Evaluation of the association of Aging-Related Immune phenotypes (ARIP) with aging outcomes using transcriptomics.

A Dissertion

SUBMITTED TO THE FACULTY OF THE UNIVERSITY OF MINNESOTA

BY

Ramya Ramasubramanian

IN PARTIAL FULLFILMENT OF THE REQUIREMENTS FOR THE DEGREE

IN DOCTOR OF PHILOSOPHY

Advisor: Dr. Bharat Thyagarajan

Co-advisor: Dr. David Jacobs

SEPTEMBER 2022

Acknowledgements

I would like to express my profound gratitude to everyone who supported me through my Ph.D. and made this achievement possible. First, I would like to acknowledge my advisor, Dr. Bharat Thyagarajan for providing me exceptional and unwavering support through my journey as a Master's and doctoral student. He has always been available to answer all my questions, provided me with the right resources to work with complex biological data which I was not very familiar with and gave great guidance, both professionally and personally.

I would also like to acknowledge my other dissertation committee members, Drs. David Jacobs, Weihua Guan and Weihong Tang. Dr, David Jacobs helped me think deeply about my methods and results which helped me become a better academic. Dr. Weihua Guan has shared his biostatistical expertise with me and given me enlightening comments for my thesis work and other projects. Dr. Weihong Tang has provided me with helpful feedback which greatly improved my thesis work.

I want to thank my wonderful co-workers Helene Barcelo, Sithara Vivek, DeVon Hunter-Schlichting and Zach Flaten for supporting me throughout my journey. Helene has always looked out for me and helped me with understanding some tricky laboratory work. Sithara has been my sounding board to discuss my thoughts and helped me shape them into more concrete ideas. I am thankful to DeVon, Zach and Helene for creating a very supportive and fun work environment when I first joined the lab, and also for our regular catch-up sessions.

I would like to thank all my friends who have been through this journey along with me. Vaishnavi Soundararajan, Subhash Sethumurugan, Karthik Srinivasan, Venkat Ram Subramanian, and Saradha Peruboina have been supporting and cheerleading my journey from the beginning of my graduate studies. I would like to thank Bhargavi Chandrasekar and Haritha Naveen who have been my long-distance support system throughout my graduate school journey.

Finally, I would like to express my deepest gratitude to my family and friends. I would like to thank my parents, Sugandhi and Ramasubramanian, for being my strongest supporters and for always putting my education above all else. I would like to express my endless appreciation to my husband, Abhinav VS, for his unconditional love, patience, and unwavering support and for always cheering me up when I need it. Finally, I would like to thank my friends and extended family for their support and encouragement throughout my journey.

Abstract

Individuals with the same chronological age exhibit different degrees of physiological dysfunction. Biomarkers that determine the underlying biological processes of age acceleration in some individuals compared to others will help understand the mechanisms that determine the dysregulation of homeostatic processes.

Senescence of the immune system characterized by an aging-related immune phenotype (ARIP) is being increasingly studied as a biological aging phenomenon. In *Manuscript 1*, we defined an ARIP marker in the Health and Retirement Study (HRS) using T-cell subsets by comparing associations between candidate ARIP markers and existing ARIP markers with multiple age-related phenotypes including multimorbidity (defined using type II diabetes, heart disease, lung disease, stroke, and cancer). We found that CD4+ naïve T (T_N) cells and CD4+ naïve (T_N) /memory (T_M) T cells ratio had the strongest association with the multimorbidity.

In *Manuscript 2*, we defined an optimal cut-off point for the ARIP marker (CD4+ T_N) and aimed to identify gene expression profiles for this ARIP marker in HRS. Two hundred ninety-two genes were differentially expressed with the cut-off and 1482 genes were differentially expressed with the continuous ARIP marker. However, after adjustment for multiple comparisons, there were no statistically enriched biological pathways among the differentially expressed genes.

In *Manuscript 3*, we aimed to identify gene expression profiles that may mediate the association of ARIP with multimorbidity and mortality. We first assessed genes associated with both ARIP and multimorbidity. We identified 1998 genes associated with

multimorbidity and 93 genes commonly associated with four of the individual chronic conditions. Five common genes (*AMIGO1, ZSCAN32, KCNK12, KRT73* and *MTRNR2L4*) were identified between ARIP and multimorbidity which were also associated with mortality. Two of the genes, *AMIGO1* and *KRT73*, partially mediated the association of CD4+ T_N cells on mortality.

Overall, this dissertation provides evidence of immune aging markers and certain genes associated with physiological dysregulation which may motivate future work for a comprehensive understanding of aging mechanisms.

Contents

List of figures

List of tables

Chapter 1 : Background

A. Executive Summary

The life expectancy in the United States has increased from 76.8 years in 2000 to 78.7 years in 2018 (1). The increase in the Healthy life expectancy (HALE), defined by the number of years spent free of disease and injury, from 2000 to 2019 has been six times lower than the increase in life expectancy (2). The number of American adults aged 65 years and older is projected to nearly double from 49 million in 2016 to 95 million in 2060 (3). The increasing aging population increases the burden of chronic age-related diseases and the prevalence of multiple chronic conditions in individuals. This necessitates a better understanding of aging mechanisms and pathways and the discovery of aging biomarkers to identify individuals with higher susceptibility to age-related diseases.

Aging-related immune phenotype (ARIP) or the age-related decline in the immune system is a global phenomenon observed in most of the common age-related diseases (4,5). Immune aging limits the ability of the immune system to mount effective immune responses against new infections, vaccinations, and antigens (6,7). Age-related decline in the immune system has been observed in age-related diseases such as type II diabetes, kidney disease, cardiovascular disease, and cancer (8,9). Despite broad agreement among biologists regarding the effect of the aging immune system on morbidity and mortality, an ARIP marker that has been associated with physiological dysregulation has not been studied before.

Multimorbidity or the occurrence of multiple chronic diseases is an aging condition and has an increased prevalence among older adults (10–12). The prevalence of multimorbidity among individuals aged ≥ 65 years is projected to increase from 54% in 2015 to 67.8% in 2035 according to a longitudinal study performed in England (13). The occurrence of multiple chronic conditions points to the dysfunction of several homeostatic mechanisms which prevent the development of disease, prevent functional decline, and preserve biochemical balance. The public health burden of multimorbidity and the integrated activity of molecular mechanisms in the development of multiple diseases in older individuals requires the development of aging biomarkers.

The emerging field of geroscience aims to understand how aging processes enable diseases at a molecular level and identify molecular mechanisms of aging which increase disease incidence (14). Transcriptomics, the study of the whole transcriptome, can be utilized to discover gene expression signatures that can explain the underlying mechanisms of aging. The transcriptome is dynamic and responds to changes in environmental and systemic stressors, allowing us to study altered physiological functioning and altered activity in age-related biological mechanisms. This dissertation will focus on identifying an immune aging signature by utilizing ARIP as an aging biomarker and multimorbidity as an aging outcome. The immune aging signature can be used to identify individuals at higher risk for diseases and individuals who warrant closer monitoring for prevention strategies.

The first manuscript of the dissertation focused on identifying ARIP profiles associated with biological aging utilizing the immunophenotyping data in the Health and Retirement Study. Candidate markers of ARIP were developed and a marker for ARIP

was chosen by comparing associations of the candidate and existing ARIP markers with chronological age, biological age, and multimorbidity. The second manuscript focused on identifying gene expression profiles of the ARIP marker and identifying enriched biological pathways of ARIP. In the third manuscript, we identified gene expression profiles and enriched biological pathways of multimorbidity. Shared biological pathways between immune aging and multimorbidity will support the theory of a common aging signature. As gene expression is modifiable, identifying an aging signature can potentially be used to modify or slow the occurrence of diseases in older individuals.

B. Healthy Aging

Aging is a complex phenomenon and is a major risk factor for most chronic diseases and mortality. The United States has a fast-growing aging population and the US Census Bureau estimates that the number of Americans aged 65 and older is projected to nearly double from 49 million in 2016 (15 percent) to 95 million in 2060 (23 percent) (3). With advancing medical care and technologies, life expectancy has increased from 76.8 years in 2000 to 78.7 years in 2018 (1). The Healthy life expectancy (HALE), which is the average number of years that a person can expect to live in "full health" calculated by taking into account years lived in less than full health due to disease and/or injury has increased from 65.8 years in 2000 to 66.1 years in 2019 (2). Life expectancy is increasing about 6 times faster than the increase in HALE and hence addressing healthy aging is necessary. Due to the increasing aging population and the increasing number of years that individuals spend with disease and disability, understanding aging processes and devising prevention strategies to increase the number of disease-free years will be helpful.

The objective of healthy aging is to reduce the number of years that individuals spend with disease and disability to improve their quality of life (15). Chronological age is a major risk factor for the most common health-related diseases but is not a sufficient and all-encompassing marker for susceptibility to diseases. As people can age differently from one another, biomarkers which determine the underlying biological processes of acceleration of aging in some individuals compared to others will be helpful in understanding aging(16,17). Biomarkers of healthy aging can be utilized as surrogate endpoints or outcome measures in trials of interventions designed to extend health span and can be used for risk stratification to identify individuals at risk for accelerated aging.

Developing therapeutics and diagnostics for specific diseases has proven a successful endeavor in studying morbidity and chronic diseases. However, addressing common biological mechanisms and the physiological nature shared between chronic aging diseases will be helpful to develop global markers for aging-related chronic diseases. The trans-NIH geroscience interest group (GSIG) is focused on identifying strategies to develop a systematic understanding of aging mechanisms and elaborate mechanistic links between aging and chronic diseases. Seven pillars of aging mechanisms were described by this group including adaptation to stress, epigenetics, inflammation, macromolecular damage, metabolism, proteostasis, and stem cells and regeneration (14).

Immune mechanisms are increasingly being studied as a biological phenomenon in the aging process and the development of age-related conditions. Over the course of life, the constant interaction between environmental stressors and systemic changes in the body leads to accumulated changes in the immune system, contributing to an aging-related immune phenotype (ARIP) (4–6). These changes have been studied in the context

of age-related diseases such as type II diabetes, cancer, cardiovascular diseases, and rheumatoid arthritis (8,9,18). Multimorbidity, defined as the occurrence of two or more chronic conditions in individuals, is an outcome of the manifestation of the pace of biological aging and the development of pathologies that increase susceptibility to diseases. The presence of multimorbidity is associated with outcomes including reduced quality of life, reduced functional status, physical disability, and mortality (19,20).

This dissertation focuses on ARIP as a biomarker of dysfunctional immune function that progresses with aging and multimorbidity as an adverse outcome that results from the accumulation of aging-related physiological dysfunction. The biological mechanisms and pathways which underlie immune aging and multimorbidity were investigated by utilizing gene expression signatures.

C. Aging-related immune phenotypes (ARIP)

One of the central pillars that accompanies aging in individuals is the changes in the immune system as they get older. The phenomenon of aging is accompanied by the dynamic changes in the immune system due to environmental interactions and systemic changes in the body, contributing to an Aging-related immune phenotype (ARIP). The immune changes accompanied by aging reduce the ability of the immune system to mount effective immune responses against new infections, vaccinations, and antigens and contribute to age-related diseases such as cardiovascular diseases, cancer, lung diseases, and type II diabetes. Immunological profiles including the Immune risk profile (IRP) have been described previously and point to cytomegalovirus (CMV) infection as a major driving force behind the immunological profiles in older individuals (21,22).

C.1. Determinants of the immune system

The human immune system exhibits variation between individuals through different functional states exhibited by a range of cell types in the body. The stability of the immune system within individuals is over a week to months in healthy individuals (23,24). Immune responses to acute stimuli such as vaccinations demonstrate changes involving expansion in cell populations but the immune system returns to the baseline state similar to before the vaccination (25). The stability of intra-individual variation can help us determine the inter-individual variability of the immune system. The determinants can be classified as (i) Intrinsic factors including age and sex; (ii) Genetic factors and (iii) Environmental factors (as shown in figure 1-1).

1. *Intrinsic factors*

(a) Age: With advancing age, the immune system undergoes remodeling and decline with an impact on health and survival (7,26). Immune system changes with age are characterized by a decrease in the number of naïve cells and an increase in the number of cells with memory phenotype (4,5,27,28). The interaction between the immune system and the environmental stressors including various microorganisms plays a role in shaping the aging human immune system by an increase in memory cell types and inducing an environment of low-grade inflammation by increased release of pro-inflammatory cytokines.

(b) Sex: Immune disorders such as autoimmune diseases have different prevalence rates between men and women. Diseases such as Systemic Lupus erythematosus, Sjogren syndrome, and autoimmune thyroid disease are more common among women(29) whereas ankylosing spondylitis is higher among men (30). Women also develop higher

antibody titers after vaccination (31). Women have higher CD4+ T cells and CD4/CD8 T cell ratios compared to men (32). Consistent differences in cell counts are not observed between females and males over the course of life suggesting sex playing a major role in disease outcome compared to the immune response.

(ii) Genetic factors

Several studies have found an influence of genetic variants in the inter-individual variation in immune cell distributions (33,34). The ImmVar project designed to assess variability in functional responses in the immune system between individuals found that 22% of the overall variation in gene expression profiling of the immune cells was attributable to heritable factors (35).

(iii) Environmental factors

The most important environmental factor that determines immune system variability is the chronic persistence of pathogens including cytomegalovirus (CMV) infections. Though CMV is thought to be a latent virus, it is known to reactivate regularly, changing the composition of T cells, and increasing the number of CMV-specific T cells (36). CMV seropositivity was associated with increased immune diversity in a study conducted among monozygotic twins, especially in T cell populations (37). Environmental factors including smoking have been associated with increased total leukocyte counts (38) and reduced Natural Killer (NK) cell functional activity (39). Obesity and increased adiposity have also been known to influence immune responses and increase pro-inflammatory cytokines production (40,41).

C.2. Characteristics of ARIP

Naïve T cells (CD45RA+CCR7+) (T_N) are initially produced in the thymus and undergo positive and negative selection in the thymus before they migrate to the periphery and form the naïve T cell pool. The primary mediator of activation in T cells is the T cell receptor (TCR) which defines the specificity of each T cell. Functionally, naïve T cells expand into effector and memory T (T_M) cells upon antigen encounter in the periphery (42,43). A major change in the immune system consistently observed with aging is the reduction of naïve T cells. The reduction in naïve T cells is due to thymic involution and chronic antigenic stimulation. Thymic involution, which is the shrinkage of the thymus, is a consequence of aging where the thymic epithelial space is reduced by approximately 90% by the time a person reaches 50 years of age (44). Thymic involution reduces the naïve T cell production in the body. Chronic antigenic stimulation is the persistent infections of pathogens such as viruses that exist for longer durations and are not cleared by the immune system. Many viruses such as Epstein-Barr virus (EBV), Cytomegalovirus (CMV), and other herpes viruses are latent viruses that persist for the lifetime of the host (45). CMV has been studied extensively in this context and episodes of CMV reactivation in the body can lead to the differentiation of naïve T cells into effector and memory T cells. Among elderly individuals, the number of CD4+ T cells specific to CMV increased from a mean of 2.2% below the age of 50 to a mean of 4.7% aged over 65 years, at the expense of the reduction of naïve T cells (46).

Memory T cells (T_M) are divided into *Central Memory T (T_{CM}) cells (CCR7+CD45RA-), Effector Memory T cells (T_{EM}) (CCR7-CD45RA-),* and *Effector T cells (CCR7-CD45RA+) (T_{EFF})*. Upon encounter with a cognate antigen, naïve T cells

proliferate and differentiate into effector cells which migrate to the peripheral tissues and damaged sites to facilitate the removal of the targets. After clearing the antigen, more than 95% of the effector cells die and a small pool of antigen-specific T cells are alive forming the T_M cell pool. T_{CM} produces high amounts of IL-2 and low levels of effector cytokines such as IFN-γ, IL-4 and IL-5, and does not have an immediate killing function (28). T_{EM} produces IFN-γ, TNF, IL-4, and IL-5, and also contain preformed perforin granules for an immediate effector/killing action. T_{CM} is a less differentiated progenitor which can produce more differentiated cells upon encounter with its specific antigen (47). Hence, T_{CM} is said to maintain its memory status in a stem cell-like fashion. T-cell population imbalance with aging leads to an increase in memory cells due to chronic antigen stimulation and differentiation of T_N to T_M (48).

In a previous study published using the HRS immune cells data, there is demonstrated evidence that age was the more important factor that determined reduction in T_N cells and CMV was the factor that promoted changes in T_M cells(49). The study showed that CD4+ and CD8+ T_N decreased with age and CD4+ T_N cells were substantially higher among CMV seronegative individuals and women. We observed that CD8+ T_{EFF}, CD4+ $T_{EFF,}$ and CD4+ T_{EM} cells were strongly influenced by CMV seropositivity, and CD8+ T_{EM} increased with age. CD8+ and CD4+ T_{CM} decreased with age although CD4+ T_{CM} was higher among women compared to men and CD8+ T_{CM} was higher among men. These results show that the inter-individual variability in T cell subsets is determined by several factors such as age, sex, and CMV serostatus and individual T cell subsets are differentially influenced by the various factors(49).

Despite broad agreement among biologists as to the changes in immune cell distribution that occurs with aging, defining an ARIP measure in terms of physiological dysregulation has not been done previously. Several population studies have reported ARIP defined by a reduction of T_N and an increase in T_M cells. A previous study demonstrated that the reduction of T_N cells was observed in both CMV seropositive and CMV seronegative individuals indicating that increase in age is the important factor for naïve T cell reduction.(50). T_{EM} and T_{EFF} cells have been shown to increase with age in many studies but there is a substantial difference of increase of these cell types between CMV-seropositive and CMV-seronegative individuals (50,51) . T_{CM} has been shown to increase with age in a study conducted among 101 healthy individuals aged between 5 to 96 years however there was a decrease in T_{CM} among individuals who were older than 75 (27). Several attempts to define ARIP by using measures including CD4/CD8 ratio have been proposed previously but these measures do not account for the accumulated changes in T_N and T_M subsets. Thus, other nuanced elements of immune aging may need to be incorporated to obtain a reliable and accurate measure of ARIP that is relevant in human populations and is associated with physiological dysregulation. Optimal cut-off points to identify immunosenescent individuals have not been defined for ARIP markers previously as well. Defining a threshold may be helpful with biological systems including immune systems and gene expression as larger changes in these systems may have a substantial effect on physiological functioning. Previous work in prostate cancer used outlier genes overexpressed in prostate cancer cases using a method called cancer outlier profile analysis (COPA) to identify novel gene fusions and the subsequent impact of these extremely overexpressed genes on clinical outcomes(52), showing that larger

changes in genes may contribute to substantial physiological effects. Hence, in this dissertation, we have attempted to define a cut-point for the ARIP marker. Immunophenotyping data in HRS is ideal for creating a measure of ARIP as the immune cells were measured using a standardized protocol published by the Human Immunology Project and among a large sample size of about 9000 individuals.

C.3. Immune aging in age-related diseases

Studying ARIP as a biomarker is imperative as the aging immune system is observed in several age-related diseases. The most common age-related diseases have been known to share some of the ARIP features such as the expansion of T_{EFF} cells, decrease in naïve T (T_N) cells, and increased CMV serology but data on the effect of ARIP on age-related disease remains sparse.

A main consequence of the aging immune system is the reduction of vaccine effectiveness as older individuals are not able to induce effective immune responsiveness to vaccines. A study performed to measure serum IgM and IgA after administering pneumococcal vaccines among individuals of different ages found that the humoral responses were significantly impaired in the older group of individuals(53). Another study found that after administering inactivated influenza vaccine, the production of vaccine-specific antibodies including plasmablast-derived polyclonal antibodies was lower in the elderly compared to young adults(54). Other vaccines including the Japanese Encephalitis vaccine(55), yellow fever vaccine(56), and herpes zoster vaccine(57) have been shown to have impaired responses in the elderly compared to younger individuals affirming that the immune system dysfunctions with age.

Circulating CD4+ T_{EM} was associated with atherosclerosis in carotid and coronary vascular disease in a study conducted among 313 individuals, and the association between T_{EM} and carotid atherosclerosis was independent of cardiovascular risk factors, indicating the relevance of adaptive immunity in cardiovascular disorders (58). Late differentiated CD4+CD28-T cells are also present in acute coronary events (59). In a meta-analysis performed to assess the relationship between CMV seropositivity and risk of future cardiovascular events, CMV seropositivity was associated with a 22% increase in the risk of cardiovascular diseases (60). In a study performed with seventy-one hypertension patients, CD8+ T cells were found to be higher among hypertension patients compared to controls (61).

In the Multi-Ethnic Study of Atherosclerosis (MESA), among 929 participants, each standard deviation higher CD4+ memory T cells (defined by $CD4^{+}CD45RO^{+}$) was associated with 21% higher odds of type 2 diabetes (T2D) and one standard deviation higher CD4+ naïve T cells (defined by $CD4^{+}CD45RA^{+}$) was associated with 22% lower odds of type 2 diabetes (T2D) cross-sectionally after adjustment for age, gender, race/ethnicity, and BMI(62). However, these results could not be replicated in the longitudinal associations between immune cell types and incident T2D over a follow-up of 9.1 years in MESA(63). Among 109 T2D patients in the SUMMIT cohort study, T_{EM} (defined by CD4+ CD45RO+CD62L-) and T_{CM} (defined by CD4+ CD45RO+CD62L+) were elevated and naïve T cells (defined by CD4+ CD45RO-CD62L+) were reduced compared to 89 non-T2D participants(64). Multiple studies have also reported higher CMV seroprevalence among T2D patients, indicating an important role of CMV in the development of diabetes (65–67).

Some features of accelerated immune aging such as decreased thymic functionality, and expansion of late differentiated effector T cells (defined by CD4+CD28- or CD8+CD28-) have been observed in rheumatoid arthritis (68). Senescent T cells (defined by CD27-CD28-) have been shown to be higher in number among women recently diagnosed with breast cancer in parallel with a reduction of recent differentiated T cells (CD27+CD28+) (69). The expansion of CD8+CD28− T cells has been observed in women with breast cancer during chemotherapy and in lung cancer (70,71).

ARIP is associated with several age-related diseases indicating immune aging as an important aging mechanism. Studying ARIP in multiple age-related disease outcomes simultaneously would allow us to determine underlying aging mechanisms and identify aging signatures. ARIP can be used as an aging biomarker which is informative for predicting disease and providing insights into disease etiologies. In this dissertation, we will focus on identifying aging signatures for multimorbidity by utilizing ARIP as an aging biomarker.

D. Aging and gene expression

Transcriptomics is the study of a complete set of transcripts in a cell and the level of their expression in a cell. The transcriptome of a cell includes all the RNAs transcribed from that particular cell at a certain functional or developmental stage, including messenger RNA (mRNA), non-coding RNA (ncRNA), and small RNAs. RNA-sequencing technology is a powerful deep sequencing technology where a population of RNA fragments is converted to a library of cDNA fragments which are sequenced in a high-throughput manner to obtain short sequences of reads (typically between 30-400

bp). Following sequencing, the reads are aligned to a reference genome to obtain the level of expression of each gene expressed as counts of the genes.

Transcriptome has previously been used to study aging. A whole blood gene expression meta-analysis study performed on 14,983 individuals of European ancestry identified 1497 genes that are differentially expressed with chronological age and the gene expression profiles were used to estimate a "transcriptomic clock" of individuals. Individuals with higher transcriptomic age compared to chronological age had higher blood pressure and total cholesterol levels (72). Several previous studies have also identified aging gene expression profiles but have been limited by small sample sizes and the non-generalizability of samples (73–75).

The transcriptome is dynamic and responds to changes in environmental and systemic stressors. Hence, the transcriptome allows us to study the altered physiological functioning and altered activity in age-related biological mechanisms. Identification of specific genes and coherent gene sets representing a specific biological function or mechanism associated with immune aging and health outcomes including multimorbidity and mortality in humans will provide key insights into the pathways mediating the association of immune aging with multimorbidity and mortality. Utilizing gene expression may provide an aging biomarker signature to monitor susceptibility to age-related diseases.

E. Multimorbidity and mortality

With increasing life expectancy in the United States and globally, the number of people living with multiple chronic conditions and well into old age is increasing as well. Multimorbidity is defined as the presence of two or more chronic medical conditions or

diseases (76). According to a report by RAND, 81 percent of adults aged 65 years and older had multiple chronic conditions from 2008 to 2014 (10). Data from the 2018 National Health Interview Study (NHIS) estimated that 27.2% above 18 years had multiple chronic conditions (11). According to a study performed among individuals from three cohorts (Understanding Society, the English Longitudinal Study of Ageing, and the Cognitive Function and Ageing Study II) in England, the prevalence of multimorbidity among individuals aged ≥ 65 years is projected to increase from 54% in 2015 to 67.8% in 2035 (13). The rate of mortality has been increasing in the United States as well from 715.2 deaths per 100,000 standard population in 2019 to 835.4 deaths per 100,000 standard population in 2020, and the life expectancy in 2020 was 77.0 years which reduced by 1.8 years from 2019(77,78).

The prevalence of multimorbidity increases progressively with age (79,80) due to the unbalanced functioning of biological mechanisms which maintain homeostasis in the human body. Multimorbidity is associated with greater disability than expected, based on the disability observed for chronic conditions individually (12). Patients with multimorbidity require greater medical care including annual visits to a higher number of physicians and are more likely to suffer from adverse drug effects due to the higher number of medications consumed (12). Reduced functional status including aspects of physical function, cognitive function, daily activities, and reduced quality of life is observed among multimorbidity patients (80). An association between multimorbidity and mortality with a hazard ratio of 1.44 (95%CI: 1.34- 1.55), compared to no multimorbidity, was detected in a meta-analysis. Due to the overwhelming evidence of the healthcare burden of multimorbidity, it will be helpful to develop a broader approach

to healthcare for patients with multimorbidity in addition to a specialized and disease-specific approach currently practiced in medical settings.

Multimorbidity can be studied as an outcome of the biological aging process as it represents a manifestation of fading homeostatic mechanisms which prevent the development of diseases, prevent functional decline, and preserve biochemical balance. The momentum of biological aging and the development of subclinical pathologies are chief factors that increase the propensity to disease. The rate of accumulated changes in the aging process due to environmental stressors and systemic changes in the body such as gene expression changes are expressed as multimorbidity. The emerging field of geroscience aims to understand how aging processes enable diseases at a molecular level and how the molecular mechanisms of aging are connected to the increasing incidence of diseases. The goal is to develop interventions to delay the incidence of multiple chronic diseases and conditions. In line with the approach of studying multimorbidity at a molecular level, it is imperative to identify aging biomarkers that characterize multimorbidity. The mortality rate also increases with higher age and is a feature of biological aging due to accumulated dysfunctional biological structures and systems leading to mortality.

A limited number of studies have assessed the association between immune aging biomarkers and multimorbidity. In a nine-year long longitudinal study of 1018 participants aged 60 years and older, multimorbidity was associated with higher IL-6, IL-1ra, TNF-α receptor II (TNFAR2), and lower dehydroepiandrosterone sulfate at baseline, independent of age, sex, BMI, and education. Further, higher IL-6 levels and steeper increase of IL-6 levels were associated with a steeper increase in multimorbidity over

time(81). In an investigation based on the ESTHER study with 2547 participants in Germany, C-reactive protein (CRP), oxidative stress markers (d-ROM), and metabolic conditions such as obesity were associated with multimorbidity independent of age, sex, education, smoking, alcohol consumption, and physical activity(82). Associations of another aging biomarker, telomere length, with multimorbidity were investigated in the Health and Retirement Study where telomere length was related to reduced likelihood of multimorbidity in men (OR=0.884, 95% CI: 0.628 to 1.246) and women (OR=0.972, 95% CI: 0.721 to 1.311) but the associations were not statistically significant(83). Previous studies have evaluated biomarkers in predicting all-cause mortality. A metabolomics platform was used to identify fourteen biomarkers that were independently associated with mortality associated in processes such as lipoprotein and fatty acid metabolism, glycolysis, fluid balance, and inflammation(84). Inflammation markers including IL-6 and CRP have also been found to be associated with all-cause mortality(85). However, specific mechanisms and biological pathways involving the association between immune aging and all-cause mortality have not been studied previously.

