Summer 8-11-2023

GENOME-SCALE METHYLATION ANALYSIS IN BLOOD AND TUMOR IDENTIFIES IMMUNE PROFILE, AGE ACCELERATION, AND DNA METHYLATION ALTERATIONS ASSOCIATED WITH BLADDER CANCER OUTCOMES

Ji-Qing Chen
ji-qing.chen.gr@dartmouth.edu

GENOME-SCALE METHYLATION ANALYSIS IN BLOOD AND TUMOR IDENTIFIES IMMUNE PROFILE, AGE ACCELERATION, AND DNA METHYLATION ALTERATIONS ASSOCIATED WITH BLADDER CANCER OUTCOMES

A Thesis

Submitted to the Faculty

in partial fulfillment of the requirements for the

degree of

Doctor of Philosophy

in

Cancer Biology

by Ji-Qing Chen

Guarini School of Graduate and Advanced Studies

Dartmouth College

Hanover, New Hampshire

August 2023

Examining Committee:

_____

Brock C. Christensen, Ph.D., Chair.

_____

Margaret R. Karagas, Ph.D.

_____

Michael N. Passarelli, Ph.D.

_____

Lucas A. Salas, M.D., Ph.D.

_____

Carmen J. Marsit, Ph.D.

_____

F. Jon Kull, Ph.D.

Dean of the Guarini School of Graduate and Advanced Studies

Abstract

Bladder cancer patients receive frequent screening due to the high tumor recurrence rate (more than 60%). Nowadays, the conventional monitoring method relies on cystoscopy which is highly invasive and increases patient morbidity and burden to the health care system with frequent follow-up. As a result, it is urgent to explore novel markers related to the outcomes of bladder cancer. Immune profiles have been associated with cancer outcomes and may have the potential to be biomarkers for outcomes management. However, little work has been conducted to investigate the associations of immune cell profiles with bladder cancer outcomes. Here, I utilized the Illumina HumanMethylationEPIC array to measure DNA methylation profiles of peripheral blood and matched tumor tissues of bladder cancer cases recruited in a population-based cohort study in New Hampshire. Then, cell-type deconvolution was applied to quantify immune cell-type proportions, and three epigenetic clocks were estimated for calculating age acceleration. Cox proportional hazard models were performed to test the association of methylation-derived profiles with bladder cancer outcomes. The *partDSA* algorithm and a semi-supervised recursively partitioned mixture model were conducted to determine overall survival groups based on immune cell profiles, clinical variables, and DNA methylation level. We used an epigenome-wide association study approach adjusting for immune cell profiles to identify CpG sites associated with the hazard of bladder cancer outcomes, and then, those identified CpG sites were used for enrichment analyses. Finally, we evaluated

the association between circulating immune cell-type proportions with the cell-type proportions in the tumor microenvironment. We demonstrated that highly circulating CD4T memory and CD8T memory cell proportions were significantly associated with a decreased hazard of tumor recurrence or death, whereas high neutrophil cell proportion, NLR, and age acceleration were associated with an increased hazard of tumor recurrence or death. Collectively, we identified associations of methylation-derived immune profiles and age acceleration with bladder cancer outcomes that may facilitate the development of bladder cancer prognostic biomarkers.

Preface

Completion of this dissertation has been made possible by the invaluable support and contributions from various individuals and institutions throughout my journey as a member of the Molecular & Cellular Biology program at Dartmouth. I am deeply grateful for the opportunities and resources provided to me during this time, including the support from the NIH grant R01CA216265.

First and foremost, I would like to express my deepest gratitude to my advisor, Dr. Brock Christensen, whose unwavering support and guidance have been instrumental throughout my doctoral studies. His mentorship has been invaluable, and I am grateful for the numerous opportunities he provided to enrich my graduate experience and scientific training. I extend my heartfelt appreciation to my thesis committee members, Margaret Karagas, Michael Passarelli, Lucas Salas, and Carmen Marsit, for their support, feedback, and invaluable advice over the past few years.

I am profoundly grateful to the past and present members of the Christensen and Salas lab, and Methylation Cytometry group, whose assistance and support have been essential to the completion of this work. Special thanks go to Lucas Salas for his guidance during my transition into the quantitative sciences and for his invaluable feedback throughout my research. I would also like to acknowledge the Cancer Biology Department for their support during my graduate career.

Beyond the academic realm, I am profoundly grateful to my family for their unending support and unwavering encouragement throughout this journey. Their

love and belief in me have been an endless source of strength and motivation.

Lastly, I would like to express my appreciation to all those who have supported and believed in me along the way. Your encouragement, guidance, and contributions have played an integral role in shaping the outcome of this dissertation.

Contents

List of Tables

List of Figures

Chapter 1

Introduction

1.1 Bladder cancer

1.1.1. Clinical and pathological basis

Bladder cancer, also known as urological cancer or urinary bladder cancer, is a highly heterogeneous epithelial malignancy and urothelial carcinoma accounts for 90% of bladder cancers. The remaining 10% of bladder cancers are squamous cell carcinoma, adenocarcinoma, and small cell carcinoma [1]. Urothelial carcinoma represents 3% of new cancer diagnoses making bladder cancer the twelfth most common cancer worldwide [2]. In the United States, an estimated 17,100 people will die from bladder cancer and there will be 81,180 new cases in 2023. In the United States, bladder cancer is the fourth most common cancer in men and the twelfth most common cancer in women [3]. The major risk factor for bladder cancer is tobacco, accounting for two-thirds of all bladder cancer. Smokers have a two to three-fold increased risk of developing the disease [4]. Since the smoking prevalence is higher in men than in women, bladder cancer is approximately three-fold more common in men [5]. Other risk factors include age, occupational exposure, diet, BMI, pathogens, genetics, and heredity [6], [7].

Bladder cancers are categorized as either muscle-invasive bladder cancer (MIBC) or non-muscle-invasive bladder cancer (NMIBC). Approximately 75% of patients suffer from NMIBC, while 25% are diagnosed with MIBC [8].

According to the European Association of Urology guidelines, NMIBC is characterized by the restriction to the mucosa (stages Tis and Ta) or with invasion only into the underlying lamina propria (stage T1), without invasion into other muscle layers. Conversely, MIBC is defined as bladder cancer that has spread into or through the deeper layers of the bladder, such as muscle (stage T2), or perivesical organs (stages T3 and T4; stages are under the TNM classification system) [9]. Typically, patients with NMIBC have a better prognosis compared with MIBC patients (5-year survival rate is around 70% and 50% respectively) [10], [11]. Despite high survival rates in NMIBC patients, they have a high recurrence rate (> 50%) even with treatments [12].

In addition to pathological characteristics, with the development of sequencing, genetic and transcription-based molecular tumor subtyping has been reported to improve disease classification [13]. For example, bladder cancer patients, including NMIBC and MIBC, have been classified into six molecular subtypes based on gene signaling (*WNT*, *ERBB2*, *FGFR3*, *ALK*, *MAPK*, and *PD1*), with significant differences in overall survival [14]. A combination of pathological parameters with molecular information is now believed to benefit the comprehensive evaluation for the diagnosis, treatment, and prognosis of bladder cancer.

1.1.2. Diagnosis, treatment, and monitoring

Early detection is the best way to improve bladder cancer survival if the cancer has not spread to other tissues. Hematuria, blood in the urine, is the most

common symptom of bladder cancer, present in nearly 85% of bladder cancer patients [15]. Diagnosis is only done with a biopsy and histopathological review. During the process, the urologist uses a cystoscope to examine the bladder to see if any abnormal tissue can be observed. If abnormal tissue is found, transurethral resection of bladder tumor (TURBT) would be performed and tissue samples would be histologically evaluated by a pathologist to determine the need for further surgery or other treatments [16]. Cystoscopy is a highly invasive surveillance method and patients often receive repeated screening to manage recurrences, resulting in discomfort, high morbidity, and the burden of health care costs for patients [7]. Hence, it is urgent to find out an ideal monitoring test that is less invasive, affordable, and high-performance in clinical practice.

According to the examination of clinical factors, such as tumor grade, size, and the location of the primary tumor site, patients receive different treatments. Generally, the initial treatment step is performing TURBT of visible tumors [17]. Following TURBT, depending on how deeply tumors have grown into the layers of the bladder or possible side effects, patients would receive other treatments, including intravesical immunotherapy, chemotherapy, radiation therapy and/or cystectomy, to control tumor recurrence and progression [18]. Intravesical Bacillus Calmette-Guérin (BCG) immunotherapy is the gold-standard treatment after surgery for patients with early-stage or non-muscle-invasive cancers [19]–[21]. Through a catheter, BCG is delivered into the bladder and affects the cells lining the bladder directly. The main effect of BCG is to induce the activation of immune cells to attack bladder cancer cells [22]. Despite response to initial

therapy, around 50% of NMIBC patients still suffer from recurrent tumors [12], [23]. When patients present with muscle-invasive bladder cancer or their tumors are BCG-unresponsive, radical cystectomy and immunotherapy with checkpoint inhibitors would be offered to lower the risk of recurrence and progression [24], [25].

There are many monitoring tests used for bladder cancer recurrence after treatments, such as urinalysis (urine cytology and urinary molecular marker tests), cystoscopy, and computed tomography urography. Urine cytology uses a microscope to identify cancer cells in bladder-washing specimens or urine samples from normal urination, and is often accompanied by other urine tests using molecular analysis [9]. Although urinalysis can assist to detect bladder cancers, it has not been recognized as the only routine monitoring test due to its unstable accuracy and sensitivity [26], [27].

1.2 Association of immune cells with cancer

1.2.1. Immune cell profiles

Among all cells in the peripheral blood, around 0.1 percent of blood cells are leukocytes comprised of granulocytes, monocytes, and lymphocytes. The leukocytes in peripheral blood derive from the hematopoietic stem cells (HSCs) in the bone marrow, where many of them mature. HSCs are multipotent primitive cells that can produce and differentiate into all lineage blood cells. Initially, HSCs give rise to intermediate progenitor cells of platelets, red blood cells,

myeloid progenitors, and lymphoid progenitors. These intermediate progenitors have lineage-committed properties before maturation. After release from the bone marrow, HSCs and intermediate progenitor cells circulate throughout the body by blood and lymph to reach the thymus, spleen, liver, or lymph nodes for further development and maturation [28]. The lineage relationships of the different types of immune cells are shown in Figure 1.1.

Myeloid progenitors are the precursor of monocytes and granulocytes. Monocytes represent around 10% of the nucleated cells in the blood in the healthy human body, and their main function includes phagocytosis, antigen presentation, cytokine secretion, and differentiation into dendritic cells and macrophages [29]. Dendritic cells play a critical role in detecting infection, presenting antigens, and inducing a primary immune response by the activation of naïve T cells [30]. Macrophages are responsible for engulfing and digesting pathogens, such as cellular debris and cancer cells [31]. Besides, macrophages can regulate immunity, for instance, initiating adaptive immunity by presenting antigens to T cells or releasing cytokines to suppress immune reactions [32].

There are four types of granulocytes: mast cells, basophils, eosinophils, and neutrophils. Mast cells involve in bacterial, parasitic, and viral clearance and immune tolerance and are able to present antigens and regulate the activation and migration of dendritic cells to lymph nodes [33]. Basophils protect human body against allergens, bacteria, and parasites through releasing histamine and prostaglandins to induce the inflammatory response [34]. Eosinophils are a circulating granulocyte and own multi-function, including parasites prevention,

lymphocyte recruitment, antigens presentation, and helper T cell polarization [35]. Neutrophils, the most common type of leukocyte, are the first immune cells to respond to eliminate the invading pathogens by phagocytosis, degranulation, and the formation of neutrophil extracellular traps. They can both upregulate and downregulate inflammation and the immune system by producing cytokines under different situations [36].

Lymphoid progenitors differentiate into three types of lymphocytes, T cells, B cells, and natural killer (NK) cells. This process is called lymphopoiesis. Once lymphocytes complete maturation, they enter circulation and migrate to protect tissues from disease-causing pathogens and abnormal cells, such as cancer cells. Peripheral lymphocytes are approximately 70 to 85 percent T cells and 5 to 10 percent B cells, and 5 to 20 percent NK cells [37]. These three types of lymphocytes correspond to innate and adaptive immune cells.

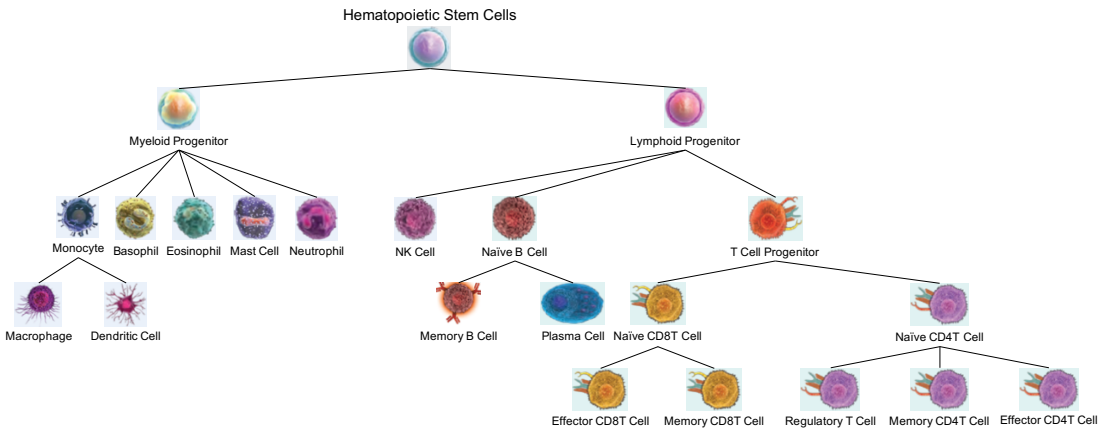


Figure 1.1: Immune cell lineage. Adapted from Peprotech Inc, 2021. Schematic illustration displayed hematopoiesis of leukocyte. Hematopoietic stem cells have differentiation capability into leukocyte lineages.

T cells play a crucial role in the adaptive immune. There are two main classes of T lymphocytes, cytotoxic T cells (CD8T) and helper T cells (CD4T). CD8T cells are important for the defense against virus-infected cells as well as tumor cells through the secretion of cytotoxic granules and cytokines, such as TNF-α [38]. Resting CD8T naïve cells normally circulate between the blood and lymphoid organs. During viral infection or tumor formation, naïve CD8T cells are primed and activated by antigen-presenting cells, differentiating into effector CD8T cells and memory CD8T cells [38]. After the clearance of pathogens and tumor cells, most effector CD8 T cells die but memory CD8T cells that are antigen-specific remain in the circulation and tissues and provide an enhanced protective response to control secondary infections more rapidly [39]. The main function of CD4T cells is immune response regulation, such as the activation of memory B cells and effector CD8T cells. With antigenic stimulation through antigen-presenting cells, naïve CD4T cells proliferate and differentiate into specific effector cells based on the cytokine milieu [40]. For now, at least seven functionally distinct subsets of effector CD4T cells have been observed to balance adaptive immunity. After the clearance of threats, a small proportion of CD4T cells survive and differentiate into memory CD4T cells for expanding to large numbers of effector cells once re-exposure to threats [41]. Besides the activation of memory and effector cells, one of the subsets of CD4T cells, regulatory T cells, is able to suppress T cell-mediated immunity for the maintenance of immunologic tolerance [42].

B cells are responsible for the humoral immunity of the adaptive immune

system. When harmful pathogens or abnormal cells emerge in the body, naïve B cells are able to recognize antigens on harmful agents and then differentiate into plasma cells and memory B cells with the assistance of follicular dendritic cells and T follicular helper cells [43]. Plasma cells produce protective antibodies in response to antigens and memory B cells are able to respond to reinfection with pathogens and tumor recurrence in the future [43]. NK cells are lymphocyte-like cells and are part of the innate immune system. They can recognize and kill abnormal cells, such as tumor cells, and intracellular pathogens by releasing cytotoxic granules, limiting the spread of tumors and microbial infections [44].

1.2.2. Immune system with bladder cancer development

Immunotherapy, such as BCG and immune checkpoint inhibitors (ICI), are the most common treatments after TURBT surgery to suppress the tumor relapse. BCG, a live attenuated strain of *Mycobacterium bovis*, activates the immune system to prevent or delay bladder cancer from becoming invasive or relapsing after surgery [45]. Many mechanisms of BCG have been proposed. For example, BCG generates pathogen-associated molecular patterns, such as toll-like receptors expressed on monocytes, B and, T cells, recognized by host pattern recognition receptors, inducing the production of proinflammatory cytokines [46]. In addition, innate immune cells, such as macrophage and NK cells, demonstrate increased cytotoxicity against BCG-infected bladder cancer cells [47], [48]. With BCG treatment, dendritic cells can process antigens and present them to recruit and activate tumor-specific CD4 and CD8T cells [49]. A recent

study has demonstrated that BCG stimulation induces epigenetic changes in innate immune cells, resulting in enhanced cytokine secretion [50].

Several studies have demonstrated that resistance to intravesical BCG therapy is associated with upregulated PD-L1 expression at the surface of cancer cells, including bladder cancer [51], [52]. The interaction of PD-1 and its ligand, PD-L1, downregulates the immune response to protect from autoimmune damage in normal conditions [53], [54]. Tumor cells exploit PD-1- PD-L1 pathway to inhibit proapoptotic and escape immunity, such as CD8T cell-mediated cytotoxicity [55]. The function of ICI drugs is to block the interaction of PD-1 and its ligand, promoting the immune system to attack cancer cells [56]. However, only 30% of metastatic urothelial carcinoma patients are sensitive to the checkpoint inhibitor and BCG treatments, and the remaining 70% of patients have cancer cells that are able to escape immune detection eventually leading to poor outcomes. Further understanding of the interaction between the immune system and bladder cancer cells is needed to improve prognostication and for developing novel treatments [56].

During tumor development, bladder cancer cells have been shown to modulate the immune system, creating a tolerant microenvironment. Bladder cancer cells overexpress sphingosine-1-phosphate receptor-1 to promote the expression of TGF-β and IL-10, inducing regulatory T-cell (Treg) expansion in the tumor microenvironment [57]–[59]. Abundant infiltrated Treg cells suppress the activation of cytotoxic T cells [60]. Similar to Treg cells, myeloid-derived suppressor cells (MDSCs) are one of the suppressor cell subsets that exhibit

immunosuppressive effects [61] and assist tumor maintenance [62]. Bladder cancers are able to secret CXC-chemokine ligand 2 (CXCL2) to recruit MDSCs [63]. Both abundant Treg and MDSC infiltration was reported to be associated with a poor prognosis [60], [63]–[65].

Multiple studies show that bladder cancer cells produce CCL2, IL-10, and TGF-β polarizing tumor-associated macrophages (TAM) into M2-like macrophages that are characterized by an immunosuppressive and anti-inflammatory phenotype [66]–[69]. High infiltration of M2-like TAM was associated with a poor outcome in bladder cancer patients [70], [71]. In addition to infiltrating immune profiles, a few immune cell types in blood have also been indicated to involve in immune modulation. Neutrophils have been found in higher proportions in peripheral blood and tumor microenvironment and have immunosuppressive properties in bladder cancer patients [72]. Also, high neutrophil-to-lymphocyte ratio (NLR) in peripheral blood has been associated with worse outcomes for bladder cancer patients [73]–[75]. Based on the above studies, future investigations are needed to uncover mechanisms of immunosuppression involved in bladder cancer development and recurrence to improve and/or design therapeutic strategies.

1.3 DNA methylation

1.3.1. Regulation of gene expression

DNA methylation is an epigenetic mark and a covalent modification of DNA

involving the addition of a methyl group to the C-5 position of the cytosine ring in the cytosine-phosphate-guanine (CpG) of DNA to form 5-methylcytosine through DNA methyltransferases (DNMTs) (Figure 1.2). Without changing DNA sequences, DNA methylation can recruit proteins related with gene expression or inhibit transcription factors binding to regulate biological phenotypes and activities [76].

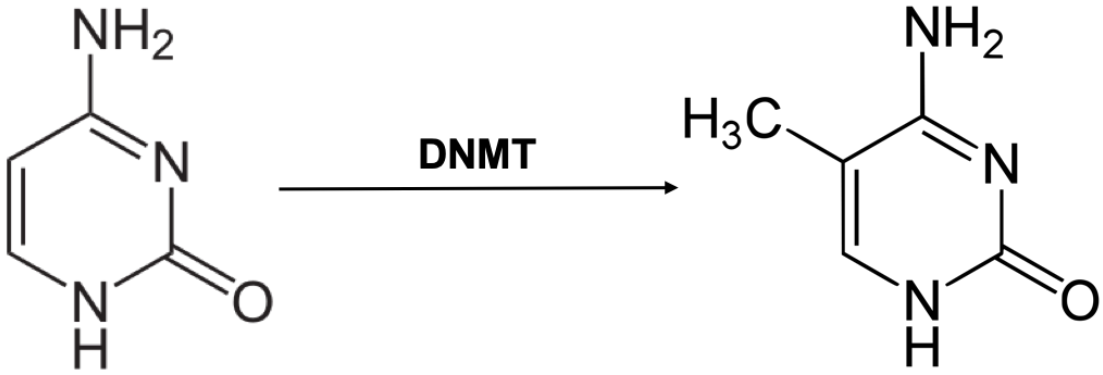


Figure 1.2: Cytosine is modified to 5-methylcytosine by DNA methyltransferase

In mammals, DNA methylation occurs at CpG dinucleotides distributed unevenly throughout the entire genome [77]. Regions with a high frequency of CpG sites, known as the CpG islands, are stretches of DNA with a higher than expected concentration of CpG sites and are typically hypomethylated [78]. In addition to CpG islands, CpG genomic context can be defined into a few regions. Shores are the regions flanking CpG islands and are located as far as 2 kb from CpG islands; shelves are regions from 2 to 4 kb from CpG islands; open seas are the rest of the genome and commonly overlap with enhancer regions [79]. Most of the known genes have a CpG island within the promoter region.

The impact of DNA methylation on gene regulation is dependent on genomic context. Generally, DNA methylation of gene promoter and enhancer regions leads to the repression of gene expression due to impeding the binding of activating transcription factors [80]. Contrary to promoter and enhancer methylation, the methylation of the gene body might not only stimulate transcription but impact splicing [81]. However, some studies have reported that gene body methylation is not associated with increased gene expression in non-dividing and slowly dividing cells [82]. As a result, the association of DNA methylation with gene regulation at different regions relative to genes still needs further exploration, and may be cell-type-dependent.

Because DNA methylation modulates gene expression, many studies have investigated the association of methylation with various phenotypic traits and clinical characteristics. For instance, DNA repair genes with hypermethylated promoters result in the accumulation of DNA damage, promoting early steps in cancer progression [83]. In recent years, age has been associated with DNA methylation [84], [85], and improvements in high-throughput arrays measuring methylation led to the development of accurate age estimators [86]–[88]. Due to the cell type-specific differences in DNA methylation profiles, each cell type has developed its specific phenotypes and functions. With this understanding, several studies had utilized the differentially methylated DNA regions (DMRs) to estimate distinct cell lineages with high sensitivity and specificity [89], [90].

1.3.2. Measurement of DNA methylation

DNA methylation can be measured through different methods. The well-known approaches are dependent on the conversion of genomic DNA by using sodium bisulfite. In brief, during sodium bisulfite treatment, unmethylated cytosines are deaminated to uracils, while methylated cytosines are preserved. After bisulfite conversion, whole genome amplification results in uracils being replaced by thymine, whereas methylated cytosines remain cytosines [91], [92].

The signal of the amplified products of bisulfite-treated DNA can be quantified using whole genome bisulfite sequencing and high-density microarray. In this thesis, we rely on the microarray, Illumina Infinium MethylationEPIC BeadChip (EPIC) [91], to measure methylation. The EPIC array is able to interrogate around 860,000 CpG sites. The fluorescence signals from the EPIC array are converted into methylation beta values representing the methylation levels for further analyses.

1.4 DNA methylation-derived variables

1.4.1. Cell type deconvolution from DNA methylation

Even though cells throughout the body have the same DNA sequences, they generally display different phenotypes. Numerous studies show that, in mammals, a specific cell type expresses a specific subset of protein coding genes for required functions and silences others through DNA methylation [93], [94], controlling cell identity in development. Hence, each tissue and cell type exhibits differences in epigenetic modifications of DNA at genomic regions that result in

the unique DNA methylation profile of each cell type [95]. Therefore, researchers can leverage this biological feature to differentiate cells types from different tissues or within a tissue [96].

Bulk tissue specimens comprise a heterogeneous mixture of different cell types. Since the DNA methylation measurement of a specimen is an average level across all cells in the sample, the measured results reflect cell-type specific methylation profiles according to the proportion of component cell types. If we investigate differentially methylated regions (DMRs) associated with phenotypes or outcomes of interest without controlling for the cellular composition, detected changes will be confounded by cell-type variation among subjects [97]. Therefore, it is important to adjust for cellular composition during the analyses of DNA methylation profiles. The cellular composition can be estimated using DNA methylation profiles from bulk specimens [98], [99]. The cell-type deconvolution method was first developed in 2012 [90].

Compared with traditional methods, such as flow cytometry, DNA methylation-based cell type deconvolution has the following advantages: (1) it does not require fresh samples because it is a DNA-based method, (2) it is objective, and (3) DNA methylation-based deconvolution can estimate several distinct immune cell types at once [98]. With cell-type deconvolution approaches, it is possible to explore the cellular composition in archived samples [100]. For cancer studies, peripheral blood and tumor tissue are the two types of specimens used most. In recent years, a lot of advanced deconvolution methods were developed to infer the cellular composition of blood and tumor

microenvironment based on methylation profiles and the latest microarray platform, making researchers able to compare cell-type proportions between cases and controls and control for cell-type mixtures in outcome analyses [89], [101], [102].

1.4.2. DNA methylation age clock

Understanding why and how we age is a critical question in biology. A lot of mechanisms and molecular hallmarks have been identified that are associated with aging, including telomere length, metabolomic variations, and epigenetic changes [103], [104]. Among these biomarkers, DNA methylation is the most promising to estimate an individual's age since methylation alterations are conspicuous across an individual's lifespan [105]–[107] and the changes are measured accessibly with the development of high-throughput arrays [84], [108]. With high-solution data, highly correlated ($r > 0.8$) age estimators, called DNA methylation clocks, were constructed and may reflect biological aging [87], [88], [109], [110]. There are two types of DNA methylation clocks, chronological DNA methylation clock and biological DNA methylation clock.

For chronological DNA methylation clock, the algorithm selects CpG sites associated with chronological age and this type methylation clock reflects age-related DNA methylation alterations shared between individuals [86]. The well-known chronological DNA methylation clocks are the Horvath clock [87] and Hannum clock [88]. The Horvath clock composed of 353 CpG sites was trained on various cell and tissue types and hence not confounded by cell and tissue-

specific alterations [106]. The Hannum clock was constructed on blood samples and consists of 71 CpG sites. These chronological DNA methylation clocks are powerful for predicting actual age, however, they only can capture associations with mortality at a population level, not an individual level [111].

Although chronological DNA methylation clocks can predict forensic age with high accuracy, they lose some variabilities for estimating associations of biological aging [112]. As a result, biological DNA methylation clock was developed to investigate biological age more directly. Unlike chronological clocks, the algorithm of biological clocks selects CpG sites associated with aging phenotype and biological clocks reflect the inter-individual variability in DNA methylation alterations related to aging decline and disease [86]. The most popular biological clock is PhenoAge which is a blood-based biological clock comprised of ten measures, such as chronological age, mean cell volume, and lymphocyte percent [113]. In general, biological DNA methylation clocks outperform chronological DNA methylation clocks in predicting mortality risk by cancer, health span, and physical function [113], [114].

Each reported DNA methylation clock has its own strengths and weaknesses. It is critical to understand which DNA methylation clock would be suitable for interpreting the association of aging with health outcomes. While the DNA methylation clock is a new research field and is much promise for improved human health, there are more challenges, such as no gold standard of biological aging, limited sample size, and under-exploration of the association between other DNA modifications and age, waiting for us to overcome [115].

1.5 Summary

In the subsequent dissertation, I present an independent investigation of DNA methylation and its derived factors in bladder cancer with an urgent need for less-invasive prognostic biomarkers. First, I assess the association of six methylation-derived immune cell-type proportions, including monocyte, neutrophil, NK, CD4T, CD8T, and B cell, in the circulating system with overall survival and recurrence-free survival in NMIBC patients. Then, an epigenome-wide association study (EWAS) is conducted to identify CpG sites associated with tumor outcomes. Next, I investigate the association of age acceleration and immune cell-type proportions, such as memory, naïve T and B cells, and granulocyte subtypes, estimated from the latest cell-type deconvolution library with bladder cancer outcomes. Finally, I explore the association between immune profiles from both peripheral blood and tumor tissue. This work provides implications for improving the development of bladder cancer prognostic biomarkers.

Chapter 2

Immune profiles and DNA methylation alterations related with non-muscle-invasive bladder cancer outcomes

This work was published on January 21, 2022 as a *Research Article* in *Clinical Epigenetics*, Volume 14, Issue 1, Article number 14. PMID: 35063012.

The following authors have contributed to the content in this chapter:

Chen, JQ, Salas, LA, Wiencke, JK, Koestler, DC, Molinaro, AM, Andrew, AS, Seigne, JD, Karagas, MR, Kelsey, KT, Christensen, BC.

Conceptualization: JQC, BCC, LAS, JKW, DCK, AMM, MRK, KTK; Formal Analysis: JQC, BCC, LAS; Data Curation: ASA, JDS; Methodology: JQC, BCC, LAS, JKW, DCK, AMM, MRK, KTK; Writing–Original Draft: JQC; Writing–Review and Editing: JQC, BCC, LAS, JKW, DCK, AMM, MRK, KTK; Funding Acquisition: BCC

2.1 Abstract

Non-muscle-invasive bladder cancer (NMIBC) patients receive frequent monitoring because ≥ 70% will have recurrent disease. However, screening is invasive, expensive, and associated with significant morbidity making bladder

cancer the most expensive cancer to treat per capita. There is an urgent need to expand the understanding of markers related to recurrence and survival outcomes of NMIBC. We used the Illumina HumanMethylationEPIC array to measure peripheral blood DNA methylation profiles of NMIBC patients (N = 603) enrolled in a population-based cohort study in New Hampshire and applied cell type deconvolution to estimate immune cell-type proportions. Using Cox proportional hazard models, we identified that increasing CD4T and CD8T cell proportions were associated with a statistically significant decreased hazard of tumor recurrence or death (CD4T: HR = 0.98, 95% CI = 0.97 – 1.00; CD8T: HR = 0.97, 95% CI = 0.95 – 1.00), whereas increasing monocyte proportion and methylation-derived neutrophil-to-lymphocyte ratio (mdNLR) were associated with the increased hazard of tumor recurrence or death (monocyte: HR = 1.04, 95% CI = 1.00 – 1.07; mdNLR: HR = 1.12, 95% CI = 1.04 – 1.20). Then, using an epigenome-wide association study (EWAS) approach adjusting for age, sex, smoking status, BCG treatment status, and immune cell profiles, we identified 2528 CpGs associated with the hazard of tumor recurrence or death ($P < 0.005$). Among these CpGs, the 1572 were associated with an increased hazard and were significantly enriched in open sea regions; the 956 remaining CpGs were associated with a decreased hazard and were significantly enriched in enhancer regions and DNase hypersensitive sites. Our results expand on the knowledge of immune profiles and methylation alteration associated with NMIBC outcomes and represent a first step toward the development of DNA methylation based biomarkers of tumor recurrence.

2.2 Introduction

In 2021, the estimated number of bladder cancer deaths is projected to be 17,200, with an estimated number of new cases of 83,730 in the USA. Bladder cancer is the fourth most common cancer among men and twelfth most common among women, which may in part be due to smoking prevalence rates being higher in men than women as cigarette smoking accounts for half of all cases (47%) in the USA [116]. Seventy-five percent of bladder cancers are diagnosed as low-grade non-muscle invasive tumors, NMIBC [117]. Cystoscopy is used for diagnosis (biopsy) with transurethral excision of localized tumors as the primary treatment [118]–[120]. Although transurethral excisions can successfully control the disease and mortality from bladder cancer among patients with localized tumors is low, 45% of NMIBC cases have recurrences within 12 months of surgery [121]. In addition, frequent invasive follow-up via cystoscopy without prognostic markers leads to significant patient morbidity and comes with a considerable cost burden to the health care system, estimated at approximately $4 billion dollars annually in the USA [122]. To control the patient and healthcare burdens associated with NMIBC, there is an imminent need for biomarkers to identify those at the highest risk of tumor recurrence.

Peripheral blood immune profiles have been associated with different outcomes in bladder cancer patients and may have clinical utility for NMIBC prognosis [123]–[128]. For instance, in NMIBC, patients with elevated neutrophil-to-lymphocyte ratio (NLR) had poorer cancer-specific survival than

patients with lower NLR [124]–[126]. Thus, elevated NLR could be a potential predictor of overall survival and cancer-specific survival for this disease. Other studies have indicated that bladder cancer patients with increased lymphocyte-to-monocyte ratio had poorer overall survival and cancer-specific survival [127], [128]. In addition, Bacillus Calmette–Guérin (BCG), a commonly used intravesical immunotherapy for NMIBC administered post-surgery, has been reported to reduce the proportion of natural killer T cells, memory CD4T, CD8T, and regulatory T cells in the peripheral blood of NMIBC patients [123]. Together, circulating immune profiles might be promising markers for reducing adverse outcomes in NMIBC patients. Previous studies have relied on complete blood count differential (CBC) tests to determine immune profile variables [129], [130]. The CBC test relies on fresh blood samples and is incapable of providing proportions of specific lymphocyte subtypes [131].

Our prior work has established methods to infer the immune profiles in archival samples using immune cell-type-specific DNA methylation [89], [132]. DNA methylation plays an essential role in gene regulation for cell lineage specification [133], [134]. Differentially methylated regions (DMRs) have been used to distinguish cell types, including leukocyte subtypes, and form the basis of reference-based deconvolution methods for estimating specific immune cell-type proportions [89], [90], [135]. Compared with cytological methods for determining cell type abundances and proportions, such as flow cytometry, DNA methylation-based cell-type deconvolution does not require a fresh substrate, intact cells, or batch-sensitive reagents, is reproducible and cost-effective

relative to time-sensitive blood processing [98], [136]. With cell-type deconvolution approaches, it is possible to identify immune cell-type profiles and test their relationship with cancer outcomes in archived samples [100]. This study used the archival blood samples of NMIBC patients to test the association between the immune profiles and outcomes in bladder cancer patients.

In the present study, we sought to identify immune profiles and epigenetic features associated with disease recurrence in the hope that such information might help improve the management of NMIBC. Here, we hypothesized that CpG-specific DNA methylation and DNA methylation-derived immune cell profiles are associated with recurrence-free survival in NMIBC patients. We used archival blood samples from a population-based case–control study to obtain genome-scale DNA methylation profiles. We then investigated the association between the methylation-derived immune profiles and outcomes in NMIBC patients. Preliminary work from our group observed an association between methylation-derived NLR (mdNLR) and survival in bladder cancer patients using a smaller sample size (223 cases) and an early genome-scale methylation array (HumanMethylation27K array) [137]. In this study, we increased the sample size (603 cases), used a new more comprehensive array (HumanMethylationEPIC array) with 30 times as many measured features, and a new cell type deconvolution library [89] to estimate immune cell-type proportions. An epigenome-wide association study (EWAS) and enrichment analyses were used to determine possible CpG sites and gene sets associated with recurrence-free survival.

2.3 Methods

2.3.1. Study Subjects and Samples

The subjects and data used in this study are described in more detail in prior publications [138]–[140]. Briefly, subjects were recruited from all three phases of a New Hampshire population-based bladder cancer case–control study [141]. The first wave of this study (phase 1) collected blood samples from 331 individuals diagnosed with incident bladder cancer between July 1994 and June 1998. The second study wave (phase 2) collected blood samples from 243 individuals diagnosed between July 1998 and December 2001. Finally, the third study wave (phase 3) obtained blood samples from 194 individuals recruited and diagnosed between July 2002 and December 2004. All the subjects were identified using the New Hampshire State Cancer Registry, hospital pathology departments, and hospital cancer registries, and all blood samples were collected after the time of diagnosis (time range: 20 – 1790 days). Among patients, 40 patients received BCG treatment in Phase 1, 29 patients received BCG in Phase 2, and 19 patients received BCG in Phase 3. All patients with BCG treatment had blood drawn after receiving BCG (time range: 7 – 1542 days). An outline of data filtering and inclusion/exclusion criteria applied to these data are shown in **Figure 2.1**. Briefly, subjects without muscle-invasive status, histopathology re-review, tumor grade, smoking status, or pack-years were removed from the study. Subjects that withstood the aforementioned exclusion criteria were retained and used in downstream statistical analyses.

