF2	chr17	.	.	5	686	686	0	100%		
GENE	TET2	chr4	.	.	37	6188	6188	0	100%		
GENE	TP53	chr17	.	.	14	1383	1307	76	95%	33-56	