The currently available studies that examine biological factors associated with multimorbidity are not sufficient to explain the predisposition of individuals to multiple chronic conditions. Improved understanding of biological mechanisms, including immune agings, using gene expression which lead to multimorbidity and mortality will assist in developing personalized treatments and identifying individuals who are at higher risk for accelerated aging for early interventions.

F. Summary

The goal of healthy aging is to increase the number of days spent without the development of disease and disability among individuals. As individuals age differently from one another, studying the biological mechanisms underlying accelerated aging in some individuals compared to others is imperative. Multimorbidity is an aging outcome to investigate as it represents the increased pace of aging in individuals and represents the accumulation of dysregulation of biological processes and mechanisms. Immune aging or aging-related immune phenotypes (ARIP) is a widely studied aging phenomenon representing the accumulated changes, during normal aging, in the immune system such as reduction in naïve T cells (T_N) and increase in memory T (T_M) cells. The systemic changes manifested in the body due to the aging immune system need to be investigated as ARIP has been observed in the most common age-related diseases.

This dissertation (described in Figure 1-2) focused on 1) Developing a T-cell ARIP profile in the Health and Retirement Study using naïve T cells, memory T cells, age, and CMV seroprevalence (denoted by yellow arrows in Figure 1-2); 2) Investigate the systemic changes in the body due to ARIP by studying gene expression profiles of ARIP and the biological pathways represented by these gene expression profiles (denoted by brown arrows in Figure 1-2) ; 3) Identify genes and biological pathways which characterize multimorbidity, and examine shared biological pathways between ARIP and multimorbidity to identify a global immune aging gene signature (denoted by green arrows in Figure 1-2). The blue boxes in figure 1-2 indicate the main exposure and outcome variables in the three manuscripts. The grey boxes and arrows in figure 1-2

indicate the potential confounders for the association between the exposures and outcomes.

Figure 1-1: Determinants of the immune system

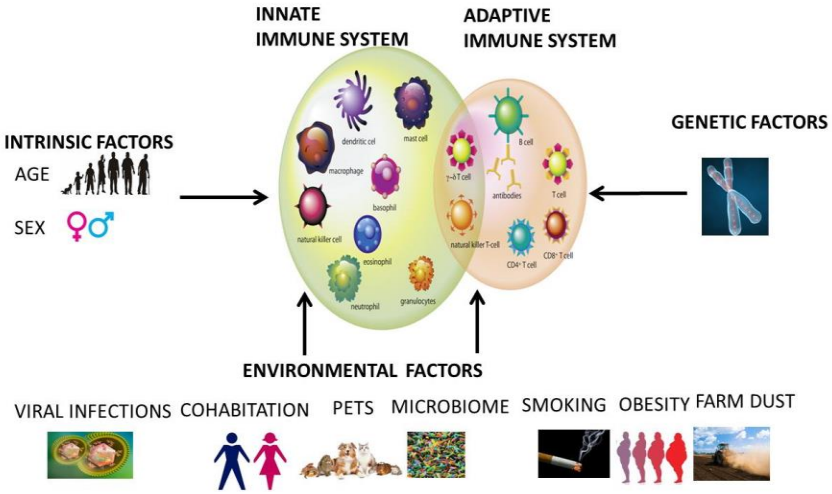
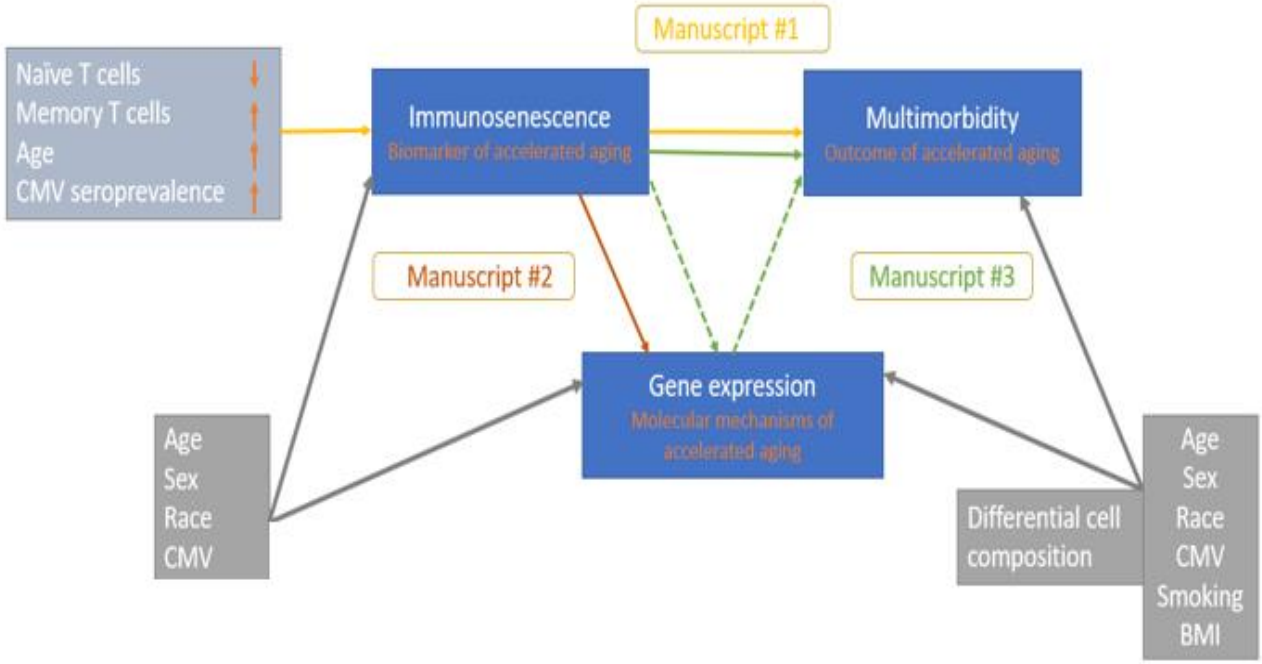


Figure 1-2: Conceptual model of the aims presented in this dissertation.



Chapter 2 : Study design and analytic approach

This dissertation constitutes three manuscripts that utilize data from Health and Retirement Study (HRS). The three manuscripts will utilize immunophenotyping data, gene expression data from RNA-sequencing, and multimorbidity data from HRS.

A. Study design of the Health and Retirement Study

The Health and Retirement Study is a nationally representative longitudinal survey with about 37,000 participants over the age of 50 supported by the National Institute of Aging (NIA). The survey, started in 1992, was established to provide a national resource for data on the changing health and economic circumstances associated with aging at both individual and population levels. HRS is based on a steady-state design where the sample is replenished every 6 years with younger cohorts not previously represented (86). The HRS sample employs a multi-stage area probability design involving geographical stratification and clustering and oversampling of African-American and Hispanic households at about twice the rate of Whites (87). Biennial data measurements are performed in HRS. The response rates for the HRS surveys have been high and the re-interview response rates have ranged from 86.1% to 89.1% till 2014 (88). A household screening interview is conducted with each sampled housing unit to determine eligibility and a primary respondent is selected from all the age-eligible household members. HRS has used a mixed-mode design for conducting interviews where biennially, half the sample is assigned a face-to-face interview with physical and biological measures and the other half of the sample completes the core interview over the telephone.

B. Venous Blood Study (VBS)

Adding to the existing wealth of information in the HRS, whole blood was collected from consenting participants in 2016 to provide a substantially fuller picture of the health of a representative sample of older people. About half of the HRS sample complete their interview face-to-face biennially, and consent for a venous blood draw was requested from all the participants who completed an HRS face-to-face interview during the 2016 wave (n=14287). The consent rate was nearly 80% for the blood collection resulting in whole blood being collected from about 10,500 participants. A biospecimen repository of serum, plasma, RNA, and DNA of the participants who have consented to future use of their samples within the broad aims of HRS to study health and aging is managed by the University of Minnesota (89). Immunophenotyping and measurement of gene expression using RNA-sequencing were performed using blood from participants in the VBS.

C. Ascertainment of multimorbidity

Multimorbidity as previously defined (90) in HRS will be used for this dissertation. Multimorbidity will be measured as a self-reported outcome from the 2016 and 2018 waves of the Health and Retirement Study. The multimorbidity measure has previously been defined in HRS as comprising five health outcomes (Stroke, Heart disease, Type II diabetes, Lung disease, and cancer). Stroke was measured by asking "Has a doctor ever told you that you have had a stroke?" to the participant. Heart disease was measured by asking "Has a doctor ever told you that you have had a heart attack, coronary heart disease, angina, congestive heart failure, or other heart problems?" to the participant. In 2016, a participant was considered diabetic if their blood glucose

measurement (performed in 2016) was >=126 mg/dL or self-report swallowing medication for diabetes or using insulin shots/pump. Diabetes was measured by asking "Has a doctor ever told you that you have diabetes or high blood sugar?" to the participant in 2018. Lung disease was measured by asking "Has a doctor ever told you that you have chronic lung disease such as chronic bronchitis or emphysema?" to the participant. Cancer was measured by asking "Has a doctor ever told you that you have cancer or a malignant tumor, excluding minor skin cancer?" to the participant. Multimorbidity is defined by a score of 0 to 5 based on the answer of the participant to the above questions and as a categorical variable with two levels (0,1 – No Multimorbidity; 2,3,4 and 5 – Multimorbidity).

Incident multimorbidity in 2018 is ascertained by defining a baseline study population with participants who had a multimorbidity score of 0 or 1 in 2016. Among the participants in this baseline population, incident multimorbidity is defined as a multimorbidity score of 2 or higher in 2018.

D. Analytical design

In Manuscript 1, the objective is to identify aging-related immune phenotype (ARIP) profiles in HRS. This will be performed by developing candidate measures of ARIP based on the role of chronological age and the role of biological mechanisms including CMV seropositivity in the immune cell subsets (naïve T cells, central memory T cells, effector memory T cells, and effector memory T cells re-expressing CD45RA). The detailed blood collection and immunophenotyping procedures are given in Manuscript 1 (Section B.6. and B.7.). To identify an ARIP biomarker, we compared associations of the candidate and existing ARIP markers with chronological age,

biological age, and multimorbidity defined by diabetes, stroke, lung disease, heart disease, and cancer, incident multimorbidity after a two-year follow-up and two-year mortality.

In Manuscript 2, we identify gene expression profiles for the ARIP marker identified in Manuscript 1. Gene expression levels were measured using RNA-sequencing and the detailed RNA-sequencing procedures are mentioned in Manuscript 2 (Section B.4.). To identify gene expression profiles, we perform associations between ARIP and gene expression levels using negative binomial models. The gene expression profiles associated with ARIP may constitute biological pathways or mechanisms which will be identified.

In Manuscript 3, we characterize the relationship between ARIP and multimorbidity by identifying global gene expression signatures. We first identify gene expression profiles and biological pathways for ARIP. We identify shared common genes between multimorbidity and ARIP. We also assess if the identified shared common genes mediate the relationship between multimorbidity and ARIP.

Chapter 3 : Evaluation of T-cell aging-related immune phenotypes in the context of biological aging and multimorbidity in the Health and Retirement Study

A. Introduction

Aging is a complex process accompanied by changes in the immune system which might result in the acceleration of biological decline and incidence of chronic diseases (91). Immune system aging reduces the ability of the immune system to mount effective immune responses against new infections, vaccinations, and antigens (4,6). Previous studies have consistently reported that aging-related changes in T cells (92–94) contribute to an aging-related immune phenotype (ARIP) that is characterized by a reduced T-cell repertoire, reduction of naïve T cells (T_N), and accumulation of memory (T_M) and effector T cells (5,95). Though the determinants of ARIP remain incompletely understood, age, sex, and exposure to cytomegalovirus (CMV) are major determinants of T cell subsets (27,32,36,96). In our previous study (97), we evaluated the associations of individual T-cell subsets measured in the Health and Retirement Study with chronological age. The study showed that CD4+ and CD8+ T_N decreased with age and CD4+ T_N cells were substantially higher among CMV seronegative individuals and women. We observed that CD8+ T_{EFF}, CD4+ $T_{EFF,}$ and CD4+ T_{EM} cells were strongly influenced by CMV seropositivity, and CD8+ T_{EM} increased with age. CD8+ and CD4+ T_{CM} decreased with age although CD4+ T_{CM} was higher among women compared to men and CD8+ T_{CM} was higher among men. These results show that the inter-individual variability in T cell subsets is determined by several factors such as age, sex and CMV serostatus and individual T cell subsets are differentially influenced by the various

factors. Optimal T-cell immunity involves harmonized action between multiple cell populations and a population level metric combining the cell populations will be biologically meaningful in understanding the aging immune system. In this regard, CD4/CD8 ratio is a well-established ARIP measure where the prevalence of an inverted CD4/CD8 ratio increases with age (98,99). However, we have shown previously that the prevalence of the inverted CD4/CD8 ratio increases with age only among CMV seropositive individuals and not among CMV seronegative individuals suggesting that this ratio is not a universal marker of ARIP as previously thought (97). Thus, other measures of ARIP, beyond the inverted CD4/CD8 ratio, that are independent of CMV serostatus need to be identified. To this end, one study identified a cellular composite measure (IMM-AGE) that was associated with mortality, and another study developed an inflammatory aging clock (iAGE) utilizing soluble systemic chronic inflammation markers or cytokines which was associated with multimorbidity (100,101). However, the interpretation of cytokine levels is complicated by the fact that circulating cytokine levels are influenced by secretion from multiple cell types and the IMM-Age measure that utilized a combination of flow cytometry and mass cytometry estimated differences in both innate and adaptive immune systems. However, since age-related and CMV-related changes are seen predominantly in the T-cell subsets, a composite measure focused on T cells and relatively easy to implement in large epidemiological studies is needed.

To address this gap, we used T-cell immunophenotyping data that measured 11 T cell subsets in a large, population-based sample and used well understood aging-related changes in T cell distribution to develop aging-related immune phenotype (ARIP) indicators that would be robust across sexes, race/ethnic groups and CMV serostatus. We

developed ARIP measures based on *a priori* knowledge of (i) the role of chronological age in the immune system and (ii) the role of underlying biological mechanisms in the immune system such as CMV which increases susceptibility to age-related chronic conditions (67,102,103).We also evaluated the correlation between individual T-cell subsets and association of individual T-cell subsets with chronological age, biological age and multimorbidity to develop composite ARIP measures. We created two candidate measures: $T_N/ (T_{CM}+ T_{EM}+ T_{EFF})$ (referred as T_N/ T_M) in CD4+ and CD8+ T cells and benchmarked them against CD4/CD8 ratio, CD8+ T_N and CD4+ T_N by evaluating associations of all these measures with chronological age, biological age, multimorbidity (that includes several aging-related diseases such as diabetes, cancer, lung disease, stroke, and heart disease), and mortality.

B. Methods

B.1. Study Population

The Health and Retirement Study (HRS), supported by National Institute of Aging (NIA), is a nationally representative longitudinal survey of adults in the United States over the age of 50. The study began in 1992 with participants 51-61 years of age and now uses a steady-state design to replenish the sample every 6 years with younger cohorts (86,104) with surveys/interviews being conducted biennially. The HRS sample employs a multi-stage area probability design with oversampling of African-American and Hispanic households at about twice the rate of Whites (86). Whole blood was collected from HRS participants in the 2016 wave as part of the Venous Blood Study (VBS) and immunophenotyping was performed on samples from 9932 participants. After removing participants with missing age, sex, race/ethnicity, CMV seroprevalence,

Venous Blood Study survey weights, and T cell data, 8603 participants were included in the study. After weighting the sample using Venous Blood Study survey weights, it is representative of a U.S older national population.

B.2. Immunophenotyping

Peripheral Blood Mononuclear cells (PBMCs) were isolated from whole blood and used for the measurement of immune cells (105). All flow cytometry measurements were performed on an LSRII flow cytometer or a Fortessa X20 instrument (BD Biosciences, San Diego, CA). Immunophenotyping data was analyzed using OpenCyto and FlowAnnotator as described previously (106). T-cells and ten T-cell subsets were evaluated in HRS. Panel 1 consisted of antibody cocktails targeting T cells and T cell subsets (Table 3-1). T cell subsets were represented as percentages of their parent population. The detailed methods and cell subset definitions have been mentioned previously (97).

B.3. Ascertainment of multimorbidity and mortality.

Prevalence of type II diabetes, stroke, lung disease, heart disease, and cancer were used to generate a 2016 multimorbidity outcome. Type II diabetes was defined as fasting glucose value >= 126 mg/dL or use of diabetes medications or insulin. Stroke, lung disease, heart disease and cancer were obtained from self-report. Multimorbidity was defined as a categorical variable with four categories (no prevalent chronic conditions (reference), a single chronic condition, two chronic conditions, and three or more chronic conditions). Incident multimorbidity in 2018 and 2020 was ascertained by defining a baseline study population with participants who had a multimorbidity score of 0 or 1 in 2016. Among the participants in this baseline population, incident

multimorbidity was defined as a multimorbidity score of 2 or higher in 2018 and 2020 using self-report of Type II diabetes, stroke, lung disease, heart disease, and cancer. Mortality assessed at 2018 was used for this analysis. HRS uses two resources to report mortality: proxy interviews and using the National Death Index (NDI) information.

B.4. Ascertainment of biological age and self-rated health

Biological age was developed using the Klemera Doubal (107) method as a combination of biomarkers which represents the decline in aging-related physiological functioning and susceptibility to disease in old age. It was previously validated in the NHANES cohort as a more reliable predictor of mortality than chronological age and other biological age algorithms (108). The biomarkers used for calculating the biological age were: Cardiovascular function (Systolic Blood pressure), Metabolic markers (Total Cholesterol and Fasting Glucose), Inflammation (CMV and C-Reactive Protein), Kidney function (Serum Creatinine and Blood Urea Nitrogen), Liver function (Alkaline phosphatase and Albumin), Lung function (Peak flow). Biological age was calculated using the R package BioAge and the in-built function kdm_calc (109). Self-rated health was measured in the Health and Retirement Study by asking the question "Would you say your health is excellent, very good, good, fair, or poor?" We combined the "excellent", "very good" and "good" categories to represent a good score for health and the "fair" and "poor" categories to represent a bad score for health.

B.4. Ascertainment of participant characteristics

Chronological age (in years), sex (female/male), race/ethnicity (Hispanic, Non-Hispanic Whites, Non-Hispanic Black, and Non-Hispanic Other) were obtained from the HRS demographic data (104). Smoking was self-reported and categorized as never

smokers, former smokers, and current smokers. Height was measured in inches and weight was measured in pounds. Height and weight for all the participants were obtained for half the sample in 2014 and the other half in 2016. BMI was calculated by using the formula (Weight in lbs/ (Height in inches)^2) *703. CMV seroprevalence was measured using the ratio of total IgG to anti-CMV IgG in serum using the Roche e411 immunoassay analyzer (Roche Diagnostics Corporation). The results are reported as non-reactive (<0.5 COI), borderline (0.5 to <1.0 COI) or reactive >=1.0 COI) where COI is cut-off interval (89). This was used as a binary variable by combining the borderline and negative groups.

B.5. Statistical analysis

We evaluated two candidate ARIP measures (T_N/T_M CD4+ and CD8+ T cells) to compare with existing CD4/CD8 ratio and the individual cell type CD8+ T_N as we have shown previously that it had a strong inverse association with chronological age (13). We studied associations of the candidate ARIP and existing ARIP measures with chronological age, biological age, and multimorbidity. CD4+ T_N/T_M, CD8+ T_N/T_M and CD4/CD8 ratio were log-transformed due to skewed distributions.

Pearson correlation coefficients were calculated between individual T-cell subsets. To study the association of ARIP measures with chronological age, and biological age, survey linear regression models were used after adjustment for chronological age, sex, race/ethnicity, and CMV status. A biological age acceleration measure was also calculated by subtracting chronological age from biological age. The participants were categorized into “higher” category if the age acceleration measure was higher than 0, and the participants were grouped into “lower” category if the age

acceleration measure was lower than 0. Using the “lower” group as a reference, logistic regression analysis was performed between ARIP measures and biological age after adjustment for age, sex, race/ethnicity and CMV status. To study the association of ARIP measures with prevalent multimorbidity in 2016, we used multinominal survey logistic regression. We studied the association between ARIP measures and the individual components of multimorbidity (type II diabetes, hypertension, lung disease, heart disease, and stroke) by using a series of survey logistic regression models after adjustment for age, sex, race/ethnicity, CMV status, smoking status, and BMI. We used survey cox regression models to study the association of ARIP measures in incident multimorbidity in 2018 and 2020. We used survey logistic regression models to study the association of the ARIP measures with self-rated health and mortality in 2018. The survey models were adjusted for survey design parameters including strata and cluster details for sampling error and participant sample weights from the Venous Blood Study to account for sample design. In sensitivity analyses, three additional characterizations of ARIP were explored: (i) CD4+ T_N/CD8+ T_N; (ii) CD4+ T_N + T_{CM}; (iii) CD8 T_N + T_{CM}.

We used SAS version 9.4 (SAS Institute, Inc., Cary, NC) and R Statistical Analysis software version 4.0.0 for all the analyses.

C. Results

C.1. Participant characteristics

Among the 8603 participants included, the average age of participants was 68.65 years. Fifty-four percent of the participants were women; 10.02 % of participants were Non-Hispanic Black and 9.01% were Hispanic; 11.61% of the participants were current smokers and the average BMI was 29.69 Kg/m^2. Sixty-four percent of the participants were seropositive for cytomegalovirus. In 2016, 48.13% of participants were without any chronic conditions, 32.82% participants with one chronic condition, 14.06% participants with two chronic conditions, and 4.98% participants with three or more chronic conditions. Heart disease was the most prevalent (24.83%) among the diseases used to define multimorbidity (Table 3-2). Eight percent of the participants had incident multimorbidity in 2018 and 2020, and four percent of the participants were not alive in 2018 (Table 3-2). As expected, CD4+ T_N was negatively correlated with CD4+ T_{CM}, CD4+ T_{EFF} and CD4+ T_{EM} (R=-0.72, -0.3 and -0.22, respectively) (Figure 3-1). CD8+ T_N was negatively correlated with CD8+ T_{EFF} and CD8+ T_{EM} (R= -0.67 and -0.26, respectively) (Figure 3-1). Based on the negative correlation measures of T_N with T_{CM}, T_{EFF}, and $T_{EM,}$ the composite ARIP measures CD4+ and CD8+ $T_N/(T_{CM}+ T_{EM}+ T_{EFF})$ (referred as T_N/ T_M) were developed.

C.2. Association of ARIP measures with chronological age

Overall, the CD8+ subsets were more strongly associated with chronological age than the CD4+ subsets. Among the individual cell subsets, CD4+ and CD8+ T_N were negatively associated with chronological age (β=-0.32; p=0.05 and β = -3.41; p<0.0001). CD4+ and CD8+ T_{EFF} (β=0.66; p<0.0001 and β = 2.58; p<0.0001), CD4+ and CD8+ T_{EM}

(β=1.81; p<0.0001 and β =1.7; p<0.0001) and CD8+ T_{CM} (β = 0.39; p=0.04) were positively associated with chronological age (Figure 3-2). Among the ARIP measures, univariate analysis showed that CD4/CD8, CD4+ T_N/T_M, and CD4+ T_N had the weakest association with chronological age whereas CD8+ T_N/T_M and CD8+ T_N had the strongest association with biological age (Figure 3-3). A one standard deviation (SD) increase in CD8+ T_N/T_M (p <0.001) was associated with a 3.61 years lower chronological age (Figure 3-4) after adjusting for sex, race/ethnicity, and CMV status. CD4/CD8 and CD4+ T_N/T_M were not associated with chronological age (p=0.13 and p=0.84, respectively). The other ARIP measures also had an inverse association with chronological age (Table 3-3).

C.3. Association of ARIP measures with biological age

Among the individual T-cell subsets, CD4+T_N was negatively associated with biological age (0.24 years decrease in biological age for one standard deviation (SD) increase in CD4+ T_N (p=0.001) (Figure 3-7), one SD increase in CD4+ T_{CM} and CD4+ T_{EFF} was associated with 0.18 years (p=0.01) and 0.13 years (p=0.02) higher biological age respectively (Figure 3-5). Among the CD8+ T cell subsets, only the CD8+ T_{CM} subset was positively associated with biological age (0.12 years higher biological age for one SD increase in CD8+T_{CM} (p=0.03)) (Figure 3-5). Univariate analysis showed that CD4/CD8, CD4+ T_N/T_M, and CD4+ T_N had the weakest association with biological age whereas CD8+ T_N/T_M and CD8+ T_N had the strongest association with biological age (Figure 6). However, after adjustment for covariates, CD4+ T_N/T_M had a strong negative association with biological age with 0.23 years lower biological age for one standard deviation increase in CD4+ T_N/T_M (p-value = 0.003) (Figure 3-7). CD4/CD8 ratio, CD8+

$T_N/T_{M,}$ and CD8+ T_N were not associated with biological age. CD4+ T_N/T_M, CD4+T_N and CD4+ T_{CM} remained associated with biological age acceleration (older category). One SD increase in CD4+T_N/T_M was associated with 8% lower odds of having higher biological age as compared to chronological age (OR: 0.92 (95% CI: 0.84 – 1.00); p-value = 0.05), one SD increase in CD4+ T_N was associated with 11% lower odds of having higher biological age (OR: 0.89 (95% CI: 0.83 – 0.95; p=0.001) and one SD increase in CD4+ T_{CM} was associated with 9% higher odds of having higher biological age (OR: 1.09 (95% CI: 1.02-1.17; p=0.01). CD8+ T_N/T_M was not associated with biological age acceleration (OR: 1.04; (95% CI: 0.98 – 1.11; p-value = 0.19) (Table 3-4). Additional adjustment for CRP did not substantially change the associations between ARIP measures and biological age (data not shown).

C.4. Association of ARIP measures with multimorbidity

CD4+ T_N had an odds ratio of 0.85 (95% CI: 0.79-0.92; p<0.0001) when comparing those with a single chronic condition (multimorbidity score = 1) vs. those without any chronic condition (multimorbidity score = 0), while the corresponding odds ratio for those with two and 3+ chronic conditions was 0.72 (95% CI: 0.65 – 0.79; p<0.0001) and 0.74 (95% CI: 0.64-0.87; p=0.0004) respectively (Figure 3-8). Among the other individual T-cell subsets, CD4+ T_{CM} was consistently positively associated across different categories of multimorbidity whereas the other T-cell subsets did not have a consistent association across multimorbidity categories (Figure 3-9). CD4+ T_N/T_M had an odds ratio of 0.89 (95% CI: 0.82-0.96; p=0.005) when comparing those with a single chronic condition vs. those without any chronic condition, while the corresponding odds ratio was 0.74 (95% CI: 0.63 – 0.86; p=0.0003) and 0.75 (95% CI: 0.63-0.90; p=0.003)

for those with two and 3+ chronic conditions, respectively (Figure 3-8). CD4/CD8, CD8+ T_N/T_M and CD8+ T_N were not consistently associated across the different categories of multimorbidity (Figure 3-8). Further stratifying the highest level of the multimorbidity score into 3 and 4+ did not change the observed associations (data not shown). Among the chronic diseases included in the multimorbidity score, ARIP measures had the strongest association with prevalent cancer after multivariate adjustment. Lung disease, diabetes, stroke, and heart disease were not individually associated with the ARIP measures after multivariate adjustment (Figure 3-10, Table 3-5). CD4+ T_{CM} was associated with incident multimorbidity over four years follow-up (OR: 1.09 (95% CI: 1.01 – 1.18; p-value = 0.03). CD4+ T_N/T_M and CD4+ T_N also had association with incident multimorbidity in the consistent direction (OR: 0.91 (95% CI: 0.82- 1.02); p-value = 0.09 and OR = 0.91 (95% CI: 0.82 – 1.01); p-value =0.09) though the associations were not statistically significant. (Table 3-5).