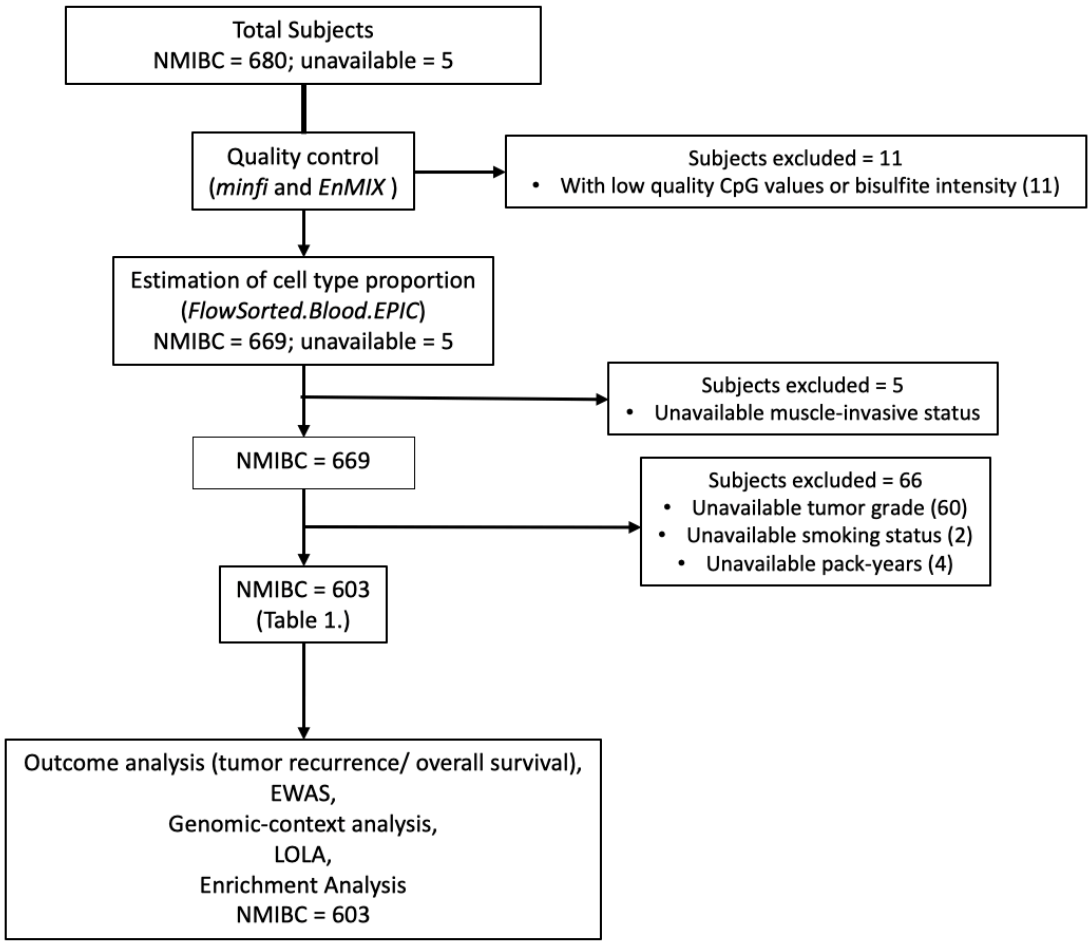


Figure 2.1: Flow chart of study

2.3.2. DNA Extraction, Quantification, and Bisulfite Modification

Each blood sample was maintained at 4°C and frozen within 24 hours of the blood draw. A hundred µl buffy coat was used to extract DNA. The QIAMP DNA blood & Tissue kit was used to extract DNA from blood samples according to the manufacturer's protocol. Extracted DNA was quantified by using the Qubit 3.0 Fluorometer. After bisulfite modification, an established 5mC microarray protocol optimized for Illumina methylation arrays was used to determine the genome-wide 5mC profile. DNA samples were subjected to bisulfite conversion (according to the manufacturer's protocol of the Zymo EZ DNA methylation Kit)

with an input of 750 ng per sample and whole-genome amplified prior to array hybridization. Recovered substrate ssDNA were submitted for DNA methylation array processing.

2.3.3. DNA Methylation Data

Bisulfite-modified DNA samples were measured for their DNA methylation status using the MethylationEPIC array, which interrogates > 860,000 CpG sites. Probe intensity data (iDAT files) from the EPIC methylation array were processed for quality control via the R package *minfi* [142] and *ENmix* [143] in R version 3.6. After quality control, 11 samples with low-quality CpG values or bisulfite intensity (threshold: 7,000) were excluded from the study. The data were then normalized and conducted background correction through *preprocessNoob* procedure from *minfi*. The *ComBat* [144] was used to adjust for potential batch effect. Probes with a detection $P > 1.0 \times 10^{-6}$ in more than 10% of the samples were excluded (32,414). Also, 98,826 probes, which are crossreactive, SNP-associated, and non-CpG (CpH) methylation [145], as well as 17,120 probes on sex chromosomes, were excluded. In total, 726,856 probes were used in downstream statistical analyses in downstream statistical analyses. *IlluminaHumanMethylationEPICanno.ilm10b4.hg19* [146] was used to annotate CpG sites. Relation to CpG island was defined by the "*Relation_to_Island*" as in the Illumina annotation used in the genomic context analysis. '5'UTR,' 'Exon,' 'Gene Body,' and '3'UTR' contexts were defined by having '5UTR,' 'ExonBnd,' 'Body' and '3UTR' in *UCSC_RefGene_Group*. 'Enhancer' context was defined

by having a record in the *Phantom5_Enhancers*. 'DHS' context was defined by finding a record in the *DNase_Hypersensitivity_NAME*. 'TFBS' context was defined by having a record in the *TFBS_NAME*.

2.3.4. Statistical Analysis

The estimation of cell-type proportions was processed through the *estimateCellCounts2* from the *FlowSorted.Blood.EPIC* package in Bioconductor (version 3.9; [89]). Methylation-derived neutrophil-to-lymphocyte ratio (mdNLR) was calculated by performing cell-mixture deconvolution to estimate the proportion of leukocyte subtypes and the ratio of neutrophil proportion to lymphocyte proportion was then computed. Individual leukocyte cell-type proportions and mdNLR were included in outcome analyses as continuous variables for Cox proportional hazard regression. In addition, mdNLR was dichotomized based on the median mdNLR for the Kaplan–Meier method.

Ten-year overall survival was defined as the time interval from the date of initial diagnosis to death within 10 years (all deaths were related with bladder cancer). Subjects who were alive or lost to follow-up were censored at the last follow-up. Ten-year recurrence-free survival was defined as the time interval from the date of initial diagnosis to the first tumor recurrence or death (all causes), whichever occurred first within 10 years. Patients alive and free of the disease or lost to follow-up were censored at the last follow-up. For both overall and recurrence-free survival, patient survival times over 10 years were truncated at 10 years, and patients were censored if the first tumor recurrence or death

occurred after 10 years. The median survival times for the two survival outcomes were estimated using the Kaplan–Meier method. In multivariable analyses, Cox proportional hazard regression models were used to examine the association of each variable on bladder cancer outcomes and were fit via *coxph* in the *survival* R package. The proportional hazards assumption was tested by using *cox.zph* from the *survival* R package. The *cox.zph* function tests the proportionality of all the predictors in Cox models by creating interactions with time. As sex and BCG treatment status violated the proportional hazards assumption, stratification on both variables was included. The linearity assumption was examined via *ggcoxfunctional* from the R *survminer* package, and methylation-derived immune cell profiles were found to violate the linearity assumption. Hence, winsorization was used on methylation-derived immune cell profiles. The winsorization cutpoint of each immune cell profile is shown in **Figure S2.1**. A P-value of < 0.05 was the significance threshold on multivariable analysis. Cox model results were presented using the *stargazer* R package.

2.3.5. Epigenome-Wide Association Study (EWAS), Enrichment Analysis, and Differentially Methylated Regions (DMRs)

An epigenome-wide association study (EWAS) was performed using *ewaff* R package (https://github.com/perishky/ewaff) to investigate the association of CpG-specific DNA methylation and bladder cancer recurrence. We fit Cox proportional hazards models independently to each CpG, controlling for age, sex, tumor grade, smoking status, Bacillus Calmette–Guérin (BCG) receiving status,

and immune cell profiles. Since all P-values adjusted for false discovery rate from EWAS results are higher than 0.05, we relaxed the threshold for genomic context and enrichment analyses using a P-value of < 0.005.

For CpG sites associated (P < 0.005) with bladder tumor recurrence, we examined whether those CpGs were enriched in CpG island-related genomic context or regulatory regions via using Mantel–Haenszel tests, adjusted for Illumina probe type to eliminate the difference in distributions on genomic context. CpG island-related genomic context includes open sea, north shelves, north shores, islands, south shores, and south shelves. Regulatory regions include enhancers, DNase hypersensitivity sites, 5'UTR, TSS1500, TSS200, 1st Exon, Exon, gene body, 3'UTR, and transcription factor binding sites. Next, the Locus Overlap analysis (*LOLA*) R package in Bioconductor (version 1.20.0; [147]) was used to investigate enrichment of genomic regions limited to tissue equal to "hematopoietic stem cell." Finally, *gometh* and *gsameth* from the *missMethyl* package in Bioconductor (version 1.6.2; [148]) were used to test the enrichment of gene sets for Gene Ontology (GO) terms and the C7: immunologic signature gene set in the Gene Set Enrichment Analysis (GSEA) Molecular Signature Database (MSigDB) *cnetplot* from the *enrichplot* package in Bioconductor (version 1.10.1; https://yulab-smu.top/biomedical-knowledge-mining-book/) was used to plot gene-concept network plots.

Differentially methylated regions were identified and extracted through the *dmrcate* and *extractRanges* from the *DMRcate* R package [149]. The inputs were logit-transform of beta values (M-values). The phenotype of interest for

comparison was “NMIBC patients with tumor recurrence within ten years or not” in our designed model. In addition, the designed model was adjusted for sex, age, tumor grade, smoking status, BCG receiving status, and immune cell profiles. We relaxed the threshold for DMRs analysis using an FDR-corrected P-value of < 0.1. Then, the *visualizeGene* from the *sesame* package [150] was used for observing the methylation levels of CpGs in regions identified by *DMRcate*.

2.4 Results

2.4.1. Characteristics of subjects

Profiles of DNA methylation were obtained from 685 peripheral blood samples using the Human MethylationEPIC array. Eighty-two subjects were excluded due to low-quality CpG value or bisulfite intensity (n = 11), or without muscle-invasive status, histopathology re-review, tumor grade, smoking status, and pack-years (n = 71) (**Figure 2.1**). The remaining subjects (N = 603) included in the study group were 75.8% men (n = 457), 82.9% ever-smokers (n = 500), and had a median age of 66 (**Table 2.1**). We estimated the cell-type proportions for each patient using methylation values by performing *FlowSorted.Blood.EPIC* (see **Figure S2.1** for the distribution). Neutrophil-to-lymphocyte ratio (NLR) was then calculated according to the ratio of neutrophil proportion to lymphocyte proportion (B cell + CD4T cell + CD8T cell + NK cell), and the median methylation-derived NLR (mdNLR) was 1.97. Further details of study population characteristics are described in **Table 2.1**.

Table 2.1: Characteristics of subjects after excluding subjects with missing values[#]

	NMIBC (n = 603)
Age	
Median (Q1, Q3)	66 (57,71)
Sex	
Male	457 (75.8%)
Female	146 (24.2%)
Tumor grade	
Low Grade	452 (75.0%)
High Grade	151 (25.0%)
Smoking status	
Never	103 (17.1%)
Ever	500 (82.9%)
BCG: Immunotherapy	
No	514 (85.2%)
Yes	89 (14.8%)
NLR	
Median (Q1, Q3)	1.97 (1.40, 2.93)
10-year dead status	
Alive	423 (70.1%)
Deceased	180 (29.9%)
10-year Survival	
Median (Q1, Q3)	120.0 (105, 120)
10-year Recurrence status	
No	193 (32.0%)
Yes	295 (48.9%)
Missing	115 (19.1%)
10-year Recurrence	
Median (Q1, Q3)	17.7 (6.9, 67.1)
Missing	115

[#]: Exclude subjects with missing values in muscle-invasive status, pathology reviewing status, tumor grade, smoking status, or pack-years

2.4.2. Associations of patient and tumor characteristics with bladder cancer outcomes

To characterize 10-year recurrence-free survival (RFS), we generated Kaplan–Meier curves for each covariate and fit a Cox proportional hazard regression model for univariate analyses and multivariable analyses, respectively. In the Kaplan–Meier analysis, the NMIBC patients aged > 65 had a worse probability of RFS compared with the NMIBC patients with age ≤ 65 (P = 0.0006). Females had a greater probability of RFS compared with males (P = 0.002), and the NMIBC patients with high-grade tumors (grade 3 + 4) had a lesser probability of RFS than those with low-grade tumors (grade 1 + 2) ($P = 8.0 \times 10^{-5}$). Ever-smokers had a worse probability of RFS than the NMIBC patients who were never smokers ($P = 3.0 \times 10^{-4}$). NMIBC patients with low mdNLR had a greater probability of RFS than the patients with high mdNLR (P = 0.002) (**Figure 2.2**). Consistent with the Kaplan–Meier results, in a multivariable Cox model, age > 65 (HR = 1.01, 95% CI = 1.00 – 1.03), high tumor grade (HR = 1.48, 95% CI = 1.17 – 1.87), ever-smoking (HR = 1.65, 95% CI = 1.22 – 2.25), and mdNLR (HR = 1.12, 95% CI = 1.04 – 1.20) were significantly associated with an increased hazard of RFS (**Table 2.2**).

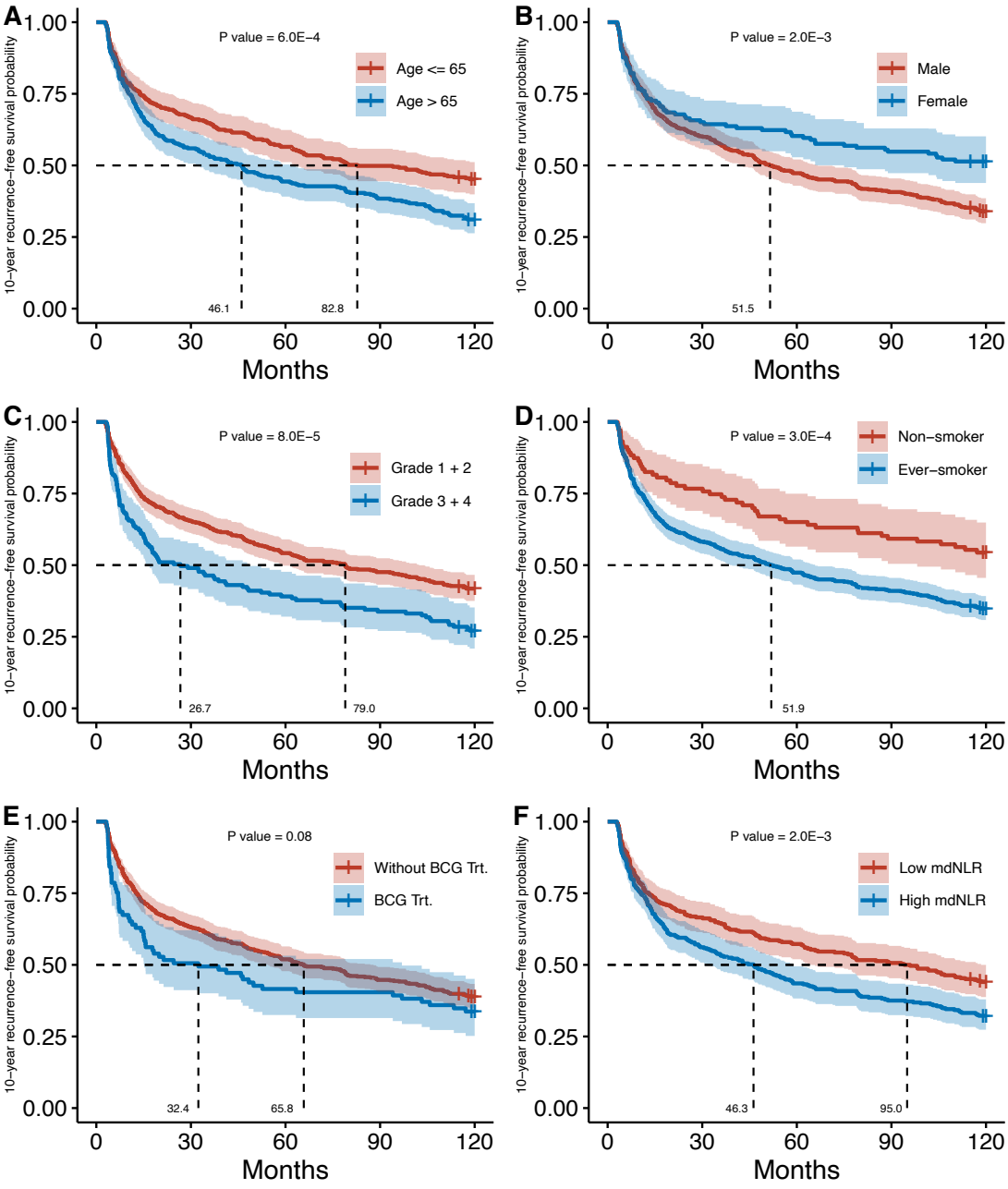


Figure 2.2: **Kaplan–Meier analysis of 10-year recurrence-free survival (RFS).** 10-year RFS curves stratified by **(A)** age, **(B)** sex, **(C)** tumor grade, **(D)** smoking status, **(E)** BCG treatment status or **(F)** mdNLR level. P-values for log-rank tests are shown

Table 2.2: Cox proportional hazards 10-year recurrence-free survival models

			Event occurrence months		Multivariable[$] model	
	n (%) event[#]	n (%) no-events[#]	Mean	Median	HR (95% CI)	P value
Age	373 (61.9)	230 (38.1)	65.6	62.0	1.01 (1.00-1.03)	0.024
Tumor grade						
Low	263 (58.2)	189 (41.8)	69.9	79.0	Referent group	
High	110 (72.8)	41 (27.2)	52.8	26.7	1.48 (1.17-1.87)	1.0E-3
Smoking status						
Non-smoker	47 (45.6)	56 (54.4)	82.3	120.0	Referent group	
Ever-smoker	326 (65.2)	174 (34.8)	62.2	51.9	1.65 (1.22-2.25)	1.0E-3
mdNLR*	373 (61.9)	230 (38.1)	65.6	62.0	1.12 (1.04-1.20)	3.0E-3

HR: hazard ratio, CI: confidence interval, mdNLR: methylation-derived neutrophil to lymphocyte ratio

Stratification was used on sex and BCG treatment status for proportional assumption.

$: The model controlling for age, sex, tumor grade, smoking status, BCG treatment status, and mdNLR.

*: Winsorization was used on the top 2% values for fitting linearity assumption.

#: initial recurrence or the death whose cause was unknown

2.4.3. Associations of circulating immune profiles with bladder cancer outcomes

We next investigated the association between immune cell-type proportions and RFS in multivariable models. CD4T (HR = 0.98, 95% CI = 0.97 – 1.00) and CD8T cell proportion (HR = 0.97, 95% CI = 0.95 – 1.00) were significantly associated with the decreased hazard of RFS. Monocyte cell proportion (HR = 1.04, 95% CI = 1.00 – 1.07) was significantly associated with the increased hazard of RFS. Both B cell and NK cell proportion hazard estimates were < 1 but not statistically significant (**Table S2.1**). We also examined the 10-year overall survival (OS) in univariate models and the multivariable models. Observed associations with RFS were consistent for OS for: age, male, high

tumor grade, ever smoking, and mdNLR (**Table 2.3** and **Figure S2.2**). In addition, CD4T, CD8T, B cell, and NK cell proportion were significantly associated with the decreased hazard of death; neutrophil cell proportion was significantly associated with the increased hazard of death (**Table S2.2**).

Table 2.3: Cox proportional hazards 10-year overall survival models

			Survival months	Multivariable model	
	n (%) deceased	*n* (%) alive	Mean[#]	HR (95% CI)	P value
Age	180 (29.9)	423 (70.1)	104.3	1.07 (1.04-1.08)	8.7E-9
Sex					
Male	156 (34.1)	301 (65.9)	102.2	Referent group	
Female	24 (16.4)	122 (83.6)	111.1	0.62 (0.40-0.96)	0.032
Tumor grade					
Low	118 (26.1)	334 (73.9)	106.6	Referent group	
High	62 (41.1)	89 (58.9)	97.7	1.54 (1.12-2.12)	8.0E-3
Smoking status					
Non-smoker	19 (18.4)	84 (81.6)	110.8	Referent group	
Ever-smoker	161 (32.2)	339 (67.8)	103.0	1.66 (1.03-2.67)	0.039
BCG treatment					
No	152 (29.6)	362 (70.4)	104.8	Referent group	
Yes	28 (31.5)	61 (68.5)	101.9	1.04 (0.68-1.60)	0.842
mdNLR*	180 (29.9)	423 (70.1)	104.3	1.46 (1.32-1.60)	6.0E-15

HR: hazard ratio, CI: confidence interval, mdNLR: methylation-derived neutrophil to lymphocyte ratio

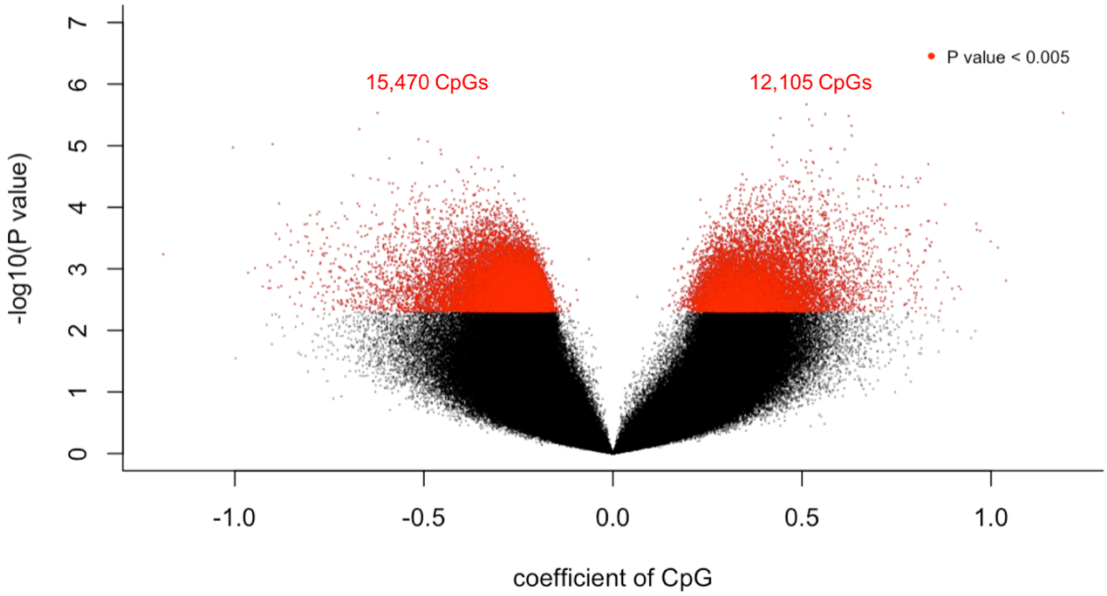
*: Winsorization was used on the top 2% values for fitting linearity assumption.

All covariates modeled met proportionality assumptions #: the median survival month for each variable is 120.

2.4.4. Epigenome-wide association study for the association of bladder cancer outcomes

Next, we assessed the relation of NMIBC patient outcomes with DNA methylation. First, Cox proportional hazards models were fit for each CpG controlling for age, stratified sex, tumor grade, smoking status, and stratified BCG receiving status. Without adjustment for immune cell proportions, we identified 27,575 CpGs whose methylation was associated with a significant ($P < 0.005$) difference in the hazard of RFS (**Figure 2.3A**, **Table S2.3**). We then conducted similar analyses controlling for age, sex, tumor grade, smoking status, BCG status, and winsorized immune cell proportions. As expected and demonstrating the importance of adjusting for cell-type proportions, the fully adjusted models were attenuated and identified 2,528 CpGs whose methylation was associated with a significant difference ($P < 0.005$) in the hazard of RFS. The 10 CpGs most strongly (with the smallest P-value) associated with the hazard of RFS corresponded to 10 genes: *TMCO4* (cg04738197), *LENG9* (cg12057190), *CDC42EP5* (cg12057190), *LNP1* (cg02540094), *TOMM70A* (cg02540094), *RUNX2* (cg08012149), *TBXAS1* (cg01584377), *SSH1* (cg16237760), *SFXN2* (cg08609163), and *COG3* (cg06172950). 1,572 CpGs were associated with the increased hazard of RFS, and 956 CpGs were associated with the decreased hazard of RFS (**Figure 2.3B**). The complete list of RFS-associated CpGs is provided in **Table S2.4** and **Table S2.5**.

A



B

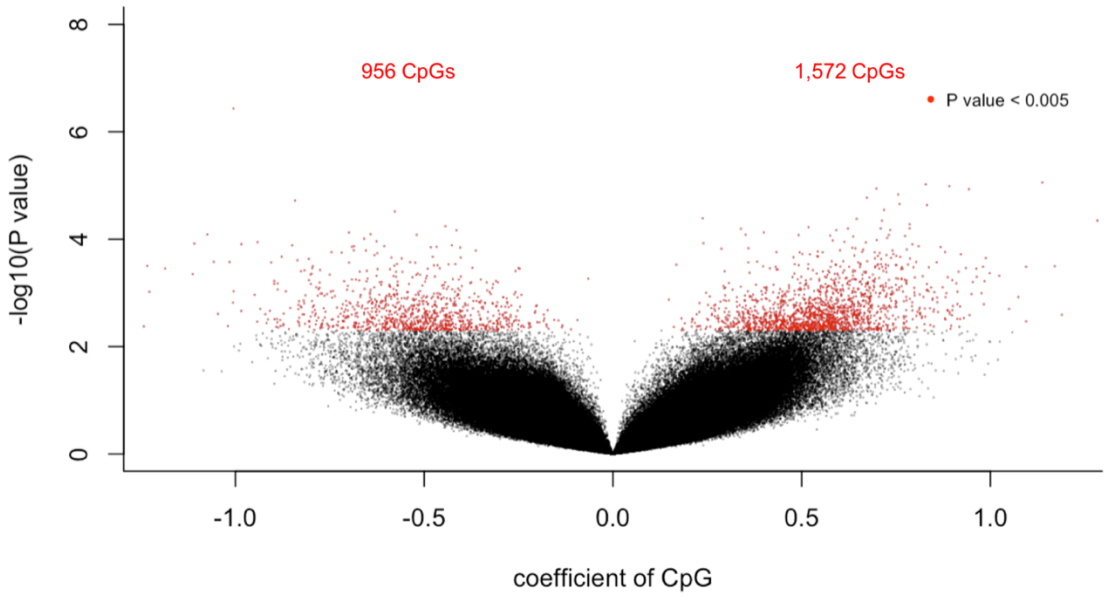


Figure 2.3: Volcano plots of recurrence-free survival (RFS) associated CpGs from the epigenome-wide association study (EWAS) analyses. The Cox multivariable model that was fitted in EWAS was shown in each plot. CpGs are colored in red **(A)** 12,105 CpGs were associated with the increased hazard of NMIBC 10-year RFS, and 15,470 CpGs were associated with the decreased hazard of NMIBC 10-year RFS. **(B)** 1,572 CpGs were associated with the increased hazard of NMIBC 10-year RFS, and 956 CpGs were associated with the decreased hazard of NMIBC 10-year RFS

2.4.5. Location of the bladder cancer outcomes-associated CpG loci

To gain a better understanding of the regions where the hazard-associated CpGs were located, we tested for enrichment of CpG island region context among CpG loci associated with a significant change in the hazard of RFS. We found that 1,572 CpGs associated with the increased hazard of RFS were significantly enriched in the open sea (OR = 1.14, 95% CI = 1.03 – 1.27) and were significantly depleted in CpG island S Shore regions (OR = 0.82, 95% CI = 0.67 – 0.99). The 956 CpGs associated with a decreased hazard of RFS were significantly enriched in CpG island N Shore regions (OR = 1.30, 95% CI = 1.06 – 1.57) and were significantly depleted in CpG island (OR = 0.60, 95% CI = 0.46 – 0.78) (**Figure 2.4A**). We also tested for enrichment of other gene regulatory regions among CpG loci associated with a significant change in the hazard of RFS.

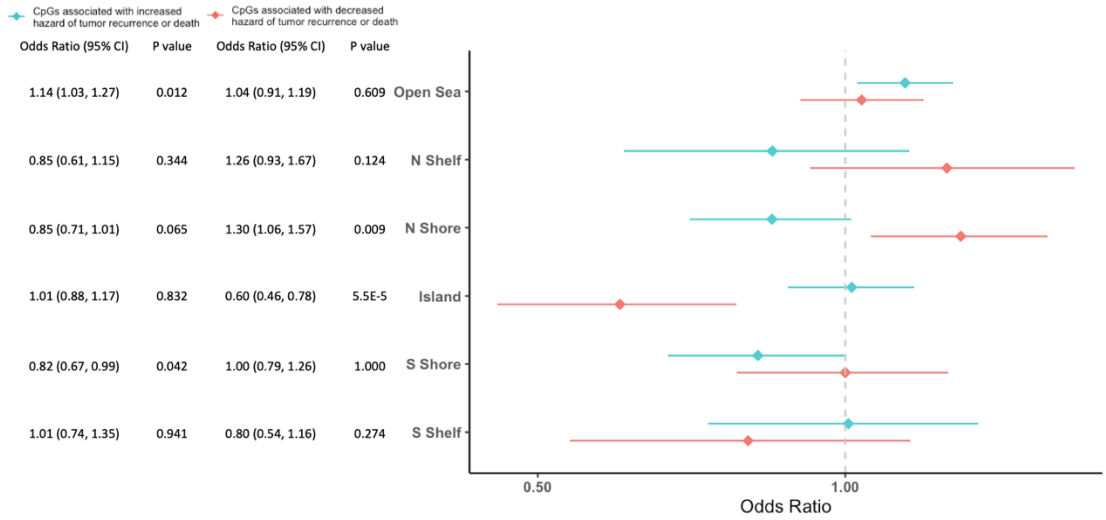
Location of CpGs in regulatory regions also was investigated and the CpGs associated with the increased hazard of RFS were significantly enriched in enhancer regions (OR = 1.78, 95% CI = 1.43 – 2.18), DNase hypersensitive sites (DHS) (OR = 1.25, 95% CI = 1.13 – 1.40), 5'UTR regions (OR = 1.37, 95% CI = 1.17 – 1.60) and gene body (OR = 1.15, 95% CI = 1.04 – 1.27); however, these CpGs were significantly depleted in regions 200–1,500 bps upstream of the transcription start site (TSS1500) (OR = 0.85, 95% CI = 0.73 – 0.99). In addition, while the CpGs associated with a decreased hazard were strongly enriched in enhancer regions (OR = 4.18, 95% CI = 3.44 – 5.05) and DHS (OR = 2.46, 95%

CI = 2.13 – 2.86), they were depleted for TSS200 (OR = 0.70, 95% CI = 0.48 – 0.99), gene body (OR = 0.85, 95% CI = 0.74 – 0.97) and transcription factor binding sites (TFBS) (OR = 0.68, 95% CI = 0.55 – 0.83) (**Figure 2.4B**). Using the CpGs associated with the hazard of RFS from cell-type unadjusted models in tests for enrichment gave results that were largely consistent with those above based on CpGs from fully adjusted models (**Figure S2.3**).

2.4.6. Gene Set Enrichment Analysis for the bladder cancer outcomes-associated CpG sites

To further understand the biological function of the hazard-associated CpGs, Gene Set Enrichment Analysis (GSEA) Molecular Signature Database (MSigDB) was used to explore the potential gene sets which might associate with the tumor recurrence or death of NMIBC patients. The input was 2,528 RFS-associated CpGs from the Cox model EWAS adjusting for immune cell composition. In the top 10 hazard-associated gene sets in gene ontology (GO) terms, some gene sets were related to neurological system processing (**Figure S2.4A**), and most related genes were associated with the increased hazard of RFS in NMIBC (**Figure S2.4B**). Since mdNLR was associated with the hazard of recurrence or death of NMIBC, we were interested in immune-related gene sets. We checked the immunologic signature gene set, and the top 10 gene sets and their genes were related to immune cell regulation. However, only one gene set (BCELL VS MDC UP: up-regulated genes in B cells compared with myeloid dendritic cells after vaccination for influenza) consisting of 41 genes was significantly associated

A



B

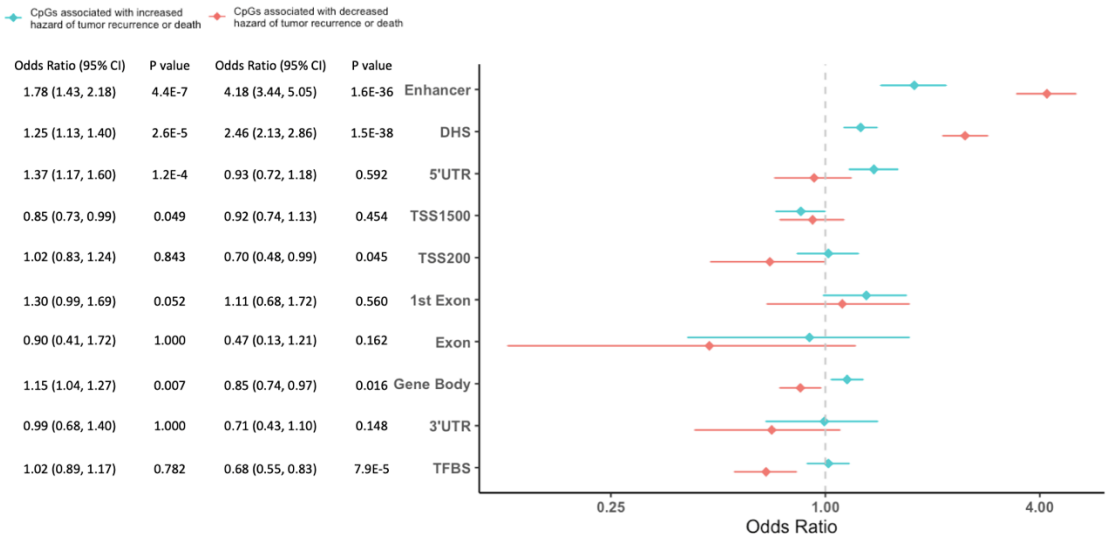


Figure 2.4: Genomic context enrichment analysis of CpG sites whose methylation state is significantly associated with recurrence-free survival. Enrichment analysis of **(A)** relation to CpG island and **(B)** genomic context of NMIBC recurrence-free survival associated CpGs. The 2,528 CpGs from EWAS ($P < 0.005$) were tested for enrichment versus all modeled CpGs. The bar represents the 95% confidence interval. Mantel–Haenszel was used to test RFS-associated CpGs enrichment of CpG island-related gnome context. An odds ratio larger than 1 means enrichment, and an odds ratio smaller than 1 indicates depletion

with RFS (FDR < 0.05) (**Figure S2.4C-D**). Results from GO term analyses using

27,575 CpGs from the cell-type unadjusted models identified 5 pathways related to immunologic regulation among the top 10 pathways. Further, in the immunologic signature gene set specifically, we observed the top 10 pathways were associated with monocytes and lymphocytes (**Figure S2.5**).

2.4.7. Locus overlap analysis for the bladder cancer outcomes-associated CpG sites

Locus overlap analysis (*LOLA*) was used to test the enrichment of CpGs in genomic regions. As our analysis was conducted on blood samples, results focus on the genomic regions within hematopoietic stem cells. When controlling for immune cell profiles, the 2,528 CpGs associated with the hazard of RFS in NMIBC patients were most significantly enriched in Histone H3 acetylated at lysine 9 and 14 (H3K9K14ac) (Q value = 1.9×10^{-20}) (**Figure S2.6A**). *LOLA* results for the 27,575 CpGs associated with the hazard of RFS from cell-type unadjusted models were most significantly enriched in cistrome of the promyelocytic leukemia protein (PML) (Q-value < 0.05) (**Figure S2.6B**).

2.4.8. Methylation levels of CpG sites of *BLCAP* found by differentially methylated regions analysis

We also ran a differentially methylated regions analysis using *DMRcate*. In this study, we found 11 CpGs in a specific genome region overlapping with the gene *BLCAP*. NMIBC patients without tumor recurrence or death within 10 years had a higher mean of methylation levels for this region compared with NMIBC

patients with tumor recurrence or death within 10 years, and one CpG in this region was significant (FDR = 2.63×10^{-14}) (**Table S2.6, Figure S2.7**).

2.5 Discussion

In this study, we tested whether immune profiles and epigenetic features are associated with NMIBC recurrence. Although previous work observed the association between NLR and overall survival in NMIBC patients, it was in a smaller study sample and used DNA methylation data from a dated (second generation) array platform with ~ 27,000 CpGs. In this study, our sample size was nearly three times larger and DNA methylation data were collected using the current genome-scale platform (fourth generation), measuring ~ 860,000 CpGs for which an optimized cell-type deconvolution library exists to determine highly accurate immune cell-type proportions. We extended tests of association with the patient outcomes beyond the methylation-derived neutrophil-to-lymphocyte ratio (mdNLR) to include leukocyte-specific cell-type proportions. Our findings suggest that elevated mdNLR increased the hazard of RFS in NMIBC patients. These findings are consistent with previous studies demonstrating that NLR was significantly higher in high-risk NMIBC patients [151], and increased NLR was positively associated with poor prognosis [152], [153]. While past studies have shown a significant association between NLR and outcomes in NMIBC patients using the conventional method of CBC tests [75], [154], [155], this study uses NLR derived from blood methylation profiles. Compared with flow cytometry, estimation using the differentially methylated

regions (DMR) library has several advantages: it does not require long sample processing time, large volume of blood, or intact cells as DNA and 5-methylcytosine are both stable [99]. The advantage of our approach is the ability to use archived samples that many investigators may already have. Moreover, utilizing blood samples to monitor patient outcomes is less invasive compared with a cystoscopy, the routine screening method. For NMIBC patients with a high risk of tumor recurrence, BCG is the standard intravesical immunotherapy to induce immune system eliminating bladder cancer cells that might be left after surgery [123], [156], [157], and therefore, blood immune profile is a potential prognostic factor. Immune profiling with DNA methylation data is a promising avenue for assessing NMIBC prognosis.