C.5. Association of ARIP measures with self-rated health and two-year mortality

All the ARIP measures were associated with self-reported health (bad score for health vs good score for health) with the strongest association in CD4+ T_N/T_M, CD4+ T_N and CD4+ T_{CM} with OR of 0.84 (95% CI: 0.78 – 0.91), 0.83 (95% CI: 0.77 – 0.89) and 1.14 (95% CI: 1.04-1.25; p=0.006), respectively (Table 3-6). The other individual T cell subsets were not associated with self-rated health (data not shown). CD4+ T_N/T_M and CD4+ T_N had the strongest association with mortality with an OR of 0.80 (95% CI: 0.67 – 0.95; p=0.01) and 0.81 (95% CI: 0.70-0.94; p=0.01), respectively. The other ARIP measures were not associated with mortality in 2018 (Figure 3-11). Among the individual

T-cell subsets, only CD8+ T_{EFF} (OR = 1.27 (95% CI: 1.09-1.49; p=0.003) was associated with mortality in 2018.

D. Discussion

This study explored the associations of ARIP measures with chronological age, biological age, and multimorbidity outcomes. CD4+ T_N/T_M and CD4+ T_N had the strongest association with biological age and mortality and was inversely associated with increasing levels of multimorbidity. Though CD8+ T_N/T_M was not associated with biological age nor mortality, CD8+ T_N/T_M had associations with individual conditions including heart disease, diabetes and cancer. More examination is needed to better understand the relationship between CD8+ T_N/T_M and the individual chronic conditions. CD4/CD8 ratio was not associated with chronological age or biological age and did not have a consistent association with multimorbidity (Table 3-7). The associations with increasing levels of multimorbidity, mortality and biological age indicates the CD4+ T_N/T_M and CD4+ T_N may be a biomarker to identify individuals at higher risk for in age acceleration and its associated morbidities and higher mortality.

CD4+ T_N/T_M and CD4+ T_N had associations of similar strengths with biological age, multimorbidity and mortality. CD4+ T_N and CD4+ T_{CM} were strongly negatively correlated. While CD4+ T_N was inversely associated with biological age and multimorbidity, CD4+ T_{CM} was positively associated with biological age and multimorbidity. These observed associations indicate antagonistic roles of CD4+ T_N and CD4+ T_{CM} in biological aging and the pathogenesis of age-related chronic conditions. Hence, a representation of ARIP as a combination of CD4+ T_N/T_M may be a better representation of the overall associations between T cell immunity and health outcomes

when compared to describing these associations with individual T-cell subsets. However, CD4+ T_N also shows similar associations with biological age, multimorbidity and mortality as CD4+ T_N/T_M.

Among the two new measures developed in this study, CD8+ T_N/T_M was associated with chronological age but not associated with biological age and higher chronic conditions. CD4+ T_N/T_M was not associated with chronological age but had a strong inverse association with biological age and with a higher number of prevalent co-occurring chronic conditions. The findings with CD4+ T_N/T_M are consistent with previous studies where CD4+ T cell subsets were associated with chronic diseases. In the Multi-Ethnic Study of Atherosclerosis (MESA), naïve and memory CD4+ T cells were cross-sectionally associated with type II diabetes and subclinical atherosclerosis (62,110). CD4+ T cells produce interferon- γ (IFN- γ) which contributes to inflammation, glucose intolerance, and insulin resistance in diet-induced obesity mice (111,112). Another study performed among type II diabetes patients found a significant reduction in the naïve pool in both CD4+ and CD8+ populations and a significant rise in T_{EM} and T_{EFF} populations of the CD4+ subset compared to age-matched controls (113). Decreased CD4+ T_N/T_M ratio has also been observed in 76 non-small cell lung cancer (NSCLC) patients compared to 28 age and sex-matched healthy volunteers (114). CD4+CD28null cells which have advanced effector functions have been found at an increased frequency in ischemic stroke patients which was also associated with stroke severity (115). These studies are consistent with our findings that the distribution of naïve and memory CD4+ T cells may be important in determining age-related outcomes. In contrast, the CD8+ T_N cells were inversely associated with chronological age after adjustment for sex, race/ethnicity, and

CMV serostatus but not associated with biological age and a higher number of chronic conditions. Previous studies have also used CD4/CD8 ratio as a measure of immune aging. An increased prevalence of the inverted CD4/CD8 ratio was associated with short-term mortality in the OCTO immune longitudinal study, possibly due to confounding by CMV serostatus (116). CD4/CD8 ratio was not associated with biological age and chronological age after adjustment for race/ethnicity, sex, and CMV status and was not consistently associated with a higher prevalence of multiple chronic diseases.

Although increasing chronological age is an important component of senescence, it does not directly measure the accelerated decline in health among individuals. Previous studies have shown that biological age predicts morbidity and mortality independent of chronological age in individuals (108,117–119). This is the first study to demonstrate an association between biological age and CD4+ T_N/T_M, CD4+ T_N and CD4+ T_{CM}. In the Rotterdam study, longitudinal associations of biological age with all-cause morbidity, stroke, cancer, and diabetes mellitus, suggests that biological age may predetermine disease occurrence (119). The strong association between multiple CD4+ subsets including CD4+ T_N/T_M and biological age suggests that CD4+ T_N/T_M might be an important risk factor for future disease occurrence and this needs to be confirmed in future studies.

Heterogeneity in strength of the association between the different components of multimorbidity and the ARIP measures were observed but the direction of the association was consistent among the diseases examined. CD4+ T_N/T_M and CD4+ T_N were lower among participants who had one chronic condition (multimorbidity score =1) compared to those without any chronic conditions (multimorbidity score =0). Since cancer was the

third most frequent chronic condition (after heart disease and diabetes), this may reflect the strong association between cancer and CD4+ T cells. The distribution of chronic conditions when participants had two chronic conditions (multimorbidity score =2) is similar to the overall prevalence of chronic conditions. Since the CD4+ T_N/T_M and CD4+ T_N CD4 subsets were associated with the three most common chronic conditions (heart disease, diabetes and cancer), this could explain the significant association between these subsests and multimorbidity score of two. However, when participants had three or more chronic conditions (multimorbidity score =3), lung disease had a higher prevalence compared to when participants had two chronic conditions (multimorbidity score =2). Though none of the ARIP measures were significantly associated with lung disease, CD4+ T_N/T_M and CD4+ T_N had the strongest association (OR = 0.88 and 0.90 respectively) with lung disease. This may explain why only CD4+ T_N/T_M and CD4+ T_N was associated with a multimorbidity score of 3. Among the disease conditions which constituted multimorbidity, cancer had the strongest association with the ARIP measures. Various theories have been proposed describing the involvement of naïve and memory CD4+ T cells in tumor immunity. The tumor microenvironment conditions promote the differentiation of memory and naïve CD4+ T cells into CD4+ T regulatory cells which suppress antitumor immunity (120–122). The concept of immunosurveillance in cancer indicates that the immune system can recognize and proactively remove precursors of cancer and naïve T cells may be involved in this process as they perform immunosurveillance roles to undergo proliferation in response to homeostatic signals (123,124). In general, T_{CM} cells have been found to produce higher levels of cytokines which have increased efficiency in an antitumor response (125,126). We also observed

that the ARIP measures were not significantly associated with incident multimorbidity after adjustment for age, sex, race/ethnicity, CMV status, smoking status, and BMI though the direction of association was consistent with what was observed in the cross-sectional analysis. This could be due to the limited number of incident multimorbidity events in the short follow-up period and the association between the ARIP measures and incident multimorbidity needs to be further evaluated in future studies. CD4+ T_N/T_M and CD4+ T_N were associated with two-year mortality as well strengthening their role in determining biological aging.

Major strengths of this study include the standardized and rigorous immunophenotyping methods used for measuring the cell populations, representative sample including Hispanic, black, and white individuals, immunophenotyping data available for a large sample size (about 9938 individuals), and concurrent measurement of cytomegalovirus seroprevalence. The novelty of this study is the creation of composite ARIP measures as a marker for an aspect of aging and characterizing the measures using chronological age, biological age, and age-related health outcomes. A limitation of this study is the absence of longitudinal immunophenotype data. Measurement of longitudinal changes in the immune cell subsets would help in studying the change in immune cells with age within the same participant and would help us to better understand the temporal relationship between immune aging and disease. The majority of the multimorbidity measure is based on the self-report of the participant and hence could be influenced by factors such as recall, social desirability, and lack of diagnosis. A previous study found that the sensitivity of self-report of hypertension was 88.9% among blacks, 82.8% for whites, and 84.0% for Hispanic ethnicity, and specificity was 92.8% for the whites and

86% for the blacks (127) informing that the measurements are robust. Another study performed to validate self-reported cancer measures with Medicare claims data had a 73.2% sensitivity and 96.2% specificity (128).

In conclusion, the results suggest that composite measures of age-related immune phenotypes are more meaningful in determining biological aging and multimorbidity. In this study, we showed that CD4+ T_N/T_M and CD4+ T_N had strong associations with biological age, multimorbidity and mortality and maybe a more consistent measure of immune aging compared to specific cell types such as CD8+ naïve T cells and the widely used CD4/CD8. CD4+ T_N/T_M and CD4+ T_N can be used in future studies with longitudinal measurements of the immune cells to better understand the role of immune aging in age-related morbidity and mortality.

Table 3-1: T-cell subset definitions measured in the Health and Retirement Study.

Cell type	**Surface markers used**	**Parent Population**
T cells	CD3+ CD19-	Single, Live lymphocytes
Cytotoxic T cells	CD3+ CD19- CD8+ CD4-	T cells
T_N (Naïve) cytotoxic T cells	CD3+ CD19- CD8+ CD4- CD45RA+ CCR7+ CD28+	Cytotoxic T cells
T_{EM} (Effector Memory) cytotoxic T cells	CD3+ CD19- CD8+ CD4- CD45RA- CCR7- CD28-	Cytotoxic T cells
T_{EMRA} cytotoxic T cells	CD3+ CD19- CD8+ CD4- CD45RA+ CCR7- CD28-	Cytotoxic T cells
T_{CM} (Central Memory) cytotoxic T cells	CD3+ CD19- CD8+ CD4- CD45RA- CCR7+ CD28+	Cytotoxic T cells
Helper T cells	CD3+ CD19- CD4+ CD8-	T cells
T_N (Naïve) helper T cells	CD3+ CD19- CD4+ CD8- CD45RA+ CCR7+ CD28+	Helper T cells
T_{EM} (Effector Memory) helper T cells	CD3+ CD19- CD4+ CD8- CD45RA- CCR7- CD28-	Helper T cells
T_{EMRA} helper T cells	CD3+ CD19- CD4+ CD8- CD45RA+ CCR7- CD28-	Helper T cells
T_{CM} (Central Memory) helper T cells	CD3+ CD19- CD4+ CD8- CD45RA- CCR7+ CD28+	Helper T cells

Table 3-2: Descriptive statistics of Health and Retirement Study participant characteristics measured in 2016 among 8603 participants with immunophenotyping data.

Characteristics	**Mean ± SE/ Frequency (%)**
Demographics	
Age (years) (Mean ±SE)	68.65 ± 0.26
Sex – Female (%)	54.12 %
Race/Ethnicity	
Hispanic (%)	9.01 %
Non-Hispanic Black (%)	10.02 %
Non-Hispanic White (%)	77.55 %
Lifestyle characteristics	
Smoking Status	
Never smokers (%)	44.29 %
Former smokers (%)	43.84 %
Current smokers (%)	11.87 %
BMI (kg/m^2) (Mean ±SE)	29.69 ± 0.10
Cytomegalovirus seroprevalence – Reactive (%)	64.54%
Biological Age (years) (Mean ±SE)	67.66 ± 0.24
Prevalence of disease conditions in 2016	
Multimorbidity score	
0	48.13%
1	32.82%
2	14.06%
3 and above	4.98%
Type II diabetes (%)	17.55 %
Heart Disease (%)	24.83 %
Stroke (%)	7.04 %
Lung Disease (%)	11.32 %
Cancer (%)	16.19 %
Percentage of participants with bad score for self-rated health in 2016	26.61%
Mortality in 2018	4.17%
Number of people with incident multimorbidity in 2018 and 2020	6.01%
Age-related immune phenotype measures	
CD4+/CD8+ (Mean ±SD) (log-transformed)	1.09 ± 0.70
T_N/T_M CD4+ (Mean ±SD) (log-transformed)	-0.03 ± 0.77
T_N/T_M CD8+ (Mean ±SD) (log-transformed)	-1.04 ± 1.14
CD8+ T_N	0.22 ± 0.16
CD4+ T_N	0.43 ± 0.18

Figure 3-1: Pearson correlation heatmap between individual T-cell subsets measured in the Health and Retirement Study.

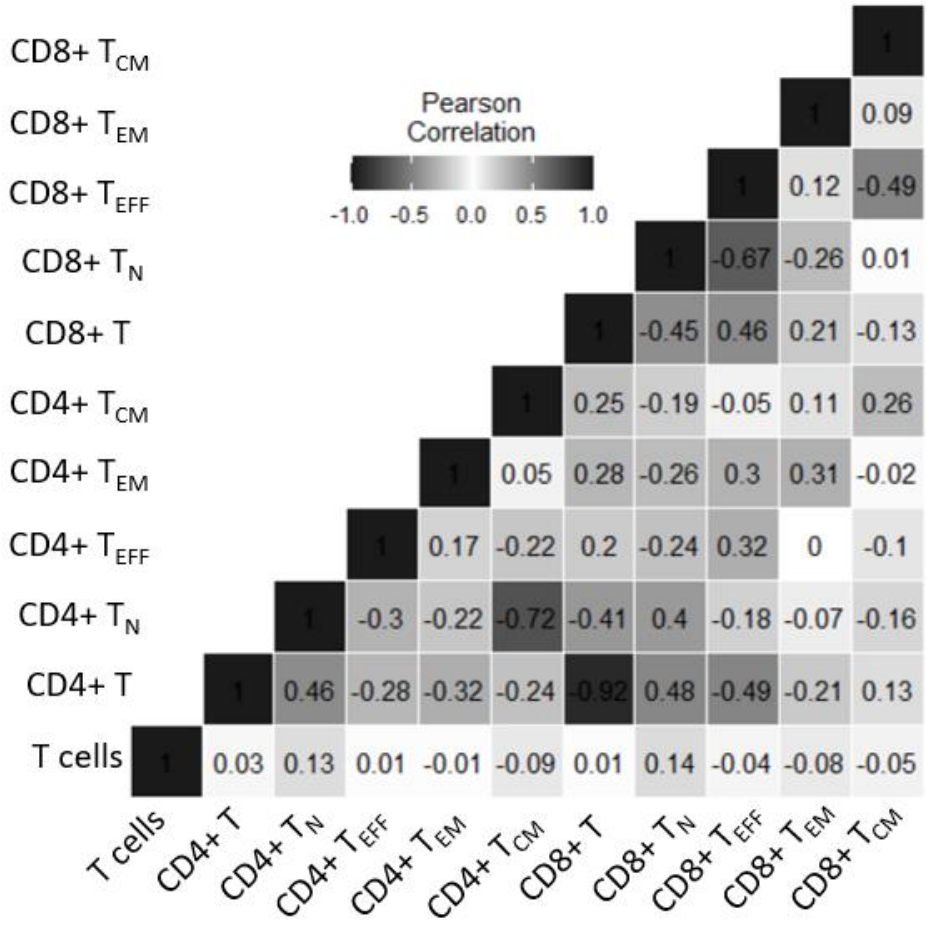


Figure 3-2: Association of individual T-cell subsets with chronological age. Age is used as the dependent variable in survey linear regression models. The beta estimates are estimated per one standard deviation unit of individual T-cell subsets. The models are adjusted for sex, race/ethnicity, and CMV status. The solid black line along 0 indicates no association.

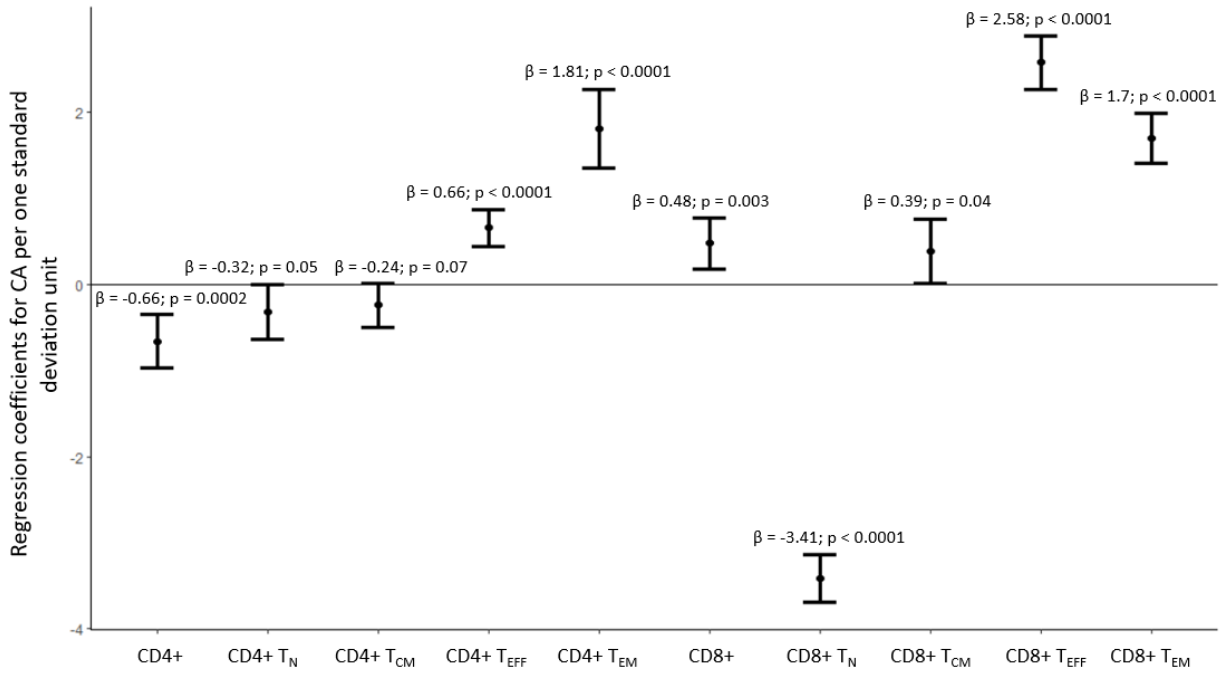


Figure 3-3: Scatterplot of ARIP measures with chronological age with a LOWESS curve.

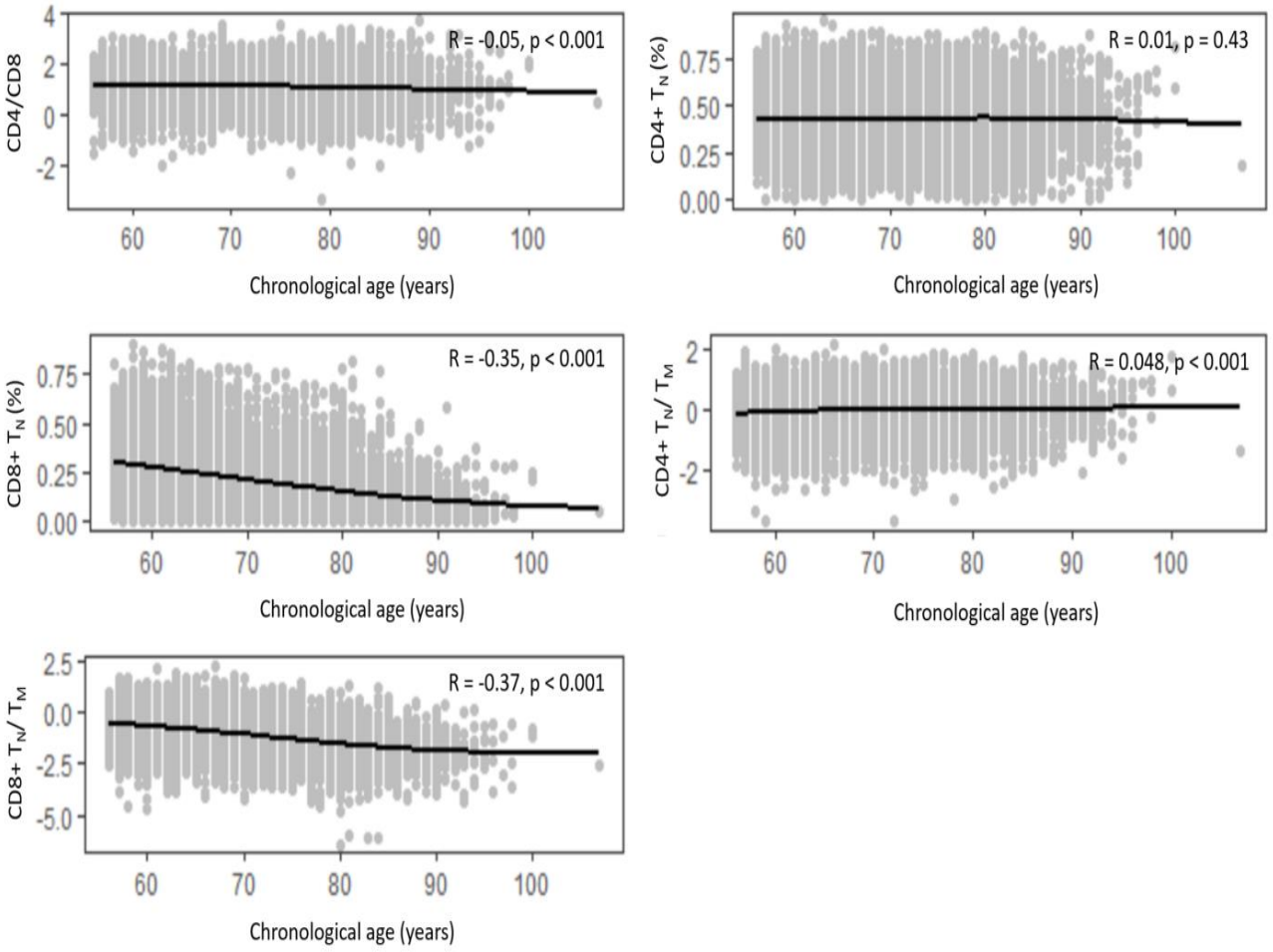


Figure 3-4: Association of ARIP measures with chronological age. Age is used as the dependent variable in survey linear regression models. The beta estimates are estimated per one standard deviation unit of the ARIP measures. The models are adjusted for sex, race/ethnicity, and CMV status. The solid black line along 0 indicates no association.

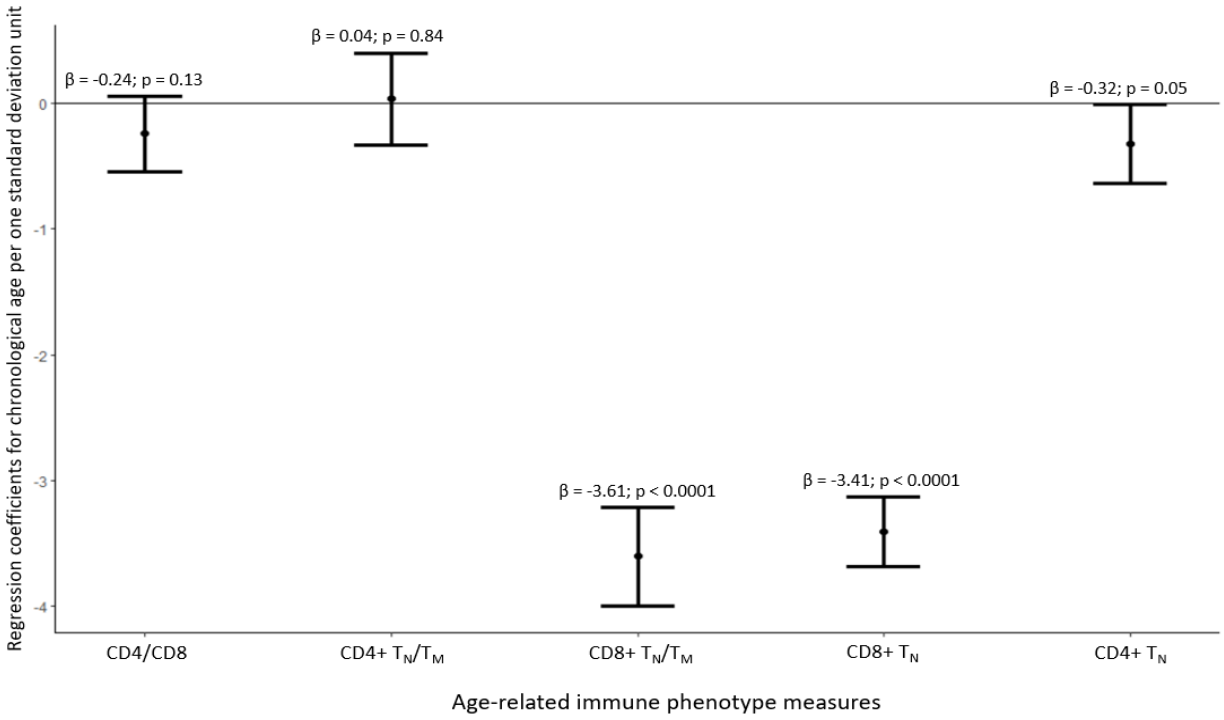


Table 3-3: Beta estimates of the association between the additional ARIP measures with chronological age and biological age.

Measure	**Chronological age (n= 9102)** (68.54 ± 10.23)	**Biological age (n= 7510)** (69.00 ± 12.35)
T_N CD4+/ T_N CD8+	2.86 ± 0.16 (<0.0001)	-0.20 ± 0.06 (0.001)
T_N + T_{CM} CD4+	-2.31 ± 0.23 (<0.0001)	-0.09 ± 0.06 (0.13)
T_N + T_{CM} CD8+	-3.71 ± 0.20 (<0.0001)	0.11 ± 0.06 (0.06)

Figure 3-5: Association of individual T-cell subsets with biological age. Biological age is used as the dependent variable in survey linear regression models. The beta estimates are estimated per one standard deviation unit of individual T-cell subsets. The models are adjusted for chronological sex, race/ethnicity, and CMV status. The solid black line along 0 indicates no association.