As NLR is composed of lymphocyte and neutrophil proportions, we also considered the association between each methylation-derived immune cell type proportion and patient outcomes. Interestingly, increased CD4T and CD8T cell proportions were associated with decreased NMIBC recurrence-free survival and overall survival. Prior work has shown that NMIBC patients with a high CD4T cell count in BCG pretreatment microenvironment have a significantly prolonged recurrence-free survival compared to patients with a low CD4T cell count [157]. In addition to the immune cell types explored in this study, other cell types have been shown to affect NMIBC development, such as GATA3 + T cells, regulatory T cells, and tumor-associated macrophages [157]. Other work showed an increased CD8+ ILT2+ T cell proportion was associated with a significantly increased hazard of NMIBC recurrence [158]. Further, peripheral

(neutrophil × platelet) / (lymphocyte) was inversely correlated with high-risk NMIBC recurrence-free survival [159]. These results demonstrate the potential of immune cell profiles in evaluating the prognosis of NMIBC patients, and future work mapping DNA methylation profiles of additional immune cell types and states could add detail to investigations of immune profiles and bladder cancer outcomes. In addition, peripheral immune cell distribution is affected by potentially residual confounding factors such as, infection [160], inflammation [161], lifestyles, treatments [162], [163], obesity [164], chronic alcohol consumption [165], and type 2 diabetes [166]. Since our study includes hundreds of subjects and has a decade or more of follow-up time, in the future, we will re-explore the effect of these residual confounding factors on the association of circulating immune cell distribution and NMIBC outcomes. In the future, we will perform higher resolution methylation cell mixture deconvolution to resolve additional immune cell types and employ blood count on prospectively collected samples. Together, these data will allow us to understand if specific subsets of neutrophils or lymphocytes contribute to high NLR in patients with poor outcomes.

Bacillus Calmette–Guérin (BCG) is the most commonly used immunotherapy for high-risk NMIBC. After transurethral resection of bladder tumors, NMIBC patients may receive BCG intravesical therapy to induce an immune response in the bladder to attack cancer cells. Previous studies have shown that tumor immune environments may interact with BCG and interfere with the efficacy of this therapy. For instance, IL-12 secreted by BCG-induced

monocytes was increased in NMIBC patients without tumor recurrence, a phenomenon that may involve the innate immune memory of circulating monocytes [167]. Nevertheless, patients receiving BCG in our study did not have peripheral immune cell profiles that differed from patients who were BCG naïve, and BCG was not significantly associated with NMIBC outcomes. As the number of patients with BCG treatment was relatively low (N = 89; 14.8%) and limited power to observe possible therapy induced changes in peripheral immune profiles, future work is needed to address the potential associations of BCG with peripheral immune profiles and patient outcomes. Most BCG-treated patients were already high-risk. In addition, blood samples were collected only after surgery or initial BCG treatment. To assess the changes in immune profiles over time, having multiple blood draws is necessary for future work.

This EWAS identified several CpGs associated with NMIBC recurrence-free survival when controlling for stratified sex, age, tumor grade, smoking status, stratified BCG receiving status, and immune cell profiles. With a P-value < 0.005, 2,528 CpGs were found to be associated with the hazard of RFS.

Among the top CpGs whose methylation was associated with NMIBC recurrence or death in the fully adjusted models, some genes have been previously associated with bladder cancer. Slingshot homolog-1 (*SSH1*) had a positive association with tumor grade, tumor invasion, and tumor recurrence in bladder cancer patients [168]. Runt-related transcription factor 2 (*RUNX2*) is a key factor of osteoblast differentiation and has been reported to be associated with epithelial-mesenchymal transition in bladder tumors. Furthermore, *RUNX2*

could predict early recurrence in bladder cancer patients with high accuracy [169], [170]. Similar to past studies, *SSH1* and *RUNX2* were associated with the increased risk of tumor recurrence in our study. In addition, NMIBC patients with tumor recurrence within ten years had significantly higher methylation levels in the gene body of these two genes compared with patients without tumor recurrence.

When not adjusting for immune cell profiles, 27,575 CpGs were associated with the decreased hazard of RFS. The top 10 CpGs with the most significant P-value corresponded to 6 genes: *BCL11A* (cg24361098), *TMCO4* (cg04738197), *MICALCL* (cg01518090), *GRAP2* (cg21012238), *TRAM2* (cg15085626), and *KIRREL* (cg10570484). While these genes have not been reported to be associated with bladder cancer outcomes, they have been reported to be involved in mechanisms promoting cancer development in tumors [171]–[174] or blood [175], [176]. Although we measured blood methylation in this study, we plan to measure tumor methylation and will explore the association of CpGs in these genes with NMIBC outcomes. What was more interesting is that the model, with or without controlling for immune cell profiles, led to different results. While immune cell profiles are usually not adjusted in the Cox model, we controlled for immune cell profiles since the immune system plays a key role in tumor development. This study presents a new perspective to demonstrate the difference between models with or without adjusting for immune profiles, indicating the need for further investigation on the involvement of immune profiles in outcome analyses.

Through DMR analysis, we found NMIBC patients with tumor recurrence or death within 10 years had a lower methylation level in the *BLCAP* region compared with patients without tumor recurrence or death. The bladder cancer-associated protein (*BLCAP*) gene encodes a protein that stimulates apoptosis. It has been reported that loss of protein expression is associated with bladder tumor progression, and the application of staining patterns for this protein could be a potential biomarker in bladder cancer [177]. In functional analysis, strong nuclear expression of *BLCAP* was associated with expression of p-*STAT3* and overall poor disease outcome. Additionally, *BLCAP* was discovered to interact with *STAT3* physically and may involve the *STAT3*-mediated progression of precancerous lesions to invasive bladder tumors [178]. These results were consistent with our finding that NMIBC patients with poor outcomes had lower methylation levels in the *BLCAP* region. Since the model we used for DMRs analysis was adjusted for immune cell profiles, we are curious whether immune cells may play roles in the interaction between *BLCAP* and *STAT3* and will investigate this in the future. The CpG site with significantly lower methylation (cg10642330) is located in the 5' UTR of *BLCAP* and the gene body of *NNAT*. The methylation beta value of this CpG site in *NNAT* had been reported significantly higher in prostate cancer tissue relative to adjacent normal tissues [179]. Though no CpG site in the *BLCAP* region was significantly associated with the hazard of tumor recurrence or death in the adjusted EWAS, five *BLCAP* CpGs (cg26083330, cg23757721, cg13790727, cg03061677, and cg04489586) were associated in the model not controlling for immune cell proportions (not

including the DMR analysis site cg10642330). In the future, we will investigate the relation of *BLCAP* tumor methylation with survival outcomes.

Our results reveal several features of peripheral blood immune profiles that are associated with outcomes in NMIBC patients. We showed that higher CD4 or CD8 proportions were associated with decreased hazard of recurrence or death and further established that high NLR is associated with an increased hazard of RFS. The EWAS portion of our study also points to epigenetic reprogramming within the immune compartment being involved in tumor recurrence of NMIBC patients. In addition, we identify preliminary evidence of discrete and regional CpG methylation associations with bladder cancer outcomes. Future study in a prospective setting will assess the clinical utility of incorporating methylation in predicting the hazard of recurrence and shaping recommendations for disease surveillance. In addition to immune cells, future work examining cell-type proportions in tumor microenvironments of NMIBC patients is needed to understand the relationship between peripheral immune profiles with tumor-infiltrating immune profiles and patient outcomes. This work contributes to our understanding of associations between methylation-derived immune profiles and NMIBC patient outcomes and could further contribute to developments in epigenetic biomarkers of cancer.

2.6 Supplement: Chapter 2

Tables S2.3, S2.4, S2.5, and S2.6 are available in the published version of this manuscript online (from *Clinical Epigenetics*), and omitted from the text here

due to size restrictions.

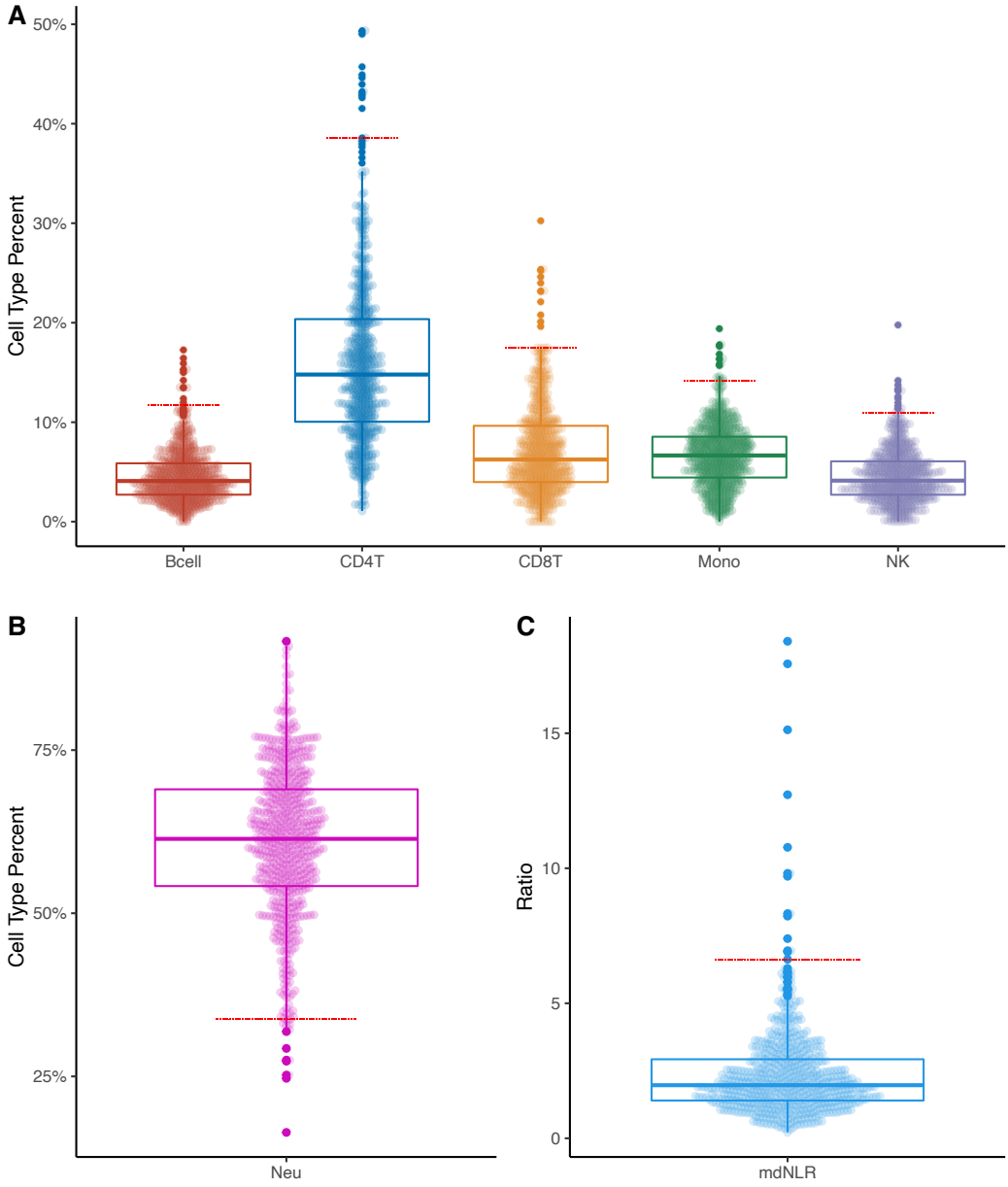


Figure S2.1: Summary of each immune cell profile from DNA methylation deconvolution (603 NMIBC patients). (A) Leukocyte proportions **(B)** Neutrophil percent **(C)** Methylation-derived neutrophil-lymphocyte ratio. For Leukocyte proportions (panel A) and NLR (panel B), winsorization was used on the top 2% value, and the red dashed line indicates the 98th percentile of each immune cell profiles. For neutrophil percent, winsorization was used on the bottom 2% value, and the red dashed line indicates the 2th percentile of neutrophil percent.

Table S2.1: Cox proportional hazards 10-year recurrence-free survival models for each immune cell type

	Multivariable model						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Age	1.01*	1.01*	1.01*	1.02**	1.02**	1.01*	1.01*
	(1.00, 1.03)	(1.00, 1.02)	(1.00, 1.03)	(1.00, 1.03)	(1.01, 1.03)	(1.00, 1.03)	(1.00, 1.03)
Grade 3 + 4 (ref.= Grade 1 + 2)	1.48**	1.47**	1.49***	1.47**	1.46**	1.48***	1.47**
	(1.17, 1.87)	(1.16, 1.85)	(1.18, 1.88)	(1.16, 1.85)	(1.16, 1.85)	(1.17, 1.87)	(1.16, 1.86)
Ever-smoker (ref.= Non-smoker)	1.65**	1.65**	1.67**	1.65**	1.64**	1.64**	1.65**
	(1.21, 2.25)	(1.21, 2.24)	(1.23, 2.27)	(1.21, 2.25)	(1.21, 2.24)	(1.20, 2.23)	(1.22, 2.25)
NLR	1.12**						
	(1.04, 1.20)						
CD4T		0.98*					
		(0.97, 1.00)					
CD8T			0.97*				
			(0.95, 1.00)				
B cell				0.99			
				(0.95, 1.03)			
NK					0.98		
					(0.94, 1.03)		
Monocyte						1.04*	
						(1.00, 1.07)	
Neutrophil							1.01
							(1.00, 1.02)
c-index	0.59	0.58	0.59	0.58	0.58	0.58	0.58
p.value (proportional assumption)	0.16	0.26	0.37	0.37	0.4	0.37	0.29
Observations	603	603	603	603	603	603	603
R^2	0.07	0.06	0.06	0.05	0.05	0.06	0.06

*p<0.05; **p<0.01; ***p<0.001

Stratification on Sex and BCG Trt. Winsorization was used for 2% value of each immune cell profile

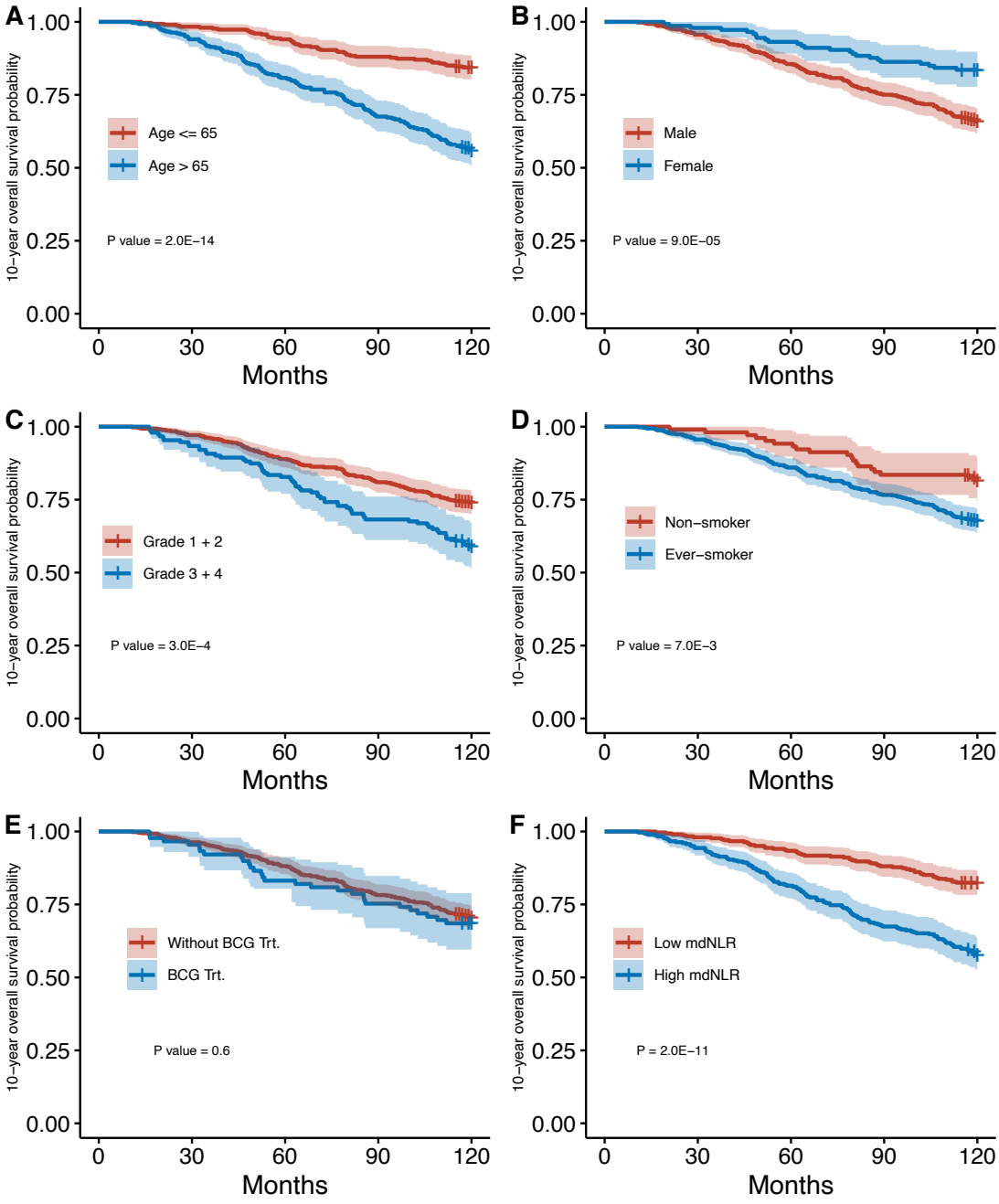


Figure S2.2: Kaplan-Meier analysis of 10-year overall survival (OS). 10-year OS curves stratified by **(A)** age, **(B)** sex, **(C)** tumor grade, **(D)** smoking status, **(E)** BCG treatment status or **(F)** mdNLR level. P-values for Log-rank tests are shown.

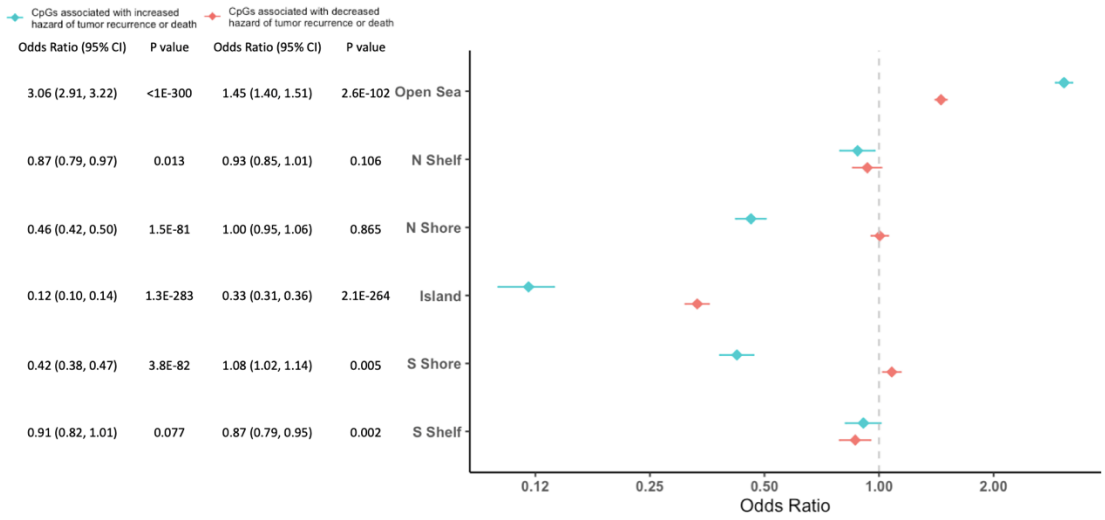
Table S2.2: Cox proportional hazards 10-year overall survival models for each immune cell type

	Multivariable model						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Age	1.06***	1.06***	1.06***	1.06***	1.08***	1.07***	1.06***
	(1.04, 1.08)	(1.04, 1.08)	(1.04, 1.09)	(1.04, 1.08)	(1.05, 1.10)	(1.05, 1.09)	(1.04, 1.08)
Female (ref.= Male)	0.62*	0.54**	0.55**	0.52**	0.48***	0.49**	0.57*
	(0.40, 0.96)	(0.35, 0.83)	(0.36, 0.85)	(0.34, 0.81)	(0.31, 0.74)	(0.32, 0.76)	(0.37, 0.88)
Grade 3 + 4 (ref.= Grade 1 + 2)	1.54**	1.47*	1.61**	1.56**	1.51*	1.51*	1.50*
	(1.12, 2.12)	(1.07, 2.03)	(1.16, 2.22)	(1.13, 2.14)	(1.09, 2.08)	(1.09, 2.08)	(1.09, 2.07)
Ever-smoker (ref.= Non-smoker)	1.65*	1.59	1.68*	1.61*	1.65*	1.64*	1.63*
	(1.03, 2.67)	(0.99, 2.57)	(1.04, 2.71)	(1.00, 2.60)	(1.02, 2.65)	(1.02, 2.65)	(1.01, 2.62)
BCG Trt. (ref.= Without BCG Trt.)	1.04	0.91	0.86	0.92	0.94	0.88	0.97
	(0.68, 1.60)	(0.60, 1.39)	(0.56, 1.31)	(0.61, 1.41)	(0.62, 1.44)	(0.58, 1.34)	(0.64, 1.49)
NLR	1.45***						
	(1.32, 1.60)						
CD4T		0.95***					
		(0.93, 0.97)					
CD8T			0.90***				
			(0.86, 0.94)				
B cell				0.88***			
				(0.82, 0.94)			
NK					0.89***		
					(0.84, 0.96)		
Monocyte						1.02	
						(0.97, 1.07)	
Neutrophil							1.05***
							(1.03, 1.06)
c-index	0.74	0.71	0.72	0.71	0.7	0.69	0.73
p.value (proportional assumption)	0.5	0.77	0.68	0.6	0.61	0.88	0.5
Observations	603	603	603	603	603	603	603
R^2	0.21	0.16	0.17	0.15	0.15	0.13	0.19

*p<0.05; **p<0.01; ***p<0.001

Winsorization was used for 2% value of each immune cell profile

A



B

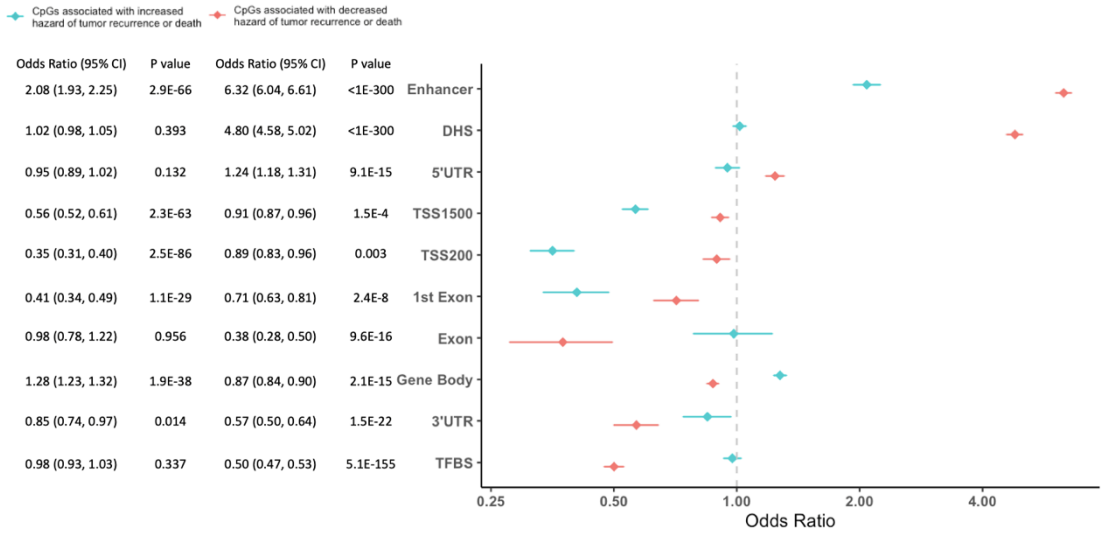
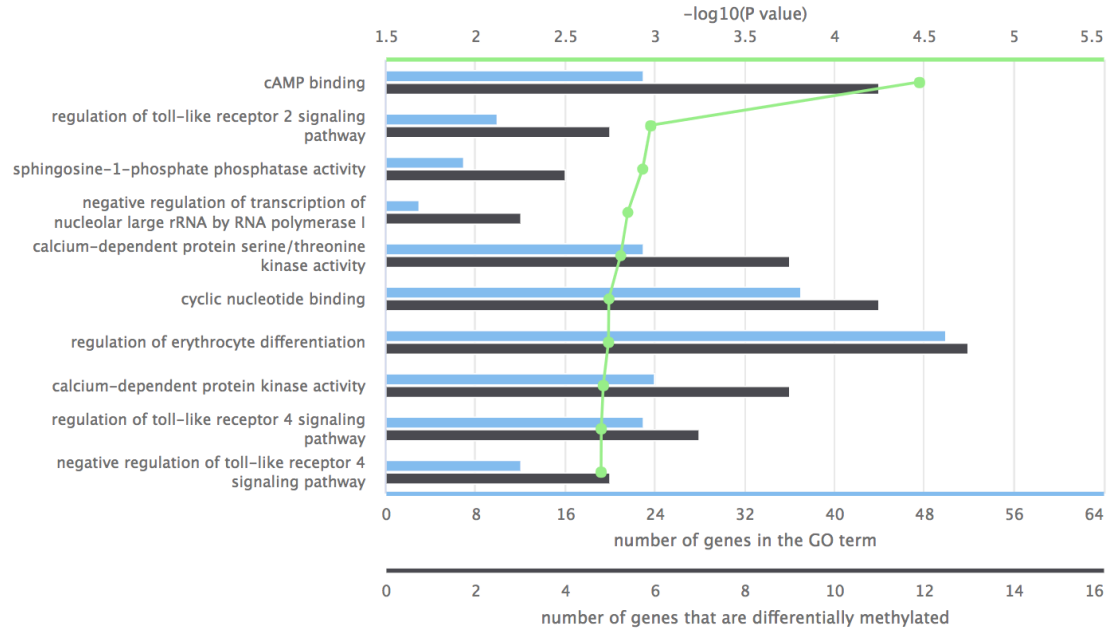
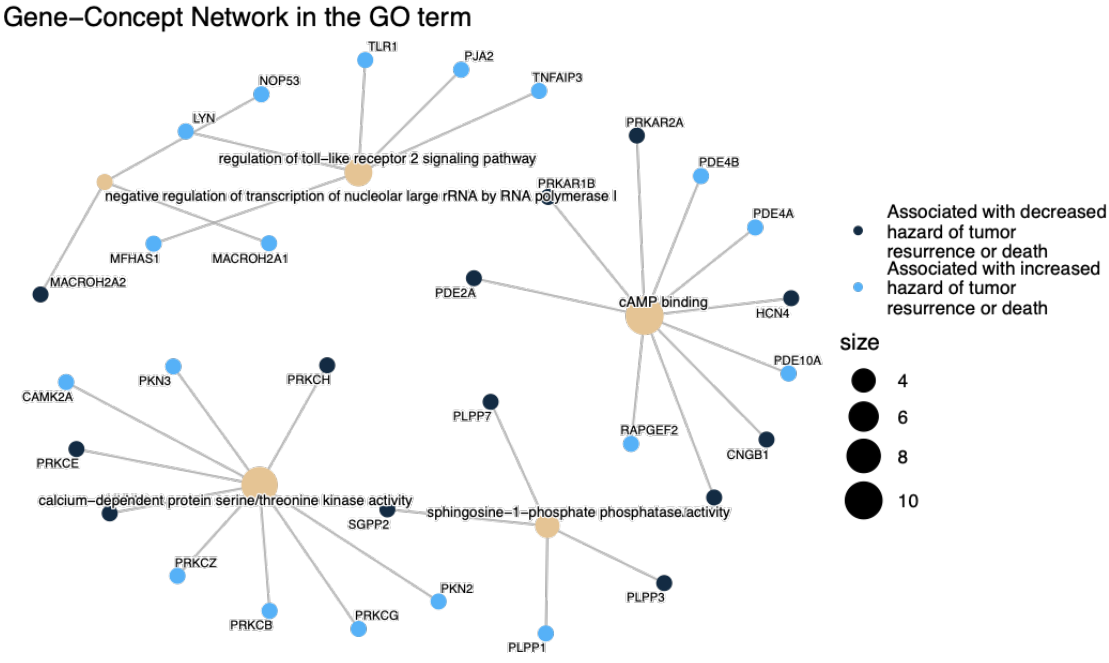


Figure S2.3: Enrichment analysis of **(A)** relation to CpG island and **(B)** genomic context of NMIBC recurrence-free survival associated CpGs. The 27,575 CpGs from EWAS (P-value < 0.005) without adjusting immune cell profiles were tested for enrichment versus all modeled CpGs. The bar represents the 95% confidence interval. Mantel-Haenszel test was used to test RFS associated CpGs enrichment of CpG island-related gnome context. Odds ratio larger than 1 means enrichment, and odds ratio smaller than 1 indicates depletion.

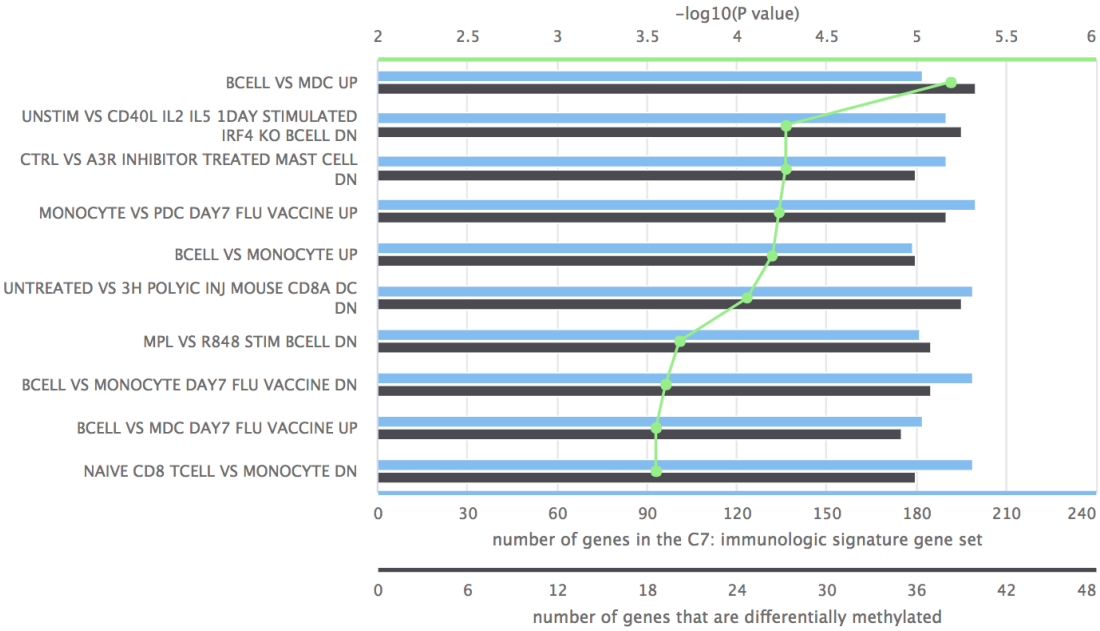
A


-log10(P value)
1.5
2
2.5
3
3.5
4
4.5
5
5.5
cAMP binding
regulation of toll-like receptor 2 signaling pathway
sphingosine-1-phosphate phosphatase activity
negative regulation of transcription of nucleolar large rRNA by RNA polymerase I
calcium-dependent protein serine/threonine kinase activity
cyclic nucleotide binding
regulation of erythrocyte differentiation
calcium-dependent protein kinase activity
regulation of toll-like receptor 4 signaling pathway
negative regulation of toll-like receptor 4 signaling pathway
0
8
16
24
32
40
48
56
64
number of genes in the GO term
0
2
4
6
8
10
12
14
16
number of genes that are differentially methylated

B


Gene–Concept Network in the GO term
TLR1
PJA2
TNFAIP3
NOP53
LYN
PRKAR2A
PDE4B
regulation of toll-like receptor 2 signaling pathway
PRKAR1B
negative regulation of transcription of nucleolar large rRNA by RNA polymerase I
PDE4A
MFHAS1
MACROH2A1
MACROH2A2
PDE2A
cAMP binding
HCN4
PDE10A
PKN3
PRKCH
CAMK2A
PLPP7
RAPGEF2
CNGB1
PRKCE
calcium-dependent protein serine/threonine kinase activity
SGPP2
sphingosine-1-phosphate phosphatase activity
PRKCZ
PKN2
PLPP3
PRKCB
PRKCG
PLPP1
Associated with decreased hazard of tumor resurrence or death
Associated with increased hazard of tumor resurrence or death
size
4
6
8
10

C



D

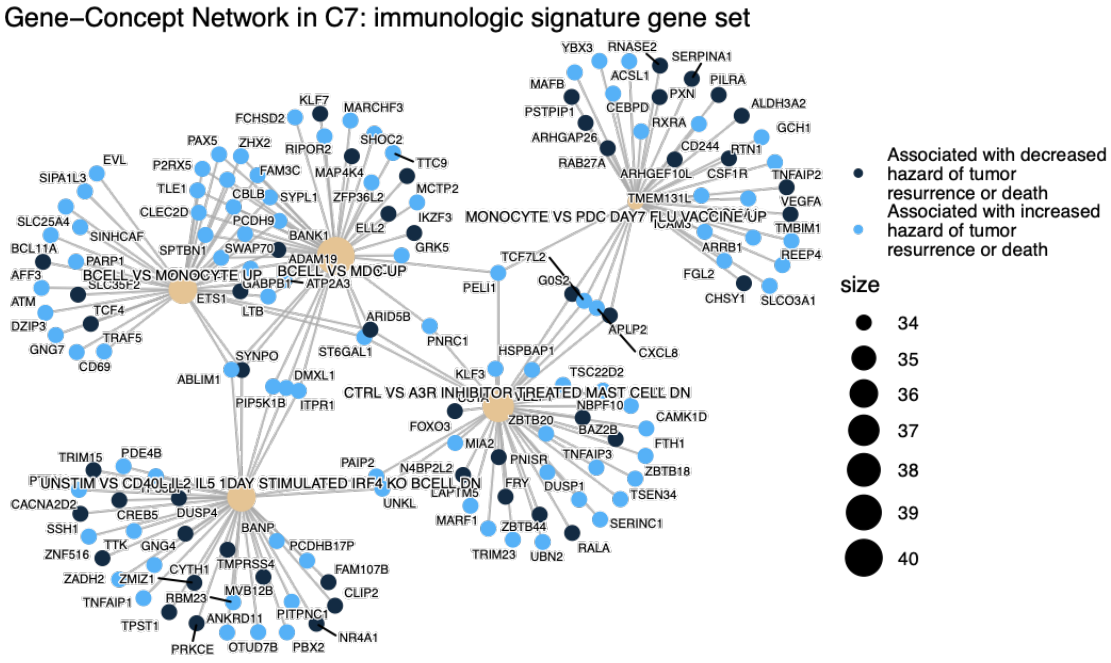
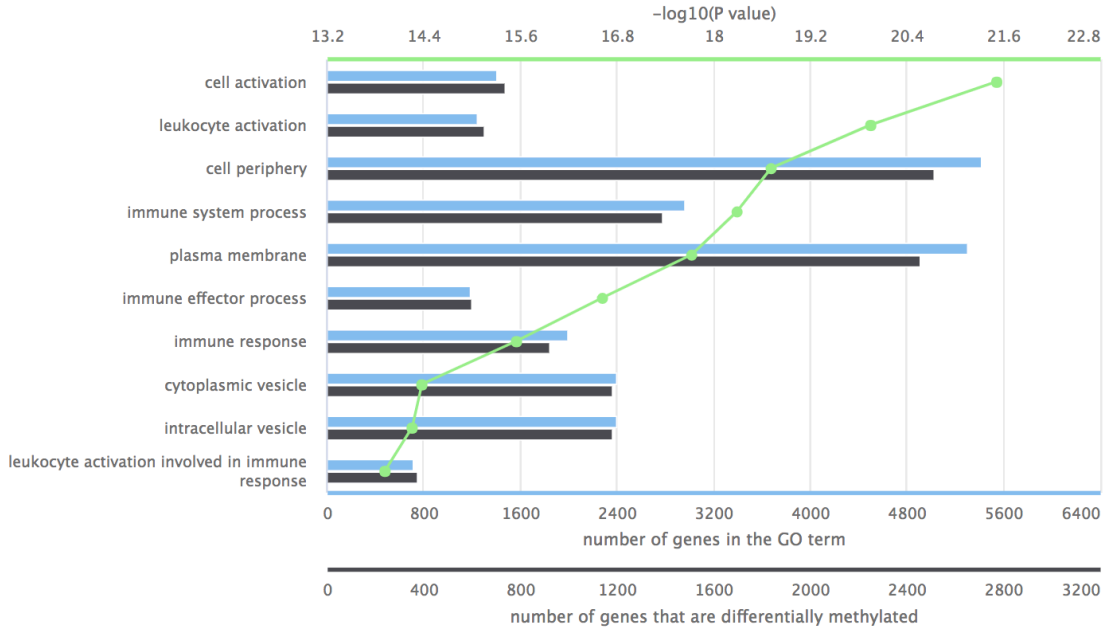


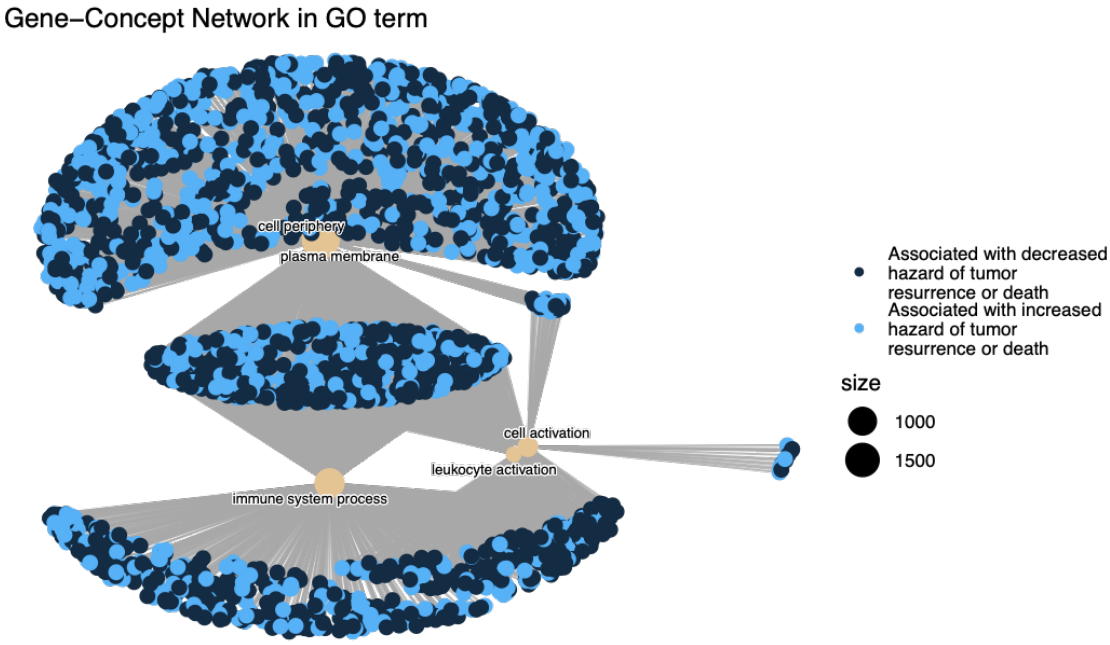
Figure S2.4: Gene set enrichment analysis (GSEA). The input was 2,528 RFS associated CpGs from the EWAS which Cox proportional multivariable model with adjusting immune cell profiles were fitted in. (A, C) The top green axis is corresponded to negative log 10 unadjusted p-value, and the further to the right, the more significant; the bottom blue axis is represented to the total number of genes in the gene set; the black axis is corresponded to the number of genes that the hazard-associated CpGs are located in the gene set. **(A-B)** The results of

GSEA for the gene ontology (GO) term. (A) Only the top 10 RFS associated pathways are shown. (B) Gene-concept network is displayed for the top 5 pathways. **(C-D)** The results of GSEA for the immunologic signature gene set. (C) the top 10 RFS associated pathways. (D) gene-concept network for the top 5 pathways.

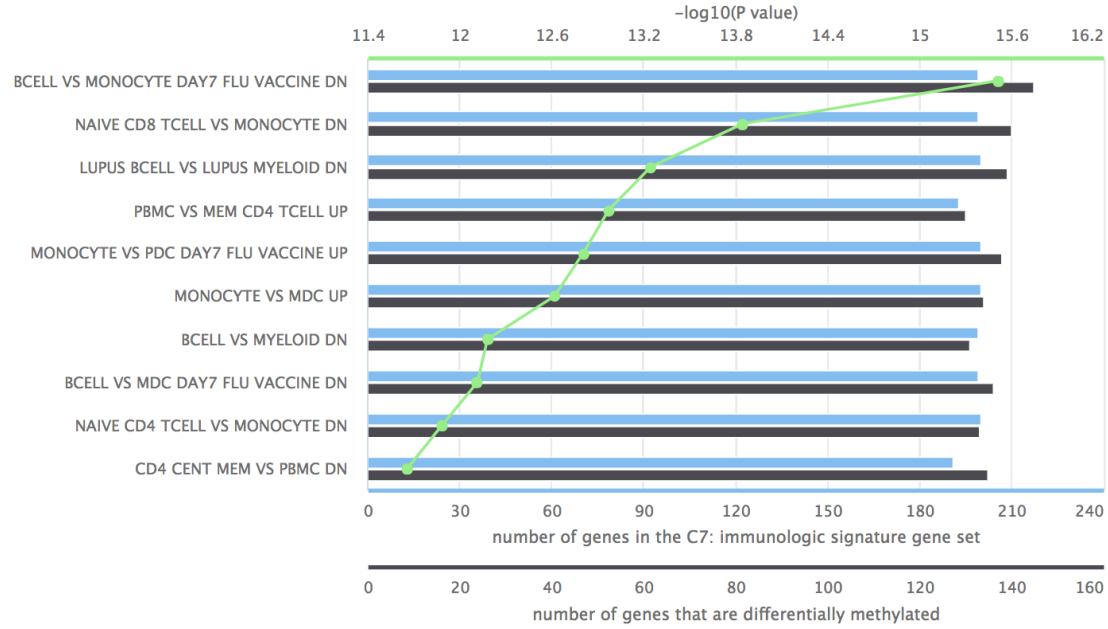
A


-log10(P value)
13.2
14.4
15.6
16.8
18
19.2
20.4
21.6
22.8
cell activation
leukocyte activation
cell periphery
immune system process
plasma membrane
immune effector process
immune response
cytoplasmic vesicle
intracellular vesicle
leukocyte activation involved in immune response
0
800
1600
2400
3200
4000
4800
5600
6400
number of genes in the GO term
0
400
800
1200
1600
2000
2400
2800
3200
number of genes that are differentially methylated

B


Gene–Concept Network in GO term
cell periphery
plasma membrane
cell activation
leukocyte activation
immune system process
Associated with decreased hazard of tumor resurrence or death
Associated with increased hazard of tumor resurrence or death
size
1000
1500

C



D

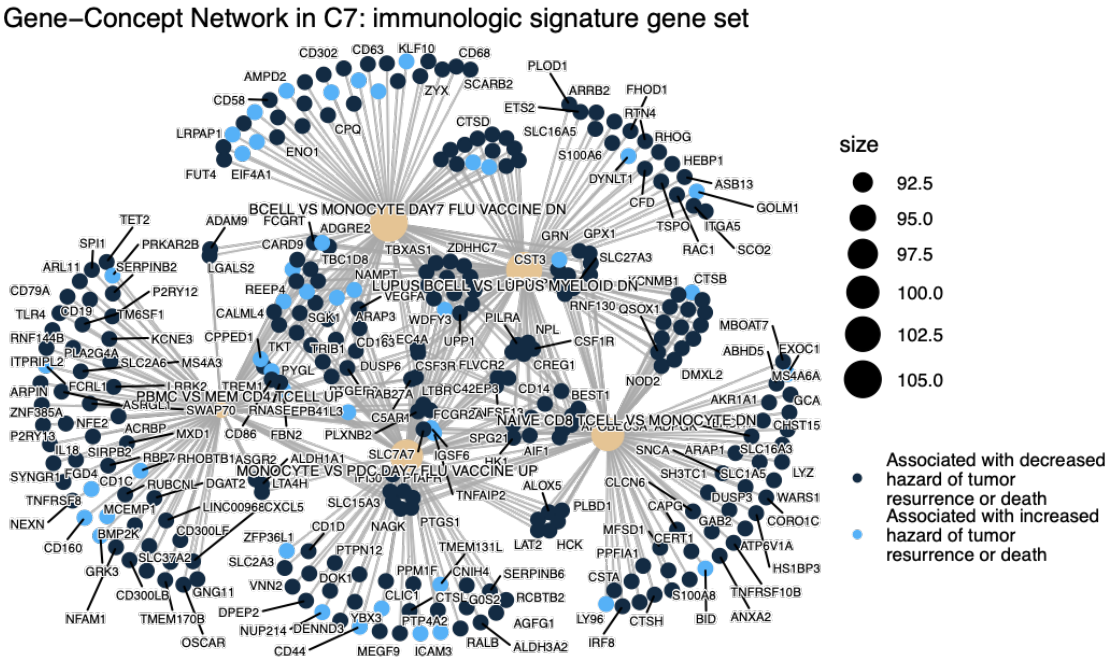


Figure S2.5: Gene set enrichment analysis (GSEA). The input was 27,575 RFS associated CpGs from the EWAS which Cox proportional multivariable model without adjusting immune cell profiles were fitted in. **(A-B)** the results of GSEA for the gene ontology (GO) term. (A) the top 10 RFS associated pathways. (B) gene-concept network for the top 5 pathways. **(C-D)** the results of GSEA for the immunologic signature gene set. (C) only the top 10 RFS associated pathways are shown. (D) gene-concept network for the top 5 pathways.

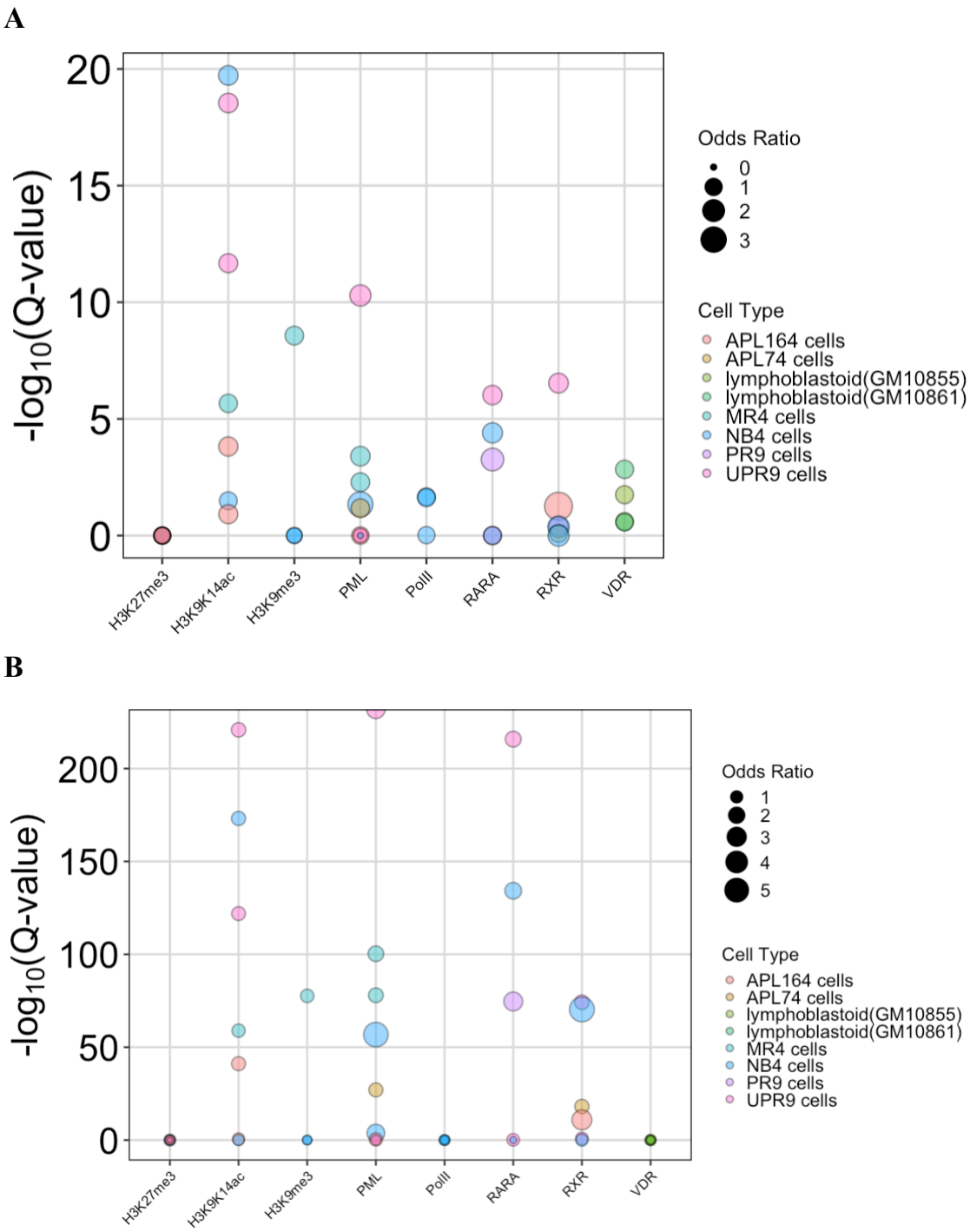
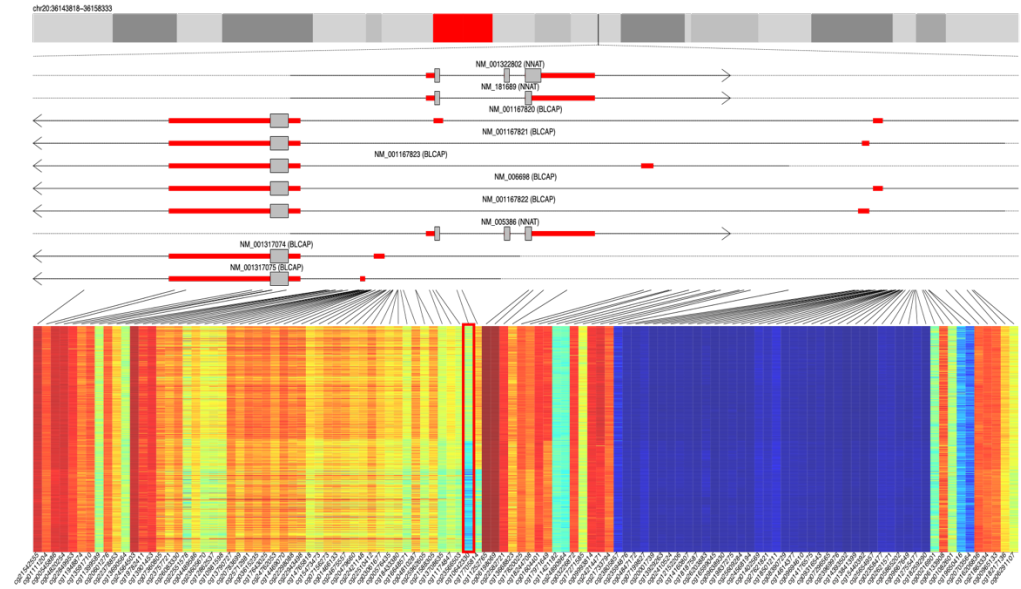
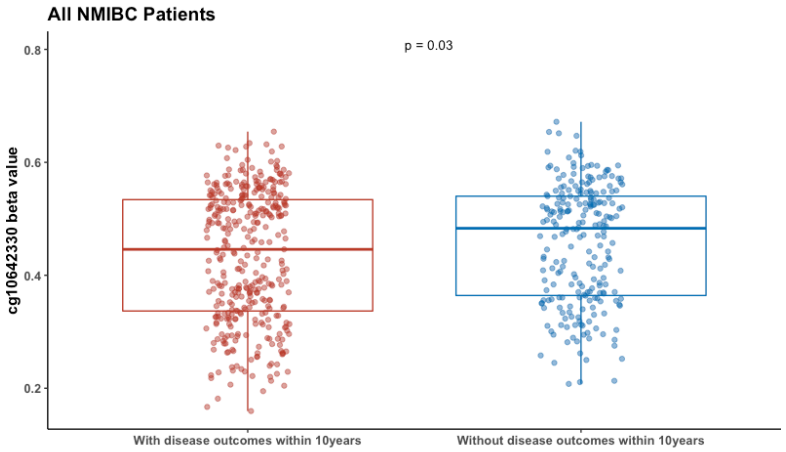


Figure S2.6: Enrichment of NMIBC recurrence-free survival associated (P-value < 0.005) CpGs from EWAS using locus overlap analysis (*LOLA*). Cell types were limited to "hematopoiet stem cell". **(A)** the input was 2,528 RFS associated CpGs from the EWAS which Cox proportional multivariable model with adjusting immune cell profiles were fitted in. **(B)** the input was 27,575 RFS associated CpGs from the EWAS which Cox proportional multivariable model without adjusting immune cell profiles were fitted in.

A


chr20:36143818–36158333
NM_001322802 (NNAT)
NM_181689 (NNAT)
NM_001167820 (BLCAP)
NM_001167821 (BLCAP)
NM_001167823 (BLCAP)
NM_006698 (BLCAP)
NM_001167822 (BLCAP)
NM_005386 (NNAT)
NM_001317074 (BLCAP)
NM_001317075 (BLCAP)

B


All NMIBC Patients
p = 0.03
cg10642330 beta value
0.8
0.6
0.4
0.2
With disease outcomes within 10years
Without disease outcomes within 10years

C

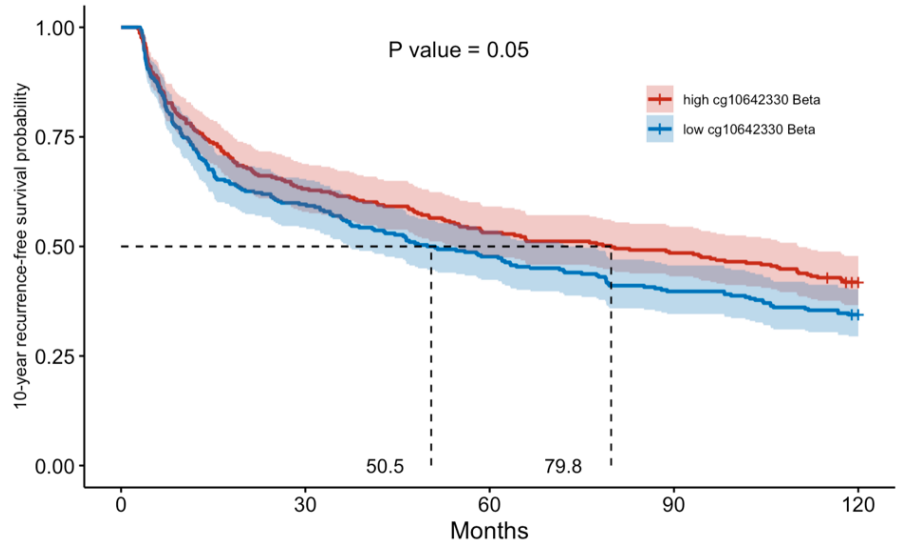

P value = 0.05
high cg10642330 Beta
low cg10642330 Beta
10-year recurrence-free survival probability
1.00
0.75
0.50
0.25
0.00
50.5
79.8
0
30
60
90
120
Months

Figure S2.7: The distribution of methylation levels in the *BLCAP* gene region of each NMIBC patients **(A)** In the heatmap, each row represents a subject; cg10642330 is labeled in the red box. **(B)** The distribution of cg10642330 beta values within NMIBC patients. Disease outcomes: death or tumor recurrence. P-values for Wilcox tests is shown. **(C)** NMIBC patients were grouped according to the median of cg10642330 methylation levels, and the KM plot for recurrence-free survival was shown.

Chapter 3

Genome-scale methylation analysis identifies immune profiles and age acceleration associations with bladder cancer outcomes

This work was published on August 1, 2023 as a *Research Article* in *Cancer Epidemiology, Biomarkers & Prevention*. PMID: 37527159.

The following authors have contributed to the content in this chapter:

Chen, JQ, Salas, LA, Wiencke, JK, Koestler, DC, Molinaro, AM, Andrew, AS, Seigne, JD, Karagas, MR, Kelsey, KT, Christensen, BC.

Conceptualization: JQC, BCC, LAS, JKW, DCK, AMM, MRK, KTK; Formal Analysis: JQC, BCC, LAS; Data Curation: ASA, JDS; Methodology: JQC, BCC, LAS, JKW, DCK, AMM, MRK, KTK; Writing–Original Draft: JQC; Writing–Review and Editing: JQC, BCC, LAS, JKW, DCK, AMM, MRK, KTK; Funding Acquisition: BCC

3.1 Abstract

Immune profiles have been associated with bladder cancer outcomes and may have clinical applications for prognosis. However, associations of detailed immune cell subtypes with patient outcomes remain underexplored and may

contribute crucial prognostic information for better managing bladder cancer recurrence and survival. Bladder cancer case peripheral blood DNA methylation was measured using the Illumina HumanMethylationEPIC array. Extended cell-type deconvolution quantified twelve immune cell-type proportions, including memory, naïve T and B cells, and granulocyte subtypes. DNA methylation clocks determined biological age. Cox proportional hazard models tested associations of immune cell profiles and age acceleration with bladder cancer outcomes. The *partDSA* algorithm discriminated 10-year overall survival groups from clinical variables and immune cell profiles, and a semi-supervised recursively partitioned mixture model (*SS-RPMM*) with DNA methylation data was applied to identify a classifier for 10-year overall survival. Higher CD8T memory cell proportions were associated with better overall survival (HR = 0.95, 95% CI = 0.93 – 0.98), while higher NLR (HR = 1.36, 95% CI = 1.23 – 1.50), CD8T naïve (HR = 1.21, 95% CI = 1.04 – 1.41), neutrophil (HR = 1.04, 95% CI = 1.03 – 1.06) proportions, and age acceleration (HR = 1.06, 95% CI = 1.03 – 1.08) were associated with worse overall survival in bladder cancer patients. *partDSA* and *SS-RPMM* classified five groups of subjects with significant differences in overall survival. We identified associations between immune cell subtypes and age acceleration with bladder cancer outcomes. The findings of this study suggest that bladder cancer outcomes are associated with specific methylation-derived immune cell-type proportions and age acceleration, and these factors could be potential prognostic biomarkers.

3.2 Introduction

Bladder cancer is a malignant urogenital neoplasm and is classified into non-muscle-invasive bladder cancer (NMIBC) and muscle-invasive bladder cancer (MIBC). In 2022, an estimated 81,000 new cases of bladder cancer and 17,000 deaths from the disease occurred in the U.S. [180]. The common risk factors of bladder cancer are age, sex, and smoking. Bladder cancer is four times more common in men compared with women [181]. About 90% of bladder cancer patients are age 55 or older, and patients younger than 60 have a higher 10-year overall survival rate than patients older than 60 [182]. Around 50% to 60% of new cases are attributed to smoking, and current smoking has a positive association with the risk of recurrence [183]. The conventional treatment for bladder cancer is surgery or surgery in combination with chemotherapy drugs or intravesical immunotherapy (Bacillus Calmette-Guérin (BCG)) [120]. Even though transurethral resection and immunotherapy generally control the disease [3], [184], the tumor recurrence rate is about 40% after treatment [185], [186]. Predictive biomarkers that alert clinicians to recurrence would help to improve the clinical management of bladder cancer.

Circulating immune cell profiles have been associated with outcomes in bladder cancer patients. For example, CD8+ cell proportions were associated with a decreased risk of tumor recurrence [187]. Also, an elevated neutrophil-to-lymphocyte ratio (NLR) has been associated with worse overall survival and higher recurrence rate [125], [126]. Previously [188], we measured peripheral blood DNA methylation profiles of NMIBC patients and applied cell-type

deconvolution to estimate the proportions of six immune cell types [89]. CD4T and CD8T cell proportions were associated with decreased risk of death and recurrence. Yet, there have been limited studies investigating the relationship of circulating immune profiles in bladder cancer with disease outcomes. Further, subtypes of each major cell type have been shown to affect cancer development distinctly. For instance, cytotoxic CD4+ T cells can kill the bladder tumor cells, and in contrast, regulatory CD4+ T can suppress the activity of cytotoxic CD4+ T cells and lead to tumor growth indirectly [189]. To broaden the scope of the effects of circulating immune profiles on bladder cancer outcomes, it is necessary to investigate the association between leukocyte subtypes and bladder cancer outcomes.

Recently, our group developed an enhanced method to perform high-resolution cell mixture deconvolution to resolve 12 immune cell types in blood using DNA methylation measures (naïve and memory B, CD4T, and CD8T, as well as regulatory T, monocyte, NK cells, neutrophils, basophils, and eosinophils) [102]. Because DNA methylation involves gene regulation for cell lineage specification [133], cell-specific differentially methylated regions (DMRs) can be utilized to distinguish cell types with reference-based deconvolution [89], [90], [96]. Also, DNA methylation cytometry is more efficient for immune profiling with high accuracy than flow cytometry and can be applied to archival specimens.

Epigenetic clocks have been developed to estimate chronological age or physiological age in regard to aging outcomes, such as cancers and all-cause

mortality [87], [88], [113]. Age acceleration derived from these clocks has been associated with prospective risk in lung, kidney, and pancreatic cancer [190]–[192]. Moreover, age acceleration has been associated with outcomes in other cancers [193], [194]. Though only a few studies have reported the association of age acceleration with bladder cancer risk, they have not exhibited consistent results nor mentioned the subtype of bladder cancer they investigated. For example, one study showed that Pheno and Grim age acceleration were positively associated with the prospective risk in bladder cancers [190]. Another study reported that Horvath and Hannum's age acceleration was not associated with bladder cancer risk [195]. Since aging is one of risks of bladder cancer [196], we investigated the association of age acceleration with bladder cancer outcomes.

Here, we hypothesized that DNA methylation-derived immune cell proportions and age acceleration are associated with bladder cancer outcomes. We applied our new methylation cytometry approach for extended immune cell resolution to DNA methylation profiles of archival blood samples from a population-based study containing NMIBC (N = 601) patients. We then tested the association of cancer outcomes with each leukocyte subtype proportion and age acceleration. We also used *partDSA* [197], a classification and regression trees method, and a semi-supervised recursively partitioned mixture model (*SS-RPMM*) [198], to group/cluster our subjects based on cancer outcomes, patients' demographics, tumor characteristics, and methylation profiles of specific CpG loci.

3.3 Methods

3.3.1. Study Subjects and Samples

A detailed description of subjects who participated in the present study is available in prior studies [138]–[140]. Briefly, bladder cancer subjects were recruited from three phases of a New Hampshire population-based case-control study [141]. The first phase collected blood samples from 331 individuals diagnosed between July 1994 and June 1998 (phase 1). The second phase collected blood samples from 243 individuals diagnosed between July 1998 and December 2001 (phase 2). The third study phase collected blood samples from 194 individuals diagnosed between July 2002 and December 2004 (phase 3). Bladder cancer patients were identified using the New Hampshire State Cancer Registry and hospital cancer registry (Patients in phase 3 were identified using the hospital cancer registry only). Patients' overall survival data were from the National Death Index, and tumor recurrence data were ascertained through chart review. We performed four comprehensive National Death Index (NDI) searches for the years 2008, 2010, 2014, and 2018 to identify cases of death. Additionally, the NH State Cancer Registry had previously reported some deaths to us through their conducted searches. During each NDI search, we included all bladder cancer case-control study participants who had not been previously matched with an NDI death. To ensure accurate matching, we followed the NDI-recommended method and utilized the code and algorithms provided by the NDI to score the matches. Furthermore, we applied the NDI score interpretation using the recommended NPCR algorithm [199] to enhance the accuracy of our findings.

Subjects without histopathology re-review, muscle-invasive status, tumor grade, or smoking status were excluded from the study. The remaining 601 patients with non-muscle-invasive bladder cancer were used in downstream statistical analyses. In addition, 40 patients received Bacillus Calmette-Guérin (BCG) in phase 1, 30 received BCG in phase 2, and 19 received BCG in phase 3. All patients with BCG treatment had their blood drawn after BCG treatment, and all blood samples were taken after the initial diagnosis. This study was approved by the Dartmouth Human Research Protection Program (IRB) (Approval Number STUDY00010107). The written informed consent was obtained from the patients and the studies were conducted in accordance with Belmont Report.

3.3.2. DNA Extraction, Qualification, and Bisulfite Modification

After the blood draw, blood samples were kept at 4°C and frozen within 24 hours. DNA was extracted from blood samples using the QIAamp DNA Blood Kit (Qiagen, CA, USA) according to the manufacturer's protocol. Extracted DNA quantity and quality were assessed with Qubit 3.0 Fluorometer (Life Technologies, CA, USA) and Fragment Analyzer (Advanced Analytical, IA, USA). Then, extracted DNA underwent bisulfite conversion using EZ DNA Methylation Kit (Zymo Research, CA, USA) according to the manufacturer's protocol. Approximately 750 ng of bisulfite-modified DNA was used as input for the DNA methylation array.

3.3.3. DNA Methylation Profiling

After DNA extraction, quantification, and bisulfite modification, Infinium MethylationEPIC Bead Chips (Illumina, Inc., CA, USA) were used to measure the DNA methylation status of bisulfite-modified DNA samples. Raw probe intensity data, iDAT files, from the methylation array were processed through *preprocessNoob* using *minfi* (RRID:SCR_012830) [142], and quality control was performed using *ENmix* [143] R package. To distinguish from background noise, samples with more than 5% of probes with a detection P $>1.0\times10^{-6}$ were not included. In addition, we dropped 32,713 probes that were not detected in more than 10% of the samples. Then, the bias of type-2 probe values was corrected using *BMIQ* (RRID:SCR_003446) [200] from *watermelon* (RRID:SCR_001296) [201] R package, and the *ComBat* (RRID:SCR_010974) [144] was used to adjust for batch effects. Next, 106,522 probes previously reported to be cross-reactive, SNP-associated, non-CpG (CpH) methylation, and sex-specific were excluded [145]. After these exclusions, 726,856 CpGs remained for downstream statistical analysis. The annotation for CpG sites was from *IlluminaHumanMethylationEPICanno.ilm10b4.hg19*.

3.3.4. Statistical Analysis

Methylation age was estimated with the function *methyAge* from the *ENmix* [143] R package implementation of the methylation age estimation. Age acceleration was defined as the residual from a regression of methylation age on chronological age. Cell-type proportions were estimated with the *projectCellType_CP* from the *FlowSorted.Blood.EPIC* (RRID:SCR_022540)

[89] R package. The neutrophil-to-lymphocyte ratio (NLR) was calculated according to the ratio of neutrophil proportion to lymphocyte proportion.

Ten-year overall survival (OS) was defined as the time interval from the date of initial diagnosis to death. Patients alive or lost to follow-up were censored at the last follow-up. Similarly, ten-year recurrence-free survival (RFS) was defined as the time interval from the date of initial diagnosis to the first tumor recurrence or death, whichever occurred first, and patients alive and without tumor recurrence or lost to follow-up were censored at the last follow-up. For OS and RFS, survival times were truncated at ten years. In univariate and multivariable analyses, *coxph* from the *survival* (RRID:SCR_021137) R package was used to fit Cox proportional hazard models to evaluate the association between bladder cancer outcomes and each variable. Only immune cell profiles significantly associated with bladder cancer outcomes in the univariate Cox model were subjected to multivariable analyses. *Cox.zph* from the *survival* R package was employed to test the proportional hazards assumption. Predictors with assumption violations were included as strata in the Cox models. The linearity assumption was examined with the *ggcoxfunctional* from the *survminer* (RRID:SCR_021094) R package. We conducted 2% winsorization on immune cell profiles identified to violate the linearity assumption. FDR-corrected P-value of < 0.05 was the significance threshold on multivariable analysis.

To explore interactions between clinical variables and immune cell proportions in survival analysis, we applied a partitioning deletion/substitution/addition algorithm (*partDSA* [197], [202]; R package) for

model building employing the inverse probability censoring weighted-L2 loss function. Those variables (age, Hannum or Pheno age acceleration, sex, tumor grade, smoking status, BCG treatment status, and immune cell type proportions) associated with 10-year overall survival were included in the multivariable model as input. The *partDSA* approach resulted in three groups of subjects that were based on neutrophil and CD8 naïve cell proportions. After the model was built, corresponding Kaplan-Meier curves were generated, and hazard ratios and 95% confidence intervals were calculated using the Cox model.

To identify a novel set of blood DNA methylation profiles associated with cancer outcomes, we applied a semi-supervised recursive partitioning mixture modeling (*SS-RPMM*) algorithm [198]. This method uses the recursive partitioning mixture model (*RPMM*), demonstrating an effective and efficient unsupervised clustering procedure for methylation data [84], [203]–[205]. To avoid overfitting and provide for validation of the model, we randomly split the total population into a training and testing set at a 2:1 ratio, stratified by deceased status (whether subjects were deceased or censored within ten years) to balance the distribution of outcome status between sets. We used the 10% most variable CpG loci in methylation beta values across all samples. After splitting subjects and subsetting CpGs, a series of Cox proportional hazard models were fit using the training set for each selected CpG loci adjusted for age, age acceleration, sex, tumor grade, smoking status, BCG treatment status, and immune cell type proportions associated with 10-year overall survival of NMIBC patients. Next, Cox-scores ($|\beta|/se(\beta)$, where β = the proportional hazards estimate of the log-

hazard ratio, and *se* = the standard error) were computed for each of the selected CpG loci, and Cpg loci were ranked based on the Cox scores. Subsequently, the top *M* (range: 5 - 50) loci with the largest absolute Cox scores were chosen using a x-fold cross-validation *RPMM* with the smallest median P-value of the log-rank test for each potentially optimal number (*M*) of CpG loci in the training set. Then, *RPMM* was fit to the testing set for clustering subjects using the optimal *M*-selected CpG loci with the largest absolute Cox score, predicting the methylation class membership for the subjects. Then, all NMIBC patients were clustered using RPMM based on the methylation levels of the optimal CpG sites. Finally, we evaluated the association of *RPMM* class membership with overall survival using Cox proportional hazard models.

3.3.5. Data Availability

All datasets generated and analyzed during this current study are available in the Gene Expression Omnibus (GEO) repository at GSE183920.

3.4 Results

3.4.1. Characteristics of subjects

DNA methylation profiles were obtained from 601 peripheral blood samples from non-muscle-invasive bladder cancer (NMIBC) patients using the Human MethylationEPIC array. The study group was 455 (75.7%) men, 306 (50.9%) former-smokers, 192 (32.0%) current-smokers, 89 (14.8%) with BCG treatment, and had a median age of 66 (**Table 3.1**). The distribution of chronological age,

methylation age, and age acceleration is shown in **Figure S3.1A**. Cell-type proportions were estimated for each patient using methylation cytometry (**Figure S3.1B**). Stratifying on median time from diagnosis to blood draw, patient and tumor characteristic summary statistics showed similar distributions. Further, we observed no significant associations of immune profile variables with time from diagnosis to blood draw. To assess the potential modification of results by time to the blood draw, we performed an analysis testing the relation of immune profile variables with patient outcome, stratifying patients into two groups based on median time to the blood draw. We did not observe differences between associations of immune variables with outcomes between the groups based on time to blood draw.

3.4.2. Risk of bladder cancer outcomes

First, we examined associations of three major methylation age clocks, Horvath age [87], Hannum age [88], and DNAmPhenoAge [113], with bladder cancer outcomes. Our findings showed that Horvath age acceleration was not associated with 10-year overall survival (OS) (**Table S3.1**). Since the Horvath clock was developed using data from a subset of CpG loci on the Illumina HumanMethylation27 (27K) CpG BeadChip (~27,000 features) compared with the Hannum and PhenoAge clocks (which were developed using 20 times more features with data from CpGs on the Illumina HumanMethylation450 (450K) and EPIC (850K) BeadChips), we focused on Hannum and Pheno age acceleration in subsequent analyses. Then, we fit multivariable Cox proportional

hazard models for demographic and tumor characteristic variables to investigate associations with 10-year OS and recurrence-free survival (RFS). For NMIBC patients, age, age acceleration, smoking, and high tumor grade were associated with worse RFS and OS. Women had a better survival outcome compared with men (**Table 3.2, Table S3.2**).

Table 3.1: Characteristics of subjects

	NMIBC (n = 601)
Age	
Median (IQR)	66 (57,71)
Pheno Age Acceleration	
Median (IQR)	-0.41 (-4.28, 3.65)
Hannum Age Acceleration	
Median (IQR)	-0.10 (-2.53, 2.54)
Sex	
Male	455 (75.7%)
Female	146 (24.3%)
Tumor grade	
Low Grade	450 (74.9%)
High Grade	151 (25.1%)
Smoking status	
Never	103 (17.1%)
Former	306 (50.9%)
Current	192 (32.0%)
BCG: Immunotherapy	
No	512 (85.2%)
Yes	89 (14.8%)
Time from diagnosis to blood draw (Days)[b]	
Median (IQR)	319 (176, 569)
NLR	
Median (IQR)	1.96 (1.38, 2.86)
10-year survival status	
Alive	423 (70.4%)
Deceased	178 (29.6%)
10-year Recurrence-free status[a]	
No	230 (38.3%)
Yes	371 (61.7%)

[a]Whether patient with tumor recurrence or deceased within 10 years.

[b]All blood samples were taken after the diagnosis.

Table 3.2: Cox proportional hazards multivariable models for demographic and tumor characteristics of 601 NMIBC patients (For Pheno Age Acceleration)

	10-year overall survival	10-year recurrence-free survival[a]
	HR (95% CI)	HR (95% CI)
Age	1.08 (1.06-1.10)	1.02 (1.01-1.03)
Pheno Age Acceleration	1.06 (1.03-1.08)	1.02 (1.00-1.03)
Sex		
Male		
Female	0.50 (0.33-0.78)	
Tumor grade		
Low Grade		
High Grade	1.58 (1.15-2.17)	1.49 (1.18-1.88)
Smoking status		
Non-smoker		
Former-smoker	1.35 (0.82-2.22)	1.56 (1.13-2.15)
Current-smoker	1.87 (1.10-3.16)	1.69 (1.20-2.38)
BCG treatment		
No		
Yes	0.89 (0.58-1.34)	

[a]Stratification was used on sex and BCG treatment status for proportional assumption.