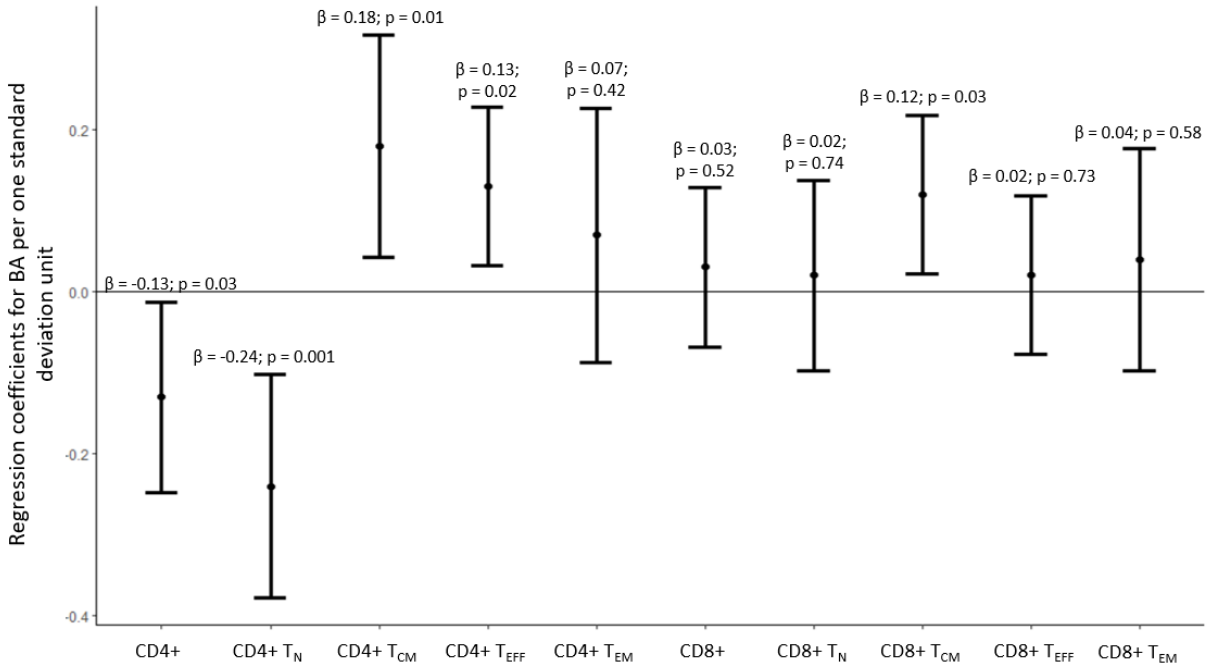


Figure 3-6: Scatterplot of ARIP measures with biological age with a LOWESS curve.

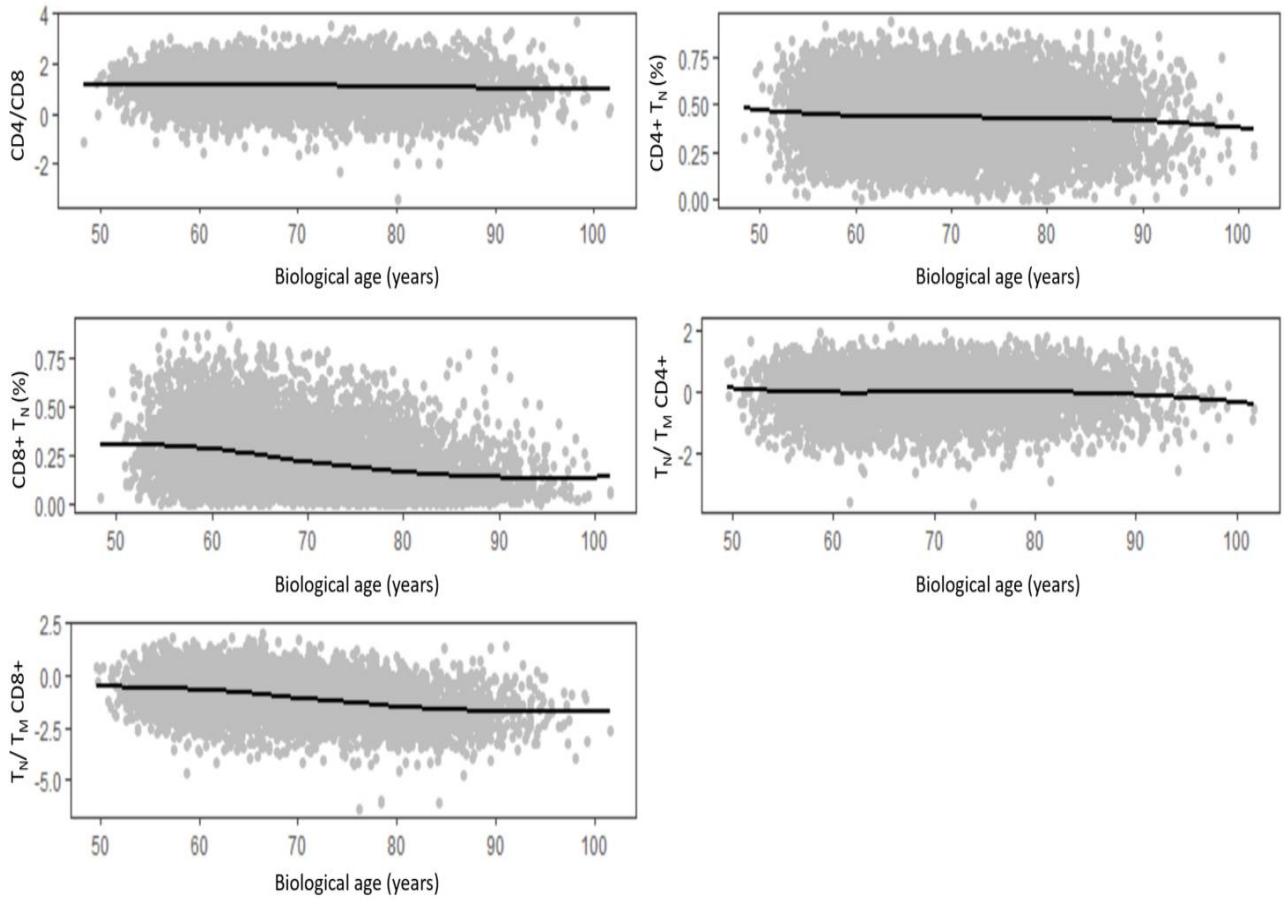


Figure 3-7: Association of ARIP measures with biological age. Biological age is used as the dependent variable in survey linear regression models. The beta estimates are estimated per one standard deviation unit of the ARIP measures. The models are adjusted for chronological age, sex, race/ethnicity, and CMV status. The solid black line along 0 indicates no association.

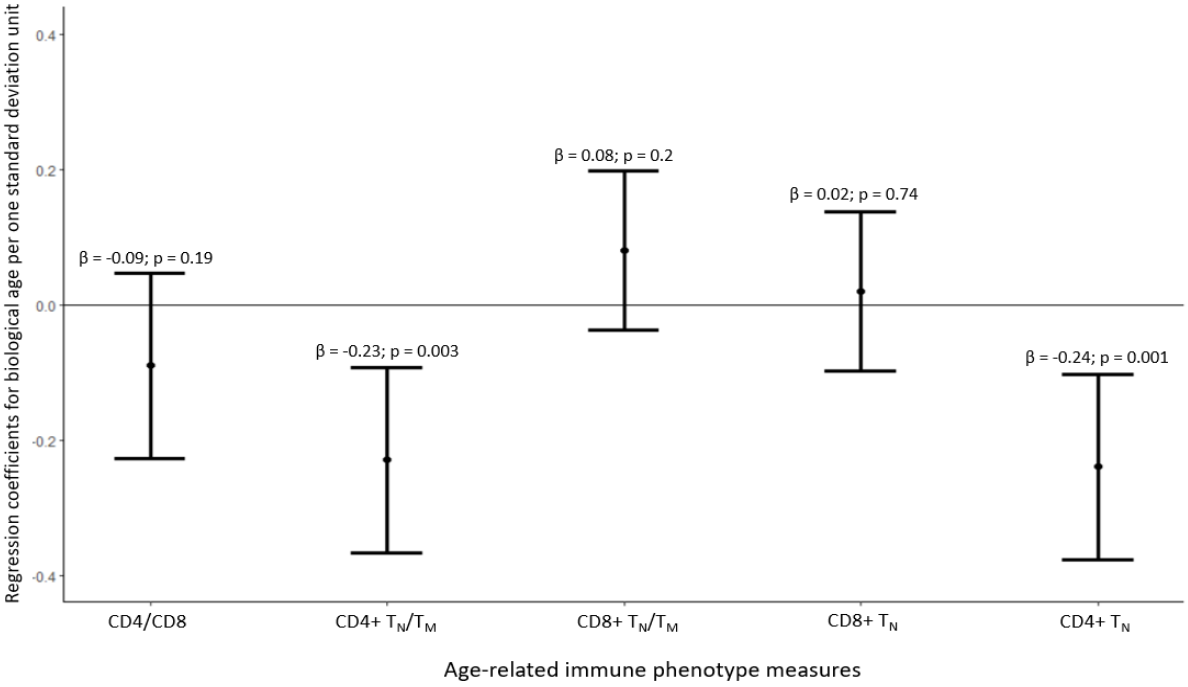


Table 3-4: Beta estimates of the association between ARIP measures and biological age acceleration measure after adjustment for age, sex, race/ethnicity, and CMV status.

Measure	**OR (95% CI) for the older (n=3189) vs younger age (n=4321) acceleration measure**
CD4/CD8	0.97 (95% CI: 0.97 – 1.04; p=0.39)
CD4+ T_N/T_M	0.92 (95% CI: 0.84-1.00; p=0.05)
CD8+ T_N/T_M	1.04 (95% CI: 0.98-1.11; p=0.19)
CD8+ T_N	0.99 (95% CI: 0.92 – 1.06; p=0.71)
CD4+ T_N	0.89 (95% CI: 0.83 – 0.95; p=0.001)
CD4+ T_{CM}	1.09 (95% CI: 1.02 – 1.17; p=0.01)

Figure 3-8: Odds ratios and 95% CI of association of multimorbidity levels with ARIP markers per one SD unit increase in ARIP marker. Adjusted for age, sex, race/ethnicity, CMV status, smoking status, and BMI. The solid black line along OR of 1 indicates no association.

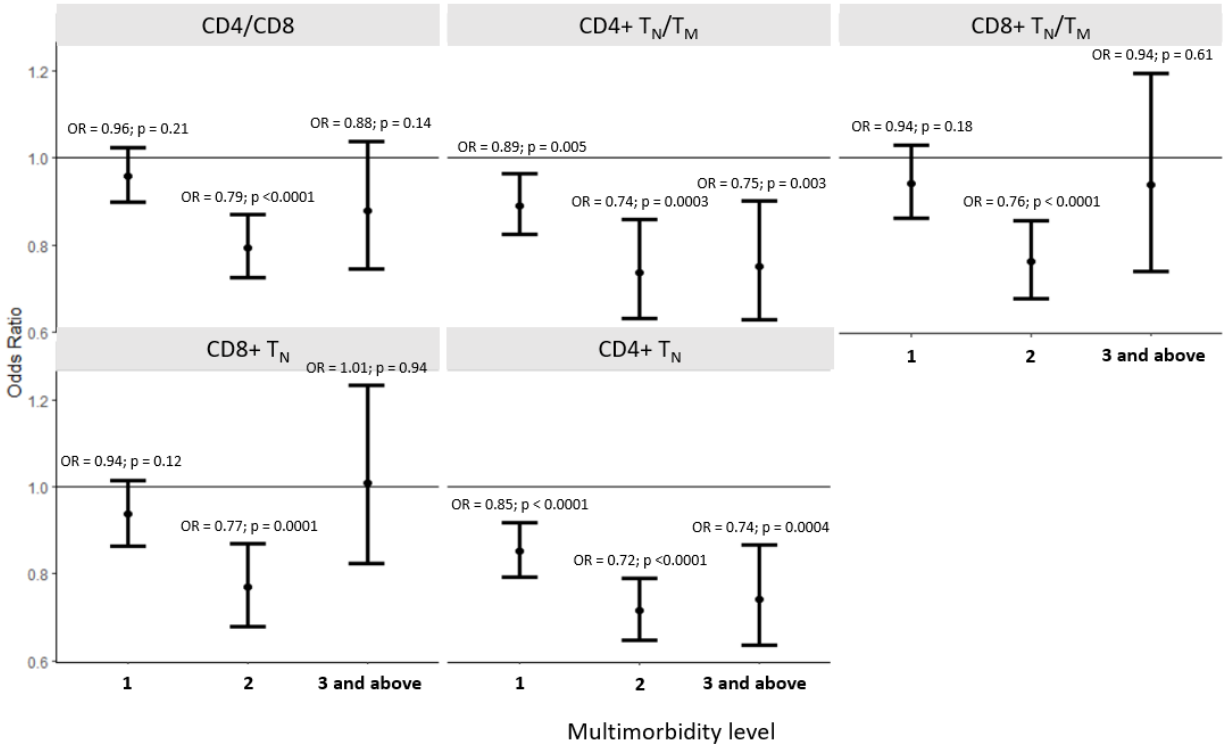


Figure 3-9: Odds ratios and 95% CI of association of multimorbidity levels with individual T-cell subsets per one SD unit increase in ARIP marker. Adjusted for age, sex, race/ethnicity, CMV status, smoking status, and BMI. The solid black line along OR of 1 indicates no association.

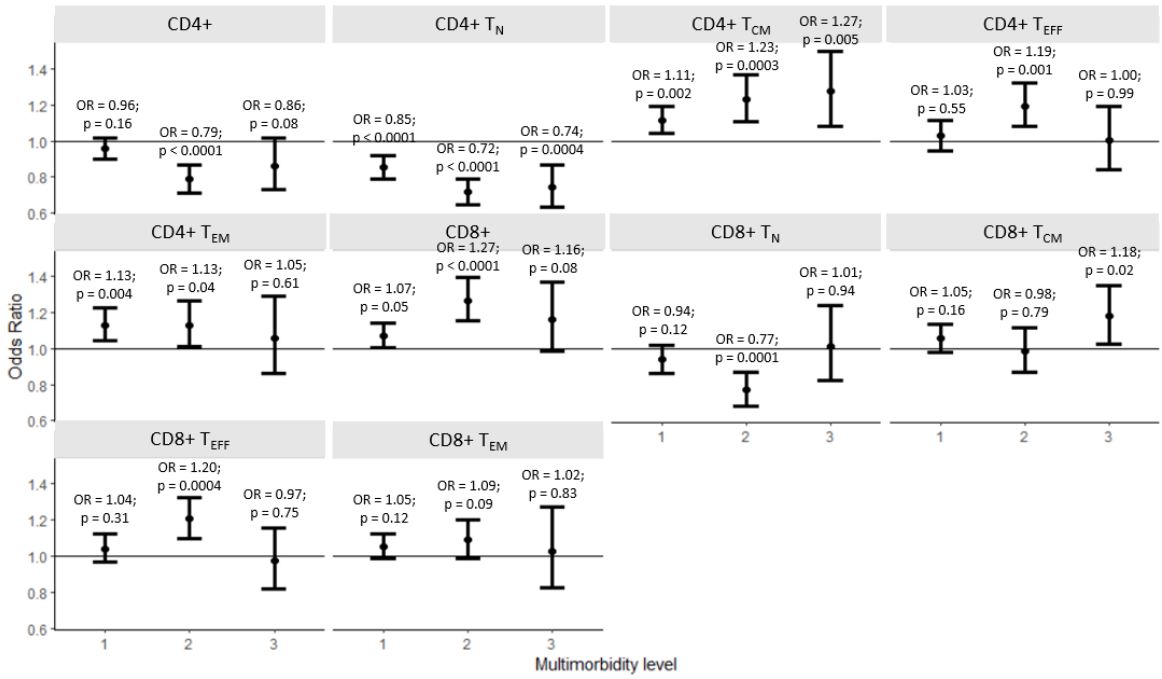


Figure 3-10: Odds ratios and 95% CI of association of individual components of multimorbidity ARIP markers per one SD unit increase in ARIP marker. Adjusted for age, sex, race/ethnicity, CMV status, smoking status, and BMI. The solid black line along OR of 1 indicates no association.

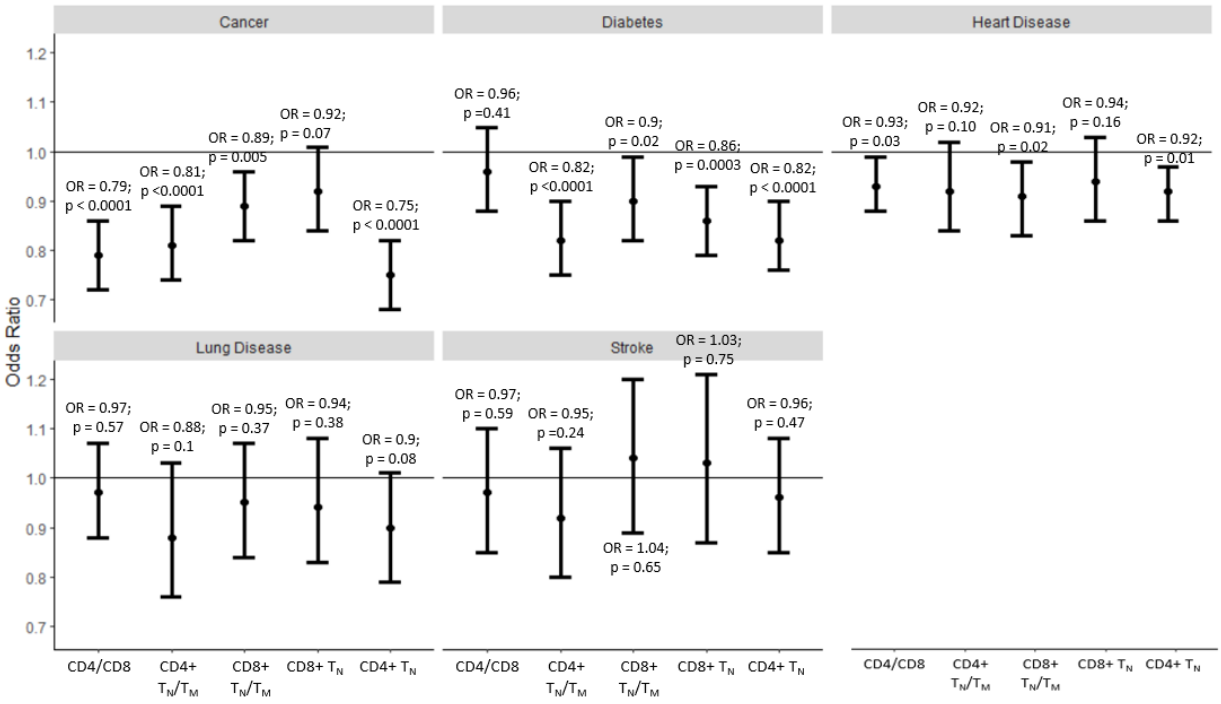


Table 3-5: Estimates of association of ARIP measures with age-related outcomes.

ARIP measures			
	Multimorbidity – 2016		
	Model 1[a]	Model 2[b]	Model 3[c]
	OR (95% CI)	OR (95% CI)	OR (95% CI)
CD4/CD8 ratio	0.80 (0.75 – 0.84)	0.84 (0.79-0.90)	0.83 (0.77 – 0.90)
CD4+ T_N / CD8+ T_N	0.97 (0.91 – 1.03)	0.87 (0.81-0.93)	0.88 (0.81-0.96)
CD4+ T_N + T_{CM}	0.80 (0.75 – 0.85)	0.90 (0.85 – 0.96)	0.85 (0.79-0.91)
CD8+ T_N + T_{CM}	0.88 (0.82 – 0.95)	1.04 (0.96 – 1.13)	1.01 (0.91 – 1.12)
CD4+ T_N/T_M	0.72 (0.64 – 0.80)	0.71 (0.64 - 0.80)	0.78 (0.68 – 0.89)
CD8+ T_N/T_M	0.72 (0.67-0.79)	0.82 (0.76-0.89)	0.82 (0.75-0.89)
CD8+ T_N	0.79 (0.75 – 0.83)	0.93 (0.88 – 1.00)	0.85 (0.77 – 0.93)
CD4+ T_N	0.74 (0.68-0.79)	0.75 (0.70 – 0.81)	0.78 (0.71 – 0.85)
CD4+ T_{CM}	1.22 (1.13 – 1.31)	1.23 (1.14 – 1.34)	1.18 (1.07 – 1.30)
Associations of immunosenescence measures with individual components of multimorbidity in 2016			
	Type II diabetes – 2016		
	Model 1[a]	Model 2[b]	Model 3[c]
	OR (95% CI)	OR (95% CI)	OR (95% CI)
CD4/CD8 ratio	0.90 (0.85 – 0.96)	0.97 (0.91 – 1.03)	0.96 (0.88 – 1.05)
CD4+ T_N / CD8+ T_N	0.92 (0.86 – 0.98)	0.93 (0.86 – 1.00)	0.92 (0.84 – 1.01)
CD4+ T_N + T_{CM}	1.02 (0.96 – 1.08)	1.10 (1.04 – 1.16)	1.05 (0.98-1.12)
CD8+ T_N + T_{CM}	1.03 (0.97 – 1.10)	1.08 (1.00 - 1.16)	1.09 (0.99 -1.20)
CD4+ T_N/T_M	0.69 (0.64 – 0.74)	0.73 (0.68 – 0.79)	0.82 (0.75 – 0.90)
CD8+ T_N/T_M	0.83 (0.78 – 0.89)	0.87 (0.80 – 0.94)	0.90 (0.82 – 0.99)
CD8+ T_N	0.94 (0.88 – 1.00)	0.99 (0.92 – 1.07)	0.86 (0.79-0.93)
CD4+ T_N	0.75 (0.69 – 0.81)	0.80 (0.74 – 0.86)	0.82 (0.76-0.90)
CD4+ T_{CM}	1.33 (1.25 – 1.41)	1.27 (1.20 – 1.35)	1.20 (1.12 – 1.29)
	Heart Disease - 2016		
	Model 1[a]	Model 2[b]	Model 3[c]
	OR (95% CI)	OR (95% CI)	OR (95% CI)
CD4/CD8 ratio	0.92 (0.88 – 0.96)	0.95 (0.90-1.00)	0.93 (0.87 – 0.99)
CD4+ T_N / CD8+ T_N	1.13 (1.07 – 1.20)	0.98 (0.93 – 1.03)	0.99 (0.93 – 1.05)
CD4+ T_N + T_{CM}	0.83 (0.78 – 0.88)	0.95 (0.90 1.00)	0.92 (0.87-0.98)
CD8+ T_N + T_{CM}	0.81 (0.76 – 0.87)	0.99 (0.92-1.06)	0.98 (0.90 – 1.06)
CD4+ T_N/T_M	0.90 (0.83 – 0.98)	0.89 (0.81 – 0.98)	0.92 (0.84 – 1.02)
CD8+ T_N/T_M	0.78 (0.73 – 0.83)	0.90 (0.84 – 0.97)	0.91 (0.83-0.98)
CD8+ T_N	0.76 (0.72 – 0.81)	0.94 (0.89 – 0.99)	0.94 (0.86 – 1.03)
CD4+ T_N	0.88 (0.83 – 0.93)	0.90 (0.84 – 0.95)	0.92 (0.86 – 0.97)
CD4+ T_{CM}	1.08 (1.02 – 1.14)	1.10 (1.04 – 1.16)	1.06 (1.00 -1.12)

[a]Model 1 assessed bivariate associations.

[b]Model 2 adjusted for age, sex, race, and CMV status.

[c]Model 3 additionally adjusted for smoking status, BMI.

Note: Effect estimates are per one SD unit increase in ARIP measures.

Table 3-5 (continued): Estimates of association of ARIP measures with age-related outcomes

	Lung Disease - 2016		
	Model 1[a]	Model 2[b]	Model 3[c]
	OR (95% CI)	OR (95% CI)	OR (95% CI)
CD4/CD8 ratio	0.93 (0.85 – 1.01)	0.97 (0.89 – 1.05)	0.97 (0.88 – 1.07)
CD4+ T_N / CD8+ T_N	0.94 (0.87 – 1.03)	0.91 (0.83 – 0.99)	0.94 (0.85 – 1.04)
CD4+ T_N + T_{CM}	0.93 (0.86 – 1.01)	0.95 (0.87 – 1.03)	0.88 (0.80 – 0.97)
CD8+ T_N + T_{CM}	1.01 (0.93 – 1.11)	1.05 (0.94 – 1.16)	0.92 (0.81 – 1.05)
CD4+ T_N/T_M	0.83 (0.74 – 0.94)	0.81 (0.71 – 0.92)	0.88 (0.76-1.03)
CD8+ T_N/T_M	0.92 (0.84 – 1.02)	0.97 (0.88 – 1.07)	0.95 (0.84 – 1.07)
CD8+ T_N	0.94 (0.87 – 1.02)	0.97 (0.88 – 1.06)	0.94 (0.83 – 1.08)
CD4+ T_N	0.84 (0.76 – 0.93)	0.83 (0.75 – 0.93)	0.90 (0.79 – 1.01)
CD4+ T_{CM}	1.12 (0.99 – 1.26)	1.16 (1.02 – 1.32)	1.09 (0.94 – 1.28)
	Stroke - 2016		
	Model 1[a]	Model 2[b]	Model 3[c]
	OR (95% CI)	OR (95% CI)	OR (95% CI)
CD4/CD8 ratio	0.90 (0.81 – 1.00)	1.00 (0.89 – 1.12)	0.97 (0.85-1.10)
CD4+ T_N / CD8+ T_N	1.03 (0.92 – 1.15)	0.97 (0.86 – 1.10)	0.98 (0.87 – 1.10)
CD4+ T_N + T_{CM}	0.86 (0.79 – 0.93)	0.98 (0.89 – 1.09)	0.95 (0.86 – 1.06)
CD8+ T_N + T_{CM}	0.86 (0.77 – 0.96)	0.99 (0.86 – 1.12)	0.99 (0.87 – 1.14)
CD4+ T_N/T_M	0.83 (0.73 – 0.95)	0.89 (0.77 – 1.02)	0.92 (0.80 – 1.06)
CD8+ T_N/T_M	0.87 (0.79 – 0.95)	1.04 (0.91 – 1.18)	1.04 (0.89 – 1.20)
CD8+ T_N	0.83 (0.77 – 0.89)	0.96 (0.86 – 1.06)	1.03 (0.87 – 1.21)
CD4+ T_N	0.83 (0.74 – 0.93)	0.93 (0.82 – 1.05)	0.96 (0.85 – 1.08)
CD4+ T_{CM}	1.17 (1.03 – 1.32)	1.12 (0.99 – 1.27)	1.09 (0.96 – 1.24)
	Cancer - 2016		
	Model 1[a]	Model 2[b]	Model 3[c]
	OR (95% CI)	OR (95% CI)	OR (95% CI)
CD4/CD8 ratio	0.83 (0.77 – 0.90)	0.81 (0.75 – 0.88)	0.79 (0.72-0.86)
CD4+ T_N / CD8+ T_N	0.94 (0.86 – 1.04)	0.78 (0.71 – 0.86)	0.79 (0.72 – 0.87)
CD4+ T_N + T_{CM}	0.65 (0.60 – 0.71)	0.69 (0.63 – 0.76)	0.66 (0.61 – 0.72)
CD8+ T_N + T_{CM}	0.79 (0.71 – 0.87)	0.92 (0.83 – 1.02)	0.89 (0.80 – 1.00)
CD4+ T_N/T_M	0.87 (0.79 – 0.95)	0.81 (0.74 – 0.89)	0.81 (0.74 – 0.89)
CD8+ T_N/T_M	0.82 (0.76 – 0.89)	0.92 (0.86 – 1.00)	0.89 (0.82 – 0.96)
CD8+ T_N	0.75 (0.70 – 0.82)	0.92 (0.84 – 1.00)	0.92 (0.84 – 1.01)
CD4+ T_N	0.80 (0.73 – 0.87)	0.76 (0.70 – 0.83)	0.75 (0.68 – 0.82)
CD4+ T_{CM}	1.07 (0.98 – 1.17)	1.12 (1.03 – 1.22)	1.12 (1.03 – 1.23)

[a]Model 1 assessed bivariate associations.
[b]Model 2 adjusted for age, sex, race, and CMV status.
[c]Model 3 additionally adjusted for smoking status, BMI.
Note: Effect estimates are per one SD unit increase in ARIP measures.