NMIBC = non-muscle-invasive bladder cancer, HR = hazard ratio, CI = confidence interval

Next, the association between immune cell-type proportions and bladder cancer patient outcomes was investigated. We fit multivariable Cox models for immune cell-type proportions associated with bladder cancer outcomes in Cox univariate models (**Table 3.3, Table S3.3**). CD4T memory, CD8T memory, and NK cell proportions were associated with a decreased risk of death. Whereas NLR, CD8T naïve, and neutrophil cell proportions were each associated with an increased risk of death. In analyses of RFS, CD4T memory was associated with a decreased risk of tumor recurrence, and NLR, monocyte, and regulatory T cell

proportions were associated with an increased risk of tumor recurrence (**Table 3.3, Table S3.3**). Hazard estimates for associations of demographic and tumor characteristic variables with outcomes were similar after adjusting for cell composition. Since there was a stronger (smaller P-value) association of immune cell profiles with 10-year overall survival compared with 10-year recurrence-free survival, subsequent analyses are focused on 10-year overall survival of NMIBC patients. Since BCG treatment may affect immune cell composition, we performed a sensitivity analysis limiting our analysis to patients not receiving BCG treatment and observed results consistent with those obtained for all NMIBC patients (**Table 3.3; Table S3.4**).

Table 3.3: Cox proportional hazards models of immune cell proportions and NMIBC patient outcomes (For Pheno Age Acceleration)

	10-year overall survival			10-year recurrence-free survival		
	Univariate model	Multivariable[a] model		Univariate model	Multivariable[b] model	
	HR (95% CI)	HR (95% CI)	FDR	HR (95% CI)	HR (95% CI)	FDR
NLR	1.49 (1.37-1.62)	1.36 (1.23-1.50)	$6.9x10^{-9}$	1.14 (1.06-1.22)	1.10 (1.01-1.18)	0.046
Memory B cell	0.82 (0.72-0.94)	0.89 (0.78-1.02)	0.13	0.94 (0.87-1.03)		
Naïve B cell	0.86 (0.79-0.94)	0.93 (0.86-1.00)	0.09	0.99 (0.94-1.04)		
Memory CD4T cell	0.92 (0.90-0.95)	0.95 (0.93-0.98)	$3.8x10^{-3}$	0.97 (0.96-0.99)	0.98 (0.96-0.99)	0.046
Naïve CD4T cell	0.87 (0.82-0.93)	0.98 (0.92-1.04)	0.48	0.95 (0.92-0.99)	0.99 (0.95-1.02)	0.46
Memory CD8T cell	0.96 (0.94-0.99)	0.95 (0.93-0.98)	$3.8x10^{-3}$	0.99 (0.97-1.01)		
Naïve CD8T cell	0.84 (0.73-0.96)	1.21 (1.04-1.41)	0.03	0.93 (0.86-1.02)		
Monocyte	1.06 (1.02-1.11)	1.01 (0.96-1.06)	0.74	1.05 (1.02-1.08)	1.03 (1.00-1.07)	0.07
Neutrophil	1.06 (1.05-1.08)	1.04 (1.03-1.06)	$2.0x10^{-6}$	1.01 (1.00-1.02)	1.01 (0.99-1.02)	0.31
Regulatory T cell	1.27 (1.08-1.49)	1.16 (0.98-1.36)	0.12	1.20 (1.06-1.35)	1.17 (1.03-1.32)	0.046
NK cell	0.92 (0.87-0.98)	0.92 (0.86-0.98)	0.03	0.98 (0.94-1.02)		
Basophil	1.51 (1.24-1.85)	1.26 (1.01-1.56)	0.07	1.16 (0.99-1.35)		
Eosinophil	1.02 (0.95-1.10)			1.03 (0.98-1.08)		

[a]The model controlling for age, sex, tumor grade, smoking status, BCG treatment status, and Pheno age acceleration.

[b]The model controlling for age, stratified sex, tumor grade, smoking status, stratified BCG treatment status, and Pheno age acceleration.

NMIBC: non-muscle-invasive bladder cancer, HR = hazard ratio, CI = confidence interval, NLR = neutrophil to lymphocyte ratio. Winsorization was used on the top 2% or the last 2% (only Neu) values for fitting linearity assumption.

3.4.3. Clinical and immune profiles recursive partitioning analysis

To partition the covariate space and explore interactions of clinical variables with immune cell proportions in 10-year overall survival, we applied a partitioning deletion/substitution/addition (*partDSA* [197]) algorithm for model building (**Figure 3.1A**). *partDSA* is an analytic algorithm employing recursive partitioning. It uses loss functions to build clinically interpretable predictors of risk for a given event based on the covariate information. In brief, this method divides the covariate space into mutually exclusive regions and stratifies patients into distinct risk groups with respect to an outcome. As a result, it can make guidelines for estimating a patient's prognosis from clinical and biological information [202]. Age, Hannum or Pheno age acceleration, sex, tumor grade, smoking status, BCG treatment status, and immune cell type proportions associated with 10-year overall survival were included in the model as input. The outcome was 10-year overall survival. The *partDSA* analysis divided subjects into three groups based on Neutrophil and CD8 naive cell proportions. The NMIBC patients with neutrophil proportion > 76.46 had the worst 10-year overall survival (Group 2, n = 17, HR = 4.93, 95% CI = 2.78 – 8.71) compared with Group 1 patients (patients with neutrophil cell proportions ≤ 76.46 and CD8 naïve cell proportions ≤ 1.76; n = 454) (**Figure 3.1B**). Although the 10- and 5-year overall survival rates (**Figure 3.1C**) for Groups 1 and 3 were not statistically significantly different, we observed that their corresponding Kaplan–Meier curves separated after five years. To further investigate the difference between 10-year OS in Group 1 and Group 3, we selected the NMIBC patients in Groups 1 and 3 whose time intervals from the date of initial diagnosis to death or the last

follow-up were greater than five years (sample size: Group 1 = 401; Group 3 = 117), using the Kaplan–Meier method to observe 10-year OS. As we expected, for patients with time intervals from the date of initial diagnosis to death or last follow-up of more than five years, Group 3 patients (neutrophil proportion ≤ 76.46, CD8T naïve proportion > 1.76) had better 10-year overall survival (HR = 0.37, 95% CI = 0.19 – 0.71) compared with Group 1 patients (**Figure 3.1D**). Next, we examined the distribution of immune cell-type proportions, methylation age, and age acceleration in the three groups. Among the groups, NMIBC patients in Group 2 had lower B naïve, eosinophil, monocyte, NK, CD4T memory, CD4T naïve, CD8T memory, and CD8T naïve proportions compared with other groups. Also, patients in Group 2 had higher neutrophil proportion, NLR, chronological age, methylation age, and age acceleration compared with other groups (**Figure S3.2**).

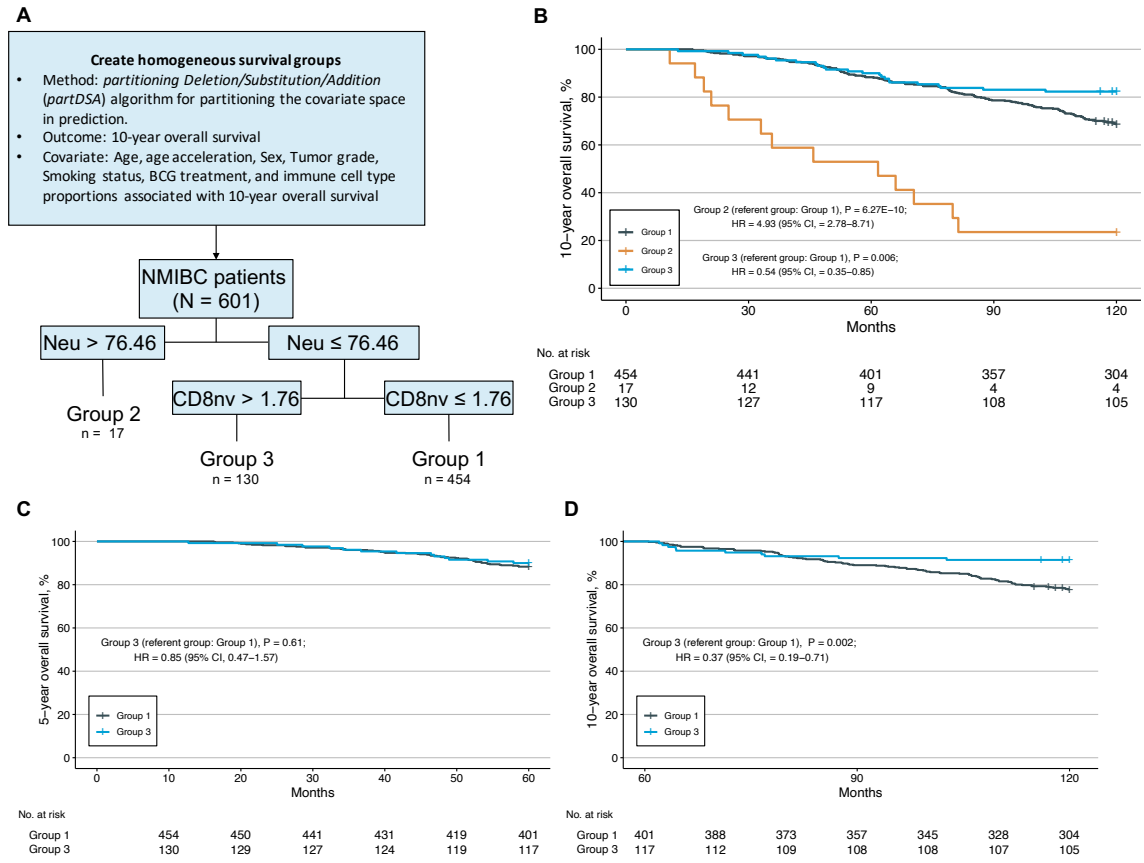


Figure 3.1: Clinical and immune profiles recursive partitioning analysis, and 10-year overall survival Kaplan-Meier curves stratified by the grouping result from *partDSA* in NMIBC patients (A) *partDSA* model setting and analysis results. For 601 non-muscle-invasive bladder cancer (NMIBC) patients, the neutrophil cell proportion in peripheral blood was the primary node, with the CD8 naïve cell proportion as the secondary node. NMIBC patients fell into 1 of 3 risk groups. Group 1 consisted of the 454 patients who had neutrophil cell proportions ≤ 76.46 and CD8 naïve cell proportions ≤ 1.76. Group 2 consisted of the 17 patients who had neutrophil cell proportions > 76.46. Group 3 consisted of the 130 patients who had neutrophil cell proportions ≤ 76.46 and CD8 naïve cell proportions > 1.76. CD4T memory, CD8T naïve, CD8T memory, NK cells, and neutrophils cell proportions were employed in the model using Pheno age acceleration; B memory, CD4T memory, CD8T naïve, CD8T memory, regulatory T, NK cells, neutrophils, and basophils cell proportions were employed in the model using Hannum age acceleration. Both models generated the same partitioning results. **(B)** Kaplan-Meier curves are shown based on clinical and immune profiles recursive partitioning analysis. **(C)** 5-year overall

survival Kaplan-Meier curves in NMIBC patients in group 1 and 3 **(D)** 5 to 10-year overall survival Kaplan-Meier curves in NMIBC patients in group 1 and 3 who were deceased or censored after 60 months. P-values for Log-rank tests are shown. All Kaplan-Meier curves are univariate analyses without adjusting for other variables. Neu = neutrophil; HR = hazard ratio; C.I. = confidence intervals

3.4.4. Semi-supervised recursively partitioned mixture model for 10-year overall survival

To investigate whether we could identify a blood DNA methylation profile associated with NMIBC survival, we applied a semi-supervised recursively partitioned mixture model (*SS-RPMM*) method. The workflow is illustrated in **Figure 3.2A** and **Figure S3.3A**. The subjects in the testing set were assigned to cluster membership using the methylation profiles of the optimal CpG sites (**Table S3.5A [Pheno]**, **and S3.5B [Hannum]**). Then, all NMIBC patients were clustered using *RPMM* based on the methylation levels of the optimal CpG sites, resulting in 2 classes, rR and rL ('R' and 'L' corresponded with branches in the dendrogram; r stands for root, **Figure 3.2B [Pheno]**, **Figure S3.3B [Hannum]**). Methylation class membership was significantly associated with 10-year overall survival; patients in cluster rR had a more favorable 10-year OS compared with those in cluster rL in both the testing set (HR = 0.35, 95% CI = 0.20 – 0.60, **Figure 3.2C [Pheno]**; HR = 0.38, 95% CI = 0.21 – 0.68, **Figure S3.3C [Hannum]**) and when using all NMIBC patients (HR = 0.35, 95% CI = 0.25 – 0.48 **Figure 3.2D [Pheno]**; HR = 0.37, 95% CI = 0.27 – 0.52 **Figure S3.3D [Hannum]**). Then, we compared the distribution of immune cell-type proportions, chronological age, methylation age, and age acceleration.

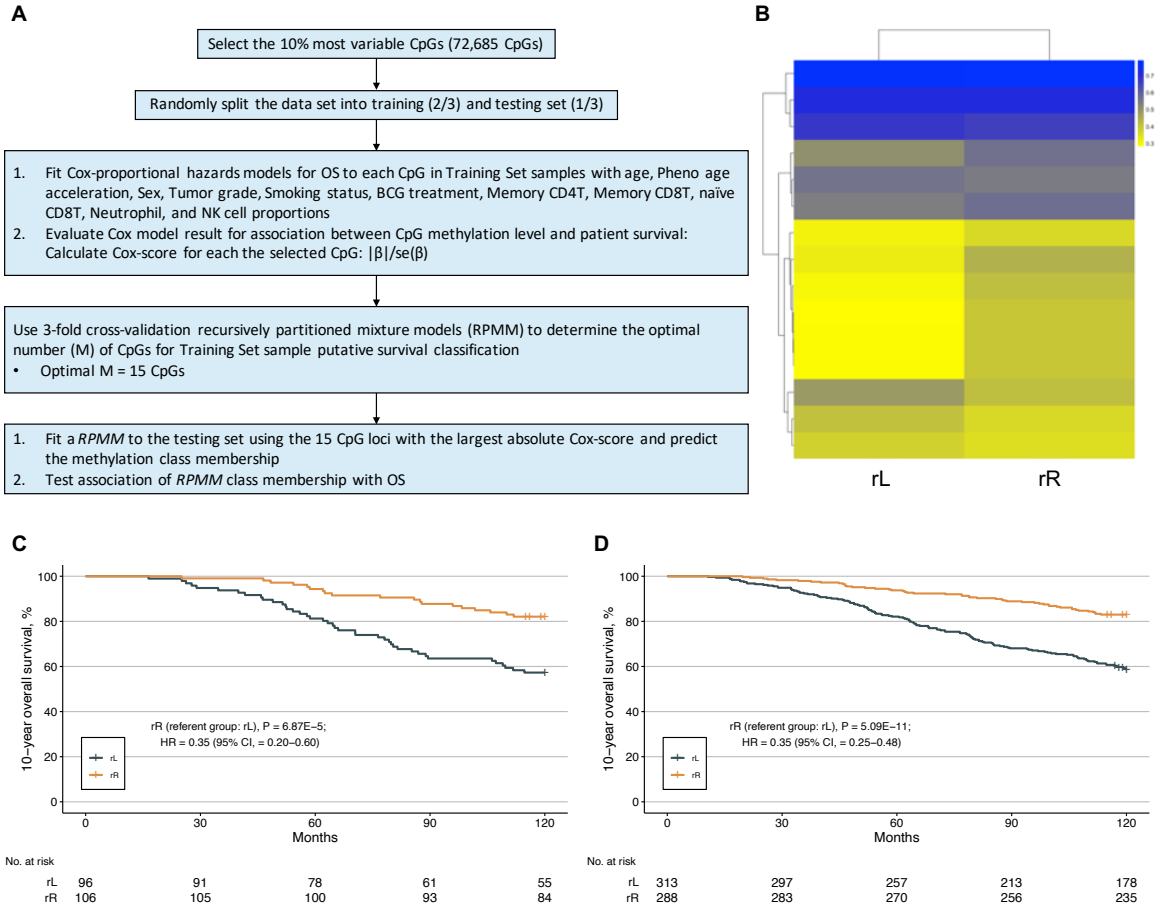


Figure 3.2: Semi-Supervised Recursively Partitioned Mixture Model (*SS-RPMM*) for 10-year overall survival (OS) in NMIBC patients [For Pheno age acceleration] **(A)** Data analysis schematic of *SS-RPMM* used for identification of blood DNA methylation profiles associated with non-muscle-invasive bladder cancer (NMIBC). **(B)** Heatmap of predicted class memberships for the observations in all NMIBC patients using the average beta values of the 15 CpG loci with the largest absolute Cox-scores. **(C)** Kaplan-Meier curves of 10-year overall survival stratified by the *SS-RPMM* classification of 202 NMIBC patients in the testing set by the 15 CpG loci. **(D)** Kaplan-Meier analysis of 10-year overall survival. 10-year overall survival curves stratified by the grouping result from *SS-RPMM* in all NMIBC patients. P-values for Log-rank tests are shown. All Kaplan-Meier curves are univariate analyses without adjusting for other variables. HR = hazard ratio; C.I. = confidence intervals; *SS-RPMM* = semi-supervised recursively partitioned mixture model

Consistent with the models in **Table 3.3** and **Table S3.3**, we observed that patients in cluster rR (n = 288) had significantly higher B memory, B naïve, CD4T memory, CD4T naïve, CD8T memory, and NK cell proportions and had significantly lower basophil, eosinophil, monocyte, neutrophil cell proportions, NLR, chronological age, methylation age, and age acceleration compared with patients in cluster rL (n = 313) (**Figure S3.4A**, **B**). The model using Hannum age acceleration had similar results shown in **Figure S3.4C**, **D**.

3.4.5. Combined results from immune cell proportions and methylation profile groups

The two methods above used immune profiles (*partDSA*) and DNA methylation profiles (*SS-RPMM*) to explore the association of these profiles with overall survival respectively. We were curious whether we could combine immune and methylation information to produce a guideline for estimating a patient's prognosis. To gain a deeper understanding of interactions between clinical variables, immune cell proportions, and blood DNA methylation profiles in 10-year overall survival, we allocated NMIBC patients based on clustering results from *partDSA* and *SS-RPMM* analyses in 5 groups (**Table S3.6**). Groups 1 and 3, from the *partDSA* analysis, were divided into two subgroups based on the *SS-RPMM* analysis (**Figure 3.3A**, **Figure S3.5A**). In Kaplan–Meier analysis, patients in Group 2 still had the worst 10-year overall survival rate among all groups. Within the Group 1 patients, the group G1_rR had better 10-year overall

survival than the group G1_rL (HR = 0.40, 95% CI = 0.27 – 0.58, **Figure 3.3B [Pheno]**; HR = 0.45, 95% CI = 0.31 – 0.65, **Figure S3.5B [Hannum]**). Consistent with Group 1 patients, in Group 3 patients, the group G3_rR had better 10-year overall survival compared with the group G3_rL (HR = 0.42, 95% CI = 0.18 – 0.95, **Figure 3.3B [Pheno]**; HR = 0.35, 95% CI = 0.15 – 0.79, **Figure S3.5B [Hannum]**).

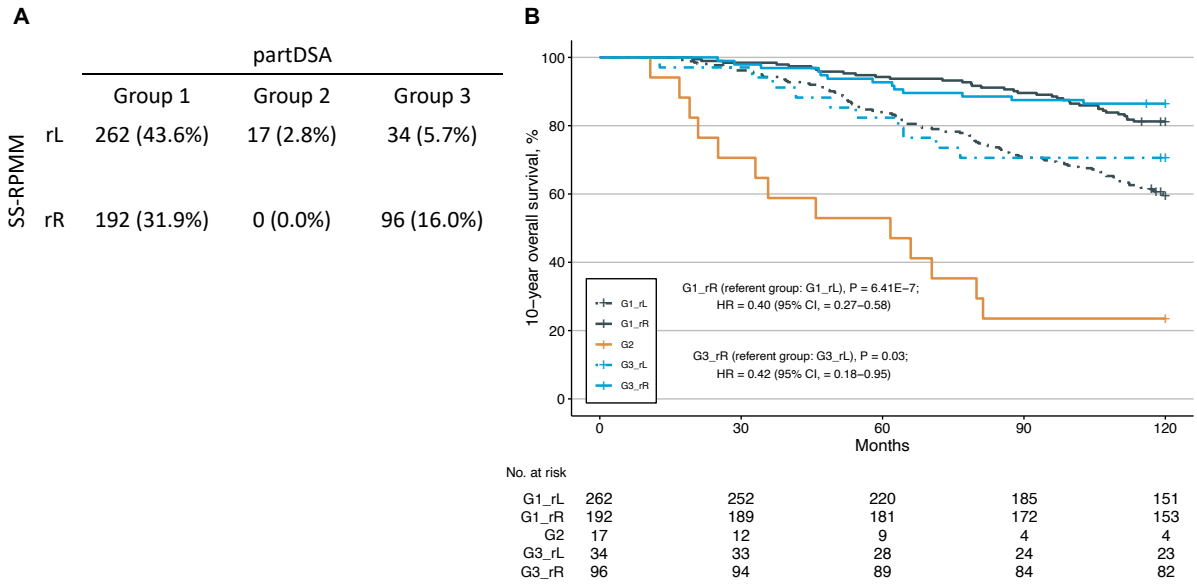


Figure 3.3: Kaplan-Meier analysis of 10-year overall survival based on the grouping results from both *partDSA* and *SS-RPMM* in all NMIBC patients [For Pheno age acceleration] (A) contingency table based on the grouping results from both *partDSA* and *SS-RPMM* in all NMIBC patients **(B)** 10-year overall survival curves of all 5 groups. P-values for Log-rank tests are shown. All Kaplan-Meier curves are univariate analyses without adjusting for other variables. HR = hazard ratio; C.I. = confidence intervals; *partDSA* = partitioning deletion/substitution/addition algorithm; *SS-RPMM* = semi-supervised recursively partitioned mixture model; NMIBC = non-muscle-invasive bladder cancer

3.5 Discussion

This study aimed to test the relation of bladder cancer outcomes with high-

resolution immune profiles using methylation cytometry and cell-independent methylation states in blood. In prior work, we observed that CD4T and CD8T proportions were associated with the decreased risk of death and tumor recurrence in non-muscle-invasive bladder cancer patients (NMIBC) patients [188]; however, the results were not generalizable to naïve and memory subtypes of CD4T and CD8T cells. With recent advances in methylation cytometry for immune profiling [102], we were able to examine the association between the proportions of regulatory T cells, eosinophils, basophils, naïve and memory subtypes of CD4T, CD8T, and B cell and bladder cancer outcomes. Consistent with our previous study, neutrophil-to-lymphocyte ratio (NLR) was associated with an increased risk of death and tumor recurrence in NMIBC patients. For CD4T and CD8T cell subsets, CD4T memory and CD8T memory cell proportions were associated with a decreased risk of death. However, only CD4T memory cell proportion was associated with a reduced risk of death and tumor recurrence. One possible explanation for the different observations is that we introduced age acceleration into the models. This might confound the association between CD8T cell proportions subtypes and bladder cancer outcomes. Here, we identified associations between immune cell subtypes and age acceleration with bladder cancer outcomes. These factors could be potentially prognostic biomarkers of bladder cancer.

Few studies have shown age acceleration from multiple age clocks to be associated with bladder cancer outcomes, and even then, they do not show consistent results [190], [195]. In our study, the direction of hazard estimates

among cell type proportions was similar. However, outcome associations with immune profile differed in Cox multivariable models controlling for Pheno and Hannum age acceleration. The observed difference was potentially due to the training methods of clocks. Unlike Hannum age using chronological age as a surrogate, Pheno age mainly focuses on aging outcomes, such as cancers, diet, and all-cause mortality, hence, Pheno age can capture age-related outcomes and perform well in predicting survival compared with chronological methylation clocks, such as Hannum and Horvath clocks. Horvath age acceleration was not associated with bladder cancer outcomes. One potential explanation is that the Horvath clock was built mainly using a subset of CpGs on the 27K methylation array platform, which had approximately twenty times fewer CpGs than the 450K and EPIC array platforms used for the developing of the Hannum and PhenoAge clocks. Furthermore, various cell and tissue types were used to develop the Horvath clock. However, Hannum and PhenoAge clocks were built based on data from blood DNA measures.

Our findings suggest that elevated NLR, neutrophil, basophil, regulatory T, and decreased CD4T memory cell proportions increased the risk of death and tumor recurrence in bladder cancer patients. These findings met our expectations, consistent with previous studies demonstrating that peripheral blood NLR levels were associated with an increased risk of NMIBC recurrence after surgery [75], [154], [206]. Moreover, basophil count was significantly associated with an increased risk of recurrence in BCG-treated bladder cancer patients [207]. One study indicated that peripheral (neutrophil x platelet) / (lymphocyte) was

inversely correlated with a high risk of tumor recurrence in NMIBC [159]. Even though some peripheral immune profiles we found had not been revealed to be associated with bladder cancer outcomes, these immune profiles in the tumor microenvironment had been reported to be associated with bladder cancer outcomes. For instance, higher regulatory T cell infiltration in the tumor microenvironment was associated with a shorter RFS [64], and regulatory T cell frequency within the tumor was inversely correlated with RFS in NMIBC patients [208]. Future work that integrates the assessment of cell-type proportions in the tumor microenvironment and periphery associated with bladder cancer outcomes would be of value.

The optimal 15 CpG loci selected by *SS-RPMM* using the model adjusting for Pheno age acceleration, and the optimal 50 CpG loci selected by *SS-RPMM* using the model adjusting for Hannum age acceleration track to several genes that have been reported to be involved in bladder cancer development. Sprouty-related EVH1 domain-containing protein 2 (*SPRED2*), is a negative regulator of the *ERK-MAPK* pathway, and has been reported to have increased mRNA and protein expression in NMIBC compared with carcinoma in situ and infiltrating urothelial carcinoma [209]. In addition, patients with higher *SPRED2* mRNA levels had better overall survival compared with low expression group [209]. Peroxisome proliferator-activated receptor gamma (*PPARG*) high-activation has been reported to promote cell cycle G2 arrest and apoptosis, leading to suppression of tumor growth and better prognosis in bladder cancer patients [210]. Fibrous sheath interacting protein 1 (*FSIP1*) was overexpressed in protein

and mRNA levels of bladder tumor tissues and cancer cell lines [211], [212]. Besides, *FSIP1* overexpression was associated with worse outcomes [211]. Phosphorylated mitogen-activated protein kinase 14 (*MAPK14*) was overexpressed in bladder cancer cell lines and tissues [213]. Furthermore, phosphorylated *MAPK14* could combine and regulate *RUNX2*, which was identified in our previous study through epigenome-wide association study (EWAS) [188], to promote the proliferation and migration of bladder cancer [213]. Consistent with our findings, patients with a worse survival rate (rL group) had lower methylation levels in the CpG site, cg16145324, located in the *MAPK14* gene region (**Table S5B**).

To further investigate the interactions between clinical variables, immune profiles, and DNA methylation levels, we performed both *partDSA* and *SS-RPMM*. Though NMIBC patients were divided into three groups using *partDSA*, the Kaplan-Meier curves for Group 1 and 3 patients were not significantly different. Interestingly, when we applied the optimal CpG loci selected by *SS-RPMM*, patients in Groups 1 and 3 were grouped into two groups, with a significant difference in 10-year overall survival. This finding illustrated the importance of methylation levels of specific CpG sites in evaluating bladder cancer prognosis.

While this study carefully evaluated the association of cancer outcomes with peripheral immune cell type proportions, there were potential study limitations. For instance, BCG treatment has been reported to affect immune cell composition [159], [207], [214] and methylation profiles [215]. We do not have

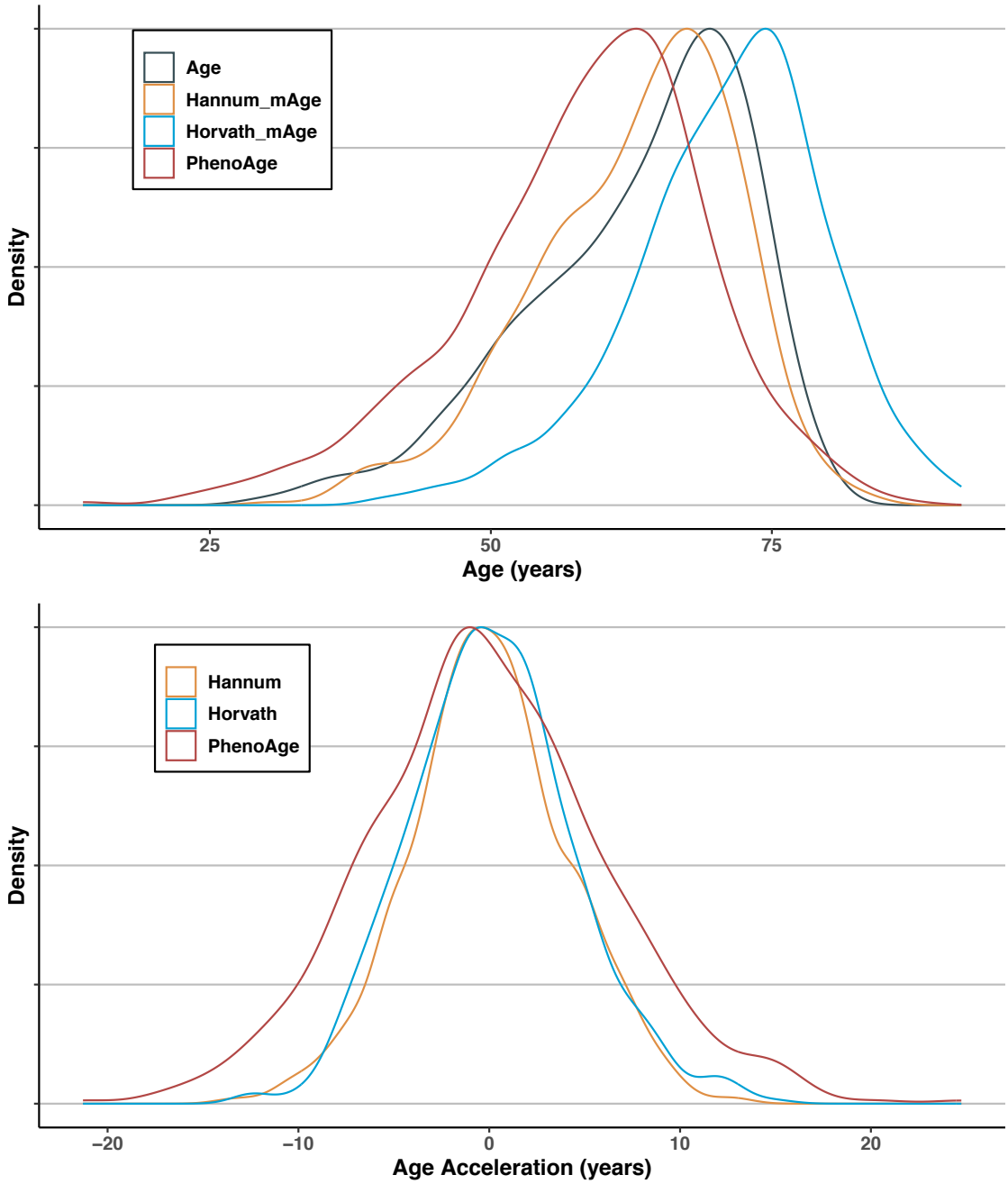
detailed information, such as cycle numbers or responsiveness for each patient, though only a few patients (14.8%) in our study received BCG treatment. In addition, our sensitivity analysis that limited to patients who didn't receive BCG treatment demonstrated consistent results with the overall NMIBC group. In addition, case ascertainment for our population-based study resulted in a median time from diagnosis to study blood draw is 319 days in our dataset. Further, some confounding factors, such as obesity [164], alcohol consumption [165], and type 2 diabetes [166], have been reported to affect peripheral immune cell distribution. However, we had incomplete information on these potential covariates.

Taken together, our analysis applied the latest differentially methylated region library and highlighted several peripheral immune profiles that were associated with bladder cancer outcomes. We assessed interactions between clinical variables, immune cell proportions, and blood DNA methylation profiles in 10-year OS using *partDSA* and *SS-RPMM* analyses, clustering NMIBC patients into 5 groups according to the methylation levels of the optimal CpG loci, neutrophil, and CD8T naïve cell proportions. While few studies have investigated the association between bladder cancer outcomes and peripheral immune profiles, our findings provide insight into the potential of peripheral immune profiles to serve as prognostic biomarkers in bladder cancer. Future work examining immune profiles in tumor tissues as well as DNA methylation profiles of bladder cancer patients is needed to integrate interactions between bladder cancer outcomes, methylation levels, peripheral immune environment, and tumor microenvironments to further validate the feasibility of methylation-

derived immune profiles for epigenetic biomarkers of bladder cancer.

3.6 Supplement: Chapter 3

A



B

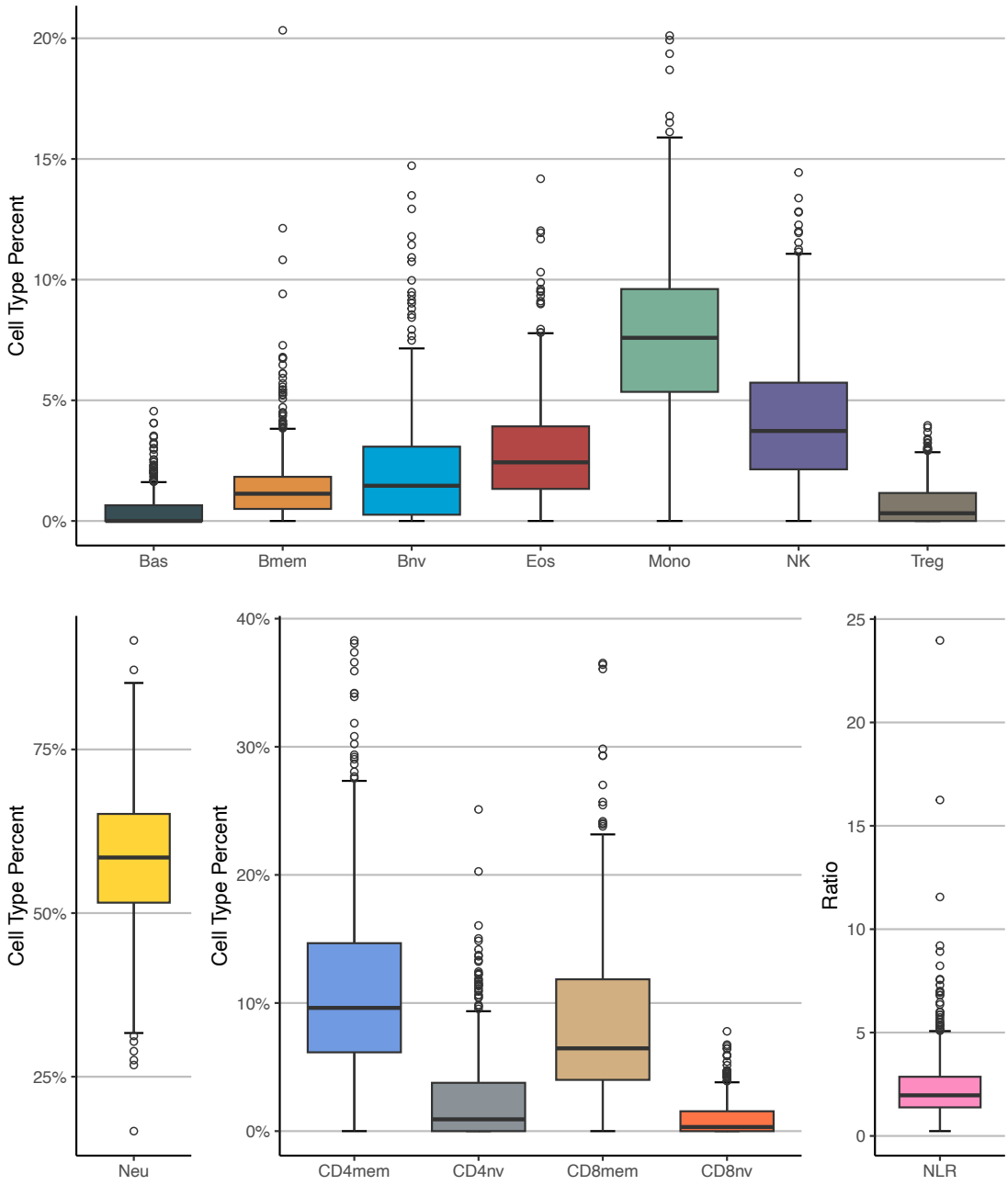


Figure S3.1: The distribution of chronological age, methylation age, age acceleration, and each immune cell profile (A) The distribution of chronological age, methylation age, age acceleration, and **(B)** each immune cell profile of 601 non-muscle-invasive bladder cancer patients

Table S3.1: Cox proportional hazards multivariable models[a] for age acceleration of 601 NMIBC patients

	10-year overall survival HR (95% CI)	10-year recurrence-free survival[b] HR (95% CI)
Horvath Age Acceleration	1.00 (0.97-1.03)	1.00 (0.98-1.02)
Hannum Age Acceleration	1.05 (1.02-1.09)	1.00 (0.97-1.02)
Pheno Age Acceleration	1.06 (1.03-1.08)	1.02 (1.00-1.03)

[a]Models was adjusted for age, sex, tumor grade, smoking status, and BCG treatment status.

[b]Stratification was used on sex and BCG treatment status for proportional assumption.

NMIBC: non-muscle-invasive bladder cancer, HR: hazard ratio, CI: confidence interval

Table S3.2: Cox proportional hazards multivariable models for demographic and tumor characteristics of 601 NMIBC patients (For Hannum Age Acceleration)

	10-year overall survival HR (95% CI)	10-year recurrence-free survival[a] HR (95% CI)
Age	1.08 (1.06-1.10)	1.02 (1.01-1.03)
Hannum Age Acceleration	1.05 (1.02-1.09)	1.00 (0.97-1.02)
Sex		
Male		
Female	0.51 (0.33-0.79)	
Tumor grade		
Low Grade		
High Grade	1.57 (1.14-2.16)	1.47 (1.16-1.86)
Smoking status		
Non-smoker		
Former-smoker	1.41 (0.86-2.32)	1.58 (1.15-2.18)
Current-smoker	2.15 (1.28-3.62)	1.77 (1.26-2.48)
BCG treatment		
No		
Yes	0.89 (0.58-1.35)	

[a]Stratification was used on sex and BCG treatment status for proportional assumption. NMIBC: non-muscle-invasive bladder cancer, HR: hazard ratio, CI: confidence interval

Table S3.3: Cox proportional hazards models of immune cell proportions and NMIBC patient outcomes (For Hannum Age Acceleration)

	10-year overall survival			10-year recurrence-free survival		
	Univariate model	Multivariable[a] model		Univariate model	Multivariable[b] model	
	HR (95% CI)	HR (95% CI)	FDR	HR (95% CI)	HR (95% CI)	FDR
NLR	1.49 (1.37-1.62)	1.40 (1.28-1.54)	$1.7x10^{-11}$	1.14 (1.06-1.22)	1.12 (1.04-1.21)	0.01
Memory B cell	0.82 (0.72-0.94)	0.86 (0.75-0.98)	0.04	0.94 (0.87-1.03)		
Naïve B cell	0.86 (0.79-0.94)	0.93 (0.86-1.01)	0.12	0.99 (0.94-1.04)		
Memory CD4T cell	0.92 (0.90-0.95)	0.95 (0.92-0.97)	$5.4x10^{-4}$	0.97 (0.96-0.99)	0.97 (0.96-0.99)	0.01
Naïve CD4T cell	0.87 (0.82-0.93)	0.96 (0.90-1.02)	0.19	0.95 (0.92-0.99)	0.97 (0.93-1.01)	0.14
Memory CD8T cell	0.96 (0.94-0.99)	0.94 (0.92-0.97)	$3.9x10^{-4}$	0.99 (0.97-1.01)		
Naïve CD8T cell	0.84 (0.73-0.96)	1.18 (1.01-1.38)	0.047	0.93 (0.86-1.02)		
Monocyte	1.06 (1.02-1.11)	1.03 (0.98-1.08)	0.26	1.05 (1.02-1.08)	1.04 (1.01-1.08)	0.02
Neutrophil	1.06 (1.05-1.08)	1.05 (1.03-1.07)	$6.1x10^{-8}$	1.01 (1.00-1.02)	1.01 (0.99-1.02)	0.08
Regulatory T cell	1.27 (1.08-1.49)	1.23 (1.05-1.46)	0.02	1.20 (1.06-1.35)	1.19 (1.05-1.34)	0.01
NK cell	0.92 (0.87-0.98)	0.90 (0.85-0.96)	0.01	0.98 (0.94-1.02)		
Basophil	1.51 (1.24-1.85)	1.36 (1.10-1.67)	0.01	1.16 (0.99-1.35)		
Eosinophil	1.02 (0.95-1.10)			1.03 (0.98-1.08)		

[a]The model controlling for age, sex, tumor grade, smoking status, BCG treatment status, and Hannum age acceleration.

[b]The model controlling for age, stratified sex, tumor grade, smoking status, stratified BCG treatment status, and Hannum age acceleration.

NMIBC: non-muscle-invasive bladder cancer, HR: hazard ratio, CI: confidence interval, NLR: neutrophil to lymphocyte ratio. Winsorization was used on the top

Table S3.4: Cox proportional hazards models of immune cell proportions and NMIBC patient outcomes (For Pheno Age Acceleration; NMIBC patients without BCG treatment; N = 512)

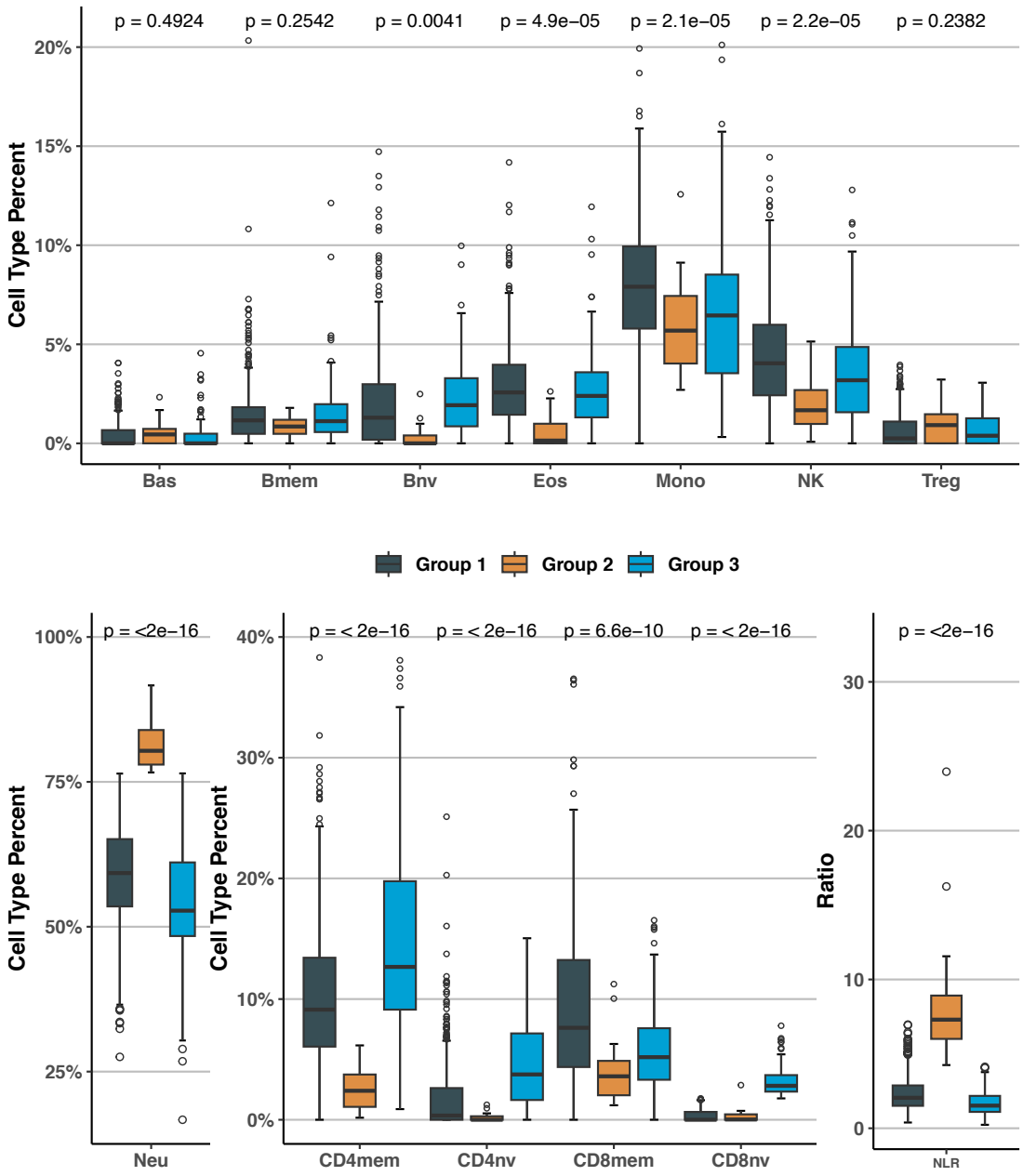
	10-year overall survival			10-year recurrence-free survival		
	Univariate model	Multivariable[a] model		Univariate model	Multivariable[b] model	
	HR (95% CI)	HR (95% CI)	FDR	HR (95% CI)	HR (95% CI)	FDR
NLR	1.49 (1.36-1.63)	1.33 (1.20-1.48)	$1.1x10^{-6}$	1.14 (1.06-1.22)	1.09 (1.00-1.19)	0.19
Memory B cell	0.80 (0.69-0.93)	0.88 (0.76-1.02)	0.12	0.96 (0.88-1.05)		
Naïve B cell	0.87 (0.80-0.96)	0.96 (0.88-1.04)	0.36	0.99 (0.94-1.05)		
Memory CD4T cell	0.92 (0.90-0.95)	0.96 (0.93-0.99)	0.01	0.98 (0.96-0.99)	0.98 (0.96-1.00)	0.19
Naïve CD4T cell	0.88 (0.82-0.94)	0.99 (0.93-1.06)	0.87	0.95 (0.92-0.99)	0.99 (0.95-1.04)	0.79
Memory CD8T cell	0.96 (0.93-0.99)	0.95 (0.92-0.98)	$2.7x10^{-3}$	0.99 (0.97-1.01)		
Naïve CD8T cell	0.83 (0.72-0.96)	1.22 (1.04-1.44)	0.03	0.93 (0.85-1.02)		
Monocyte	1.07 (1.01-1.12)	1.00 (0.95-1.06)	0.92	1.04 (1.01-1.08)	1.03 (0.99-1.06)	0.21
Neutrophil	1.06 (1.05-1.08)	1.04 (1.02-1.06)	$3.9x10^{-5}$	1.01 (1.001-1.02)	1.01 (0.99-1.02)	0.45
Regulatory T cell	1.31 (1.10-1.56)	1.21 (1.01-1.45)	0.06	1.14 (1.003-1.30)	1.12 (0.98-1.29)	0.19
NK cell	0.93 (0.87-0.99)	0.93 (0.86-0.99)	0.06	0.97 (0.93-1.02)		
Basophil	1.64 (1.33-2.02)	1.38 (1.10-1.74)	0.01	1.21 (1.03-1.43)	1.16 (0.97-1.38)	0.19
Eosinophil	1.01 (0.93-1.10)			1.01 (0.95-1.07)		

[a]The model controlling for age, sex, tumor grade, smoking status, and Pheno age acceleration.

[b]The model controlling for age, stratified sex, tumor grade, smoking status, and Pheno age acceleration.

NMIBC: non-muscle-invasive bladder cancer, HR = hazard ratio, CI = confidence interval, NLR = neutrophil to lymphocyte ratio. Winsorization was used on the top 2% or the last 2% (only Neu) values for fitting linearity assumption.

A


p = 0.4924
p = 0.2542
p = 0.0041
p = 4.9e−05
p = 2.1e−05
p = 2.2e−05
p = 0.2382
Cell Type Percent
20%
15%
10%
5%
0%
Bas
Bmem
Bnv
Eos
Mono
NK
Treg
Group 1
Group 2
Group 3
p = <2e−16
100%
75%
50%
25%
Neu
p = < 2e−16
p = < 2e−16
p = 6.6e−10
p = < 2e−16
40%
30%
20%
10%
0%
CD4mem
CD4nv
CD8mem
CD8nv
p = <2e−16
Ratio
30
20
10
0
NLR

B

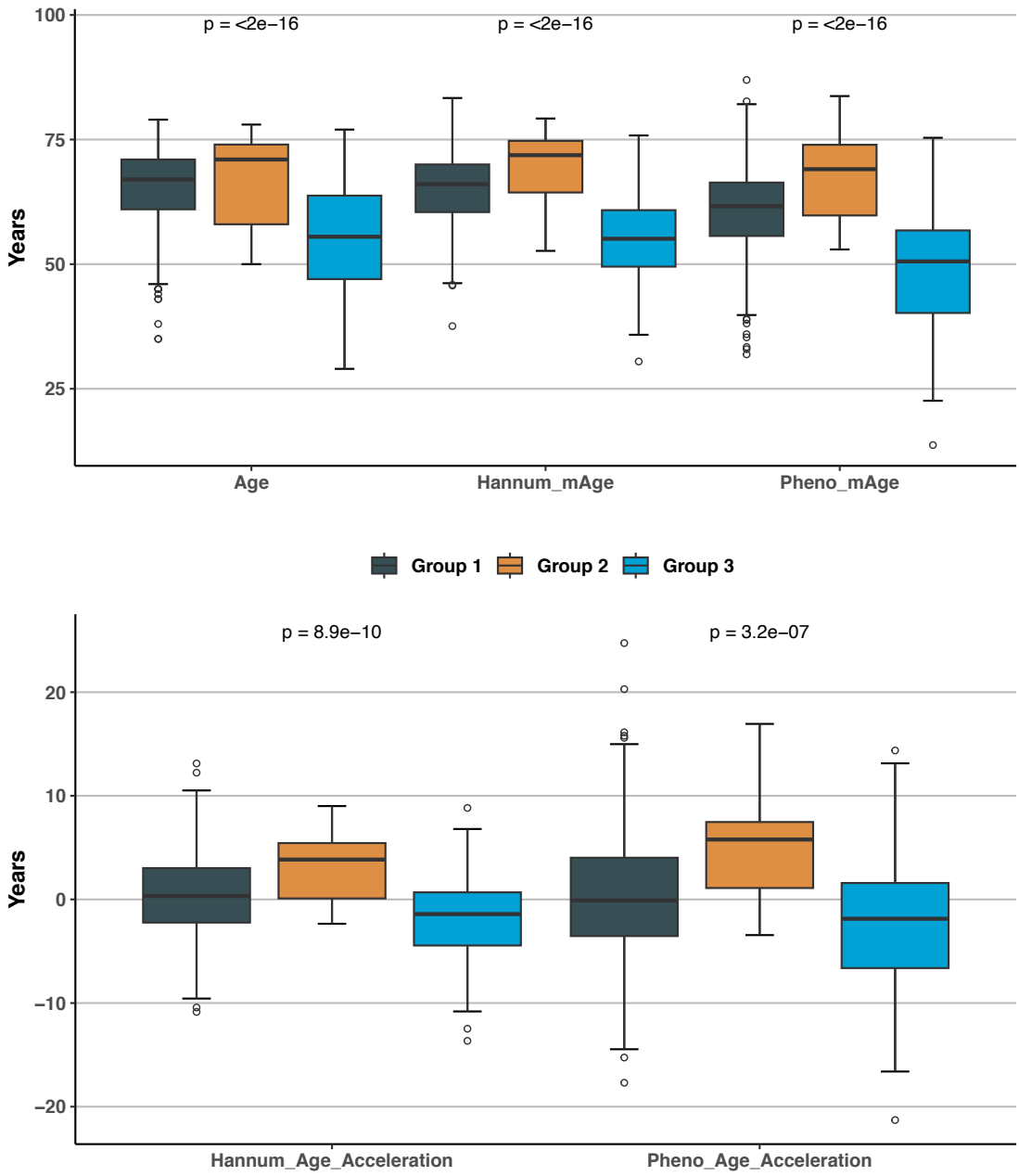


Figure S3.2: Blood immune cell profiles, age and age acceleration distribution of three groups assigned by *partDSA* algorithm in NMIBC patients (A) Blood immune cell profiles **(B)** age and age acceleration. Group 1 consisted of the 454 patients who had neutrophil cell proportions ≤ 76.46 and CD8 naïve cell proportions ≤ 1.76. Group 2 consisted of the 17 patients who had neutrophil cell proportions > 76.46. Group 3 consisted of the 130 patients who had neutrophil cell proportions ≤ 76.46 and CD8 naïve cell proportions > 1.76. Differences in cell-type proportions, age, methylation age, and age acceleration between three groups were evaluated using the ANOVA test.

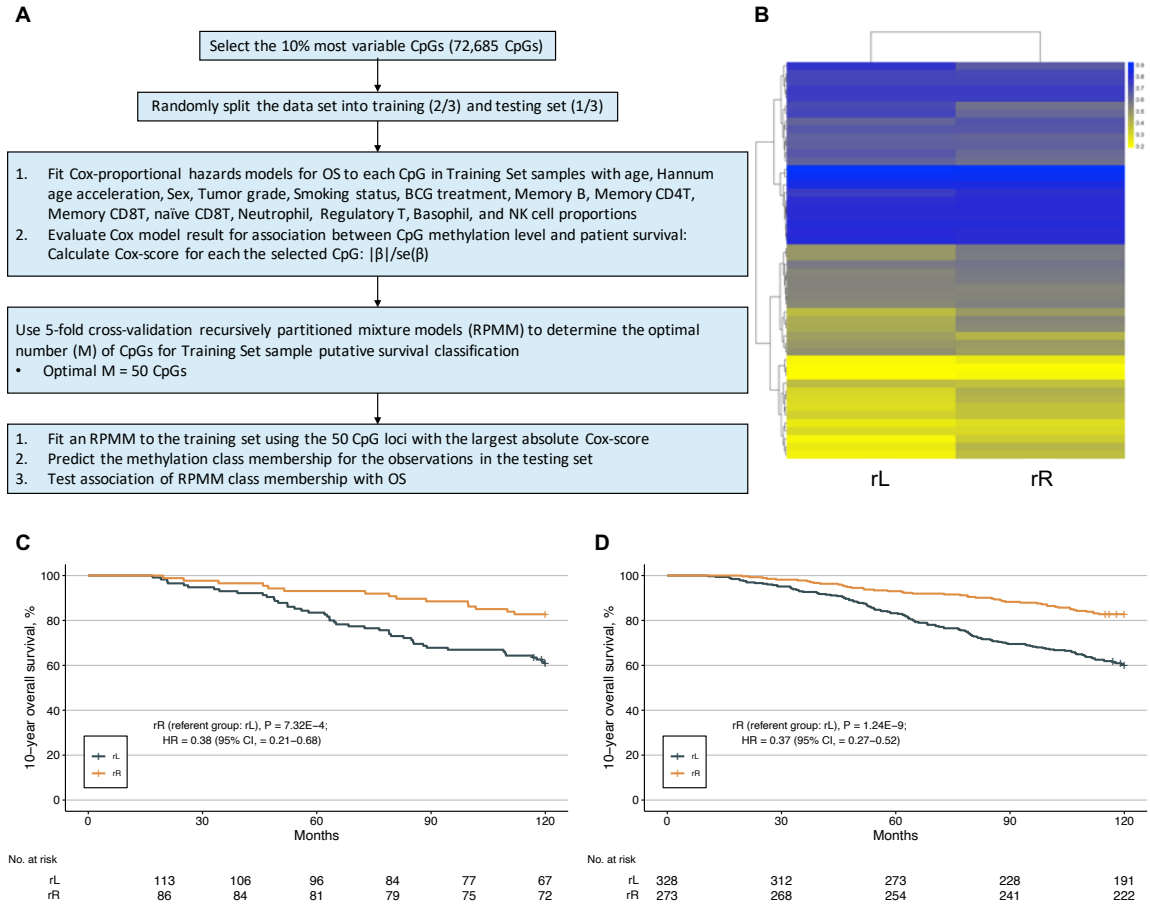


Figure S3.3: Semi-Supervised Recursively Partitioned Mixture Model (*SS-RPMM*) for 10-year overall survival (OS) in NMIBC patients [For Hannum age acceleration] **(A)** Data analysis schematic of SS-RPMM used for identification of blood DNA methylation profiles associated with non-muscle-invasive bladder cancer (NMIBC). **(B)** Heatmap of predicted class memberships for the observations in all NMIBC patients using the average beta values of the 50 CpG loci with the largest absolute Cox-scores. **(C)** Kaplan-Meier curves of 10-year overall survival stratified by the SS-RPMM classification of 202 NMIBC patients in the testing set by the 50 CpG loci. **(D)** Kaplan-Meier analysis of 10-year overall survival. 10-year overall survival curves stratified by the grouping result from SS-RPMM in all NMIBC patients. P-values for Log-rank tests are shown. All Kaplan-Meier curves are univariate analyses without adjusting for other variables. HR = hazard ratio; C.I. = confidence intervals; SS-RPMM = semi-supervised recursively partitioned mixture model

Table S3.5A: The information of optimal 15 CpGs selected by *SS-RPMM* using the model controlling for Pheno age acceleration

CpG	DHS	Enhancer	Relation_to_Island	chr	UCSC_RefGene_Name	UCSC_RefGene_Group	delta_beta
cg11187668	Yes	No	S_Shore	11			0.114
cg13077138	Yes	No	Island	15	C15orf2 (NPAP1)	1stExon	0.006
cg06864351	Yes	No	OpenSea	14			-0.015
cg16441900	Yes	No	OpenSea	17	NXN	Body	0.003
cg23510764	No	No	OpenSea	5			0.061
cg20688738	Yes	No	OpenSea	17			0.114
cg13074203	Yes	No	OpenSea	8	SNTB1	Body	0.115
cg14521240	No	No	OpenSea	14			-0.023
cg06991118	Yes	No	OpenSea	18	NDUFV2	Body	-0.050
cg20016914	Yes	No	S_Shore	14	C14orf181	TSS200	0.035
cg14768892	Yes	Yes	OpenSea	15	FSIP1	Body	-0.070
cg03872677	Yes	No	OpenSea	6			0.050
cg12086028	Yes	No	S_Shore	6	VPS52 (ARE1)	TSS1500	-0.024
cg25644567	Yes	No	OpenSea	2	SPRED2	1stExon	0.107
cg20187292	Yes	No	OpenSea	21			0.108

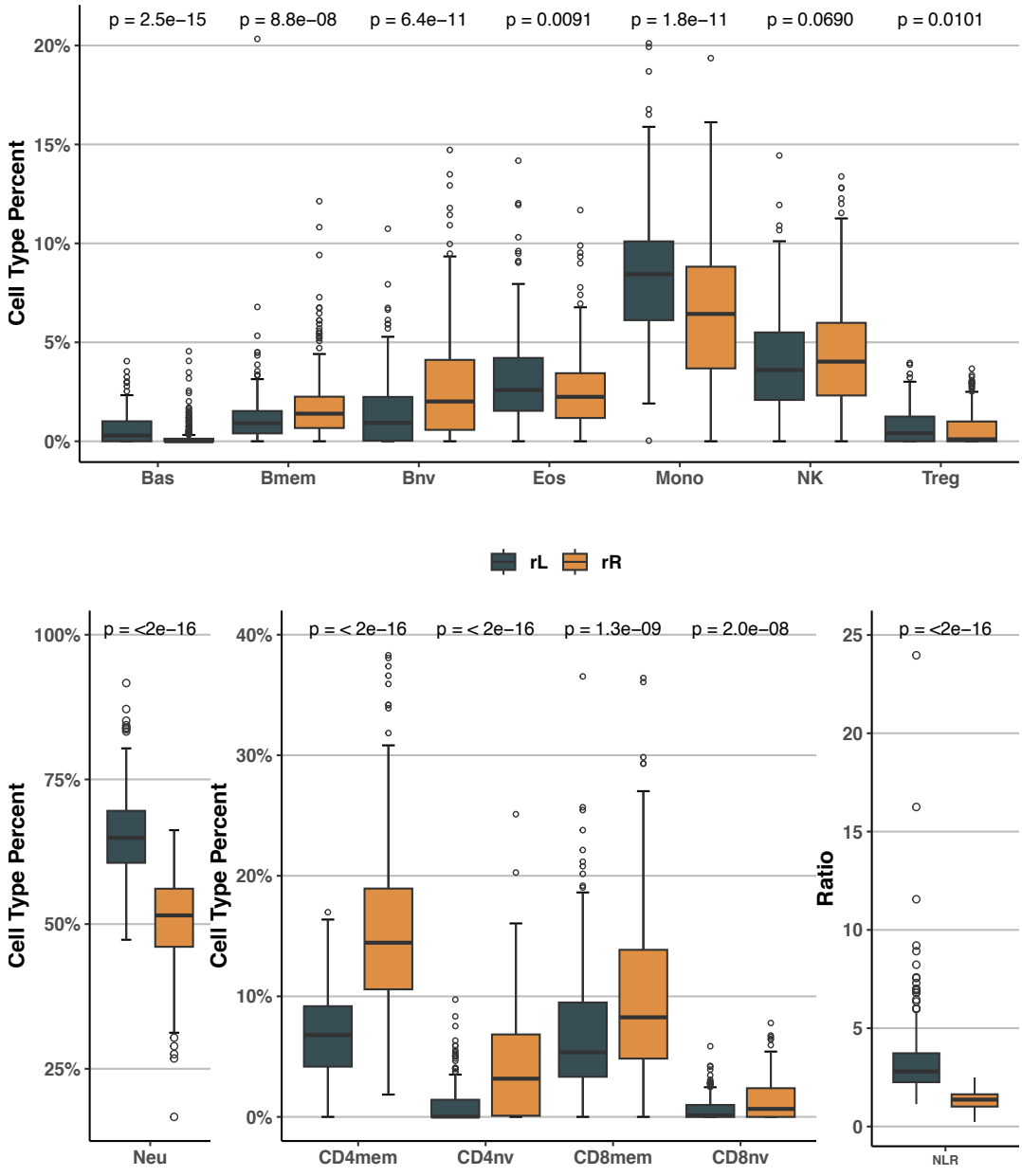
delta_beta: the difference of average beta values between rR and rL groups (subtract rL from rR)

Table S3.5B: The information of optimal 50 CpGs selected by *SS-RPMM* using the model controlling for Hannum age acceleration

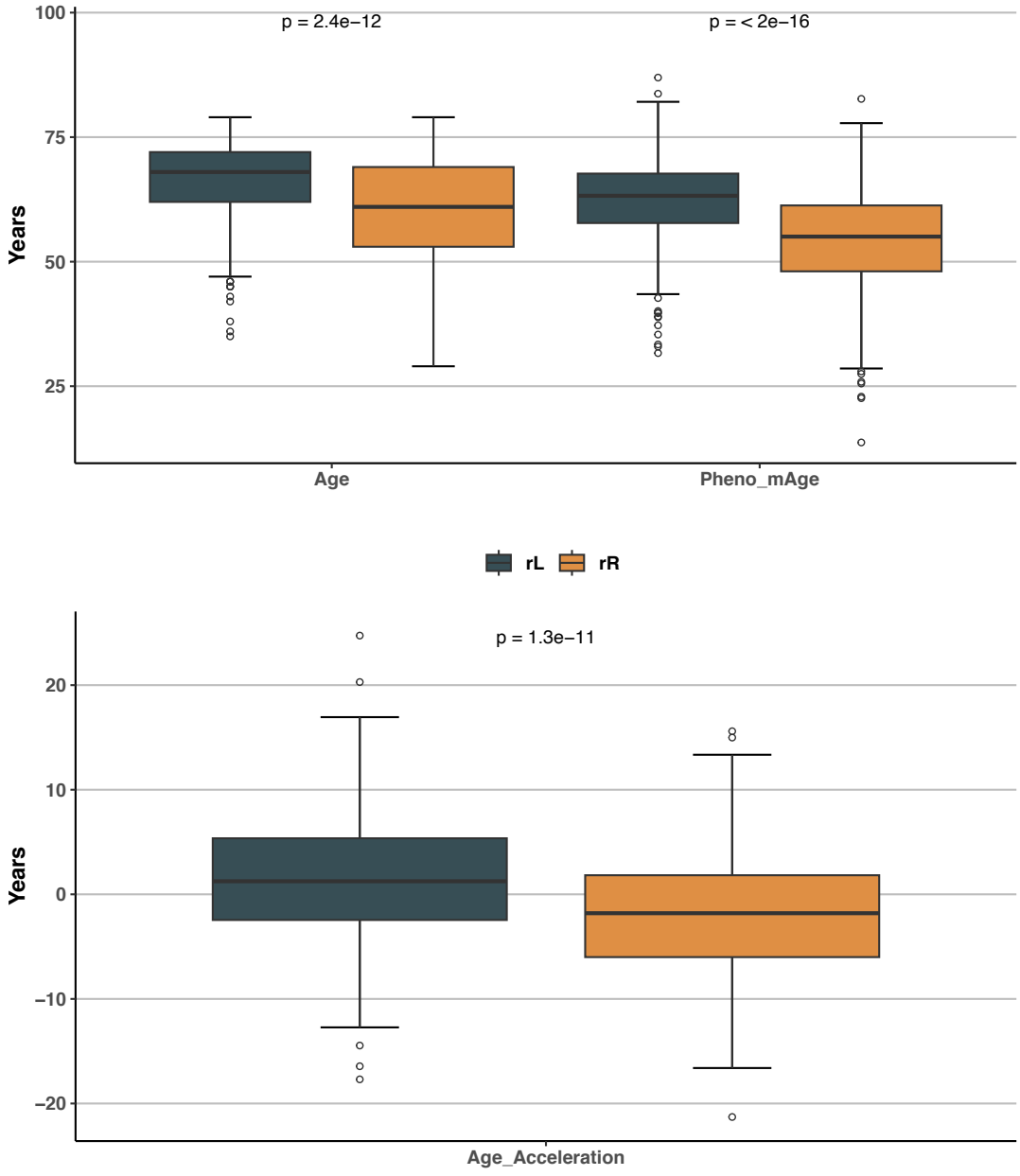
CpG	DHS	Enhancer	Relation_to_Island	chr	UCSC_RefGene_Name	UCSC_RefGene_Group	delta_beta
cg19584733	No	No	OpenSea	9	DENND4C	Body	-0.009
cg04036920	No	No	OpenSea	11	C11orf41	TSS1500	0.089
cg23614413	No	No	Island	1	TMEM53	Body	-0.008
cg02288976	Yes	No	N_Shelf	19	ARID3A	Body	-0.114
cg04524766	No	No	OpenSea	20			0.006
cg15898112	No	No	OpenSea	15			-0.008
cg06857887	Yes	No	N_Shelf	16	SEZ6L2	Body	-0.004
cg13077138	Yes	No	Island	15	C15orf2 (NPAP1)	1stExon	0.010
cg24460268	Yes	No	OpenSea	13	INTS6	Body	-0.098
cg23120101	Yes	No	OpenSea	5	MCC	Body	0.075
cg03862957	No	No	OpenSea	20			0.079
cg22048216	Yes	No	Island	18			-0.016
cg02946878	Yes	No	OpenSea	4	PPARGC1A	Body	0.047
cg22354988	Yes	No	OpenSea	15			0.004
cg16145324	Yes	No	OpenSea	6	MAPK14	Body	0.110
cg07330145	Yes	No	N_Shore	11	STARD10	3'UTR	-0.015
cg21242079	Yes	No	N_Shore	15	ADAMTS7	Body	-0.006
cg00816164	Yes	No	OpenSea	3			0.040
cg03417884	Yes	No	Island	18	GNAL	TSS200	1.416
cg01567825	No	No	Island	7	DOCK4	Body	0.002
cg11490805	Yes	No	OpenSea	5			-0.009
cg05796652	Yes	No	N_Shore	11	PSMA1	Body	-0.043
cg09430957	Yes	No	S_Shelf	19			-0.009
cg07744547	Yes	No	N_Shore	8	MYOM2	Body	0.042
cg05527920	Yes	No	OpenSea	4			0.033
cg19904425	No	No	OpenSea	14	SERPINA12	TSS1500	0.006
cg11187668	Yes	No	S_Shore	11			0.112
cg20016914	Yes	No	S_Shore	14	C14orf181	TSS200	0.033
cg09555736	Yes	No	N_Shore	16	RAB11FIP3	TSS1500	0.083
cg08235551	Yes	No	Island	22			0.002
cg03168497	Yes	No	Island	17	MYCBPAP	Body	0.022
cg15205435	Yes	No	Island	1	CHD5	Body	0.038
cg10963061	Yes	Yes	Island	1			0.007
cg07143050	No	No	N_Shelf	16			0.003
cg04232850	Yes	No	OpenSea	5	EMB	Body	0.146
cg10598596	No	No	N_Shelf	19	KLK10	3'UTR	-0.012
cg16801601	Yes	Yes	OpenSea	10	PRKCQ	5'UTR	-0.065
cg24011500	Yes	No	OpenSea	11	VWCE	Body	0.002
cg27058077	Yes	No	OpenSea	1			0.096
cg25433586	No	No	OpenSea	8			-0.108
cg07838270	Yes	No	OpenSea	16			0.015
cg26391794	Yes	No	OpenSea	13			-0.052
cg16441900	Yes	No	OpenSea	17	NXN	Body	0.004
cg18766861	Yes	No	OpenSea	2	PLCL1	Body	0.131
cg03354717	No	No	N_Shore	16	LONP2	TSS1500	0.020
cg05373255	Yes	No	OpenSea	10	CHAT	Body	-0.040
cg24401262	Yes	Yes	OpenSea	1			0.147
cg25644567	Yes	No	OpenSea	2	SPRED2	1stExon	0.102
cg22382021	No	No	Island	8	MGC70857	TSS1500	0.026
cg12414181	Yes	No	Island	15	SCAMP5	TSS200	0.005

delta_beta: the difference of average beta values between rR and rL groups (subtract rL from rR)

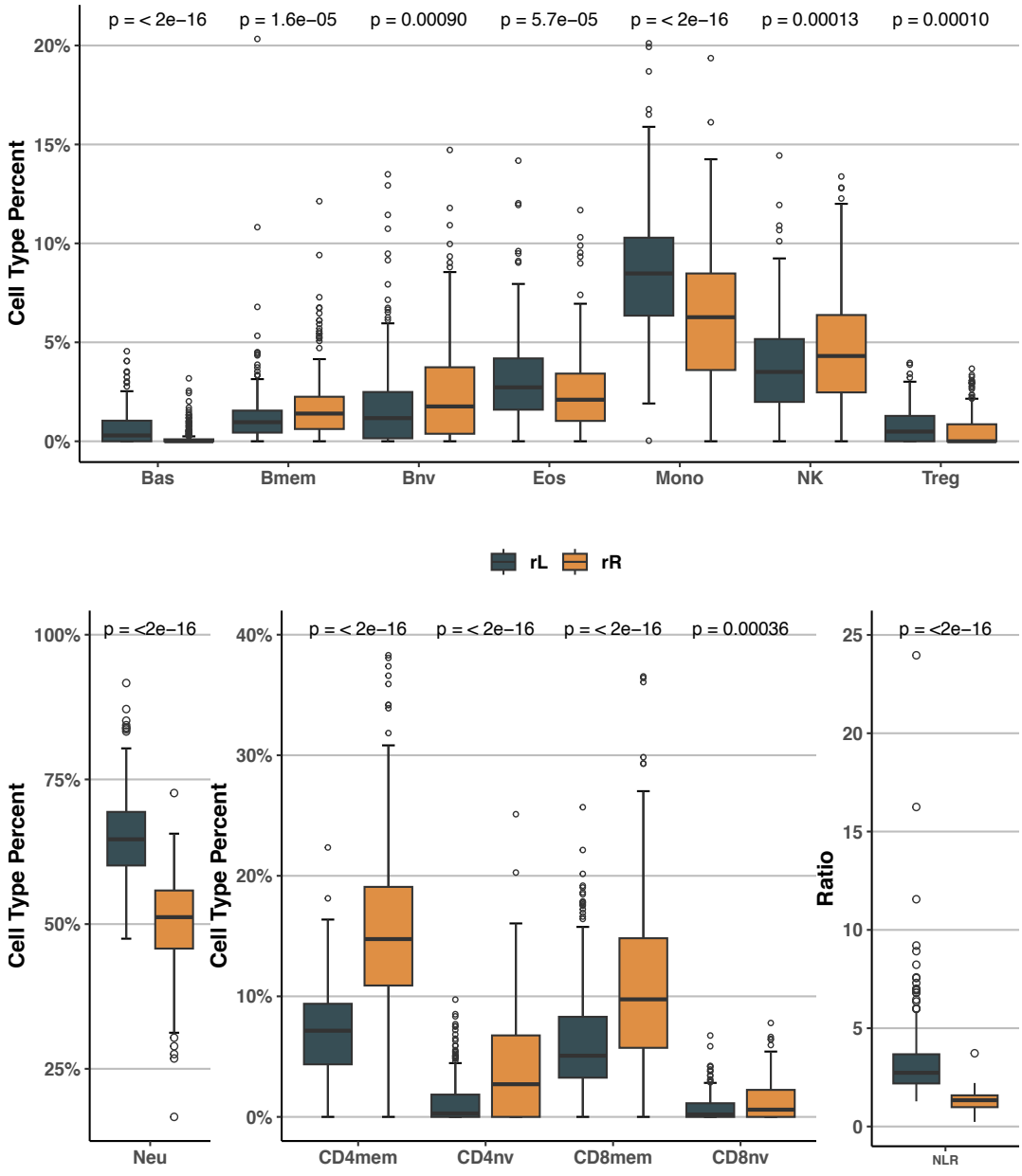
A


p = 2.5e−15
p = 8.8e−08
p = 6.4e−11
p = 0.0091
p = 1.8e−11
p = 0.0690
p = 0.0101
Cell Type Percent
20%
15%
10%
5%
0%
Bas
Bmem
Bnv
Eos
Mono
NK
Treg
rL
rR
p = <2e−16
p = < 2e−16
p = < 2e−16
p = 1.3e−09
p = 2.0e−08
p = <2e−16
100%
75%
50%
25%
40%
30%
20%
10%
0%
Neu
CD4mem
CD4nv
CD8mem
CD8nv
Ratio
25
20
15
10
5
0
NLR

B


p = 2.4e−12
p = < 2e−16
100
75
50
25
Years
Age
Pheno_mAge
rL
rR
p = 1.3e−11
20
10
0
−10
−20
Years
Age_Acceleration

C


p = < 2e−16
p = 1.6e−05
p = 0.00090
p = 5.7e−05
p = < 2e−16
p = 0.00013
p = 0.00010
Cell Type Percent
20%
15%
10%
5%
0%
Bas
Bmem
Bnv
Eos
Mono
NK
Treg
rL
rR
p = <2e−16
100%
75%
50%
25%
Neu
p = < 2e−16
p = < 2e−16
p = < 2e−16
p = 0.00036
40%
30%
20%
10%
0%
CD4mem
CD4nv
CD8mem
CD8nv
p = <2e−16
Ratio
25
20
15
10
5
0
NLR

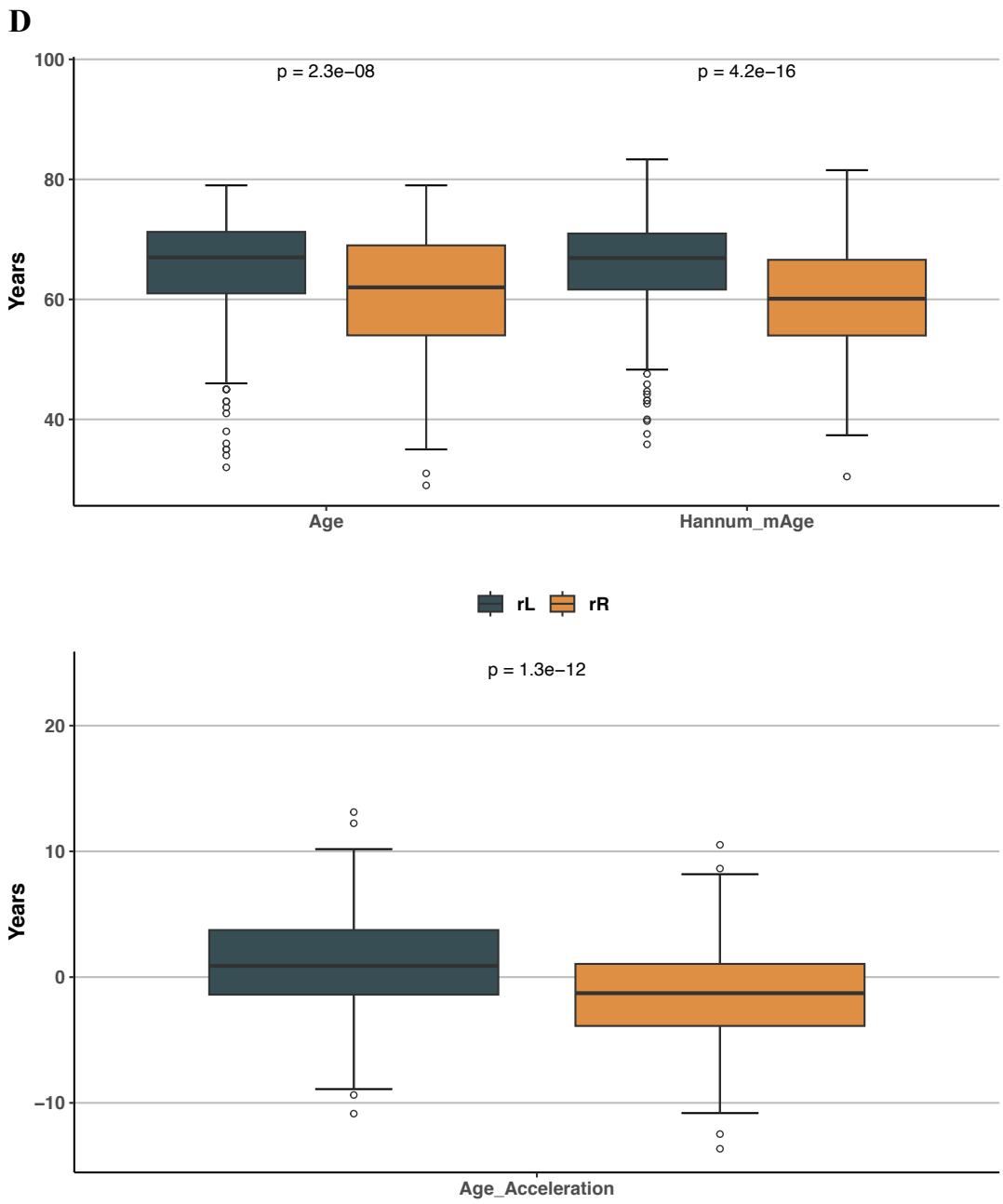


Figure S3.4: Blood immune cell profiles, age and age acceleration distribution of two clusters assigned by the SS-RPMM approach in NMIBC patients (A) Blood immune cell profiles **(B)** age and age acceleration distribution for the SS-RPMM adjusting for Pheno age acceleration. Differences in cell-type proportions, age, Pheno age, and age acceleration between two groups were evaluated using the Wilcoxon rank sum test. **(C)** Blood immune cell profiles **(D)** age and age acceleration distribution for the SS-RPMM adjusting for Hannum age acceleration. Differences in cell-type proportions, age, Hannum age, and age acceleration between two groups were evaluated using the Wilcoxon rank sum test.