Table 3-5 (continued): Estimates of association of ARIP measures with age-related outcomes

	Incident multimorbidity – 2018 and 2020		
	Model 1[a]	Model 2[b]	Model 3[c]
	Hazard Ratio (95% CI)	Hazard Ratio (95% CI)	Hazard Ratio (95% CI)
CD4/CD8 ratio	0.93 (0.85 – 1.02)	0.95 (0.86 – 1.05)	0.94 (0.85 – 1.04)
CD4+ T_N / CD8+ T_N	0.98 (0.90 – 1.08)	0.95 (0.87 – 1.05)	0.96 (0.87 – 1.06)
CD4+ T_N + T_{CM}	0.99 (0.92 – 1.06)	1.07 (0.98 – 1.16)	1.04 (0.95 – 1.13)
CD8+ T_N + T_{CM}	0.99 (0.90 – 1.11)	1.09 (0.98 – 1.22)	1.07 (0.95 – 1.19)
CD4+ T_N/T_M	0.82 (0.75 – 0.90)	0.86 (0.77 – 0.95)	0.91 (0.82 – 1.02)
CD8+ T_N/T_M	0.87 (0.82 – 0.92)	0.93 (0.87 – 1.00)	0.94 (0.87 – 1.02)
CD8+ T_N	0.92 (0.84 – 1.01)	0.99 (0.89 – 1.09)	0.99 (0.89 – 1.10)
CD4+ T_N	0.87 (0.80 – 0.95)	0.90 (0.82 – 0.99)	0.91 (0.82 – 1.01)
CD4+ T_{CM}	1.17 (1.09 – 1.26)	1.14 (1.06 – 1.23)	1.09 (1.01 – 1.18)

[a]Model 1 assessed bivariate associations.
[b]Model 2 adjusted for age, sex, race, and CMV status.
[c]Model 3 additionally adjusted for smoking status, BMI.
Note: Effect estimates are per one SD unit increase in ARIP measures.

Figure 3-11: Odds ratios and 95% CI of association of mortality with ARIP markers per one SD unit increase in ARIP marker. Adjusted for age, sex, race/ethnicity, CMV status, smoking status, and BMI. The solid black line along OR of 1 indicates no association.

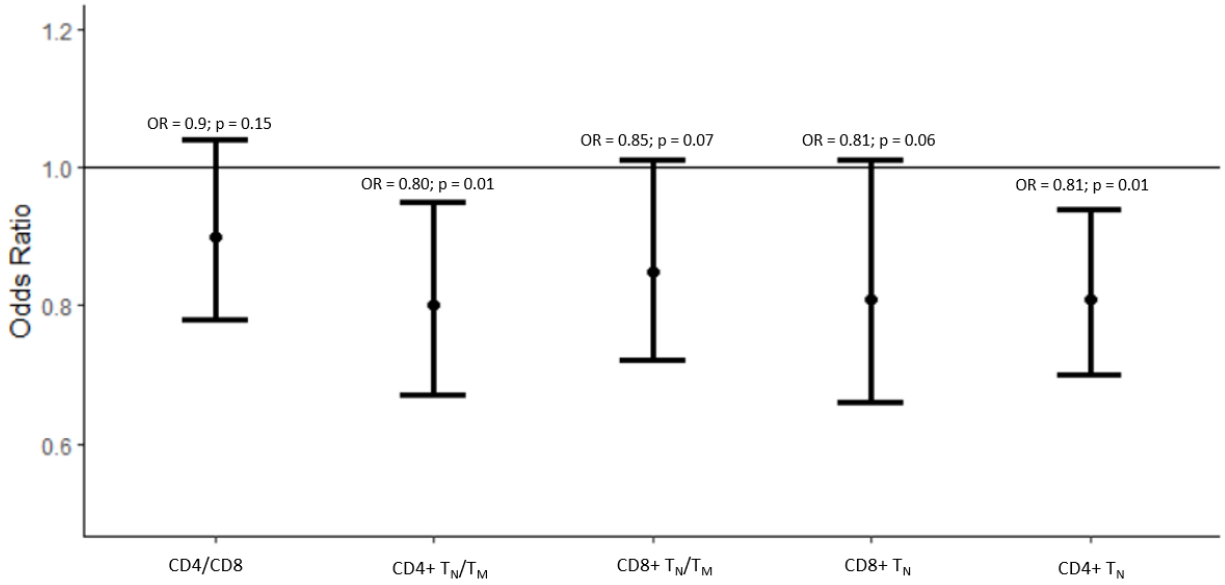


Table 3-6: Associations between self-rated health and ARIP measures after adjustment for age, sex, race/ethnicity, CMV status, smoking status, and BMI.

Measure	**Odds ratio (bad (n=2486) vs good self-rated health (6108))**
CD4/CD8	0.89 (95% CI: 0.82 – 0.96); p=0.002
T_N/T_M CD4+	0.84 (95% CI: 0.78 – 0.91); p <0.0001
T_N/T_M CD8+	0.88 (95% CI: 0.80 – 0.96); p=0.01
T_N CD8+	0.91 (95% CI: 0.83 – 1.01); p=0.07
T_N CD4+	0.83 (95% CI: 0.77 – 0.89); p<0.0001
T_{CM} CD4+	1.12 (95% CI: 1.04 – 1.20); p =0.004

Table 3-7: Summarizing the associations of the candidate and existing ARIP measures with chronological age, biological age, and multimorbidity.

	Chronological Age	Biological Age	Multimorbidity	Mortality
Existing ARIP measure				
CD4/CD8	No	No	No inverse association with increasing levels of multimorbidity	No
CD8+ T_N	Yes (inverse)	No	No inverse association with increasing levels of multimorbidity	No
CD4+ T_N	Yes	Yes	Yes (inverse association with increasing levels of multimorbidity)	Yes
CD4+ T_{CM}	No	Yes	Yes (positive association with increasing levels of multimorbidity)	No
Candidate ARIP measure				
CD8+ T_N/T_M	Yes	No	No inverse association with increasing levels of multimorbidity	No
CD4 T_N/T_M	No	Yes	Yes (inverse association with increasing levels of multimorbidity)	Yes

Chapter 4 : Gene expression for aging-related immune phenotype (ARIP) marker.

A. Background

In Chapter 3, we identified a marker for the aging-related immune phenotype (ARIP) by studying associations of existing and candidate ARIP markers with chronological age, biological age, multimorbidity, and mortality. CD4+ T_N and CD4+ T_N/T_M had similar and strongest associations with biological age, multimorbidity, and individual chronic conditions used to define multimorbidity and mortality. In this study, we will identify gene expression profiles associated with this ARIP marker.

Genome-wide gene expression profile is referred to as Transcriptomics and can be utilized to study the mechanisms by which biological pathways are associated with an aging immune system. Transcriptomics helps us understand gene activity and expression, and develop a comprehensive picture of specific biological pathways in which the activated genes are involved. Transcriptomics has been utilized to identify gene expression profiles associated with hypertension, Alzheimer's disease, type II diabetes, and Parkinson's disease (129–133). However, the age-related gene expression changes in the blood induced by immune aging remain unclear. Immune aging associated changes in the gene expression in the blood may help us identify the molecular basis and the molecular consequences of the aging immune system.

Transcriptomics is the study of a complete set of transcripts in a cell and the level of their expression in a cell. Transcriptional profiles related to aging have been studied in various cell subsets such as T cells, T-cell subsets, PBMCs, and whole blood previously.

Gene expression profiles compared between four aged and five healthy young adults in CD8+ T cells showed that the differentially expressed genes were mainly involved in response to oxidative stress and cytokines, apoptosis, and the MAPK signaling cascade (134). Gene expression patterns studied in CD4+ T cells from 31 donors with ages ranging from 25-81 years showed that NF-κB induced pathway was the most up-regulated in older individuals compared to younger individuals, and mRNA processing and splicing pathways were most down-regulated in older participants (135). Transcriptomic profiles studied in Multi-Ethnic Study of Atherosclerosis (MESA) in purified CD4+ T cells among 423 participants identified 188 differentially expressed genes with chronological age which has suggestive evidence constituting the ribonucleoprotein complex and immune response pathways (136). Gene expression associations with age have also been studied previously. A whole blood gene expression meta-analysis study performed on 14,983 individuals of European ancestry identified 1497 genes that are differentially expressed with chronological age and the gene expression profiles were used to estimate a "transcriptomic clock" of individuals. Individuals with higher transcriptomic age compared to chronological age had higher blood pressure and total cholesterol levels (72). Previous studies have identified gene expression profiles in CD4+ and CD8+ T-cell subsets which do change with age, however, identifying gene expression profiles for an optimal threshold for an ARIP marker which has been associated with biological age, multimorbidity and mortality has not been done before and will be more meaningful. This is because biological systems including the immune system and gene expression have redundancy and are resilient to smaller changes in physiological functioning. Using a threshold can help us understand

what the biological impact of larger variations in the ARIP marker could be. Identifying a gene signature for this threshold may help predict individuals at higher risk for age-related chronic diseases and identify individuals who warrant specific attention and closer monitoring for interventions.

In this chapter, we defined an optimal threshold for the ARIP biomarker (CD4+ T_N) that can be used to define ARIP at a population level from Chapter 3 and identified gene expression signatures of the ARIP marker with genes that are differentially in the Health and Retirement study in about 3749 participants. We also used gene set enrichment analysis to elucidate the biological functions of the differentially expressed genes and identified biological pathways among the differentially expressed genes.

B. Methods

B.1. Study Population

The study populations and description of the study design of the Health and Retirement Study (HRS) are given in Section A of Chapter 2. Whole blood was collected from HRS participants in the 2016 wave as part of the Venous Blood Study (Section B of Chapter 2). Among the participants from whom whole blood was collected, immunophenotyping was performed among 9932 participants. A subsample of 3749 participants was selected to perform RNA sequencing. After removing participants with missing data on age, sex, race/ethnicity, CMV seroprevalence, and missing data on CD4+ T_N, 3513 participants were included in the analysis.

B.2. Blood collection

The blood collection procedure is as described in Section B of Chapter 2. 2.5mL of blood collected in Paxgene RNA tubes is used for RNA-seq. These blood samples are processed and stored as stabilized RNA for sequencing(89).

B.3. Defining an optimal cut-off point for the ARIP marker

The ARIP marker identified in Chapter 3 was used for identifying gene expression profiles. Briefly, we defined two ARIP candidate markers and compared them with existing ARIP markers by evaluating associations with chronological age, biological age, multimorbidity, and mortality. The existing marker CD4+ naïve T cells (CD4+ T_N) and CD4+ T_N/T_M had the strongest associations with biological age, multimorbidity, and mortality. As CD4+ T_N and CD4+ T_N/T_M had a high correlation and had similar associations with biological age, multimorbidity, and morality, we used CD4+ T_N as the ARIP marker for the subsequent chapters in this dissertation. Instead of using the CD4+ TN as a continuous variable, we identified a threshold level for CD4 + T_N to determine immune aging at a population level. Since critical biological systems such as the adaptive immune system have multiple redundancies, minor changes in CD4+TN cells may not impact biological processes significantly. Identifying a threshold level for CD4+ T_N at the population level can help us evaluate a specific threshold at which the ARIP marker has a potent impact on biological systems and gene expression. As proof of concept, we have demonstrated in our previous study that dividing the gene expression levels into quartiles and splitting the fourth quartile into two categories could help us evaluate if there was a threshold effect beyond which gene expression levels in the pathogen recognition pathway could affect lung function(137). To define an optimal cut-off point

for this ARIP marker, we used two methods. For the first method, we calculated the percentile of participants who had a CD4/CD8 ratio < 1, and participants with CD4+ T_N lower than this percentile value were considered immunosenescent. We used CD4/CD8 ratio<1 to determine ARIP as previous studies have shown that CD4/CD8 ratio < 1 is associated with aging outcomes including mortality(138). For the second method, we used a ROC curve from the association between CD4+ T_N and multimorbidity to calculate the point with the shortest distance to the (0, 1) point in the ROC curve.

B.4. Library preparation, RNA-sequencing, and analysis

All the samples were sequenced as 50 base pair single read sequences with a minimum of 20 million reads per sample on NovaSeq. We also performed ribosomal RNA (rRNA) and globin mRNA depletion before sequencing. All the RNA sequencing samples were analyzed using the HRS RNA-seq analysis pipeline. This pipeline was based on the TopMed/GTEX analysis pipeline. STAR (139) was used for alignment of the RNA-seq reads to the GRCh38 (hg38) human reference genome obtained from GENCODE along with GENCODE annotations. Quality control metrics after alignment are calculated using RNASeQC (140). SAMTools (141) and RSEM (142) will be used to obtain gene read counts for further analyses.

B.5. Gene counts to differential expression analysis

Normalization/Transformation: Count normalization for mapped reads for each gene was done to scale raw count values so that the raw count reads for genes are proportional to the expression of RNA. The size factor method implemented by the R package DESeq2 (143) was used for normalization of raw counts for sequencing depth differences. DESeq2's default method to normalize read counts to account for differences

in sequencing depth is implemented by calculating size factors. Size factors were calculated by determining the geometric mean of read counts for all genes across all samples yielding a "pseudo-reference", i.e. one value per gene. Each read count for a particular gene across each sample is divided by this pseudo-reference value yielding a ratio. The median of these ratios for each sample was determined which is the size factor. Normalized count values were then obtained by dividing each raw count value by the sample's normalization factor (size factor).

Filtering: In gene expression studies, filtering for lowly expressed genes is important as they provide very little evidence for differential expression and may not be important from a biological standpoint either. They can also interfere with the multiple testing burden while estimating false discovery rates, reducing the power to detect differentially expressed genes. In this analysis, we used filtering using the log2cpm (log_2 transformed counts per million) parameter. We only used genes with log2cpm values >=5 in at least 1% of the participants. Log2cpm of 5 was used as the correlation values of counts of blinded duplicates were high after restricting the counts to log2cpm >=5. The average Pearson correlation coefficient among 33 blinded duplicate pairs was 0.96 after restricting to genes with log2cpm counts higher than 5 (n=2377 genes) compared to a correlation coefficient of 0.90 without the restriction (n=50611 genes). After applying this cut-off, there were 9442 genes included in the analysis.

Differential expression

From the differential gene expression analysis, we estimated the significance of the differential gene expression for immunosenescent participants compared to participants who are not immunosenescent. In this process, we also corrected for multiple

testing due to a large number of genes analyzed. The differential expression analysis in this chapter was performed using the functions in the R package DESeq2. The association between the gene counts and the ARIP marker was modeled using a negative binomial model where the gene counts are assumed to have a negative binomial distribution. The difference in gene expression with the ARIP marker was tested using Log-Likelihood ratio tests. Multiple testing correction was performed using the Benjamini-Hochberg (144) method. Differentially expressed genes were defined as those with an adjusted p-value ≤ 0.01 and an absolute log fold change (log2FC) of >=0.25. A cut-off of 0.25 for the log2foldchange was chosen by calculating the log2FC between each pair of technical duplicates to find the minimum variation found between similar samples in this study. The median log2FC for the technical variables was 0.21 (0.06-0.31) and hence a log2FC cut-off of 0.25 was chosen. The model was adjusted for age, sex, race/ethnicity, CMV status, and sequence run numbers. The models were not adjusted for differential cell composition to avoid over-adjustment as we were evaluating gene expression profiles for an immune cell type. Once candidate genes were identified using the genome-wide approaches, additional analyses adjusting for differential cell composition and other covariates were performed to confirm the initial associations. The differential gene expression results were visualized using an MA plot and a volcano plot. The MA plot was plotted by using log fold change versus the mean expression strength of the genes. The ability of the algorithm to detect differential expression (adjusted p-value ≤ 0.01) was also visualized in the MA plot. A volcano plot was plotted between the adjusted p-value and log fold-change and displays the statistical significance of the difference relative to the magnitude of difference for every single gene in the comparison.

B.6. Gene set enrichment and pathway analysis

Gene set enrichment analysis: Gene set enrichment analysis (GSEA) (145) was performed to elucidate the pathways constituted by the differentially expressed genes. The basic assumption of GSEA is that although large changes in individual genes can have significant effects on pathways, weaker but coordinated changes in sets of functionally related genes (i.e., pathways) can also have significant effects. Differentially expressed genes along with their log fold change values were used for this analysis. We used GSEA to study enrichments of GO (Gene Ontology) pathways, KEGG pathways (https://www.genome.jp/kegg/pathway.html), Molecular Signatures Database (MSigDB) [Broad Institute] (145) containing human gene sets divided into nine collections. This was implemented using the GSEA software for Windows (146). GSEA pre-ranked was conducted using the results from the DESeq2 where rank for the genes was calculated by using the formula: ($-\log_{10}$p-value*sign of log2FC). An FDR cut-off of 0.25 was used to identify statistically enriched biological pathways as suggested by the GSEA manual (147,148).

B.7. Statistical analysis

The first method of using the percentile value of CD4/CD8 ratio < 1 to define an optimal cut-off for CD4+ T_N yielded a cut-off of 15.3%. The second method of identifying a point with the shortest distance from (0,1) point on the ROC curve for the association between CD4+ T_N and multimorbidity also yielded a cut-off of 15.3%. Hence, participants with a CD4+ T_N value lower than 0.15 were considered immunosenescent and participants with CD4+ T_N higher than 0.15 were considered non-immunosenescent. Study characteristics of the participants measured (i.e., sex, race,

smoking status, BMI, CMV status, differential cell composition) in 2016 were examined across categories of immunosenescence in 2016 as mean ± SD or percentages.

C. Results

C.1. Study characteristics

The percentage of Non-Hispanic black participants was higher (27.75% vs 15.62%; $p<0.0001$) and the percentage of CMV seropositive participants were higher (88.04% vs. 68.16%; $p<0.0001$) among immunosenescent participants (Table 4-1). The other variables were not associated with the ARIP marker.

C.2. Differential expression analysis

DESEq2 was used to perform differential expression analysis to identify genes differentially expressed in immunosenescent participants compared to non-immunosenescent participants. The model was adjusted for age, sex, race/ethnicity, CMV, and sequence run numbers as a continuous variable. 292 genes were differentially expressed and the top 10 differentially expressed (based on adjusted p-value) protein-coding genes are shown in Table 4-2. The MA plot between the mean normalized counts of the genes and the log2fold changes showed that the mean counts of the genes were evenly distributed for the high and low fold changes (Figure 4-1). From the volcano plot, *CA6* gene has the lowest p-value and highest fold change (Figure 4-2). When CD4+ T_N was used as a continuous variable, there were 1482 differentially expressed. There were 281 genes in common between differentially expressed genes for binary CD4+ T_N and continuous CD4+ T_N (Figure 4-3).

C.3. Gene set enrichment analysis

Gene set enrichment analysis was performed using the GSEA pre-ranked feature where the rank for each gene was generated by using the formula $-\log_{10}$(adjusted p-value)*sign of log2FC. The gene sets used for GSEA included KEGG, Reactome, Wiki, and Biocarta pathways. Only one pathway was upregulated in immunosenescent participants with an FDR of 0.14 (Transient Receptor Potential channels pathway). A few of the upregulated pathways in immunosenescent participants which had a p-value lower than 0.05 were cell cycle pathways, diseases of the mitotic cell cycle, and steroid biosynthesis. Some of the downregulated pathways in immunosenescent participants which had a p-value lower than 0.05 were nitrogen metabolism, VEGF signaling pathway, and pathways in the pathogenesis of cardiovascular disease. However, none of these upregulated or downregulated pathways had an FDR cut-off lower than 0.25.

D. Discussion

In this study, we defined an optimal cut-off point for the immunosenescence marker CD4+ T_N using two methods that can be tested in further studies for its predictability of aging outcomes. Using the defined cut-off for CD4+ T_N, we identified 292 differentially expressed genes in immunosenescent participants compared to non-immunosenescent participants after adjustment for age, sex, race/ethnicity, CMV status, and batch effects. However, there was no significant enrichment of any biological pathways in these differentially expressed genes.

The differentially expressed genes with the highest association and significance were downregulated in immunosenescent participants. *CA6* had the strongest association with CD4+ T_N and it is a carbonic anhydrase (CA) that is involved in the hydration of

carbon dioxide into bicarbonate and protons (149). CAs have been associated with tumor activity and inhibitors of specific isoforms of CA have been explored previously as anti-cancer therapy(150,151). CAII was found to be increased in mitochondria from the middle-aged brain and skeletal muscle and increased in the brain of young adult mice undergoing neurodegeneration(152). Also, *C.elegans* had a shorter lifespan when exposed to CAII suggesting its role in aging(152). *CA6* is a secretory isozyme of CAs and is secreted into the saliva however previous studies have no clear roles or processes that *CA6* is involved in except for maintaining homeostasis on oral cavity surfaces and upper alimentary canal(153). Amongst the top differentially expressed genes were *TLR2* and *RFTN1* which are major components in the toll-like receptor pathway or pathogen recognition receptor pathway. *RFTN1* activates the *TICAM1* signaling by assisting the internalization of *TLR4* in immune cells including macrophages and dendritic cells (154). *TLR2* recognizes conserved patterns from microorganisms which are known as pathogen-associated molecular patterns (PAMPs) which lead to an immune response by immune cells such as dendritic cells and macrophages. TLR signaling has been associated with T-cell immune responses previously where TLRs can act as co-stimulatory molecules for T-cell proliferation, survival, and cytokine production(155). Some of the functions of the other differentially expressed genes include hydration of carbon dioxide and metabolism(149), targeting protein kinases to the plasma membrane(156), and negative regulation of T-cell differentiation in thymus (157). The differentially expressed genes having these different functions may indicate the mechanisms through which immune aging occurs in older individuals.

A study performed on peripheral blood cells of healthy donors to identify gene signatures for Helper T cells or Th cells found a gene signature for Th cells consisting of 492 genes(158) out of which 33 genes which associated with the continuous ARIP marker in this study. Although the current study focused on the naïve subset of the Th cells, the naïve subset has the highest proportion in the Th cells.

The strengths of the study include the measurement of gene expression levels on a large sample size of 3500 participants, a representative sample including Hispanic, black, and white individuals, and immunophenotyping data available for a large sample size (about 8000 individuals). The limitations of the study include the measurement of immunophenotyping and gene expression levels at the same point restricting our ability to dissect the direction of associations. The participants included in the Venous blood study (VBS) and the sub-sample of participants included for measurement of RNA sequencing comprise a smaller subset of the HRS sample.

In conclusion, we were able to identify genes that were associated with an optimal threshold level of an immune aging marker. Although there were no enriched pathways associated with these genes, some of the biological functions of the genes which had strong associations with the ARIP marker may help us understand immune aging better.

Table 4-1: Descriptive statistics of participant characteristics of 3513 participants with immunophenotyping and gene expression data in the Health and Retirement Study.

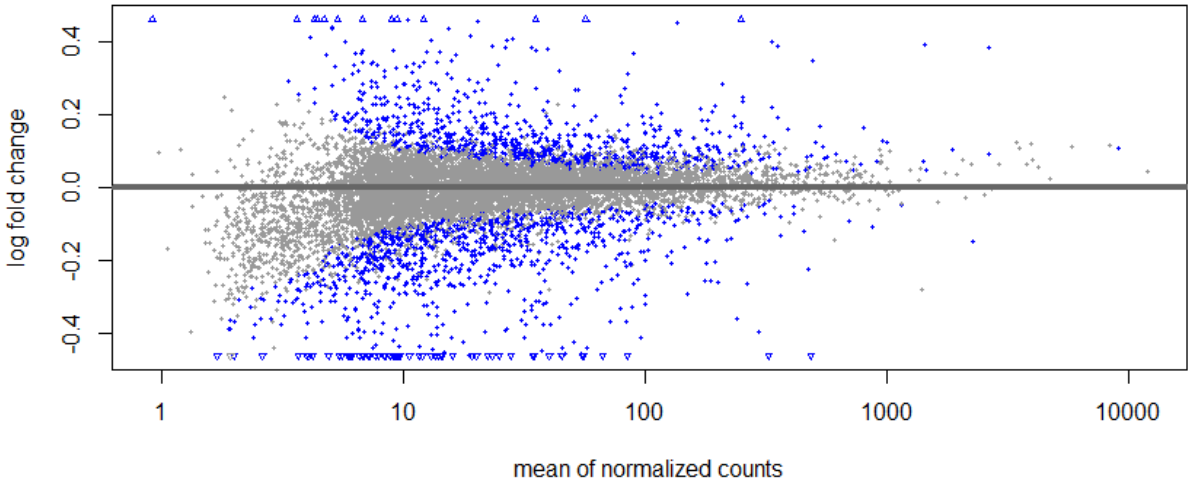
Characteristics	**Immunosenescent** (n=209) (Mean of CD4+ T_N = 0.10±0.04)	**Non-immunosenescent** (n=3304) (Mean of CD4+ T_N = 0.45±0.16)	**p-value**
Demographics			
Age (years) (Mean ±SD)	69.92 ± 9.17	70.55 ± 9.26	0.34
Sex – Female (%)	51.67%	58.38%	0.07
Race			
Hispanic (%)	12.44%	8.84%	<0.0001
Non-Hispanic Black (%)	27.75%	15.62%	
Non-Hispanic White (%)	48.80%	67.77%	
Lifestyle characteristics			
Smoking Status			
Never smokers (%)	8.69%	10.93%	0.62
Former smokers (%)	45.11%	44.93%	
Current smokers (%)	46.20%	44.15%	
BMI (kg/m^2) (Mean ±SD)	30.00±5.9	30.55±6.21	0.22
Biological characteristics			
Cytomegalovirus seroprevalence – Reactive (%)	88.04%	68.16%	<0.0001

Table 4-2: List of top 10 differentially expressed protein-coding genes based on p-values for the ARIP marker from DESeq2.