Table S3.6: Characteristics of subjects of each group based on the grouping results from both *partDSA* and *SS-RPMM* in all NMIBC patients.

	Group1_rL (n = 262)	Group1_rR (n = 192)	Group2 (n = 17)	Group3_rL (n = 34)	Group3_rR (n = 96)
Age					
Median (IQR)	68 (63, 72)	65 (57, 70)	71 (58, 74)	61 (49, 68)	55 (47, 61)
Pheno Age Acceleration					
Median (IQR)	1.16 (-2.49, 5.24)	-1.48 (-5.22, 2.33)	5.78 (1.10, 7.47)	-0.35 (-2.91, 3.05)	-2.18 (-7.74, 0.91)
Hannum Age Acceleration					
Median (IQR)	0.88 (-1.27, 3.32)	-0.83 (-3.15, 2.27)	3.85 (0.10, 5.44)	0.75 (-1.36, 2.20)	-2.44 (-5.11, -0.39)
Sex					
Male	211 (80.5%)	144 (75.0%)	15 (88.2%)	25 (73.5%)	60 (62.5%)
Female	51 (19.5%)	48 (25.0%)	2 (11.8%)	9 (26.5%)	36 (37.5%)
Tumor grade					
Low Grade	190 (72.6%)	146 (76.0%)	10 (58.8%)	25 (73.5%)	79 (82.3%)
High Grade	72 (27.4%)	46 (24.0%)	7 (41.2%)	9 (26.5%)	17 (17.7%)
Smoking status					
Never	40 (15.3%)	38 (19.8%)	4 (23.5%)	5 (14.7%)	16 (16.7%)
Former	152 (58.0%)	93 (48.4%)	8 (47.1%)	17 (50.0%)	36 (37.5%)
Current	70 (26.7%)	61 (31.8%)	5 (29.4%)	12 (35.3%)	44 (45.8%)
BCG: Immunotherapy					
No	217 (82.8%)	164 (85.4%)	15 (88.2%)	31 (91.2%)	85 (88.5%)
Yes	45 (17.2%)	28 (14.6%)	2 (11.8%)	3 (8.8%)	11 (11.5%)
NLR					
Median (IQR)	2.73 (2.19, 3.57)	1.40 (1.06, 1.67)	7.30 (6.01, 8.92)	2.78 (2.50, 3.49)	1.31 (0.97, 1.57)
10-year survival status					
Alive	156 (59.5%)	156 (81.3%)	4 (23.5%)	24 (70.6%)	83 (86.5%)
Deceased	106 (40.5%)	36 (18.7%)	13 (76.5%)	10 (29.4%)	13 (13.5%)
10-year Recurrence-free status[a]					
No	83 (31.7%)	86 (44.8%)	3 (17.6%)	15 (44.1%)	43 (44.8%)
Yes	179 (68.3%)	106 (55.2%)	14 (82.4%)	19 (55.9%)	53 (55.2%)

[a]Whether patient with tumor recurrence or deceased within 10 years.

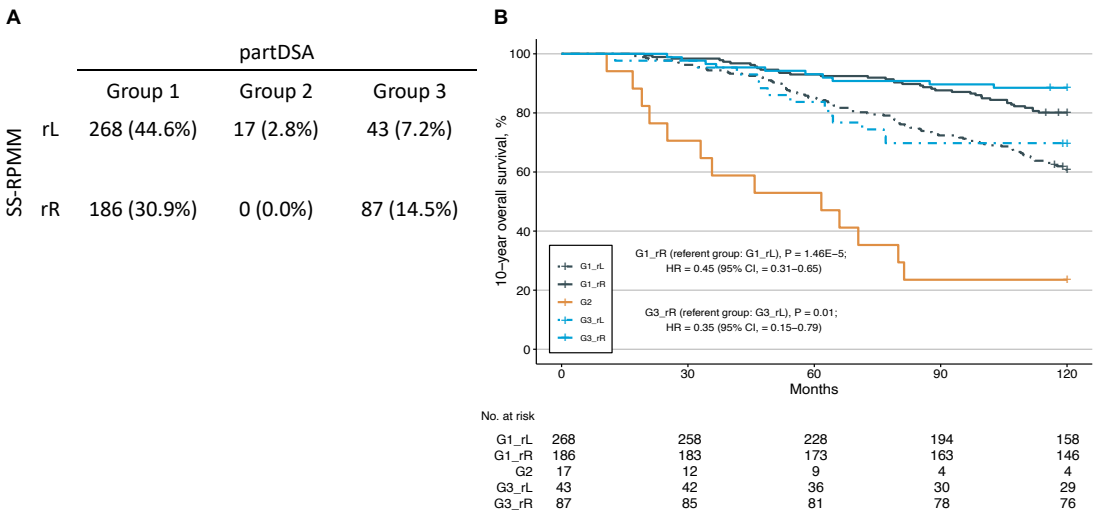


Figure S3.5: Kaplan-Meier analysis of 10-year overall survival based on the grouping results from both partDSA and SS-RPMM in all NMIBC patients [For Hannum age acceleration] (A) contingency table based on the grouping results from both partDSA and SS-RPMM in all NMIBC patients. **(B)** 10-year overall survival curves of all 5 groups. P-values for Log-rank tests are shown. All Kaplan-Meier curves are univariate analyses without adjusting for other variables. HR = hazard ratio; C.I. = confidence intervals; partDSA = partitioning deletion/substitution/addition algorithm; SS-RPMM = semi-supervised recursively partitioned mixture model; NMIBC = non-muscle-invasive bladder cancer.

Chapter 4

Integrated investigation of bladder cancer immune profiles in the tumor microenvironment and peripheral blood

The following authors have contributed to the content in this chapter:

Chen, JQ, Salas, LA, Wiencke, JK, Koestler, DC, Molinaro, AM, Andrew, AS, Seigne, JD, Karagas, MR, Kelsey, KT, Christensen, BC.

Conceptualization: JQC, BCC, LAS, JKW, DCK, AMM, MRK, KTK; Formal Analysis: JQC, BCC, LAS; Data Curation: ASA, JDS; Methodology: JQC, BCC, LAS, JKW, DCK, AMM, MRK, KTK; Writing–Original Draft: JQC; Writing–Review and Editing: JQC, BCC, LAS, JKW, DCK, AMM, MRK, KTK; Funding Acquisition: BCC

4.1 Abstract

Bladder cancer development and the response to therapy are influenced by immune profiles in the tumor microenvironment (TME) and peripheral blood. Despite the significance of immune involvement, studies investigating associations of tumor-infiltrating immune cells with immune profiles in circulating blood are limited. The Illumina HumanMethylationEPIC array was utilized to assess DNA methylation levels in tumor tissues and matched blood

samples from bladder cancer patients. DNA methylation cytometry approaches were employed to quantify the proportions of thirteen immune cell types in the TME and twelve immune cell types in peripheral blood. These cell types included dendritic cells, naïve and memory subsets of T and B cells, as well as granulocyte subtypes. Tumors were grouped by immune hot and cold status as well as using consensus clustering with immune profile data as the input. Differences in immune composition between the identified groups were assessed through the permutational multivariate analysis of variance method. The Wilcoxon rank sum test was used to evaluate differences in individual immune cell-type proportions within the TME and blood across the identified groups. Compared with non-muscle-invasive bladder cancer patients, muscle-invasive bladder cancer patients had higher dendritic cell ($P=1.9x10^{-4}$), memory B ($P=9.3x10^{-6}$), natural killer ($P=7.4x10^{-5}$), and memory CD8T ($P=1.1x10^{-5}$) cell proportions in the TME. Patients with high anti-tumor immune infiltration – immune hot tumors – had higher dendritic cell ($P=1.0x10^{-9}$), eosinophil ($P=0.02$), memory CD4T ($P=0.002$), memory CD8T ($P=8.6x10^{-13}$), natural killer ($P=2.7x10^{-10}$), and memory B ($P=9.3x10^{-14}$) cell proportions in the TME compared with low anti-tumor immune infiltration group. However, there were no differences in circulating immune profiles between patients with immune hot and cold tumors. With consensus clustering based on tumor immune profiles, two clusters of patients were identified. Patients in Cluster 1 had higher dendritic cells ($P=6.4x10^{-12}$), memory CD8T (P=1.2x10-13), natural killer (P=3.9x10-11), and memory B (P=8.0x10-13) cell proportions and had lower basophil

(P=0.0047) and monocyte (P=0.011) cell proportions in the TME compared with Cluster 2. In addition, in peripheral blood samples, patients in Cluster 1 had higher basophil (P=0.043) but lower naïve CD4T (P=0.029) and memory CD4T (P=0.037) cell proportions. Our findings showed that methylation cytometry approaches offer detailed immune profiles, and that consensus clustering provided insights into the relationship of tumor and systemic immune status in bladder cancer patients.

4.2 Introduction

The immune system in the tumor microenvironment (TME) plays a critical role in bladder cancer development and treatment. Considerable evidence has indicated that bladder cancer cells are able to shape the microenvironment, resulting in immunosuppression beneficial for tumorigenesis. For example, due to the overexpression of sphingosine 1 phosphate receptor 1 in bladder cancer cells, TGF-β and IL-10-induced regulatory T-cell (Treg) expansion has been shown to lead to suppression of cytotoxic T cells in the tumor microenvironment [58], [59]. Numerous studies have indicated that myeloid-derived suppressor cells (MDSCs), which exhibit immunosuppressive properties, are recruited by bladder cancer cells with the secretion of CXC-chemokine ligand 2, and macrophage migration inhibitory factor [63]. Infiltrating MDSCs are able to express Arginase 1, inducible nitric oxide synthases, and PD-L1 to induce Treg activation and suppress the function of T, B, and natural killer cells [61], [216], [217]. Since bladder cancer cells can escape from immune surveillance with the

assistance of immunosuppressive cells, many treatments have been developed to target immunosuppressive cells and/or stimulate the activation of anti-tumor immune cells.

After transurethral resection of the bladder tumor, intravesical Bacillus Calmette-Guérin (BCG) and immune checkpoint inhibitors (ICI) are the most common treatments to control tumor recurrence. BCG treatment uses the attenuated form of the tuberculosis vaccine, which induces an inflammatory reaction and activates anti-tumor immune cells, such as natural killer and CD8T memory cells, within the TME [47], [218]. ICIs are used to block the interaction of PD-1 with PD-L1, aiming to restore the activity of effector T cells against bladder cancer cells [56]. However, < 30% of bladder cancer patients responded to immunotherapy [219]–[221]. One potential contribution to immunotherapy resistance is the immune landscape of the TME. For instance, patients with abundant T cell infiltration in the TME, typically called immune hot, respond better to ICIs compared with patients with little immune cell infiltration, also known as immune cold tumors [222]–[225]. Currently, there is no standard definition of immune hot and cold tumors for bladder cancers, and a more complete understanding of the immune composition of the TME is required to improve treatments for bladder cancer.

Only a few studies have investigated both the tumor and peripheral blood immune status in bladder cancer. Existing work to date has been limited to measures of tumor and peripheral blood neutrophils. For example, high neutrophil proportions in both peripheral blood and the TME have been

associated with immunosuppressive effects in bladder cancer [72]. Additional works investigating other immune cell types within the bladder cancer microenvironment and their association with circulating immune profiles are needed to improve the management of bladder cancer.

Due to the critical role in gene regulation for cell lineage specification [133], [134], DNA methylation profiles can be leveraged to estimate cell-type proportions [89], [90], [96]. The major advantages of DNA methylation cytometry are high accuracy and the ability to work in archival specimens [98]–[100]. Many studies have exploited DNA methylation cytometry in the cancer research field, including bladder cancer [100], [188], [226], [227]. Currently, publicly available DNA methylation data for tumor tissue is predominantly from muscle-invasive bladder cancers (MIBC) and thus information on non-muscle-invasive bladder cancers (NMIBC) lacks as matched data from blood samples are lacking. Here, we obtained tumor tissues and matched blood samples from MIBC and non-muscle-invasive bladder cancer patients through population-based studies [138]–[140]. Since our samples were archival specimens, we performed DNA methylation cytometry to identify the immune profiles of both compartments. Recently, our group developed methods that enable the deconvolution of DNA methylation data quantifying cell types in the tumor microenvironment [101] and immune cell types in peripheral blood [89]. Here, we identified and compared the distributions of immune landscapes in bladder cancer patient tumors and peripheral blood using DNA methylation.

4.3 Methods

4.3.1. Study Subjects and Samples

The study subjects have been described in detail in prior research [138]–[140]. In brief, we included a subset of bladder cancer patients recruited in three phases from a population-based case-control study in New Hampshire [141]. For our study, we included 331 subjects diagnosed between July 1994 and June 1998 from the phase 1 study, 243 subjects diagnosed between July 1998 and December 2001 from the phase 2 study, and 194 subjects diagnosed between July 2002 and December 2004 in the phase 3 study. Subjects who were from phases 1 and 2 were identified using the New Hampshire State Cancer Registry and hospital cancer registry, while subjects in phase 3 were only those included in the hospital cancer registry. Patients with both available blood and tumor samples (N = 88) were included in the present study. Among 88 subjects, the median time interval from the initial diagnosis to the blood draw was 207 days. All blood samples were stored at 4°C and then frozen within 24 hours of blood draw. Overall 11% (N = 10) of patients received treatment with Bacillus Calmette-Guérin (BCG), and blood samples were taken after treatment (median: 583 days). The resections were generally taken around the time of initial diagnosis, and BCG was administered after the tumor resection. This study was approved by the Dartmouth Human Research Protection Program (IRB).

Formalin-fixed, paraffin-embedded (FFPE) tumor specimens and pathology reports were requested from treating physicians/pathology laboratories with initial diagnoses. Bladder tumors were reviewed by a single pathologist and

classified based on the WHO classification criteria. DNA extraction was performed using a previously reported method [228]. In brief, 20-μm sections were cut from the paraffin-embedded tissue and placed into tubes containing a digestion buffer. The paraffin was removed with microwave treatment and centrifugation, and the resulting paraffin-free tissue pellets were further processed with proteinase K digestion in a separate digestion buffer. Subsequently, supernatants containing the DNA lysate were subjected to boiling to denature any residual protease activity. Study participants underwent an extensive in-person interview to gather comprehensive data on demographic characteristics and risk factors, including information on cigarette smoking history [138]–[141].

4.3.2. DNA Extraction, Qualification, and Bisulfite Modification

Blood and tumor DNA extraction was performed using the QIAamp DNA Blood Kit and the QIAamp DNA FFPE Tissue Kit (Qiagen, CA, USA), respectively, according to the manufacturer's protocol. We quantified and qualified the extracted DNA using Qubit 3.0 Fluorometer (Life Technologies, CA, USA) and Fragment Analyzer (Advanced Analytical, IA, USA). Extracted DNA samples were conducted for bisulfite modification with EZ DNA Methylation Kit (Zymo Research, CA, USA) according to the instructions of the manufacturer.

4.3.3. DNA Methylation Profiling

Bisulfite-modified DNA samples were subjected to the Infinium

MethylationEPIC Bead Chip array (Illumina, Inc., CA, USA), measured for their methylation intensity at > 850,000 CpG sites. Probe intensity data (iDAT) files were processed via *preprocessNoob* normalization method from the R package *minfi* [142]. *ENmix* [143], an R package, was performed for the quality control of probes. For sample selection, we set that samples with more than 5% of probes with a detection P value larger than 1.0×10^{-6} would be excluded to distinguish from background noise; fortunately, all samples had good quality and were kept in the study. Probes were checked by observing any probes that were not missing in more than 10% of the samples. *BMIQ* [200] from the *watermelon* [201] R package was performed for probe-type normalization, and batch effects were corrected using *ComBat* [144]. Then, 119,258 probes previously described to be potentially cross-hybridizing, sex-specific, non-CpG (CpH) methylation, and SNP-associated were filtered [145]. In total, 746,980 were included in the final analysis (**Figure 1A**). The CpG loci were annotated with *IlluminaHumanMethylationEPICanno.ilm10b4.hg19* [146] R package.

4.3.4. Statistical Analysis

Cell-type proportions in the tumor microenvironment (TME) were inferred using hierarchical tumor immune microenvironment epigenetic deconvolution (*HiTIMED*; https://github.com/SalasLab/HiTIMED) [101], and circulating cell-type proportions of 12 cell types were estimated through the *projectCellType_CP* from the *FlowSorted.Blood.EPIC* [89] R package.

To integrate the information of cell-type proportions from tumor and blood,

we grouped subjects using two methods. The first grouping method is according to the concept of immune hot and immune cold tumors. Since there is no standard definition of immune hot and immune cold tumors in bladder cancers, we grouped subjects based on the total percentage of anti-tumor immune cell-type proportions in the TME. If the sum of dendritic cell, natural killer, and naïve and memory subsets of CD4T, CD8T, and B cell proportions was > 5%, patients were assigned to the high immune infiltration (immune hot) group; otherwise they were assigned to the low immune infiltration (immune cold) group.

To also take an agnostic approach to grouping tumors by immune infiltration status, we performed consensus clustering [229] of standardized immune cell-type proportions in TME as input using the *ConsensusClusterPlus* [230] R package with the setting of the following parameters and standardized immune cell-type proportions in TME as input. The clustering algorithm was hierarchical clustering using a distance of 1 – Spearman correlation values. The number of repetitions of subsampling and clustering was 1,000. To avoid overfitting, 80% of subjects were randomly selected for each repetition. To test if the immune composition in peripheral blood was different between groups, we performed permutational multivariate analysis of variance (PERMANOVA) using *PERMANOVA* [231] R package, and the input was Aitchison distance matrix calculated by *robCompositions* [232] R package. Differences in cell-type proportions between two groups generated via two methods (immune hot-cold and consensus clustering) were evaluated using the Wilcoxon rank sum test. Statistical comparisons between two categorical variables were performed using

Fisher's exact test (two-tailed).

4.3.5. Data Availability

Datasets generated and analyzed in the present study are available in the Gene Expression Omnibus (GEO) repository at GSE237100.

4.4 Results

4.4.1. Characteristics of subjects

The study group consisted of 75 non-muscle-invasive bladder cancer (NMIBC) patients and 13 MIBC patients. The median age was 67 years, 65 (74%) were men, 59 (67%) had tumors in low grade, 46 (52%) and 26 (30%) were former and current smokers respectively, and 10 (11%) patients received BCG treatment (**Table 4.1**). DNA methylation profiles obtained from the tumor microenvironment (TME) of 88 patients were used to estimate cell-type proportions in TME through *HiTIMED*, and the distribution of cell proportions was displayed in **Figure 4.1B**. The distribution of the matched methylation-derived immune cell profiles in the blood is shown in **Figure 4.1C**.

Table 4.1: Characteristics of subjects

	NMIBC (n = 75)	MIBC (n = 13)	All (n = 88)
Age			
Median (Q1, Q3)	67 (56,71)	69 (64,76)	67 (57,72)
Sex			
Male	57 (76%)	8 (62%)	65 (74%)
Female	18 (24%)	5 (38%)	23 (26%)
Tumor grade			
Low Grade	56 (75%)	3 (23%)	59 (67%)
High Grade	19 (25%)	10 (77%)	29 (33%)
Tumor stage			
Stage 0a	61 (81%)	0 (0%)	61 (69%)
Stage I	14 (19%)	0 (0%)	14 (16%)
Stage II	0 (0%)	8 (62%)	8 (9%)
Stage III	0 (0%)	2 (15%)	2 (2%)
Stage IV	0 (0%)	3 (23%)	3 (4%)
Smoking status			
Never	12 (16%)	2 (15%)	14 (16%)
Former	40 (53%)	6 (46%)	46 (52%)
Current	21 (28%)	5 (39%)	26 (30%)
Missing	2 (3%)	0 (0%)	2 (2%)
BCG: Immunotherapy			
No	65 (87%)	13 (100%)	78 (89%)
Yes	10 (13%)	0 (0%)	10 (11%)

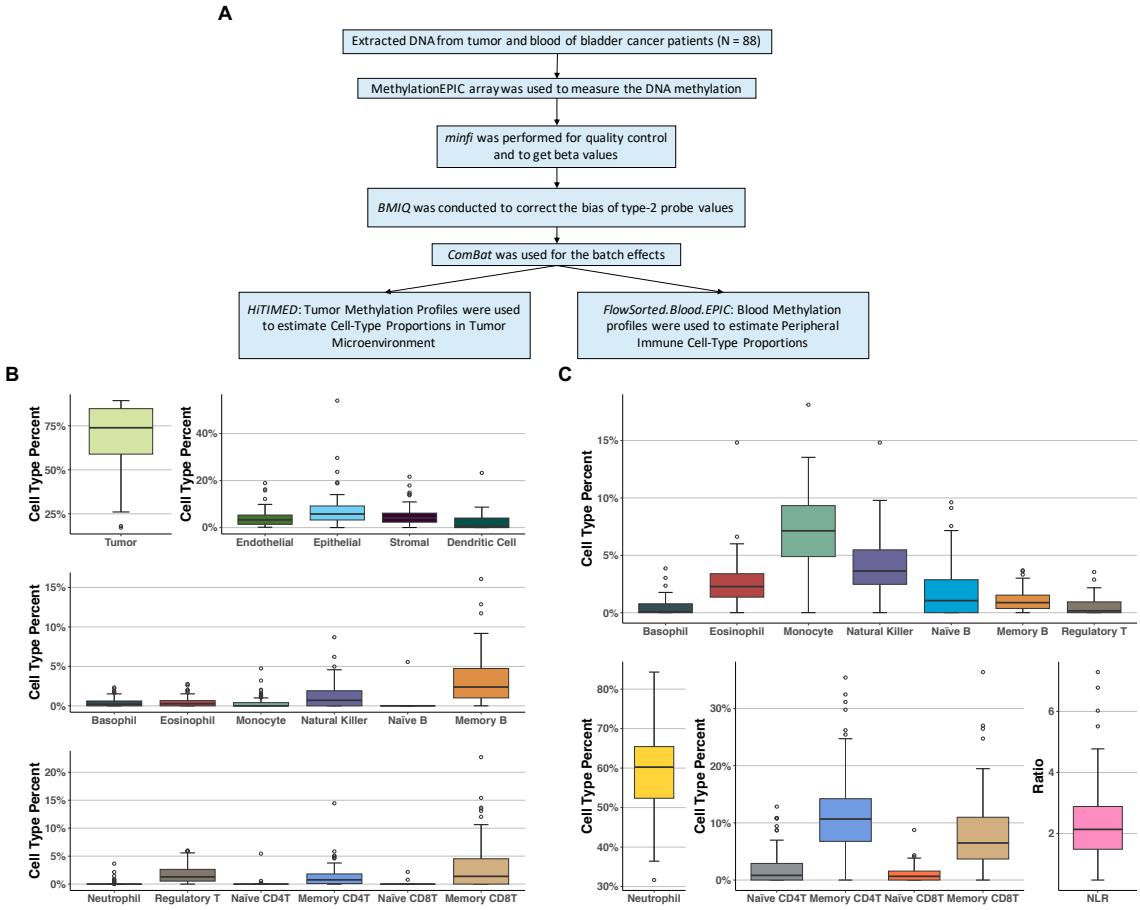


Figure 4.1: Data processing and cell distribution in tumor microenvironment and peripheral blood (A) Data processing schematic; DNA was extracted from tumor tissues and matched blood samples of 88 bladder cancer patients. After serial pre-processing steps, DNA methylation profiles were applied to estimate cell-type proportions in **(B)** tumor tissues and **(C)** blood using *HiTIMED* and *FlowSorted.Blood.EPIC* methods respectively.

4.4.2. Differences in tumor cell distribution between non-muscle-invasive and muscle-invasive bladder cancer patients

We first investigated the association of tumor muscle invasive status with the distribution of cell-type proportions in both the tumor microenvironment (TME) and peripheral blood. Within TME, we observed that NMIBC patients had significantly lower memory B, natural killer, neutrophil, memory CD8T,

endothelial, stromal, and dendritic cell proportions, and a higher epithelial cell proportion compared with MIBC patients (**Figure 4.2**); however, no significant differences in circulating immune profiles was observed between NMIBC and MIBC patients (**Figure S4.1**).

4.4.3. The impact of tumor immune infiltration on peripheral blood immune cell distribution

In recent years, several studies have reported that cellular immune composition within a tumor plays a critical role in tumor development and the response to immunotherapy, generating the concept of immune hot and cold tumors [224], [225], [233], [234]. Since tumor-infiltrated immune cells translocate to tumor sites through the bloodstream, we assessed the association between immune cell-type proportions in TME and blood. Due to the lack of a standard definition of immune hot and cold tumor for bladder cancer, we clustered subjects based on the infiltrated anti-tumor immune cell proportions (the sum of dendritic cells, natural killer, and naïve and memory subsets of lymphocytes). Anti-tumor immune cell proportion of < 5% in the TME was defined as low immune infiltration (immune cold, N = 38), and tumors with infiltrating immune cell proportions >5% were in the high immune infiltration group (immune hot, N = 50).

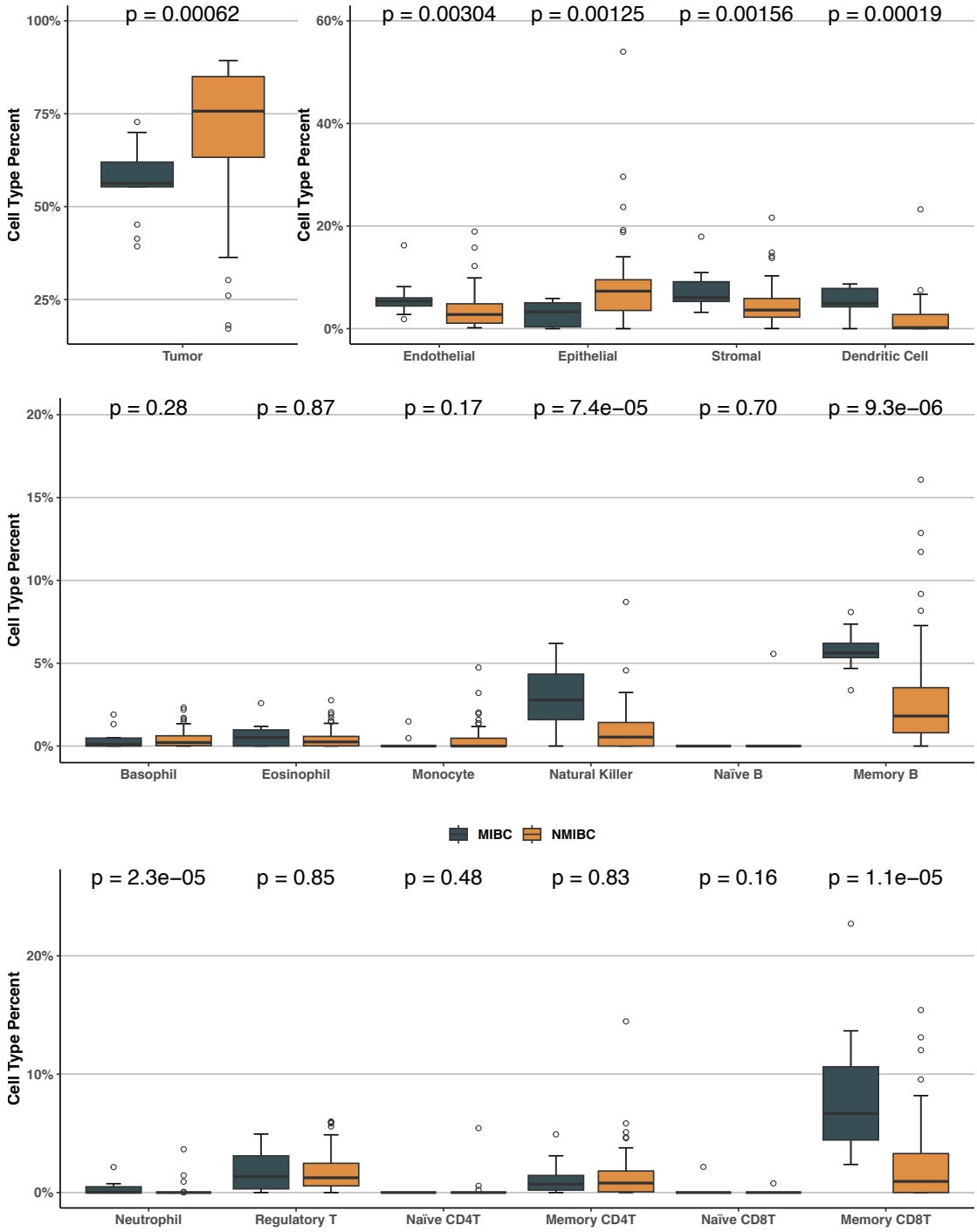


Figure 4.2: Cell profiles of tumor microenvironment in NMIBC and MIBC patients Differences in cell-type proportions between NMIBC and MIBC patients were evaluated using the Wilcoxon rank sum test.

While our findings showed patients in the high immune infiltration group had higher eosinophil, memory CD4T, memory CD8T, natural killer, memory B,

and dendritic cell proportions in TME compared with the low immune infiltration group, we did not observe any statistical differences in the immune composition in peripheral blood (P = 0.13; PERMANOVA test) and individual cell types between immune hot and immune cold patient groups (**Figure 4.3**).

We conducted Fisher's exact test to determine if there are associations between tumor infiltration and tumor characteristics. We observed that high anti-tumor immune infiltration was associated with tumor grade, stage, and muscle-invasiveness, but not with BCG treatment (**Table S4.1**).

4.4.4. Comparing tumor and blood immune cell distributions using consensus clustering

We next used an unbiased approach for grouping patients based on tumor immune infiltration with consensus clustering which resulted in two groups (Group 1, n = 43; Group 2, n = 45). Among non-immune cell types, Group 1 patients had higher endothelial and stromal cell proportions and lower epithelial cell proportion in TME compared with Group 2 patients. Additionally, Group 1 patients had significantly higher memory CD8T, natural killer, memory B, and dendritic cell proportions and lower basophil and monocyte proportions in TME (**Figure 4.4**).

Unlike the more basic approach using anti-tumor immune cell percentage in TME, with groupings from consensus clustering based on tumor immune cell proportions, we observed statistical differences in the immune composition in peripheral blood (P = 0.01; PERMANOVA test). In addition, individual

circulating immune profiles had significant differences between Groups including basophil, naïve CD4T, and memory CD4T cell types (**Figure 4.4**). Interestingly, some peripheral blood cell proportions with significant differences between groups had the opposite direction of cell proportion differences observed in the tumor. For example, patients in Group 1 had a lower basophil proportion in TME, but a higher basophil proportion in blood compared with Group 2. Despite not reaching statistical significance, patients with a higher monocyte proportion in TME (Group 2) had a lower monocyte proportion in their blood (**Figure 4.4**). In addition, we found the grouping result was associated with tumor grade, stage, muscle-invasive status, and BCG treatment (**Table S4.2**). Interestingly, the findings showed that Cluster 1 patients demonstrated significantly increased odds of having an advanced tumor stage (Stage I – IV vs 0a; odds ratio = 17.68; 95% CI = 4.74 – 66.00) and substantially elevated odds of having a high tumor grade (odds ratio = 14.24; 95% CI = 4.32 – 46.90) compared to patients in Cluster 2.