Gene name	**Base Mean**	**Log_2fold change[a]**	**p-value**
CA6	45.22	-1.52	$3.35x10^{-124}$
RFTN1	327.34	-0.51	$3.83x10^{-49}$
DNAJC16	40.42	-0.52	$2.55x10^{-43}$
AKAP7	45.69	-0.61	$7.45x10^{-38}$
AMIGO1	96.41	-0.41	$1.84x10^{-36}$
ZC3H8	54.77	-0.46	$4.68x10^{-28}$
PLB1	9.06	-0.82	$9.02x10^{-27}$
TLR2	7.23	-1.09	$4.93x10^{-25}$
GIMAP8	23.52	-0.49	$9.21x10^{-25}$
KRT73	4.22	-1.06	$1.74x10^{-23}$

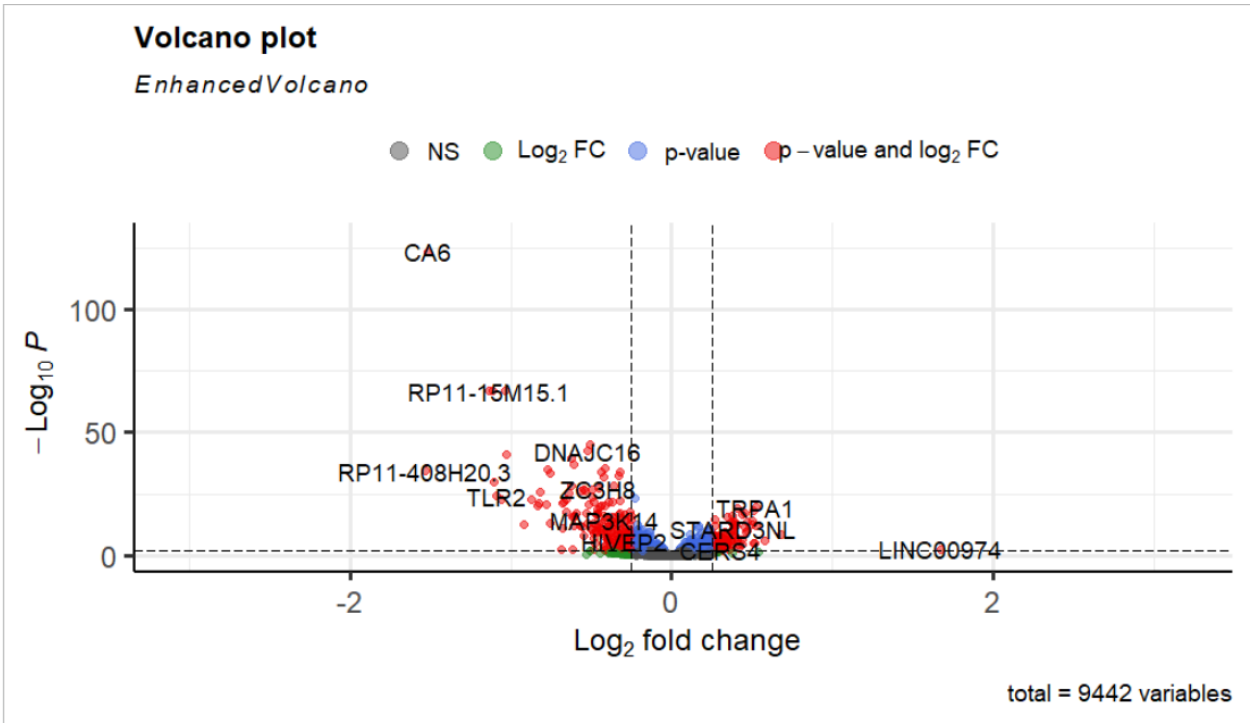
[a]Log_2 fold change is calculated for immunosenescent participants vs non-immunosenescent participants

Figure 4-1: MA/Bland Altman plot between log fold change and average normalized expression of the genes[a].



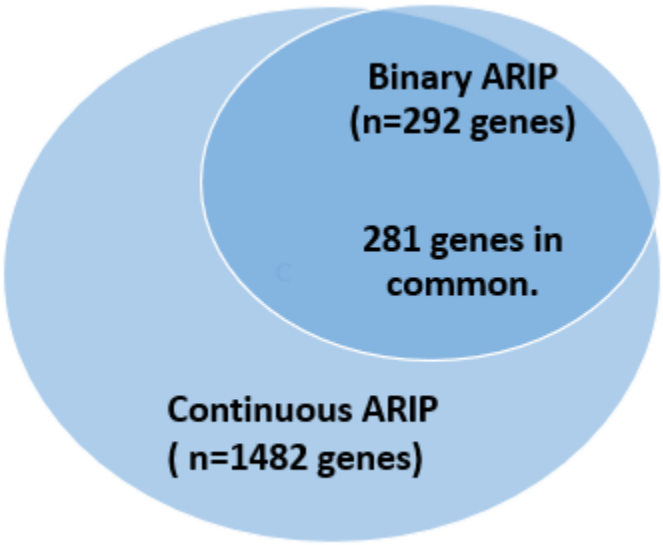
[a]The blue dots represent genes which have an adjusted p-value <0.05.

Figure 4-2: Volcano plot between p-values and log fold changes of the genes.



Note: A -log10 p-value of 2 represents an adjusted p-value (FDR) of 0.01. Red dots represent genes with adjusted p-value < 0.01 and absolute log2FC > 0.25.

Figure 4-3: Differentially expressed genes that were common between binary CD4+ TN and continuous CD4+ TN



Chapter 5 : Gene expression profiling of multimorbidity and aging-related immune phenotypes

A. Background

Age related non-communicable chronic diseases such as cardiovascular diseases, lung diseases and metabolic disorders have been the major public health challenge of the 21st century. Multimorbidity, the co-occurrence of two or more chronic diseases in the same individual that increases disability, healthcare utilization, psychological distress and lowers quality of life (12,79,80), is also increasing among aging individuals with a prevalence of more than 60% in people aged 65 years and above (159,160). An underlying cause for the increased incidence of these diseases in the elderly population is thought to be the acceleration of physiological dysregulation in some individuals compared to others. Multimorbidity as an aging outcome can be used to study age acceleration as the increased susceptibility to multiple chronic diseases may represent the manifestation of dysfunctional homeostatic mechanisms. The adequate functioning of these homeostatic mechanisms is essential for maintaining normal physiological process and the dysfunctions are thought to be the underlying cause of several age-related chronic diseases. Age-related decline in both baseline levels of physiological process and its ability to respond to acute stressors is considered the major risk factor for the increased incidence of chronic diseases, higher functional decline, and lower biochemical balance. The public health burden of multimorbidity and the necessity to understand the molecular mechanisms of development of multiple diseases in older individuals requires development of new aging biomarkers which help understand the underlying physiological dysfunction associated with aging. An important component of aging is the

declining immune system represented by reduction in naïve T (T_N) cells and increase in memory T (T_M) cells contributing to an aging-related immune phenotype (ARIP). The correlation of ARIP with chronological age in known due to the declining immune system observed even in centenarians (161). Gene expression levels allow us to study, in an unbiased fashion, all age-related changes and physiological decline (72,75,162). We have identified gene expression profiles for CD4+ T_N (ARIP marker identified in Chapter 3). In this chapter 5, we assessed if the ARIP marker gene expression profiles are also associated with multimorbidity and if the ARIP marker and multimorbidity share commonly associated genes. The shared common genes would support the theory of a common "aging signature" which can be used to identify individuals at higher risk for diseases. Since gene expression is modifiable, identifying an aging signature can potentially be used to modify or slow occurrence of diseases in older individuals. Understanding shared mechanisms between aging immune system and multimorbidity will also help understanding the aging process better.

There are few studies conducted to study molecular biomarkers for multimorbidity. In the InCHIANTI longitudinal study from Italy, inflammatory and hormonal biomarkers examined among older individuals showed that higher baseline IL-6 and steeper increase of IL-6 levels were associated with steeper increase in multimorbidity (defined as number of diagnosed diseases from a predefined list of 15 candidate chronic conditions) over nine-year longitudinal follow-up (81). However, the study was limited to a specific demographic, including only people living in the Chianti area in Italy, and IL-6 is a specific inflammatory biomarker which does not represent an integrated aging signature. Among age-associated molecular biomarkers, the association

of telomere length and multimorbidity was examined in the Health and Retirement study where longer telomeres were associated with lower risk of multimorbidity but prospective analysis suggested no association (83). Transcriptome has previously been used to study aging. A whole blood gene expression meta-analysis study performed in 14,983 individuals of European ancestry identified 1497 genes which are differentially expressed with chronological age and the gene expression profiles were used to estimate a "transcriptomic clock" of individuals. Individuals with higher transcriptomic age compared to chronological age had higher blood pressure and total cholesterol levels (72). Several previous studies have also identified aging gene expression profiles but have been limited by small sample sizes and non-generalizability of the sample (73–75).

Gene signatures for multimorbidity have not been defined previously in a large population study. An improved understanding of biological mechanisms which are involved in multimorbidity and which characterize the relationship between aging immune system and multimorbidity will help us understand the mechanisms by which an aging immune system influences multimorbidity. In this chapter, the primary objective is to identify gene expression signatures associated with multimorbidity and identify biological pathways and mechanisms which the gene expression signatures constitute. We also identified gene expression signatures for each of the chronic conditions separately to assess if there were overlapping genes with the gene expression signature for the ARIP marker (described in Chapter 4). We also performed mediation analysis to assess if the association between ARIP and multimorbidity was mediated by the overlapping genes between the ARIP marker and the chronic conditions. We also evaluated if these genes were associated with mortality in 2018.

B. Methods

B.1. Study population

The study population and study design of the Health and Retirement Study (HRS) is given in Section A of Chapter 2. Whole blood was collected from HRS participants in the 2016 wave as part of the Venous Blood Study (Section B of Chapter 2). Among the participants from whom whole blood was collected, immunophenotyping was performed among 9932 participants. A subsample of about 3749 participants were selected to perform RNA sequencing. After removing participants with missing data on age, sex, race/ethnicity, CMV seroprevalence, multimorbidity and CD4+ T_N, 3138 participants were included in this study.

B.2. Blood collection

The blood collection procedure is as described in Section B.6. of Chapter 3. 2.5mL of blood collected in Paxgene RNA tubes are used for RNA-seq. These blood samples were processed and stored as stabilized RNA for sequencing(89). The details of the RNA-sequencing procedure have been mentioned in Section B.4. of Chapter 4.

B.3. ARIP marker

As described in Chapter 4, ARIP was defined using the cut-off for CD4+ T_N as 1.53%. Participants whose CD4+ T_N values were lower than 15.3% were considered immunosenescent and participants whose CD4+ T_N were higher than 15.3% were considered non- immunosenescent.

B.4. Ascertainment of multimorbidity and mortality

Ascertainment of multimorbidity is described in section C of Chapter 2. Briefly, type II diabetes (defined using measured blood glucose levels among participants), and self-reported stroke, lung disease, heart disease and cancer were used to describe multimorbidity outcome in 2016. Multimorbidity was used as a continuous variable with score of 0 to 5 and as a categorical variable with two levels (Score of 0 or 1 – No Multimorbidity; Score of 2 or 3 or 4 or 5 – Multimorbidity). Mortality assessed at 2018 was used for this analysis. HRS uses two resources to report mortality: proxy interviews and using the National Death Index (NDI) information.

B.5. Measurement of covariates

Variables that are used as covariates are age, sex, race, smoking status, Cytomegalovirus (CMV) status, Body Mass Index (BMI) and percentage neutrophils and monocytes. Age (in years), sex (female/male), race (Hispanic, Non-Hispanic Whites, Non-Hispanic Black and Other) were obtained from the demographics data. Smoking status has three categories: never smokers, former smokers, and current smokers. Height was measured in inches by having the respondent stand against a wall, without shoes. A mark was made on a post-it on the wall by the interviewer, who then measured the distance from the floor to the mark. Weight was measured in pounds by asking a respondent to step on a Healthometer 830KL scale. BMI was calculated by using the formula (Weight in lbs/ (Height in inches)^2) *703. CMV seroprevalence was measured using IgG antibodies to CMV in serum using the Roche e411 immunoassay analyzer (Roche Diagnostics Corporation). The results were reported as non-reactive (<0.5 COI), borderline (0.5 to <1.0 COI) or reactive >=1.0 COI) where COI is cut-off interval.

Differential cell composition, which included % neutrophils and % monocytes was measured in EDTA whole blood using a Sysmex XE-2100 instrument, (Sysmex America, Inc., Lincolnshire, IL).

B.6. Statistical Analysis

We used SAS version 9.4 (SAS Institute, Inc., Cary, NC) and R Statistical Analysis software version 4.0.0 for all the analyses.

Study characteristics of the participants measured (i.e., sex, race, smoking status, BMI, CMV status, differential cell composition) in 2016 were examined across categories of multimorbidity in 2016 (Score of 0 or 1 and Score of 2-5) as mean ± SD or percentages.

To analyze the association between gene expression levels and multimorbidity in 2016, we used a multivariate logistic regression model. Multimorbidity, the outcome measure, was used as a binary variable (Score of 0 or 1 and Score of 2 and above). The main exposure is the gene expression counts which were represented as $\log_2$ transformed counts per million (CPM) and genes with log2cpm >=5 in at least 1% of the participants were included for the analysis. Log2cpm of 5 was used as the correlation values of counts of blinded duplicates were high (average correlation among 33 blinded duplicate pairs = 0.85) after restricting the counts to log2cpm >=5. After applying this cut-off, there were 9442 genes included in the analysis. The logistic regression analysis was adjusted for age, sex, race/ethnicity, CMV status and sequence run numbers. These models were also used to study associations between the individual chronic conditions and gene expression. Multiple testing correction was performed using the Benjamini-Hochberg (144) method and adjusted p-values (False Discovery Rate) ≤ 0.01 was used to select genes for

downstream analysis. Gene set enrichment analysis was performed after studying these associations to identify any statistically enriched pathways with multimorbidity. An FDR cut-off of 0.25 was used to identify statistically enriched biological pathways as suggested by the GSEA manual (147,148). Once the associated genes for each of the chronic conditions and multimorbidity were identified, the overlapping genes associated with each of the chronic conditions and associated with the ARIP marker were used for mediation analysis. The mediation analysis estimated the natural indirect effect or the mediation effect, defined as the fraction of total effect between ARIP marker and multimorbidity explained by gene expression. Figure 5-1 depict the sequential steps to perform the mediation analysis. Total effect between the ARIP marker and multimorbidity is represented by b (as shown in figure 5-1a) has been estimated in Chapter 3. The direct effect represented by "b1" in figure 1b was estimated by the coefficient of the ARIP marker in a model with multimorbidity as outcome and ARIP as predictor after adjustment for gene expression. The indirect or mediation effect was estimated as a product of" b2" and "b3" where "b2" was estimated in Chapter 4 and "b3" was estimated in Chapter 5. The association between these overlapping genes and mortality were also assessed validate that they could be used as an aging signature and meditation analysis was performed with mortality as described before.

C. Results

C.1. Study characteristics

Among the 3138 participants included in the study, 642 (20.5%) participants had multimorbidity. Participants who had multimorbidity had higher age (73.14 vs 69.27 years), lower female participants (52.49% vs. 58.93%), higher BMI (31.1 vs 29.76 kg/m^2), higher former smokers (52.02% vs. 43.19%), higher percentage of CMV seroprevalence (75.86% vs. 67.27%), higher percentage neutrophils (60.13% vs.56.67%) compared to participants who did not have multimorbidity in 2016 (Table 5-1).

C.2. Gene expression profiles for multimorbidity

Logistic regression models were used to identify genes associated with multimorbidity after adjustment for age, sex, race/ethnicity, CMV status and batch effects. After using a cut-off for adjusted p-value of < 0.01, there were 1998 genes associated with multimorbidity. The top ten associated genes based on adjusted p-values are shown in Table 5-2. After using gene set enrichment analysis (GSEA) software to identify biological pathways, there was no statistically enriched pathways with an FDR value < 0.25.

C.3. Gene expression profiles of individual chronic conditions used to define multimorbidity

Gene expression profiles for heart disease, diabetes, lung disease, stroke and cancer were identified individually. After applying a cut-off for adjusted p-value of <0.01, there were 1558 genes associated with heart disease, 802 genes associated with diabetes, 1172 genes associated with lung disease, 621 genes associated with cancer and no genes associated

with stroke. There were 93 overlapping genes associated with these chronic conditions (Figure 5-2).

C.4. Genes associated with ARIP marker and multimorbidity

There were ten genes commonly associated with each of the chronic conditions and the ARIP marker CD4+ T_N. Out of these 10 genes, five genes were protein coding genes which were *AMIGO1, ZSCAN32, KRT73, KCNK12* and *MTRNR2L4*. We further adjusted the models with multimorbidity with smoking status, BMI, percent neutrophils and monocytes and showed that all the five genes still had consistent association with multimorbidity after further adjustment (Table 5-3). The five genes are inversely associated with multimorbidity. For every unit increase in log2cpm counts of *AMIGO1, ZSCAN32, KCNK12, KRT73* and *MTRNR2L4*, the odds of multimorbidity is lower by 47%, 30%, 29%, 24% and 20%, respectively (Table 5-3), after adjustment for age, sex, race/ethnicity, CMV status, smoking status, BMI, percent neutrophils, percent monocytes and batch effects. *AMIGO1* had the strongest association with multimorbidity. While using gene expression profiles of CD4+ T_N as a continuous variable, there were 33 genes in common with the genes associated with each of the chronic conditions. Among these 33 genes, 22 were protein coding and included *AMIGO1, ZSCAN32, KCNK12, KRT73* and *MTRNR2L4*.

C.5. Association with mortality

There were 3133 participants included for this analysis, there were 139 (4.4%) participants who were not alive at 2018. The five genes were also associated with mortality in 2018. After adjustment for age, sex, race/ethnicity, CMV status, smoking status, BMI, percent neutrophils and monocytes, *AMIGO1* had the strongest inverse

association with mortality. For every one unit increase in *AMIGO1*, the odds of mortality is lower by 52% (Table 4). The other genes were also inversely associated with mortality. For every one unit increase in log2cpm counts of *ZSCAN32, KCNK12, KRT73* and *MTRNR2L4*, the odds of mortality is lower by 37%, 35%, 43% and 37%, respectively (Table 5-4).

C.6. Characteristics of the five genes

The five genes were associated with age, sex, race, CMV status and smoking status. The mean expression level of the five genes was higher among women compared to men, higher among CMV seronegative participants, higher among current smokers and had an inverse association with age (Table 5-5). *AMIGO1* was higher amongst Non-Hispanic Black participants, *ZSCAN32* and *MTRNR2L4* was lowest amongst Non-Hispanic Black participants, *KCNK12* was lowest amongst Non-Hispanic White and Hispanic participants, *KRT73* was lowest amongst Non-Hispanic White participants (Table 5-5). The genes had an inverse correlation with C-reactive protein levels and IL-6 levels (Table 5-5).

C.7. Mediation analysis with multimorbidity and mortality

We performed mediation analysis to estimate indirect effects of association of CD4+ T_N with multimorbidity and mortality by the genes. The mediation analysis for multimorbidity was performed with using CD4+ T_N as a binary variable defined using an optimal cut-off point where the mediation analysis for mortality was done using CD4+ T_N as a continuous variable as the binary CD4+ T_N was not associated with mortality. After including the genes in the model with CD4+ T_N as predictor and multimorbidity as the outcome, the direct effects were estimated as shown in Table 5-6. All the direct effects were significant with multimorbidity indicating that there could be incomplete mediation by the genes between CD4+ T_N and multimorbidity. The strongest indirect effect was observed in *AMIGO1* where the proportion of association between CD4+ T_N and multimorbidity mediated by *AMIGO1* was 38.3% with an indirect effect of 0.22 (95% CI: 0.13-0.32). With mortality, the direct effects of *AMIGO1* and *KRT73* were not significant suggesting a significant mediation between CD4+ T_N and mortality by these two genes. *AMIGO1* and *KRT73* had the strongest indirect effect with 41.3% and 51.9% of the association between CD4+ T_N and mortality mediated by these genes (Table 5-6).

D. Discussion

In this study, we have identified an aging signature using genes associated with an immune aging marker (CD4+ T_N) and genes associated with individual chronic conditions which constitute multimorbidity. These genes which constitute the signature are also associated with two-year mortality and two of the genes mediate the association between CD4+ T_N and mortality. Since the five genes, *AMIGO1, ZSCAN32, KCNK12, KRT73* and *MTRNR2L4*, have been associated with heart disease, diabetes, lung disease, cancer and the immune aging marker (CD4+ T_N), their expression could be important in determining accelerated physiological dysfunction in aging individuals and can be validated in other cohort studies.

Current literature on the five genes is sparse. The five genes were inversely associated with immune aging and multimorbidity. The function of *AMIGO1* (Adhesion Molecule With Ig Like Domain 1) has been largely observed in the development of neural axons and regulating dendritic growth (163). *AMIGO1* has also been found to be a marker gene which was significantly downregulated in dorsolateral later prefrontal cortex obtained from Alzheimer's patients in a study using transcriptome data from ROSMAP (Religious Orders Study and Rush Memory and Aging Project)(164). In another study where the relationship of genetic variants in eleven gastric acid secretion pathway genes and gastric cancer risk was assessed, *AMIGO1* had a lower expression in gastric cancer tissue than in normal tissues(165). These results indicate an inverse association between *AMIGO1* and aging-associated chronic conditions. *KCNK12* (Potassium Two Pore Domain Channel Subfamily K Member 12) gene is part of the two-pore domain potassium channels family of genes (166). Some previous studies have shown that

changes in function of the genes in the potassium channels can have an effect on cancer progression(167), and *KCNK12* overexpression has been associated with triple negative breast cancer subtype(168). According to functional annotation provided by the Gene Ontology (GO), *MTRNR2L4* (Mitochondrially Encoded 16S RRNA Like 4) is involved in negative regulation of apoptotic processes, *KRT73* (Keratin 73) has a function in the control of spatial organization of filaments and is a structural component of skin epidermis and *ZSCAN32* (Zinc Finger And SCAN Domain Containing 32) has some DNA-binding transcription factor activity which regulates the transcription by RNA polymerase II(169).

The gene expression levels of the five genes reduced as participants' age increased indicating that they are inversely associated with aging. Consistent with this, the genes were inversely associated with multimorbidity and inversely associated with immune aging. The mean gene expression level was higher in CMV negative compared to CMV positive participants which is consistent with previous studies where CMV seropositivity has been associated with immune aging and accumulation of memory T cells(170). The genes are also inversely correlated with inflammatory markers including C-reactive protein and IL-6 which is consistent with previous studies where increased inflammation is associated with aging-related chronic conditions(171–174).

PIK3CD, CACNA2D2, CACNB2, MAGI3 were among the genes which associated with both continuous CD4+ T_N and individual chronic conditions. They were positively associated with CD4+ T_N and was inversely associated with multimorbidity. Mutations in the gene *PIK3CD* have previously been associated with immunodeficiency disease(175) . This gene encodes p110δ subunit and deletion of this subunit has been associated with

several immune diseases(176,177). The gain of function *PIK3CD* also leads to a rare genetic disorder called PASLI or p110 delta activating mutation causing senescent T cells, lymphadenopathy, and immunodeficiency(178). This disorder also manifests in recurrent respiratory disorders which can potentially lead to damage in the airways (179). These immune implications of *PIK3CD* could be important in understanding how immune aging could affect multimorbidity. *CACNA2D2* and *CACNB2* are involved in calcium signaling which has implications in neurological and cardiac syndromes(180), and in tumorigenesis(181). *MAGI3* was found to be a potential glioma suppressor in previous studies where loss of *MAGI3* led to the proliferation of glioma cells and *MAGI3* overexpression led to the suppression of malignant phenotypes of glioma cells(182,183). The biological functions of the other genes associated with both continuous CD3+ T_N and individual chronic conditions include having a role in cytokinesis, catalyzing the interconversion of citrate to isocitrate in the TCA cycle, regulating cell cycle by acting as a growth suppressor in confluent cells, hydrolyzing sphingomyelin to release ceramide and phosphoryl-choline(184,185).

Cell composition (percentage neutrophils, percentage monocytes) was not used as a covariate in the initial analysis to identify associated genes with multimorbidity or the individual chronic conditions. Adjusting for neutrophils may indicate that we are studying the associations of gene expression as transcripts per each cell or neutrophils, in this case, as opposed to the overall expression of transcripts in an individual. As the interpretation of our findings was focused on studying how the change in gene or transcript expression was associated with health outcomes, we did not include neutrophils to study the transcripts per cell initally.

The strengths of the study include measurement of gene expression levels on a large sample of 4000 participants, representative sample including Hispanic, black, and white individuals and immunophenotyping data available for large sample size (about 8000 individuals). We are also identifying gene expression profiles for multimorbidity and characterizing the relationship between immunosenescence and multimorbidity for the first time. The limitations of the study include the measurement of multimorbidity using self-report of participants. The participants included in the Venous blood study (VBS) and the sub-sample of participants included for measurement of RNA sequencing comprise a smaller subset of the HRS sample which could lead to non-generalizability of the results. Gene expression levels measured in blood may be confounded by differential cell counts and the aging gene signature may not be conserved across various tissues in the body. The cross-sectional association between gene expression and multimorbidity may not completely identify aging signatures as the temporality of association is not certain. We also acknowledge that gene expression represents only a component of an aging signature which needs to be combined with other epigenetic, proteomic, and genetic biomarkers to identify direct drivers of the aging process.

In conclusion, we have identified a gene expression signature of accelerated physiological dysregulation with five genes which could help us understand how aging in the immune system could influence multimorbidity and mortality. These genes need to be further validated in future studies to confirm the associations of these genes with aging-related chronic conditions and immune aging markers.

Figure 5-1: Steps of the mediation analysis.

Figure 5-1a: Step 1 of the mediation analysis

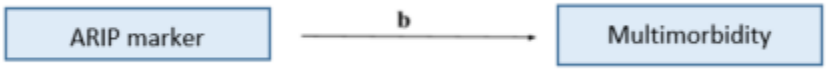


Figure 5-1b: Step 2 of the mediation analysis

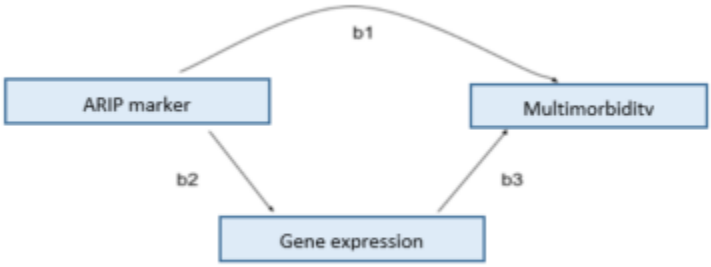


Table 5-1: Descriptive statistics of participants across multimorbidity levels

	Multimorbidity		
Baseline characteristics	Yes (N =642)	No (N =2496)	P-value
Demographic			
Age, years (SD)	73.14±9.03	69.27±9.07	<0.0001
Female (%)	52.49%	58.93%	0.004
Race			
Hispanic (%)	11.68%	14.66%	0.12
Non-Hispanic Black (%)	15.89%	16.23%	
Non-Hispanic White (%)	70.25%	66.11%	
Lifestyle/Medical factors			
BMI, kg/m^2 (SD)	31.10±6.54	29.76±6.07	<0.0001
Smoking status			
Current smoker	37.38%	46.03%	0.0002
Former smoker	52.02%	43.19%	
Never smoker	10.59%	10.78%	
CMV status (% reactive)	75.86%	67.27%	<0.0001
Percentage neutrophils	60.13±10.06	56.67±9.52	<0.0001
Percentage monocytes	8.49±2.60	8.60±2.36	0.30

Table 5-2: Top 10 significant protein- coding genes based on p-values associated with multimorbidity.

Gene name	Base Mean of log2cpm counts	Odds ratio	Adjusted p-value
RNF175	9.10	2.22	$1.55x10^{-19}$
CTBP2	9.72	3.40	$1.55x10^{-19}$
NECAB1	6.24	1.85	$2.43x10^{-15}$
SYCP3	8.30	2.23	$2.43x10^{-15}$
PFKFB2	9.49	1.56	$8.99x10^{-15}$
LHFPL5	7.21	2.24	$8.99x10^{-15}$
KCNH7	8.87	1.74	$1.70x10^{-14}$
IAH1	8.71	3.69	$3.34x10^{-14}$
SLC9B1	7.38	1.61	$6.02x10^{-14}$
COG1	10.44	1.71	$1.32x10^{-12}$

Figure 5-2: Overlap of genes associated with the individual chronic conditions.

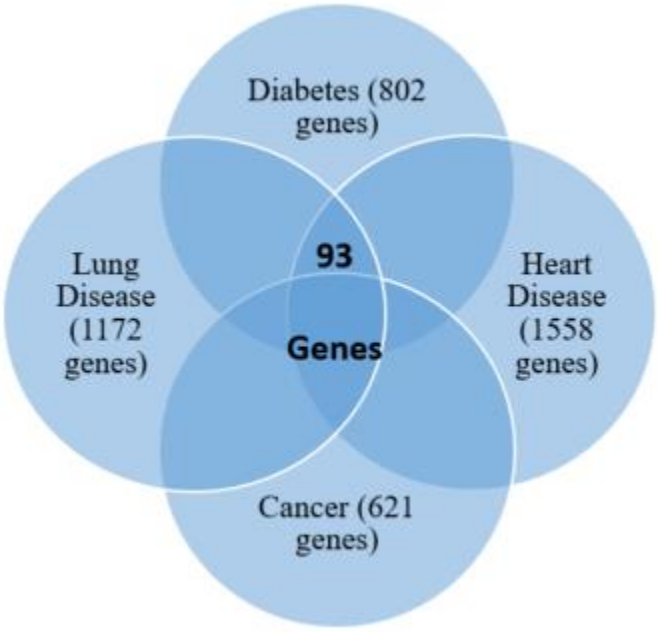


Table 5-3: Association of overlapping genes (between CD4+ TN and individual chronic conditions) with multimorbidity after adjustment for age, sex, race/ethnicity, CMV status, batch effects, smoking status, BMI, percent neutrophils and monocytes.

Gene name	Mean of the gene ± SD	Odds ratio (95% CI)	p-value
AMIGO1	7.26±0.50	0.53 (0.43-0.66)	<0.0001
ZSCAN32	4.94±0.69	0.70 (0.61-0.80)	<0.0001
KCNK12	5.21±0.54	0.71 (0.59-0.85)	0.0002
KRT73	3.10±0.84	0.76 (0.67-0.85)	<0.0001
MTRNR2L4	3.67±0.73	0.80 (0.70-0.91)	0.0008

Table 5-4: Association of overlapping genes (between CD4+ TN and individual chronic conditions) with mortality after adjustment for age, sex, race/ethnicity, CMV status, batch effects, smoking status, BMI, percent neutrophils and monocytes.

Gene name	Odds ratio (95% CI)	p-value
AMIGO1	0.48 (0.33-0.70)	0.0001
ZSCAN32	0.63 (0.49-0.83)	0.0007
KCNK12	0.65 (0.47-0.91)	0.01
KRT73	0.57 (0.44-0.73)	<0.0001
MTRNR2L4	0.73 (0.57-0.94)	0.01

Table 5-5: Association of the five genes with demographics and participant characteristics. The means represent the log2cpm counts.

Characteristics	*AMIGO1* (Mean/ Correlation)	*ZSCAN32* (Mean/ Correlation)	*KCNK12* (Mean/ Correlation)	*KRT73* (Mean/ Correlation)	*MTRNR2L4* (Mean/ Correlation)
Age	p<0.0001	p<0.0001	p<0.0001	p<0.0001	p<0.0001
56-65	7.4	5.0	5.3	3.3	3.7
66-75	7.3	5.0	5.2	3.1	3.7
76-85	7.1	4.8	5.1	2.9	3.6
86+	7.0	4.8	5.1	2.7	3.5
Sex	p<0.0001	p<0.0001	p<0.0001	p<0.0001	p<0.0001
Female	7.4	5.0	5.3	3.2	3.8
Male	7.1	4.8	5.2	3.1	3.6
Race/ethnicity	p<0.0001	p=0.002	p=0.0001	p<0.0001	p=0.0003
Non-Hispanic White	7.2	5.0	5.2	3.1	3.7
Non-Hispanic Black	7.4	4.8	5.3	3.2	3.6
Hispanic	7.2	5.0	5.2	3.2	3.7
Other	7.4	5.0	5.4	3.3	3.8
CMV status	p<0.0001	p<0.0001	p=0.57	p<0.0001	p=0.0007
Positive	7.2	4.9	5.2	3.1	3.6
Negative	7.3	5.0	5.2	3.2	3.8
Smoking status	p<0.0001	0.002	p=0.02	p=0.007	p=0.004
Current smoker	7.4	5.0	5.3	3.1	3.7
Former smoker	7.2	4.9	5.2	3.1	3.6
Never smoker	7.3	5.0	5.2	3.2	3.7
BMI (kg/m^2) (Correlation)	-0.01 (p=0.59)	-0.04 (p=0.02)	-0.01 (p=0.46)	-0.02 (p=0.26)	-0.03 (p=0.08)
C-reactive protein (Correlation)	-0.15 (p<0.0001)	-0.12 (p<0.0001)	-0.11 (p<0.0001)	-0.07 (p<0.0001)	-0.10 (p<0.0001)
IL-6 (Correlation)	-0.06 (p=0.002)	-0.04 (p=0.02)	-0.03 (p=0.05)	-0.04 (p=0.03)	-0.03 (p=0.05)
IL-10 (Correlation)	-0.09 (p<0.0001)	-0.07 (p<0.0001)	-0.06 (p=0.0009)	-0.08 (p<0.0001)	-0.04 (p=0.01)

Table 5-6: Mediation effect of commonly identified genes between ARIP marker and multimorbidity.

Genes	Multimorbidity (CD4+ T_N as a binary predictor)			Mortality (CD4+ T_N as a continuous predictor)		
	Direct effect (b1)	Indirect effect (b2*b3)	Proportion of total effect mediated	Direct effect (b1)	Indirect effect (b2*b3)	Proportion of total effect mediated
AMIGO1	0.36 (95% CI: 0.01-0.72)	0.22 (95% CI=0.13-0.32)	38.3%	-0.81 (95% CI: -1.94-0.33)	-0.57 (95% CI: -1.00 – -0.14)	41.3%
ZSCAN32	0.44 (95% CI: 0.09-0.78)	0.15 (95% CI: 0.08-0.23)	25.4%	-0.98 (95% CI: -2.07-0.10)	-0.41 (95% CI: -0.68--0.17)	29.5%
KCNK12	0.51 (95% CI: 0.16-0.85)	0.07 (95% CI: 0.03-0.12)	12.1%	-1.23 (95% CI: -2.29--0.17)	-0.18 (95% CI: -0.38-0.01)	12.8%
KRT73	0.44 (95% CI: 0.10-0.79)	0.15 (95% CI: 0.07-0.22)	27.0%	-0.75 (95% CI: -1. 85-0.35)	-0.81 (95% CI: -1.28--0.34)	51.9%
MTRNR2L4	0.52 (95% CI: 0.18-0.87)	0.06 (95% CI: 0.02-0.12)	10.3%	-1.24 (95% CI: -2.30--0.17)	-0.19 (95% CI: -0.43-0.04)	13.3%

Chapter 6: Summary

A. Summary of results

Aging-related immune phenotype (ARIP) or the age-related decline in the immune system is a global phenomenon observed in most of the common age-related diseases which limit the ability of the immune system to mount effective immune responses (4,5,8,9). Despite broad agreement among biologists regarding the effect of the aging immune system on morbidity and mortality, an ARIP marker that has been associated with systemic dysregulation of the human body has not been studied before. The mechanisms and pathways behind the association of immune aging with morbidity and mortality have also not been strongly established yet. Gene expression and transcriptomics can be used as a resource to study mediating biological pathways between immune aging and morbidity. The objective of this dissertation is to address these questions by utilizing immune cells, RNA-sequencing, and health outcomes data from the Health and Retirement study.

The main research aims are as listed here. First, we sought to define an ARIP marker that was associated with biological age, multimorbidity, and mortality by creating candidate ARIP markers and comparing them with existing ARIP markers. Second, we aimed to define an optimal cut-off point for the identified ARIP marker, and identify gene expression profiles and biological pathways associated with the ARIP marker. Third, the objective was to identify gene expression profiles for multimorbidity and for the individual chronic conditions which defined multimorbidity. Genes associated in common with the ARIP marker and each of the individual chronic conditions were used

to assess whether specific biological pathways mediated the association between immune aging, multimorbidity, and mortality.

In the first manuscript, we found that the candidate ARIP marker CD4+ T_N/T_M and existing marker CD4+ T_N had a similar strength of association with biological age, multimorbidity, individual chronic conditions which defined multimorbidity, and mortality. The other markers did not have a strong association with biological age or multimorbidity. Overall, these findings indicate that the CD4+ T cells were involved in the mechanisms involved in immune dysregulation with aging.

In the second manuscript, we defined an optimal threshold point for the ARIP marker (CD4+ T_N) using two methods which converged at 15.3% (of overall CD4+ T cells). Using this optimal cut-off, we identified 300 genes that were differentially expressed with the ARIP marker and 1482 genes differentially expressed with the continuous CD4+ T_N marker. After adjustment for multiple comparisons, there were no statistically enriched pathways present. Some of the biological functions of the highly associated genes include pathogen pattern recognition, hydration of carbon dioxide, and negative regulation of T-cell differentiation in the thymus. These findings help us narrow down to a set of genes which could influence immune aging.

In the third manuscript, we identified 1998 genes associated with multimorbidity and 93 genes which were associated in common with each of the individual chronic conditions. Among these 93 genes, there were 5 protein-coding genes associated with the threshold point of CD4+ T_N and 22 protein-coding genes associated with continuous CD4+ T_N. The five genes, which were *AMIGO1, ZSCAN32, KCNK12, KRT73,* and *MTRNR2L4,* associated with the threshold point of CD4+ T_N were also associated with

mortality, and *AMIGO1* and *KRT73* had a significant mediating effect with mortality. *AMIGO1* and *KCNK12* were found to be associated with Alzheimer's disease, gastric cancer, and triple-negative breast cancer subtype in previous studies.

B. Strengths and limitations

A primary strength of this dissertation is the measurement of immune cells and gene expression amongst the same participants, which can be useful for identifying biological pathways which mediate the association between immune aging and other health outcomes. Another important strength of this dissertation is the large sample size used for the measurement of immune cells and the measurement of gene expression levels in the participants. Other strengths include the usage of standardized immunophenotyping methods and a representative sample including Hispanic, black, and white individuals. The novelty of the study includes characterizing the relationship between immunosenescence and multimorbidity using gene expression data. The results from this dissertation provide initial evidence involving a few genes which can illustrate the effect of immune aging on physiological dysregulation and can be explored further in future studies.

There are also some important limitations in our study. The measurement of immune cells was done only for a single time point and measurement of the longitudinal change of immune cells can help us study the temporality of associations. Longitudinal data also helps us understand how the change in the certain immune cell populations within an individual influences health outcomes. The majority of the multimorbidity measure is based on the self-report of the participant and hence could be influenced by factors such as recall, social desirability, and lack of diagnosis. However, the question

posed to participants included having been informed previously by a physician about a health condition. Another important limitation to consider was the impact of medication on immune aging and multimorbidity conditions, which we were not able to assess due to the non-availability of these data in HRS.

C. Future directions

Our findings point to an ARIP marker that was strongly associated with physiological dysregulation with aging and also indicate specific genes which could mediate the association of immune aging with multimorbidity and mortality. These findings can be used to design future studies to validate the results.

Measurement of the immune cells among the same participants at future time points can help us further investigate and validate if changes in the ARIP marker chosen in this dissertation influence the incidence of health outcomes. In this dissertation, we have focused only on the T-cell repertoire to study immune aging due to the existing firm evidence in the literature. However, it might be useful to include the distribution of the other immune cells along with T-cell subsets to define a more comprehensive immune aging marker associated with physiological dysregulation represented by multimorbidity and mortality. Combining the longitudinal changes of the immune cell distribution along with other immune cell subsets in a large cohort study such as HRS can help us understand immune aging better.

Another direction of research that can be explored using the findings of this dissertation is the validation of the gene expression results in other cohort studies which have also performed whole-genome RNA-sequencing. The identified five genes that were associated with immune aging, multimorbidity, and mortality can be studied in other

studies to confirm the obtained results. Another method of validation for these genes is using a real-time polymerase chain reaction technique (qPCR). RNA-sequencing approaches in studies can possibly suffer from bias and non-reproducibility and hence validating the results on qPCR can help confirm the results.

Although our results may not have direct clinical indications at this point, these results confirm the important role of immune aging in the systemic dysregulation of the body. These results can provide us a path forward to focus on certain genes and immune markers which can be targeted in experimental studies and identifying high-risk individuals undergoing accelerated aging.

References:

1. Health, United States 2019. National Center for Health Statistics. 2021.
2. Global Health Observatory data repository.
3. Vespa J, Medina L, Armstrong D. Demographic turning points for the United States: Population projections for 2020 to 2060. Current Population Reports. 2020;(P25-1144):1–15.
4. Pawelec G. Age and immunity: What is "immunosenescence"? Exp Gerontol. 2018;105(September 2017):4–9.
5. Bischof J, Gärtner F, Zeiser K, Kunz R, Schreiner C, Hoffer E, et al. Immune Cells and Immunosenescence. Folia Biol (Praha). 2019;65(2):53–63.
6. Oh SJ, Lee JK, Shin OS. Aging and the immune system: The impact of immunosenescence on viral infection, immunity and vaccine immunogenicity. Immune Netw. 2019;19(6):1–18.
7. Weiskopf D, Weinberger B, Grubeck-Loebenstein B. The aging of the immune system. Transplant International. 2009 Nov 1;22(11):1041–50.
8. Barbé-Tuana F, Funchal G, Schmitz CRR, Maurmann RM, Bauer ME. The interplay between immunosenescence and age-related diseases. Semin Immunopathol. 2020;42(5):545–57.
9. Costantini E, D'Angelo C, Reale M. The role of immunosenescence in neurodegenerative diseases. Mediators Inflamm. 2018;2018:84–91.
10. Buttorff C, Ruder T BM. Multiple chronic conditions in the United States. Santa Monica (CA): RAND Corporation; 2017.
11. Ward BW, Schiller JS, Goodman RA. Multiple chronic conditions among US adults: a 2012 update. Prev Chronic Dis. 2014 Apr;11:E62.
12. Vogeli C, Shields AE, Lee TA, Gibson TB, Marder WD, Weiss KB, et al. Multiple chronic conditions: Prevalence, health consequences, and implications for quality, care management, and costs. J Gen Intern Med. 2007;22(SUPPL. 3):391–5.
13. Kingston A, Robinson L, Booth H, Knapp M, Jagger C. Projections of multi-morbidity in the older population in England to 2035: estimates from the Population Ageing and Care Simulation (PACSim) model. Age Ageing. 2018 May;47(3):374–80.

14. Kennedy BK, Berger SL, Brunet A, Campisi J, Cuervo AM, Epel ES, et al. Geroscience: linking aging to chronic disease. Cell. 2014 Nov 6;159(4):709–13.

15. World Health Organization: Healthy ageing.

16. Lara J, Cooper R, Nissan J, Ginty AT, Khaw KT, Deary IJ, et al. A proposed panel of biomarkers of healthy ageing. BMC Med. 2015;13(1):222.

17. Sebastiani P, Thyagarajan B, Sun F, Schupf N, Newman AB, Montano M, et al. Biomarker signatures of aging. Aging Cell. 2017 Apr 1;16(2):329–38.

18. Fulop T, Larbi A, Dupuis G, Page A Le, Frost EH, Cohen AA, et al. Immunosenescence and inflamm-aging as two sides of the same coin: Friends or Foes? Front Immunol. 2018;8(JAN).

19. Zheng DD, Loewenstein DA, Christ SL, Feaster DJ, Lam BL, McCollister KE, et al. Multimorbidity patterns and their relationship to mortality in the US older adult population. PLoS One. 2021 Jan 20;16(1):e0245053.

20. Barnett K, Mercer SW, Norbury M, Watt G, Wyke S, Guthrie B. Epidemiology of multimorbidity and implications for health care, research, and medical education: a cross-sectional study. The Lancet. 2012 Jul 7;380(9836):37–43.

21. Wikby A, Johansson B, Olsson J, Löfgren S, Nilsson BO, Ferguson F. Expansions of peripheral blood CD8 T-lymphocyte subpopulations and an association with cytomegalovirus seropositivity in the elderly: the Swedish NONA immune study. Exp Gerontol. 2002;37(2):445–53.

22. Olsson J, Wikby A, Johansson B, Löfgren S, Nilsson BO, Ferguson FG. Age-related change in peripheral blood T-lymphocyte subpopulations and cytomegalovirus infection in the very old: the Swedish longitudinal OCTO immune study. Mech Ageing Dev. 2000 Dec;121(1–3):187–201.

23. Brodin P, Jojic V, Gao T, Bhattacharya S, Angel CJL, Furman D, et al. Variation in the Human Immune System Is Largely Driven by Non-Heritable Influences. Cell. 2015;160(1):37–47.

24. Carr EJ, Dooley J, Garcia-Perez JE, Lagou V, Lee JC, Wouters C, et al. The cellular composition of the human immune system isshaped by age and cohabitation. Nat Immunol. 2016;17(4):461–8.

25. Tsang JS, Schwartzberg PL, Kotliarov Y, Biancotto A, Xie Z, Germain RN, et al. Global Analyses of Human Immune Variation Reveal Baseline Predictors of Postvaccination Responses. Cell. 2014;157(2):499–513.

26. Jiang N, He J, Weinstein JA, Penland L, Sasaki S, He XS, et al. Lineage Structure of the Human Antibody Repertoire in Response to Influenza Vaccination. Sci Transl Med. 2013 Feb 6;5(171):171ra19 LP-171ra19.

27. Saule P, Trauet J, Dutriez V, Lekeux V, Dessaint JP, Labalette M. Accumulation of memory T cells from childhood to old age: Central and effector memory cells in CD4+ versus effector memory and terminally differentiated memory cells in CD8+ compartment. Mech Ageing Dev. 2006;127(3):274–81.

28. Li M, Yao D, Zeng X, Kasakovski D, Zhang Y, Chen S, et al. Age related human T cell subset evolution and senescence. Immunity and Ageing. 2019;16(1):1–7.

29. Whitacre CC. Sex differences in autoimmune disease. Nat Immunol. 2001;2(9):777–80.

30. Lee W, Reveille JD, Davis JC, Learch TJ, Ward MM, Weisman MH. Are there gender differences in severity of ankylosing spondylitis? Results from the PSOAS cohort. Ann Rheum Dis. 2007 May 1;66(5):633 LP – 638.

31. Fischinger S, Boudreau CM, Butler AL, Streeck H, Alter G. Sex differences in vaccine-induced humoral immunity. Semin Immunopathol. 2019;41(2):239–49.

32. Klein SL, Flanagan KL. Sex differences in immune responses. Nat Rev Immunol. 2016;16(10):626–38.

33. Nalls MA, Couper DJ, Tanaka T, van Rooij FJA, Chen MH, Smith A V, et al. Multiple Loci Are Associated with White Blood Cell Phenotypes. PLoS Genet. 2011 Jun 30;7(6):e1002113.

34. Pilia G, Chen WM, Scuteri A, Orrú M, Albai G, Dei M, et al. Heritability of Cardiovascular and Personality Traits in 6,148 Sardinians. PLoS Genet. 2006 Aug 25;2(8):e132.

35. De Jager PL, Hacohen N, Mathis D, Regev A, Stranger BE, Benoist C. ImmVar project: Insights and design considerations for future studies of “healthy” immune variation. Semin Immunol. 2015;27(1):51–7.

36. Sylwester AW, Mitchell BL, Edgar JB, Taormina C, Pelte C, Ruchti F, et al. Broadly targeted human cytomegalovirus-specific CD4+ and CD8+ T cells dominate the memory compartments of exposed subjects . Journal of Experimental Medicine. 2005 Sep 6;202(5):673–85.

37. Yan Z, Maecker HT, Brodin P, Nygaard UC, Lyu SC, Davis MM, et al. Aging and CMV discordance are associated with increased immune diversity between monozygotic twins. Immunity & Ageing. 2021;18(1):5.

38. Higuchi T, Omata F, Tsuchihashi K, Higashioka K, Koyamada R, Okada S. Current cigarette smoking is a reversible cause of elevated white blood cell count: Cross-sectional and longitudinal studies. Prev Med Rep. 2016;4(2016):417–22.

39. Ferson M, Edwards A, Lind A, Milton GW, Hersey P. Low natural killer-cell activity and immunoglobulin levels associated with smoking in human subjects. Int J Cancer. 1979 May 15;23(5):603–9.

40. Milner JJ, Beck MA. The impact of obesity on the immune response to infection. Proceedings of the Nutrition Society. 2012/03/14. 2012;71(2):298–306.

41. de Heredia FP, Gómez-Martínez S, Marcos A. Obesity, inflammation and the immune system. Proceedings of the Nutrition Society. 2012/03/20. 2012;71(2):332–8.

42. van den Broek T, Borghans JAM, van Wijk F. The full spectrum of human naive T cells. Nat Rev Immunol. 2018;18(6):363–73.

43. Pennock ND, White JT, Cross EW, Cheney EE, Tamburini BA, Kedl RM. T cell responses: naive to memory and everything in between. Adv Physiol Educ. 2013 Dec;37(4):273–83.

44. Montecino-Rodriquez E, Min H, Dorshkind K. Reevaluating current models of thymic involution. Semin Immunol. 2005;17(5):356–61.

45. Fülöp T, Larbi A, Pawelec G. Human T cell aging and the impact of persistent viral infections. Front Immunol. 2013 Sep 13;4:271.

46. Pourgheysari B, Khan N, Best D, Bruton R, Nayak L, Moss PAH. The Cytomegalovirus-Specific CD4⁺ T-Cell Response Expands with Age and Markedly Alters the CD4⁺ T-Cell Repertoire. J Virol. 2007 Jul 15;81(14):7759 LP – 7765.

47. Mahnke YD, Brodie TM, Sallusto F, Roederer M, Lugli E. The who's who of T-cell differentiation: Human memory T-cell subsets. Eur J Immunol. 2013 Nov 1;43(11):2797–809.

48. Roberts AD, Ely KH, Woodland DL. Differential contributions of central and effector memory T cells to recall responses. Journal of Experimental Medicine. 2005;202(1):123–33.

49. Thyagarajan B, Faul J, Vivek S, Kim JK, Nikolich-Žugich J, Weir D, et al. Age-Related Differences in T-Cell Subsets in a Nationally Representative Sample of People Older Than Age 55: Findings From the Health and Retirement Study. The Journals of Gerontology: Series A [Internet]. 2022 May 1;77(5):927–33. Available from: https://doi.org/10.1093/gerona/glab300

50. Weltevrede M, Eilers R, de Melker HE, van Baarle D. Cytomegalovirus persistence and T-cell immunosenescence in people aged fifty and older: A systematic review. Exp Gerontol. 2016;77:87–95.

51. Wertheimer AM, Bennett MS, Park B, Uhrlaub JL, Martinez C, Pulko V, et al. Aging and cytomegalovirus infection differentially and jointly affect distinct circulating T cell subsets in humans. J Immunol. 2014 Mar;192(5):2143–55.

52. Tomlins SA, Rhodes DR, Perner S, Dhanasekaran SM, Mehra R, Sun XW, et al. Recurrent Fusion of TMPRSS2 and ETS Transcription Factor Genes in Prostate Cancer. Science (1979) [Internet]. 2005 Oct 28;310(5748):644–8. Available from: https://doi.org/10.1126/science.1117679

53. Ademokun A, Wu YC, Martin V, Mitra R, Sack U, Baxendale H, et al. Vaccination-induced changes in human B-cell repertoire and pneumococcal IgM and IgA antibody at different ages. Aging Cell [Internet]. 2011 Dec 1;10(6):922–30. Available from: https://doi.org/10.1111/j.1474-9726.2011.00732.x

54. Sasaki S, Sullivan M, Narvaez CF, Holmes TH, Furman D, Zheng NY, et al. Limited efficacy of inactivated influenza vaccine in elderly individuals is associated with decreased production of vaccine-specific antibodies. J Clin Invest. 2011 Aug;121(8):3109–19.

55. Wagner A, Garner-Spitzer E, Jasinska J, Kollaritsch H, Stiasny K, Kundi M, et al. Age-related differences in humoral and cellular immune responses after primary immunisation: indications for stratified vaccination schedules. Sci Rep. 2018;8(1):9825.

56. Roukens AH, Soonawala D, Joosten SA, de Visser AW, Jiang X, Dirksen K, et al. Elderly subjects have a delayed antibody response and prolonged viraemia following yellow fever vaccination: a prospective controlled cohort study. PLoS One. 2011;6(12):e27753.

57. Li S, Sullivan NL, Rouphael N, Yu T, Banton S, Maddur MS, et al. Metabolic Phenotypes of Response to Vaccination in Humans. Cell. 2017 May 18;169(5):862-877.e17.

58. Enrico A, Domenico C, Viviana V, Michela B, C. VA, Monica DM, et al. Effector Memory T cells Are Associated With Atherosclerosis in Humans and Animal Models. J Am Heart Assoc. 2021 May 5;1(1):e000125.

59. Dumitriu IE, Araguás ET, Baboonian C, Kaski JC. CD4+CD28null T cells in coronary artery disease: when helpers become killers. Cardiovasc Res. 2009 Jan 1;81(1):11–9.

60. Haoran W, Geng P, Jing B, Bing H, Kecheng H, Xinrong H, et al. Cytomegalovirus Infection and Relative Risk of Cardiovascular Disease (Ischemic Heart Disease, Stroke, and Cardiovascular Death): A Meta-Analysis of Prospective Studies Up to 2016. J Am Heart Assoc. 2021 May 5;6(7):e005025.

61. Youn JC, Yu HT, Lim BJ, Koh MJ, Lee J, Chang DY, et al. Immunosenescent CD8+ T clls and C-X-C cemokine receptor type 3 chemokines are increased in human hypertension. Hypertension. 2013;62(1):126–33.

62. Olson NC, Doyle MF, de Boer IH, Huber SA, Jenny NS, Kronmal RA, et al. Associations of Circulating Lymphocyte Subpopulations with Type 2 Diabetes: Cross-Sectional Results from the Multi-Ethnic Study of Atherosclerosis (MESA). PLoS One. 2015 Oct 12;10(10):e0139962.

63. Olson NC, Doyle MF, Sitlani CM, de Boer IH, Rich SS, Huber SA, et al. Associations of Innate and Adaptive Immune Cell Subsets With Incident Type 2 Diabetes Risk: The MESA Study. J Clin Endocrinol Metab. 2020 Mar 1;105(3):e848–57.

64. Rattik S, Engelbertsen D, Wigren M, Ljungcrantz I, Östling G, Persson M, et al. Elevated circulating effector memory T cells but similar levels of regulatory T cells in patients with type 2 diabetes mellitus and cardiovascular disease. Diab Vasc Dis Res. 2018 Dec 21;16(3):270–80.

65. Yoo SG, Han K Do, Lee KH, La Y, Kwon DE, Han SH. Impact of Cytomegalovirus Disease on New-Onset Type 2 Diabetes Mellitus: Population-Based Matched Case-Control Cohort Study. Diabetes Metab J. 2019/01/21. 2019 Dec;43(6):815–29.

66. Roberts BW, Cech I. Association of type 2 diabetes mellitus and seroprevalence for cytomegalovirus. South Med J. 2005 Jul;98(7):686–92.

67. Chen S, de Craen AJM, Raz Y, Derhovanessian E, Vossen ACTM, Westendorp RGJ, et al. Cytomegalovirus seropositivity is associated with glucose regulation in the oldest old. Results from the Leiden 85-plus Study. Immun Ageing. 2012 Aug 28;9(1):18.

68. Bauer ME. Accelerated immunosenescence in rheumatoid arthritis: impact on clinical progression. Immunity & Ageing. 2020;17(1):6.

69. Trintinaglia L, Bandinelli LP, Grassi-Oliveira R, Petersen LE, Anzolin M, Correa BL, et al. Features of Immunosenescence in Women Newly Diagnosed With Breast Cancer . Vol. 9, Frontiers in Immunology . 2018. p. 1651.

70. Song G, Wang X, Jia J, Yuan Y, Wan F, Zhou X, et al. Elevated level of peripheral CD8+CD28− T lymphocytes are an independent predictor of progression-free survival in patients with metastatic breast cancer during the course of chemotherapy. Cancer Immunology, Immunotherapy. 2013;62(6):1123–30.

71. Meloni F, Morosini M, Solari N, Passadore I, Nascimbene C, Novo M, et al. Foxp3 Expressing CD4+ CD25+ and CD8+CD28− T Regulatory Cells in the Peripheral Blood of Patients with Lung Cancer and Pleural Mesothelioma. Hum Immunol. 2006;67(1):1–12.

72. Peters MJ, Joehanes R, Pilling LC, Schurmann C, Conneely KN, Powell J, et al. The transcriptional landscape of age in human peripheral blood. Nat Commun. 2015;6(1):8570.

73. Harries LW, Hernandez D, Henley W, Wood AR, Holly AC, Bradley-Smith RM, et al. Human aging is characterized by focused changes in gene expression and deregulation of alternative splicing. Aging Cell. 2011 Oct 1;10(5):868–78.

74. Kent JW, Göring HHH, Charlesworth JC, Drigalenko E, Diego VP, Curran JE, et al. Genotype×age interaction in human transcriptional ageing. Mech Ageing Dev. 2012;133(9):581–90.

75. Glass D, Viñuela A, Davies MN, Ramasamy A, Parts L, Knowles D, et al. Gene expression changes with age in skin, adipose tissue, blood and brain. Genome Biol. 2013 Jul 26;14(7):R75–R75.

76. van den Akker M, Buntinx F, Knottnerus JA. Comorbidity or multimorbidity. European Journal of General Practice. 1996 Jan 1;2(2):65–70.

77. Vital Statistics Reports N. National Vital Statistics Reports Volume 70, Number 8 July 26, 2016 Deaths: Final Data 2019 [Internet]. 2021. Available from: https://www.cdc.gov/nchs/products/index.htm.

78. Murphy SL, Kochanek KD, Xu J, Arias E. Mortality in the United States, 2020. NCHS Data Brief. 2021 Dec;(427):1–8.

79. Salisbury C, Johnson L, Purdy S, Valderas JM, Montgomery AA. Epidemiology and impact of multimorbidity in primary care: a retrospective cohort study. Br J Gen Pract. 2011 Jan;61(582):e12-21.

80. Salive ME. Multimorbidity in Older Adults. Epidemiol Rev. 2013 Jan 1;35(1):75–83.

81. Fabbri E, An Y, Zoli M, Simonsick EM, Guralnik JM, Bandinelli S, et al. Aging and the Burden of Multimorbidity: Associations With Inflammatory and Anabolic

Hormonal Biomarkers. The Journals of Gerontology: Series A. 2015 Jan 1;70(1):63–70.

82. Schöttker B, Saum KU, Jansen EHJM, Holleczek B, Brenner H. Associations of metabolic, inflammatory and oxidative stress markers with total morbidity and multi-morbidity in a large cohort of older German adults. Age Ageing. 2016 Jan 1;45(1):127–35.

83. Niedzwiedz CL, Katikireddi SV, Pell JP, Smith DJ. Sex differences in the association between salivary telomere length and multimorbidity within the US Health & Retirement Study. Age Ageing. 2019 Sep 1;48(5):703–10.

84. Deelen J, Kettunen J, Fischer K, van der Spek A, Trompet S, Kastenmüller G, et al. A metabolic profile of all-cause mortality risk identified in an observational study of 44,168 individuals. Nat Commun [Internet]. 2019;10(1):3346. Available from: https://doi.org/10.1038/s41467-019-11311-9

85. Singh-Manoux A, Shipley MJ, Bell JA, Canonico M, Elbaz A, Kivimäki M. Association between inflammatory biomarkers and all-cause, cardiovascular and cancer-related mortality. CMAJ. 2017 Mar 13;189(10):E384–90.

86. Heeringa SG, Connor J. Technical Description of the Health and Retirement Study Sample Design. Ann Arbor, Michigan: Institute for Social Research, University of Michigan; 1995.

87. Ofstedal MB, Weir DR. Recruitment and retention of minority participants in the health and retirement study. Gerontologist. 2011 Jun;51 Suppl 1(Suppl 1):S8–20.

88. Staff HRS. Sample Sizes and Response Rates. Ann Arbor, MI: Survey Research Center, Institute for Social Research, University of Michigan; 2011.

89. Crimmins EM, Faul JD, Thyagarajan B, Weir DR. Venous Blood Collection and Assay Protocol in the 2016 Health and Retirement Study. Ann Arbor, Michigan: Survey Research Center, Institute for Social Research, University of Michigan; 2017.

90. Crimmins EM, Thyagarajan B, Kim JK, Weir D, Faul J. Quest for a summary measure of biological age: the health and retirement study. Geroscience. 2021;43(1):395–408.

91. Gruver AL, Hudson LL, Sempowski GD. Immunosenescence of ageing. J Pathol. 2007 Jan;211(2):144–56.

92. Weng NP. Aging of the immune system: how much can the adaptive immune system adapt? Immunity. 2006 May;24(5):495–9.

93. Hakim FT, Gress RE. Immunosenescence: deficits in adaptive immunity in the elderly. Tissue Antigens. 2007 Sep 1;70(3):179–89.

94. Crooke SN, Ovsyannikova IG, Poland GA, Kennedy RB. Immunosenescence and human vaccine immune responses. Immunity & Ageing. 2019;16(1):25.

95. Aiello A, Farzaneh F, Candore G, Caruso C, Davinelli S, Gambino CM, et al. Immunosenescence and its hallmarks: How to oppose aging strategically? A review of potential options for therapeutic intervention. Front Immunol. 2019;10(SEP):1–19.

96. Li M, Yao D, Zeng X, Kasakovski D, Zhang Y, Chen S, et al. Age related human T cell subset evolution and senescence. Immunity & Ageing. 2019;16(1):24.

97. Thyagarajan B, Faul J, Vivek S, Kim JK, Nikolich-Žugich J, Weir D, et al. Age-Related Differences in T-Cell Subsets in a Nationally Representative Sample of People Older Than Age 55: Findings From the Health and Retirement Study. The Journals of Gerontology: Series A. 2021 Oct 11;glab300.

98. Strindhall J, Skog M, Ernerudh J, Bengner M, Löfgren S, Matussek A, et al. The inverted CD4/CD8 ratio and associated parameters in 66-year-old individuals: the Swedish HEXA immune study. Age (Omaha). 2013;35(3):985–91.

99. Vasson MP, Farges MC, Goncalves-Mendes N, Talvas J, Ribalta J, Winklhofer-Roob B, et al. Does aging affect the immune status? A comparative analysis in 300 healthy volunteers from France, Austria and Spain. Immunity & Ageing. 2013;10(1):38.

100. Alpert A, Pickman Y, Leipold M, Rosenberg-Hasson Y, Ji X, Gaujoux R, et al. A clinically meaningful metric of immune age derived from high-dimensional longitudinal monitoring. Nat Med. 2019;25(3):487–95.

101. Sayed N, Huang Y, Nguyen K, Krejciova-Rajaniemi Z, Grawe AP, Gao T, et al. An inflammatory aging clock (iAge) based on deep learning tracks multimorbidity, immunosenescence, frailty and cardiovascular aging. Nat Aging. 2021;1(7):598–615.

102. Spyridopoulos I, Martin-Ruiz C, Hilkens C, Yadegarfar ME, Isaacs J, Jagger C, et al. CMV seropositivity and T-cell senescence predict increased cardiovascular mortality in octogenarians: results from the Newcastle 85+ study. Aging Cell. 2016 Apr 1;15(2):389–92.

103. Maussang D, Verzijl D, van Walsum M, Leurs R, Holl J, Pleskoff O, et al. Human cytomegalovirus-encoded chemokine receptor US28 promotes tumorigenesis. Proc Natl Acad Sci U S A. 2006 Aug;103(35):13068–73.

104. Sonnega A, Faul JD, Ofstedal MB, Langa KM, Phillips JWR, Weir DR. Cohort Profile: the Health and Retirement Study (HRS). Int J Epidemiol. 2014 Mar 25;43(2):576–85.

105. Barcelo H, Faul J, Crimmins E, Thyagarajan B. A Practical Cryopreservation and Staining Protocol for Immunophenotyping in Population Studies. Curr Protoc Cytom. 2018 Apr;84(1):e35.

106. Hunter-Schlichting D, Lane J, Cole B, Flaten Z, Barcelo H, Ramasubramanian R, et al. Validation of a hybrid approach to standardize immunophenotyping analysis in large population studies: The Health and Retirement Study. Sci Rep. 2020;10(1):8759.

107. Klemera P, Doubal S. A new approach to the concept and computation of biological age. Mech Ageing Dev. 2006 Mar;127(3):240–8.

108. Levine ME. Modeling the rate of senescence: can estimated biological age predict mortality more accurately than chronological age? J Gerontol A Biol Sci Med Sci. 2012/12/03. 2013 Jun;68(6):667–74.

109. Kwon D, Belsky DW. A toolkit for quantification of biological age from blood chemistry and organ function test data: BioAge. Geroscience. 2021;43(6):2795–808.

110. Olson NC, Doyle MF, Jenny NS, Huber SA, Psaty BM, Kronmal RA, et al. Decreased Naive and Increased Memory CD4+ T Cells Are Associated with Subclinical Atherosclerosis: The Multi-Ethnic Study of Atherosclerosis. PLoS One. 2013 Aug 23;8(8):e71498.

111. Winer S, Chan Y, Paltser G, Truong D, Tsui H, Bahrami J, et al. Normalization of obesity-associated insulin resistance through immunotherapy. Nat Med. 2009;15(8):921–9.

112. Rocha VZ, Folco EJ, Sukhova G, Shimizu K, Gotsman I, Vernon AH, et al. Interferon-gamma, a Th1 cytokine, regulates fat inflammation: a role for adaptive immunity in obesity. Circ Res. 2008 Aug;103(5):467–76.

113. Lau EYM, Carroll EC, Callender LA, Hood GA, Berryman V, Pattrick M, et al. Type 2 diabetes is associated with the accumulation of senescent T cells. Clin Exp Immunol. 2019/07/08. 2019 Aug;197(2):205–13.

114. Yang P, Ma J, Yang X, Li W. Peripheral CD4+ naïve/memory ratio is an independent predictor of survival in non-small cell lung cancer. Oncotarget. 2017 Jul 18;8(48):83650–9.

115. Tuttolomondo A, Pecoraro R, Casuccio A, Di Raimondo D, Buttà C, Clemente G, et al. Peripheral Frequency of CD4+ CD28− Cells in Acute Ischemic Stroke: Relationship With Stroke Subtype and Severity Markers. Medicine. 2015;94(20).

116. Wikby A, Maxson P, Olsson J, Johansson B, Ferguson FG. Changes in CD8 and CD4 lymphocyte subsets, T cell proliferation responses and non-survival in the very old: the Swedish longitudinal OCTO-immune study. Mech Ageing Dev. 1998;102(2):187–98.

117. Liu Z, Kuo PL, Horvath S, Crimmins E, Ferrucci L, Levine M. A new aging measure captures morbidity and mortality risk across diverse subpopulations from NHANES IV: A cohort study. PLoS Med. 2019 Dec 31;15(12):e1002718.

118. Bahour N, Cortez B, Pan H, Shah H, Doria A, Aguayo-Mazzucato C. Diabetes mellitus correlates with increased biological age as indicated by clinical biomarkers. Geroscience. 2022;44(1):415–27.

119. Waziry R, Gras L, Sedaghat S, Tiemeier H, Weverling GJ, Ghanbari M, et al. Quantification of biological age as a determinant of age-related diseases in the Rotterdam Study: a structural equation modeling approach. Eur J Epidemiol. 2019 Aug;34(8):793–9.

120. Tay RE, Richardson EK, Toh HC. Revisiting the role of CD4+ T cells in cancer immunotherapy—new insights into old paradigms. Cancer Gene Ther. 2021;28(1):5–17.

121. Valzasina B, Piconese S, Guiducci C, Colombo MP. Tumor-Induced Expansion of Regulatory T Cells by Conversion of CD4⁺CD25[−] Lymphocytes Is Thymus and Proliferation Independent. Cancer Res. 2006 Apr 15;66(8):4488 LP – 4495.

122. Liu VC, Wong LY, Jang T, Shah AH, Park I, Yang X, et al. Tumor Evasion of the Immune System by Converting CD4⁺CD25[−] T Cells into CD4⁺CD25⁺ T Regulatory Cells: Role of Tumor-Derived TGF-β. The Journal of Immunology. 2007 Mar 1;178(5):2883 LP – 2892.

123. Larbi A, Fulop T. From “truly naïve” to “exhausted senescent” T cells: When markers predict functionality. Cytometry Part A. 2014 Jan 1;85(1):25–35.

124. Zitvogel L, Tesniere A, Kroemer G. Cancer despite immunosurveillance: immunoselection and immunosubversion. Nat Rev Immunol. 2006;6(10):715–27.

125. Sallusto F, Lenig D, Förster R, Lipp M, Lanzavecchia A. Two subsets of memory T lymphocytes with distinct homing potentials and effector functions. Nature. 1999;401(6754):708–12.

126. Klebanoff CA, Gattinoni L, Restifo NP. CD8+ T-cell memory in tumor immunology and immunotherapy. Immunol Rev. 2006 Jun;211:214–24.

127. White K, Avendaño M, Capistrant BD, Robin Moon J, Liu SY, Maria Glymour M. Self-reported and measured hypertension among older US- and foreign-born adults. J Immigr Minor Health. 2012 Aug;14(4):721–6.

128. Mullins MA, Kler JS, Eastman MR, Kabeto M, Wallner LP, Kobayashi LC. Validation of Self-reported Cancer Diagnoses Using Medicare Diagnostic Claims in the US Health and Retirement Study, 2000-2016. Cancer Epidemiol Biomarkers Prev. 2021 Nov;

129. Huan T, Esko T, Peters MJ, Pilling LC, Schramm K, Schurmann C, et al. A meta-analysis of gene expression signatures of blood pressure and hypertension. PLoS Genet. 2015 Mar;11(3):e1005035.

130. Nho K, Nudelman K, Allen M, Hodges A, Kim S, Risacher SL, et al. Genome-wide transcriptome analysis identifies novel dysregulated genes implicated in Alzheimer's pathology. Alzheimer's & Dementia. 2020;16(9):1213–23.

131. Santiago JA, Bottero V, Potashkin JA. Transcriptomic and Network Analysis Identifies Shared and Unique Pathways across Dementia Spectrum Disorders. Int J Mol Sci. 2020 Mar 17;21(6):2050.

132. Kelly J, Moyeed R, Carroll C, Albani D, Li X. Gene expression meta-analysis of Parkinson's disease and its relationship with Alzheimer's disease. Mol Brain. 2019;12(1):16.

133. Das UN, Rao AA. Gene expression profile in obesity and type 2 diabetes mellitus. Lipids Health Dis. 2007 Dec 14;6:35.

134. Cao JN, Gollapudi S, Sharman EH, Jia Z, Gupta S. Age-related alterations of gene expression patterns in human CD8+ T cells. Aging Cell. 2010 Feb 1;9(1):19–31.

135. Bektas A, Zhang Y, Lehmann E, Wood WH, Becker KG, Madara K, et al. Age-associated changes in basal NF-κB function in human $CD4^+$ T lymphocytes via dysregulation of PI3 kinase. Aging. 2014 Jan 1;6(11):957–69.

136. Reynolds LM, Ding J, Taylor JR, Lohman K, Soranzo N, de la Fuente A, et al. Transcriptomic profiles of aging in purified human immune cells. BMC Genomics. 2015;16(1):1–17.

137. Ramasubramanian R, Kalhan R, Jacobs DR, Washko GR, Hou L, Gross MD, et al. Lung Function and Gene Expression of Pathogen Recognition Pathway Receptors: the Cardia Lung Study. Sci Rep [Internet]. 2020;10(1):9360. Available from: https://doi.org/10.1038/s41598-020-65923-z

138. Domínguez-Domínguez L, Rava M, Bisbal O, Lopez-Cortés L, Portilla J, Podzamczer D, et al. Low CD4/CD8 ratio is associated with increased morbidity and mortality in late and non-late presenters: results from a multicentre cohort study, 2004–2018. BMC Infect Dis [Internet]. 2022;22(1):379. Available from: https://doi.org/10.1186/s12879-022-07352-z

139. Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, et al. STAR: ultrafast universal RNA-seq aligner. Bioinformatics. 2013 Jan 1;29(1):15–21.

140. DeLuca DS, Levin JZ, Sivachenko A, Fennell T, Nazaire MD, Williams C, et al. RNA-SeQC: RNA-seq metrics for quality control and process optimization. Bioinformatics. 2012 Jun;28(11):1530–2.

141. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, et al. The Sequence Alignment/Map format and SAMtools. Bioinformatics. 2009 Aug;25(16):2078–9.

142. Li B, Dewey CN. RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. BMC Bioinformatics. 2011;12(1):323.

143. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol. 2014;15(12):550.

144. Benjamini Y, Hochberg Y. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. Journal of the Royal Statistical Society Series B (Methodological). 1995 Apr 24;57(1):289–300.

145. Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, et al. Gene set enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles. Proceedings of the National Academy of Sciences. 2005 Oct 25;102(43):15545 LP – 15550.

146. Alhamdoosh M, Ng M, Wilson NJ, Sheridan JM, Huynh H, Wilson MJ, et al. Combining multiple tools outperforms individual methods in gene set enrichment analyses. Bioinformatics. 2017 Feb 1;33(3):414–24.

147. Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, et al. Gene set enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles. Proceedings of the National Academy of Sciences [Internet]. 2005 Oct 25;102(43):15545–50. Available from: https://doi.org/10.1073/pnas.0506580102

148. Mootha VK, Lindgren CM, Eriksson KF, Subramanian A, Sihag S, Lehar J, et al. PGC-1α-responsive genes involved in oxidative phosphorylation are coordinately downregulated in human diabetes. Nat Genet [Internet]. 2003;34(3):267–73. Available from: https://doi.org/10.1038/ng1180

149. Supuran CT, de Simone G. Chapter 1 - Carbonic Anhydrases: An Overview. In: Supuran CT, de Simone G, editors. Carbonic Anhydrases as Biocatalysts [Internet]. Amsterdam: Elsevier; 2015. p. 3–13. Available from: https://www.sciencedirect.com/science/article/pii/B9780444632586000019

150. Supuran CT. Carbonic anhydrases: novel therapeutic applications for inhibitors and activators. Nat Rev Drug Discov [Internet]. 2008;7(2):168–81. Available from: https://doi.org/10.1038/nrd2467

151. Mboge MY, Mahon BP, McKenna R, Frost SC. Carbonic Anhydrases: Role in pH Control and Cancer. Metabolites. 2018 Feb 28;8(1).

152. Pollard A, Shephard F, Freed J, Liddell S, Chakrabarti L. Mitochondrial proteomic profiling reveals increased carbonic anhydrase II in aging and neurodegeneration. Aging. 2016;8(10):2425–36.

153. Kivela J, Parkkila S, Parkkila AK, Leinonen J, Rajaniemi H. Salivary carbonic anhydrase isoenzyme VI. J Physiol. 1999 Oct 15;520 Pt 2:315–20.

154. Tatematsu M, Yoshida R, Morioka Y, Ishii N, Funami K, Watanabe A, et al. Raftlin Controls Lipopolysaccharide-Induced TLR4 Internalization and TICAM-1 Signaling in a Cell Type–Specific Manner. The Journal of Immunology [Internet]. 2016 May 1;196(9):3865. Available from: http://www.jimmunol.org/content/196/9/3865.abstract

155. Rahman AH, Taylor DK, Turka LA. The contribution of direct TLR signaling to T cell responses. Immunol Res [Internet]. 2009;45(1):25–36. Available from: https://doi.org/10.1007/s12026-009-8113-x

156. Potet F, Scott JD, Mohammad-Panah R, Escande D, Baró I. AKAP proteins anchor cAMP-dependent protein kinase to KvLQT1/IsK channel complex. American Journal of Physiology-Heart and Circulatory Physiology [Internet]. 2001 May 1;280(5):H2038–45. Available from: https://doi.org/10.1152/ajpheart.2001.280.5.H2038

157. Hwang ES, Ho IC. Regulation of thymocyte homeostasis by Fliz1. Immunology [Internet]. 2002 Aug 1;106(4):464–9. Available from: https://doi.org/10.1046/j.1365-2567.2002.01455.x

158. Höllbacher B, Duhen T, Motley S, Klicznik MM, Gratz IK, Campbell DJ. Transcriptomic Profiling of Human Effector and Regulatory T Cell Subsets Identifies Predictive Population Signatures. Immunohorizons [Internet]. 2020 Oct 1;4(10):585. Available from: http://www.immunohorizons.org/content/4/10/585.abstract

159. Makovski TT, Schmitz S, Zeegers MP, Stranges S, van den Akker M. Multimorbidity and quality of life: Systematic literature review and meta-analysis. Ageing Res Rev. 2019 Aug;53:100903.

160. Nunes BP, Flores TR, Mielke GI, Thumé E, Facchini LA. Multimorbidity and mortality in older adults: A systematic review and meta-analysis. Arch Gerontol Geriatr. 2016;67:130–8.

161. Wikby A, Månsson IA, Johansson B, Strindhall J, Nilsson SE. The immune risk profile is associated with age and gender: findings from three Swedish population studies of individuals 20–100 years of age. Biogerontology. 2008;9(5):299–308.

162. Frenk S, Houseley J. Gene expression hallmarks of cellular ageing. Biogerontology. 2018/02/28. 2018 Dec;19(6):547–66.

163. Chen Y, Hor HH, Tang BL. AMIGO is expressed in multiple brain cell types and may regulate dendritic growth and neuronal survival. J Cell Physiol [Internet]. 2012 May 1;227(5):2217–29. Available from: https://doi.org/10.1002/jcp.22958

164. Bayraktar A, Lam S, Altay O, Li X, Yuan M, Zhang C, et al. Revealing the Molecular Mechanisms of Alzheimer's Disease Based on Network Analysis. Int J Mol Sci. 2021 Oct 26;22(21).

165. Liu M, Li S, Du M, Chu H, Wang M, Zhang Z. Novel CpG-SNPs in the gastric acid secretion pathway GNAI3 and susceptibility to gastric cancer. Gene [Internet]. 2020;736:144447. Available from: https://www.sciencedirect.com/science/article/pii/S0378111920301165

166. Girard C, Duprat F, Terrenoire C, Tinel N, Fosset M, Romey G, et al. Genomic and Functional Characteristics of Novel Human Pancreatic 2P Domain K+ Channels. Biochem Biophys Res Commun [Internet]. 2001;282(1):249–56. Available from: https://www.sciencedirect.com/science/article/pii/S0006291X01945620

167. Comes N, Serrano-Albarrás A, Capera J, Serrano-Novillo C, Condom E, Ramón y Cajal S, et al. Involvement of potassium channels in the progression of cancer to a more malignant phenotype. Biochimica et Biophysica Acta (BBA) - Biomembranes [Internet]. 2015;1848(10, Part B):2477–92. Available from: https://www.sciencedirect.com/science/article/pii/S0005273614004362

168. Dookeran KA, Zhang W, Stayner L, Argos M. Associations of two-pore domain potassium channels and triple negative breast cancer subtype in The Cancer Genome Atlas: systematic evaluation of gene expression and methylation. BMC Res Notes [Internet]. 2017;10(1):475. Available from: https://doi.org/10.1186/s13104-017-2777-4

169. Gaudet P, Livstone MS, Lewis SE, Thomas PD. Phylogenetic-based propagation of functional annotations within the Gene Ontology consortium. Brief Bioinform [Internet]. 2011 Sep 1;12(5):449–62. Available from: https://doi.org/10.1093/bib/bbr042

170. Jergović M, Contreras NA, Nikolich-Žugich J. Impact of CMV upon immune aging: facts and fiction. Med Microbiol Immunol [Internet]. 2019;208(3):263–9. Available from: https://doi.org/10.1007/s00430-019-00605-w

171. C-reactive protein concentration and risk of coronary heart disease, stroke, and mortality: an individual participant meta-analysis. The Lancet [Internet]. 2010;375(9709):132–40. Available from: https://www.sciencedirect.com/science/article/pii/S0140673609617177

172. Agassandian M, Shurin G v, Ma Y, Shurin MR. C-reactive protein and lung diseases. Int J Biochem Cell Biol [Internet]. 2014;53:77–88. Available from: https://www.sciencedirect.com/science/article/pii/S1357272514001745

173. Volpato S, Guralnik JM, Ferrucci L, Balfour J, Chaves P, Fried LP, et al. Cardiovascular Disease, Interleukin-6, and Risk of Mortality in Older Women. Circulation [Internet]. 2001 Feb 20;103(7):947–53. Available from: https://doi.org/10.1161/01.CIR.103.7.947

174. Hart PC, Rajab IM, Alebraheem M, Potempa LA. C-Reactive Protein and Cancer—Diagnostic and Therapeutic Insights. Vol. 11, Frontiers in Immunology. Frontiers Media S.A.; 2020.

175. Vogan K. PIK3CD mutations cause immunodeficiency. Nat Genet [Internet]. 2013;45(12):1417. Available from: https://doi.org/10.1038/ng.2840

176. Jou ST, Carpino N, Takahashi Y, Piekorz R, Chao JR, Carpino N, et al. Essential, Nonredundant Role for the Phosphoinositide 3-Kinase p110δ in Signaling by the B-Cell Receptor Complex. Mol Cell Biol. 2002 Dec 15;22(24):8580–91.

177. Okkenhaug K, Bilancio A, Farjot G, Priddle H, Sancho S, Peskett E, et al. Impaired B and T cell antigen receptor signaling in p110delta PI 3-kinase mutant mice. Science. 2002 Aug 9;297(5583):1031–4.

178. Lucas CL, Kuehn HS, Zhao F, Niemela JE, Deenick EK, Palendira U, et al. Dominant-activating germline mutations in the gene encoding the PI(3)K catalytic subunit p110δ result in T cell senescence and human immunodeficiency. Nat Immunol. 2014 Jan;15(1):88–97.

179. Condliffe AM, Chandra A. Respiratory manifestations of the activated phosphoinositide 3-kinase delta syndrome. Vol. 9, Frontiers in Immunology. Frontiers Media S.A.; 2018.

180. Bidaud I, Mezghrani A, Swayne LA, Monteil A, Lory P. Voltage-gated calcium channels in genetic diseases. Biochimica et Biophysica Acta (BBA) - Molecular Cell Research [Internet]. 2006;1763(11):1169–74. Available from: https://www.sciencedirect.com/science/article/pii/S0167488906002692

181. Warnier M, Roudbaraki M, Derouiche S, Delcourt P, Bokhobza A, Prevarskaya N, et al. CACNA2D2 promotes tumorigenesis by stimulating cell proliferation and angiogenesis. Oncogene [Internet]. 2015;34(42):5383–94. Available from: https://doi.org/10.1038/onc.2014.467

182. Ma Q, Yang Y, Feng D, Zheng S, Meng R, Fa P, et al. MAGI3 negatively regulates Wnt/β-catenin signaling and suppresses malignant phenotypes of glioma cells. Oncotarget. 2015 Nov 3;6(34):35851–65.

183. Ma Q, Zhang Y, Meng R, Xie KM, Xiong Y, Lin S, et al. MAGI3 Suppresses Glioma Cell Proliferation via Upregulation of PTEN Expression. Biomed Environ Sci. 2015 Jul;28(7):502–9.

184. Chen XJ, Wang X, Kaufman BA, Butow RA. Aconitase Couples Metabolic Regulation to Mitochondrial DNA Maintenance. Science (1979) [Internet]. 2005 Feb 4;307(5710):714–7. Available from: https://doi.org/10.1126/science.1106391

185. Gaudet P, Livstone MS, Lewis SE, Thomas PD. Phylogenetic-based propagation of functional annotations within the Gene Ontology consortium. Brief Bioinform. 2011 Sep;12(5):449–62.