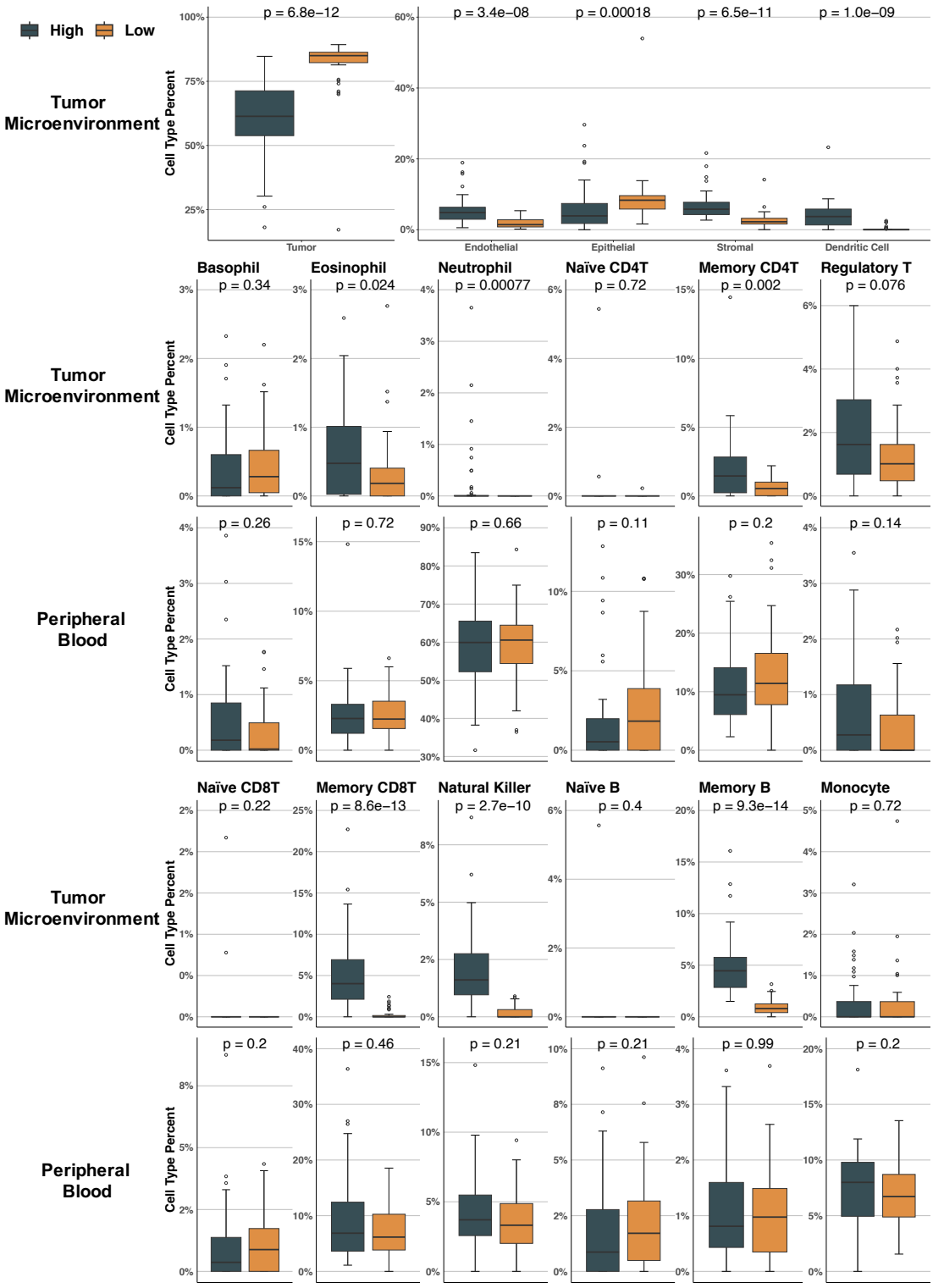


Figure 4.3: Tumor and blood cell profiles distribution of two groups assigned by the proportion of anti-tumor immune infiltration High immune infiltration group (High) consisted of the 50 patients who had the sum of B, CD8T, CD4T, natural killer, and dendritic cell proportions > 5% in tumor microenvironment. Low immune infiltration group (Low) consisted of the 38 patients who had the sum of B, CD8T, CD4T, natural killer, and dendritic cell proportions ≤ 5% in tumor microenvironment. Differences in cell-type proportions between two groups were evaluated using the Wilcoxon rank sum test.

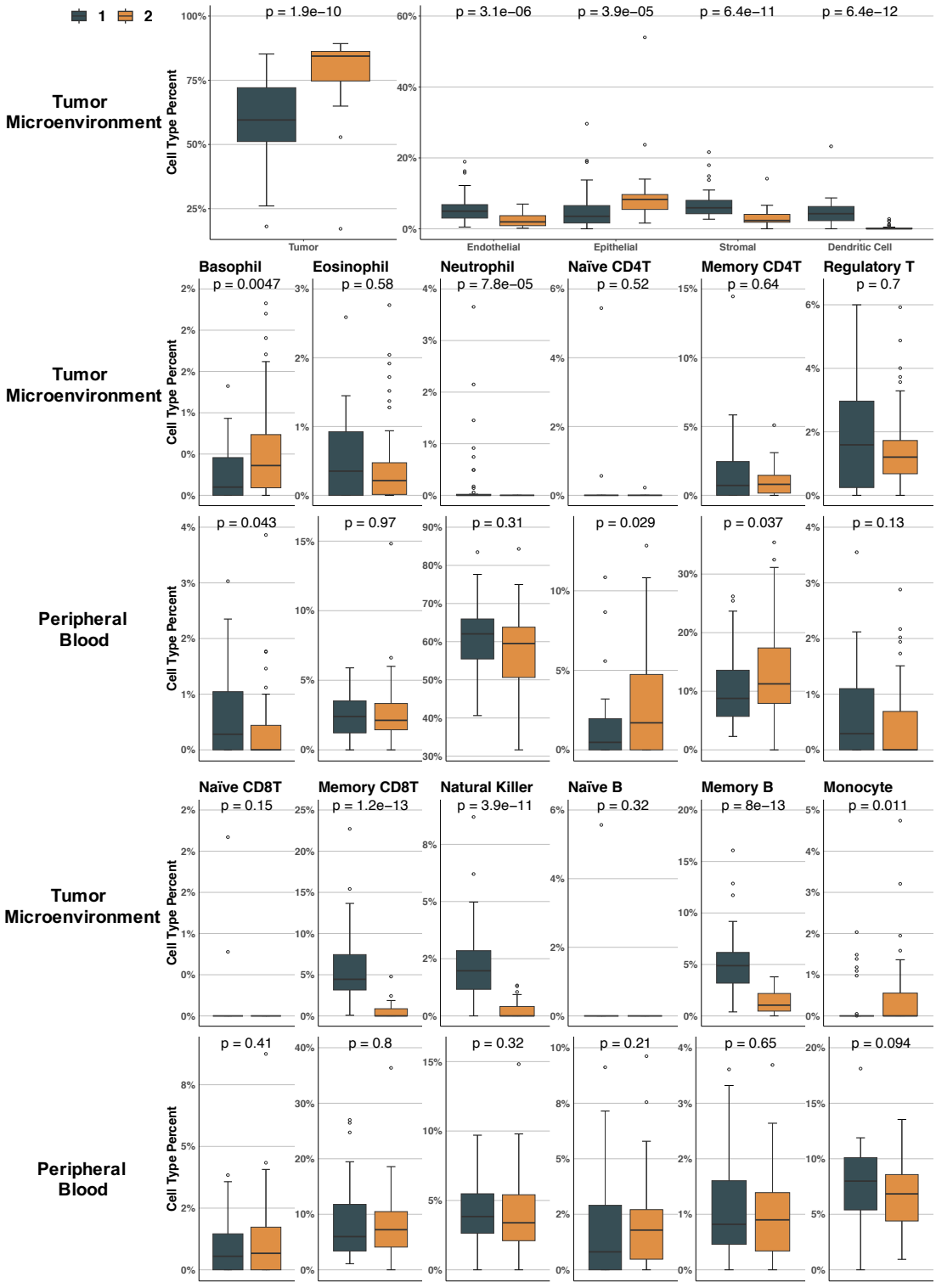


Figure 4.4: Tumor and blood cell profiles distribution of two groups assigned by consensus clustering algorithm using tumor immune profiles as input Group 1 and Group 2 consisted of 43 and 45 patients respectively. Differences in cell-type proportions between two groups were evaluated using the Wilcoxon rank sum test.

4.5 Discussion

The objective of this study was to investigate the distribution of immune cells within the tumor microenvironment (TME) of bladder cancer and explore its potential association with circulating immune profiles. Initially, we employed the *HiTIMED* [101] approach to estimate the cell distribution in the TME. Our findings revealed a scarcity of naïve lymphocytes (approximately 0%) in tumor tissues, aligning with their rarity in non-lymphoid tissues except peripheral blood [235]–[238]. Furthermore, we utilized the *FlowSorted.Blood.EPIC* [89] method to infer immune profiles within matched blood samples. Unsurprisingly, the mean neutrophil-to-lymphocyte ratio (NLR), a prognostic marker in various cancer types, was found to be 2.3 (Q1-Q3: 1.5-2.9). This value slightly exceeded the normal range of 1-2, as NLR values above 2.3 in adults serve as an early indicator of pathological conditions and cancer development [239], [240]. In past studies, peripheral blood NLR has been demonstrated to be associated with an increased risk of death and tumor recurrence in bladder cancer patients [154], [188], [241].

To address a knowledge gap regarding the association of immune profiles in bladder cancer between two compartments, we conducted several comparisons using different clustering approaches. While no significant differences were observed in circulating immune profiles between patients with MIBC and NMIBC, higher proportions of natural killer cells, neutrophils, memory B cells, memory CD8T cells, and dendritic cells were found in the TME of MIBC compared to NMIBC patients. These findings were consistent with a

previous study that reported significantly elevated levels of CD3 and CD8 tumor-infiltrating lymphocytes in MIBC compared to NMIBC patients [242].

In order to investigate the association of immune distribution between two compartments, we initially adopted the concept of immune hot and cold tumors to categorize the groups, although no standardized definitions were available in previous studies. Significant variations in immune profiles within the TME were observed, indicating distinctive immune characteristics among the categorized groups. In contrast, no significant differences were detected in immune profiles within the peripheral blood between the groups, suggesting that the conventional classification of immune hot and cold tumors might be inadequate in capturing the intricate relationship of immune profiles between the two compartments.

Interestingly, when applying the consensus clustering approach to categorize subjects based on more detailed immune infiltration profiling data, significant differences in immune profiles emerged not only within the TME but also in peripheral blood across the clustering groups. Of particular note, we observed an inverse relationship between basophil proportions in the TME and peripheral blood (**Figure 4**). Previous research exploring the association of basophils with bladder cancer demonstrated that high basophil counts in peripheral blood were linked to an increased hazard of recurrence in high-grade T1 bladder cancer patients undergoing BCG treatment [207]. In our study, patients with a high circulating basophil proportion (Cluster 1) exhibited 17.68 times the odds of having an advanced tumor stage (Stage I – IV vs 0a) compared to patients in Cluster 2. Similarly, patients with a high circulating basophil

proportion (Cluster 1) had 14.24 times the odds of having a high tumor grade compared to patients in Cluster 2. Our previous study also demonstrated that elevated basophil proportions in blood were associated with an increased risk of death in NMIBC patients, with patients exhibiting poor overall survival showing higher circulating basophil proportions [243]. Although there is no study surveilling basophils in the TME of bladder cancer currently, tumor-infiltrating basophils in other cancer types have been reported to play a critical role in tumor progression [244]–[246]. In addition to basophils, Cluster 1 patients exhibited significantly lower monocyte proportions in the TME, but slightly higher monocyte and significantly lower CD4T cell proportions in peripheral blood. Numerous studies have indicated that a high monocyte-to-lymphocyte ratio in the blood is associated with poor bladder cancer outcomes [247]–[250]. Circulating monocytes can be recruited into the TME through tumor-derived chemokines [251], where they differentiate into tumor-associated macrophages (TAMs), promoting tumor cell survival, local invasion, and metastasis [252]–[254]. The lower monocyte proportion in the TME of Cluster 1 patients may be attributed to the tumor's ability to facilitate the differentiation of infiltrating monocytes into TAMs. Further investigations are warranted to validate this inference.

Several potential limitations should be acknowledged in this study. Firstly, all blood samples were collected after diagnosis, with a median time interval of 207 days (Q1 – Q3: 120 – 517). Although the elapsed time may impact the distribution of circulating immune profiles, we found that patient and tumor

characteristic summary statistics, as well as peripheral immune profiles, exhibited similar distributions when stratified based on the median time interval. Secondly, BCG treatment has been shown to influence immune profiles in both the tumor microenvironment and the circulating system [255]–[257]. Despite lacking the information for responsiveness and the number of treatment cycles for each patient with treatment (N = 10; 13%), we did a comparison for immune distribution between the dataset, including patients with BCG treatment (N = 88) or not (N = 78), demonstrating similar distribution. Furthermore, the sample size ratio between NMIBC and MIBC is imbalanced. To address this issue, future data collection efforts should focus on including more MIBC as well as NMIBC patients as well as NMIBCs. However, currently, publicly available datasets with DNA methylation data from tumor tissues and matched blood samples are lacking. Additionally, due to incomplete information, we were unable to investigate the influence of other confounders, such as alcohol consumption [165], [258] and obesity [164], [259], on immune cell distribution. Finally, further studies with larger and independent patient cohorts are needed to validate and confirm these preliminary results.

Taken together, further investigations are warranted to unravel the intricate interplay between local and systemic events in immune modulation against bladder cancer cells. Our study presents a potential opportunity facilitated by a detailed immune profiling approach that utilizes cell-specific methylation data to deconvolute the tumor microenvironment. This work has provided additional insights into the analysis of data encompassing both tumor and peripheral blood

immune profiles. We believe that our findings will provide valuable contributions to the further understanding of immunotherapy response and the underlying mechanisms of bladder cancer development.

4.6 Supplement: Chapter 4

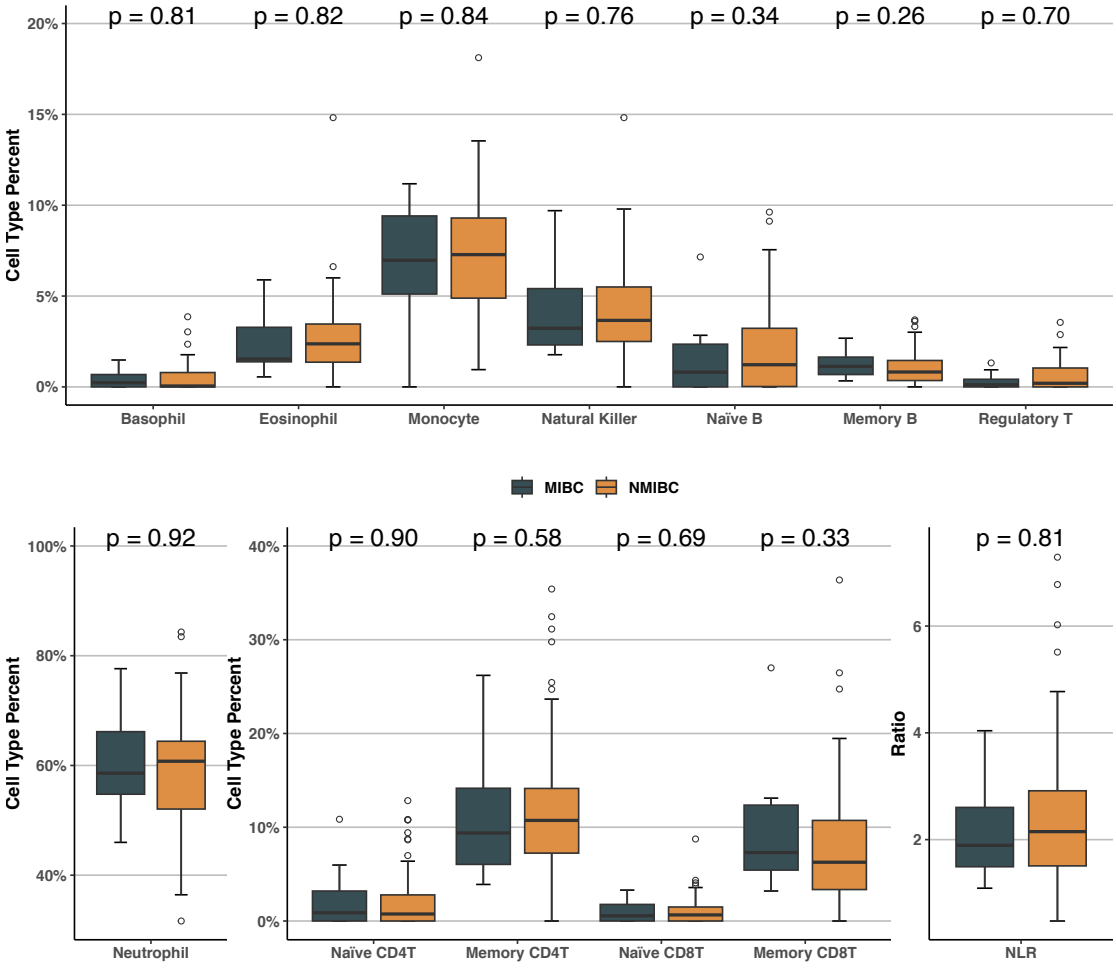


Figure S4.1: Circulating cell profiles in NMIBC and MIBC patients Differences in cell-type proportions between NMIBC and MIBC patients were evaluated using the Wilcoxon rank sum test

Table S4.1: The distribution of subject characteristics within anti-tumor immune infiltration groups

	Low Infiltration (N = 38)	High Infiltration (N = 50)	P (Fisher's Exact Test; Two-tailed)
Tumor Grade			6.22E-5
Low Grade	35 (92%)	24 (48%)	
High Grade	3 (8%)	26 (52%)	
Tumor Stage			5.73E-8
Stage 0a	38 (100%)	23 (46%)	
Stage I	0 (0%)	14 (28%)	
Stage II	0 (0%)	8 (16%)	
Stage III	0 (0%)	2 (4%)	
Stage IV	0 (0%)	3 (6%)	
Muscle Invasive			0.0004
No	38 (100%)	37 (74%)	
Yes	0 (0%)	13 (26%)	
BCG Treatment			0.177
No	36 (95%)	42 (84%)	
Yes	2 (5%)	8 (16%)	

Table S4.2: The distribution of subject characteristics within groups derived from consensus clustering

	Cluster 1 (N = 43)	Cluster 2 (N = 45)	P (Fisher's Exact Test; Two-tailed)
Tumor Grade			2.76E-6
Low Grade	18 (42%)	41 (91%)	
High Grade	25 (58%)	4 (9%)	
Tumor Stage			4.06E-6
Stage 0a	19 (44%)	42 (94%)	
Stage I	12 (28%)	2 (4%)	
Stage II	7 (16%)	1 (2%)	
Stage III	2 (5%)	0 (0%)	
Stage IV	3 (7%)	0 (0%)	
Muscle Invasive			0.0006
No	31 (72%)	44 (98%)	
Yes	12 (28%)	1 (2%)	
BCG Treatment			0.047
No	35 (81%)	43 (96%)	
Yes	8 (19%)	2 (4%)	
Immune Infiltration			0.013
Low	2 (5%)	36 (80%)	
High	41 (95%)	9 (20%)	

Chapter 5

Discussion and Conclusions

5.1 Overview of Findings

5.1.1. Chapter 2: Immune profiles and DNA methylation alterations related with non-muscle-invasive bladder cancer outcomes

Bladder cancer ranks as the sixth most prevalent malignancy in the USA, with approximately seventy percent of patients diagnosed with non-muscle-invasive bladder cancer (NMIBC) [3]. Despite the initial utilization of transurethral resection and immunotherapy for disease control, nearly forty percent of NMIBC patients experience tumor recurrence within one year of surgery. [121]. Furthermore, cystoscopy, the current invasive and standard screening method for monitoring recurrence and progression, imposes a substantial burden on patient morbidity and healthcare costs [260]. Therefore, there is an urgent need to identify less invasive prognostic markers for effective surveillance of tumor recurrence. Potential markers include circulating immune profiles and blood DNA methylation patterns.

The work presented in Chapter 2 focuses on evaluating the association of immune profiles and epigenetic features with tumor recurrence using DNA methylation levels obtained from archived blood samples of NMIBC patients. By employing Cox proportional hazard analysis, we observed that CD4T and

CD8T cell proportions, estimated using the *FlowSorted.Blood.EPIC* approach, were associated with a reduced risk of death and tumor recurrence, whereas the neutrophil-to-lymphocyte ratio (NLR) and monocyte cell proportion were associated with an increased risk of death and tumor recurrence. Of particular interest, adjusting for immune cell proportions had an impact on the results of the epigenome-wide association study (EWAS). Through the EWAS results, we identified 1,572 CpG loci associated with increased risk of tumor recurrence or death, as well as 956 CpG loci associated with decreased risk of tumor recurrence or death. Among the CpGs associated with an increased risk of poor outcome, we found enrichment in open sea regions with low CpG density, and depleted in the CpG island S Shore regions. Meanwhile, CpGs associated with decreased risk of poor outcomes were enriched in CpG island N Shore regions but depleted in CpG islands. Notably, after correcting for cell type variation, CpG loci related with tumor recurrence and death were associated with immune cell regulation and showed enrichment for Histone H3 acetylated at lysine 9 and 14 (H3K9K14ac).

The findings of this study reveal multiple immune profiles in peripheral blood that are significantly associated with outcomes in NMIBC and underscore the significance of incorporating immune profile adjustments in analysis models. Crucially, these results hold the potential for enhancing and establishing a novel biomarker for the surveillance of tumor relapse in NMIBC.

5.1.2. Chapter 3: Genome-scale methylation analysis identifies immune

The findings presented in Chapter 2 demonstrate significant associations between circulating immune profiles derived from DNA methylation profiles, such as CD4T and CD8T cell proportions, and the risk of death and tumor recurrence in NMIBC patients. However, these findings were not generalizable to naïve and memory subtypes of CD4T and CD8T cells, as well as subtypes of granulocytes. To expand on our initial findings, we employed the latest deconvolution approach [102] to obtain high-resolution immune cell profiles in peripheral blood. This approach allowed for a more comprehensive investigation of immune profiles beyond the previously studied cell types. In addition to immune profiles, another DNA methylation-derived variable, age acceleration, has been implicated in cancer outcomes. However, limited studies have focused on age acceleration and bladder cancer specifically [190]–[192], [195]. Therefore, we investigated the association of various immune cell types and age acceleration with NMIBC outcomes to broaden our understanding of their prognostic relevance.

Chapter 3 presents the utilization of DNA methylation data from blood samples to estimate twelve circulating immune cell-type proportions and age acceleration variables from three DNA methylation clocks, the Horvath, Hannum, and Pheno age clocks. We observed significant associations of Hannum and Pheno age accelerations with worse overall survival (OS) and recurrence-free survival (RFS). Moreover, among the immune profiles assessed in the blood samples, high proportions of CD4T memory cells, CD8T memory

cells, and natural killer (NK) cells were associated with a decreased risk of death. Conversely, elevated neutrophil cell proportions, CD8T naïve cell proportions, and neutrophil-to-lymphocyte ratio (NLR) were associated with an increased risk of death.

Furthermore, employing a partitioning deletion/substitution/addition (*partDSA*) algorithm, we identified that neutrophil and CD8T naïve cell proportions partitioned the covariate space, enabling the stratification of patients into three distinct groups. Patients assigned to the group exhibiting the worst 10-year OS demonstrated lower proportions of B naïve, eosinophil, monocyte, NK, CD4T memory, CD4T naïve, CD8T memory, and CD8T naïve cells. In contrast, these patients displayed higher neutrophil proportions, NLR, and age acceleration compared to the other groups. We further employed a semi-supervised recursively partitioned mixture model (*SS-RPMM*) method to assign subjects into two clusters based on optimal CpG sites. Although based on cell-type adjusted models, methylation-derived classes identified that patients with better 10-year OS exhibited higher proportions of B memory, B naïve, CD4T memory, CD4T naïve, CD8T memory, and NK cells. Conversely, patients with better outcomes displayed lower proportions of basophil, eosinophil, monocyte, and neutrophil cells, along with lower NLR and lower age acceleration, in comparison to the other methylation-derived cluster from *SS-RPMM*. Lastly, utilizing the information derived from circulating immune profiles and DNA methylation profiles obtained from the previous clustering approaches, we successfully classified subjects into five distinct groups that exhibited significant

differences in OS.

Collectively, the study represents one of the few investigations to leverage high-resolution immune profiles in blood, age acceleration, and DNA methylation profiles for the evaluation of bladder cancer prognosis, suggesting these factors hold promise as potentially epigenetic biomarkers of bladder cancer.

5.1.3. Chapter 4: Integrated investigation of bladder cancer immune profiles in the tumor microenvironment and peripheral blood

In Chapters 2 and 3, we identified several circulating immune profiles that were associated with bladder cancer outcomes. Besides the circulating immune system, infiltrating immune status has been reported to modulate bladder cancer development [47], [58], [63]. As a result, it is very valuable to understand the immune landscape in tumor tissues and its association with circulating immune profiles, especially for the improvement of immunotherapy. Despite the acknowledged significance of both infiltrating immune status and circulating immune profiles, limited studies have explored the association between immune cell distribution in the tumor microenvironment and peripheral blood in the context of bladder cancer [72]. Fortunately, the recent development of hierarchical tumor immune microenvironment epigenetic deconvolution provides a valuable opportunity to discern cell types in tumor microenvironment using DNA methylation data [89], [90], [96].

Work presented in Chapter 4 leverages DNA methylation cytometry approaches to estimate cell-type proportions in tumors and blood according to

their methylation profiles. Analysis revealed that muscle-invasive bladder cancer patients had a higher dendritic cell, B memory, natural killer, and CD8T memory cell proportions in the tumor microenvironment compared with non-muscle-invasive bladder tumors. Despite the observed differences in immune profiles in tumors, no significant differences were observed in circulating immune profiles between muscle-invasive and non-muscle-invasive bladder cancer patients. Subsequently, subjects were grouped according to conventional immune hot and cold tumor classification, and here too, no significant differences in peripheral immune cell types were observed based on tumors with high and low immune infiltration. To gain additional insights, we applied a consensus clustering approach using tumor immune cell profiles as input data and identified two clusters of tumors. Patients in Cluster 1 had significantly higher proportions of tumor infiltrating dendritic cells, CD8T memory cells, natural killer cells, and B memory cells compared to Cluster 2 patients. Furthermore, Cluster 1 patients, characterized by higher odds of having a high tumor stage and grade compared to Cluster 2 patients, exhibited significantly higher circulating basophil proportions and lower circulating CD4T naïve and memory cell proportions.

This study not only elucidates the potential relation of immune profiles between two compartments but also highlights the promising opportunity to leverage the detailed immune profiling approach based on methylation information. Such an approach holds the potential to enhance our understanding of bladder cancer development and improve the assessment of immunotherapy response.

5.2 Perspectives and Future Directions

Currently, post-surgery and immunotherapy, bladder cancer patients routinely undergo an invasive monitoring approach, cystoscopy, imposing substantial patient morbidity and placing a burden on the healthcare system [260]. Given the emerging understanding of the pivotal role of the immune system in bladder cancer pathogenesis and response to immunotherapies, immune profiling holds promise as a potential prognostic biomarker [261], [262]. Understanding the intricate interplay between the immune system and bladder cancer relapse, as well as the development of resistance to immunotherapy, presents a complex challenge. Despite extensive research in the field of bladder cancer immunology, comprehensive investigations focusing on the distribution of lymphocyte and granulocyte subsets within both tumor and blood, and their association with treatment outcomes, remain limited. One potential constraint is the requirement for fresh specimens and a sufficient sample quantity. The work presented in this thesis utilized DNA methylation cytometry, which allows for the analysis of archived samples and does not necessitate large sample volumes, to identify the association between immune status and cancer outcomes.

An important question arising from the findings presented in this dissertation remains unanswered: does the immune composition of peripheral blood undergo changes during bladder cancer relapse or progression? To date, no studies have investigated alterations in lymphocyte subtypes and granulocyte subtypes over time during cancer development. By utilizing DNA methylation

cytometry, we can simultaneously estimate multiple distinct immune cell types, providing a time and cost-efficient alternative to conventional methods such as the CBC test and flow cytometry. Given the potential for changes in the immune composition of blood over time during cancer development within an individual, future studies incorporating frequent and repeated measures of blood are necessary to better understand the optimal timing and frequency of these assessments. To further elucidate the potential of methylation-derived immune profiles as prognostic biomarkers, it is crucial to investigate whether circulating immune profiles undergo changes following surgery. Consequently, it is imperative to collect preoperative and postoperative blood samples within a specific time frame. Subsequently, a comparison of immune cell types before and after treatment will be conducted, followed by an assessment of the association between these changes and bladder cancer outcomes. Furthermore, addressing immunotherapy resistance is a critical concern in the management of bladder cancer. In the case of the most commonly used immunotherapy, BCG treatment, patients typically undergo treatment once a week for six weeks [263], [264]. By incorporating more frequent blood draws during treatment, changes in circulating immune profiles can be monitored and integrated with recommendations for follow-up time to assess potential tumor relapse. Such an approach would enable the development of a model to distinguish patients who are likely to respond to BCG treatment, facilitating early intervention with alternative therapies and minimizing unnecessary costs and patient discomfort. In addition to BCG treatment, other treatment modalities and lifestyle factors,

including chemotherapy [265], cystectomy [266], obesity [164], alcohol consumption [258], [267], and type II diabetes [166], [268], may introduce confounding effects on our analysis results. Therefore, it is crucial to gather information regarding these factors in future studies to enhance our understanding of predictive prognostic factors. Conducting studies that evaluate the impact of interventions such as surgery and alternative treatment methods on the stability and trajectory of peripheral immune status would be instrumental in furthering our understanding of immune dynamics in bladder cancer.

Urine cytology is a non-invasive diagnostic test with high specificity but limited average sensitivity in detecting cancer cells in bladder-washing specimens or urine samples obtained during normal urination [269]. Consequently, several studies have explored the utilization of DNA methylation profiles from urine samples to enhance the sensitivity of early-stage bladder cancer diagnosis and facilitate tumor recurrence monitoring [270]–[272]. Also, considering the substantial presence of granulocytes and monocytes in urine samples from bladder cancer patients, as well as the association between a robust inflammatory response and increased cell numbers in the urine, the development of DNA methylation cytometry for urine samples holds promise [273], [274]. The integration of urine methylation cytometry could allow for the combination of urine immune profiles with established DNA methylation markers, leading to the creation of a rapid, high-throughput, and non-invasive approach for surveillance in bladder cancer. Recently, there have been notable advancements in utilizing deep learning techniques for urine cytology analysis [275]–[277].

One such example is the development of AutoParis-X [277], [278], a novel method that demonstrates improved performance in determining the level of atypia in specific cells and detecting abnormal cells, including bladder cancer cells, compared to conventional cytological assessments. Moving forward, there is a potential opportunity for integrating urine methylation profiles with feature representations derived from AutoParis-X analysis results. This integration could offer a promising approach for predicting bladder cancer recurrence and identifying CpG loci within urine samples that are associated with recurrence. We expect that this integrated approach holds promise for achieving enhanced performance compared to existing methods.

In Chapter 4, we examined the potential association between immune profiles in tumor tissue and blood. However, the regulatory role of DNA methylation levels in the tumor microenvironment was not investigated. The major challenge lies in the sample size limitation. In future studies, our aim is to increase the number of tumor specimens with matched blood samples or explore the utilization of external datasets, although currently, there are no publicly available datasets for this purpose. Subsequently, we plan to conduct epigenome-wide association studies to identify differentially methylated regions between groups categorized using the methods outlined in Chapter 4. Furthermore, we will investigate whether the identified CpG sites are involved in bladder cancer progression or immune regulation.

Machine learning is a discipline that leverages task-specific objective functions to extract valuable information from large datasets [279]. In recent

years, with advancements in hardware capabilities and the accumulation of data, deep learning models have emerged as powerful tools widely applied in cancer pathology [280], [281]. While previous studies have demonstrated the potential of convolutional neural networks in bladder cancer pathology, these findings have been limited to the analysis of whole slide images (WSIs) [282], [283]. To further enhance cancer prognostication, we propose the integration of DNA methylation profiles and derived immune cell-type proportions, into the model. Our next objective is to develop a multimodal neural network that incorporates WSIs, DNA methylation levels, immune profiles estimated through DNA methylation cytometry, and clinical attributes, such as age and tumor grade. Specifically, graph neural networks and autoencoders will be applied to WSIs to extract morphological features and spatial locality, generating an embedding (vector representation). Moreover, additional autoencoders will be trained using DNA methylation levels, immune profiles, and clinical attributes to generate corresponding vector representations. Finally, multiple vector representations will be combined and then the fused representation passes through several fully connected layers, culminating in a final prognosis prediction output layer. This work primarily aims to quantify prognosis as a scalar hazard ratio and alleviate the burden on practitioners while optimizing resource utilization.

5.3 Concluding Remarks

The collection of research presented in this study employs innovative approaches to enhance our understanding of immune status and DNA methylation levels in

bladder cancer. Through this work, we have identified several methylation-derived immune profiles that are significantly associated with risk and prognosis in bladder cancer with the urgent need for less-invasive methods to monitor relapse. Furthermore, our findings shed light on the relationship between immune profiles in the tumor microenvironment and peripheral blood, providing valuable insights into immune dynamics in bladder cancer. Together with future investigations, these findings will contribute to the development of novel prognostic biomarkers for bladder cancer outcomes and immunotherapy, ultimately advancing our ability to personalize patient management and improve treatment strategies.

Bibliography

[1] D. J. Grignon, “The current classification of urothelial neoplasms,” *Mod. Pathol.*, vol. 22, pp. S60–S69, 2009.

[2] B. S. Chhikara and K. Parang, “Global Cancer Statistics 2022: the trends projection analysis,” *Chem. Biol. Lett.*, vol. 10, no. 1, pp. 1–16, 2023.

[3] Atlanta: American Cancer Society, “Cancer Facts & Figures 2022,” *American Cancer Society*. 2022.

[4] M. G. Cumberbatch, M. Rota, J. W. F. Catto, and C. La Vecchia, “The Role of Tobacco Smoke in Bladder and Kidney Carcinogenesis: A Comparison of Exposures and Meta-analysis of Incidence and Mortality Risks,” *Eur. Urol.*, vol. 70, no. 3, pp. 458–466, 2016.

[5] H. Sung *et al.*, “Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries,” *CA. Cancer J. Clin.*, vol. 71, no. 3, pp. 209–249, 2021.

[6] S. A. Halaseh, S. Halaseh, Y. Alali, M. E. Ashour, and M. J. Alharayzah, “A Review of the Etiology and Epidemiology of Bladder Cancer: All You Need To Know,” *Cureus*, vol. 14, no. 7, 2022.

[7] L. M. C. van Hoogstraten, A. Vrieling, A. G. van der Heijden, M. Kogevinas, A. Richters, and L. A. Kiemeney, “Global trends in the epidemiology of bladder cancer: challenges for public health and clinical practice,” *Nat. Rev. Clin. Oncol.*, vol. 20, no. May, pp. 287–304, 2023.

[8] B. W. G. van Rhijn *et al.*, “Recurrence and Progression of Disease in Non-Muscle-Invasive Bladder Cancer: From Epidemiology to Treatment

Strategy," *Eur. Urol.*, vol. 56, no. 3, pp. 430–442, 2009.

[9] Z. Kirkali *et al.*, "Bladder cancer: Epidemiology, staging and grading, and diagnosis," *Urology*, vol. 66, no. 6 SUPPL. 1, pp. 4–34, 2005.

[10] A. K. Das, D. K. Mishra, and S. S. Gopalan, "Effect of Intravesical Chemotherapy on the Survival of Patients with Non-Muscle-Invasive Bladder Cancer Undergoing Transurethral Resection: A Retrospective Cohort Study Among Older Adults," *Open Urol. Nephrol. J.*, vol. 14, no. 1, pp. 20–25, 2021.

[11] K. Ogawa, Y. Shimizu, S. Uketa, N. Utsunomiya, and S. Kanamaru, "Prognosis of patients with muscle invasive bladder cancer who are intolerable to receive any anti-cancer treatment," *Cancer Treat. Res. Commun.*, vol. 24, p. 100195, 2020.

[12] S. Cambier *et al.*, "EORTC Nomograms and Risk Groups for Predicting Recurrence, Progression, and Disease-specific and Overall Survival in Non-Muscle-invasive Stage Ta-T1 Urothelial Bladder Cancer Patients Treated with 1-3 Years of Maintenance Bacillus Calmette-Guérin," *Eur. Urol.*, vol. 69, no. 1, pp. 60–69, 2016.

[13] S. Zhu, W. Yu, X. Yang, C. Wu, and F. Cheng, "Traditional Classification and Novel Subtyping Systems for Bladder Cancer," *Front. Oncol.*, vol. 10, no. February, pp. 1–13, 2020.

[14] T. Z. Tan, M. Rouanne, K. T. Tan, R. Y. J. Huang, and J. P. Thiery, "Molecular Subtypes of Urothelial Bladder Cancer: Results from a Meta-cohort Analysis of 2411 Tumors," *Eur. Urol.*, vol. 75, no. 3, pp. 423–432,

2019.

[15] C. L. Pashos, M. F. Botteman, B. L. Laskin, and A. Redaelli, “Bladder cancer: epidemiology, diagnosis, and management.,” *Cancer Pract.*, vol. 10, no. 6, pp. 311–322, 2002.

[16] M. G. K. Cumberbatch *et al.*, “Epidemiology of Bladder Cancer: A Systematic Review and Contemporary Update of Risk Factors in 2018,” *Eur. Urol.*, vol. 74, no. 6, pp. 784–795, 2018.

[17] A. Guevara *et al.*, “The Role of Tumor-Free Status in Repeat Resection Before Intravesical Bacillus Calmette-Guerin for High Grade Ta, T1 and CIS Bladder Cancer,” *J. Urol.*, vol. 183, no. 6, pp. 2161–2164, 2010.

[18] A. C. Goh and S. P. Lerner, “Application of new technology in bladder cancer diagnosis and treatment,” *World J. Urol.*, vol. 27, no. 3, pp. 301–307, 2009.

[19] S. Guallar-Garrido and E. Julián, “Bacillus Calmette-Guérin (BCG) Therapy for Bladder Cancer: An Update.,” *ImmunoTargets Ther.*, vol. 9, pp. 1–11, 2020.

[20] J. Palou, P. Laguna, F. Millán-Rodríguez, R. R. Hall, J. Salvador-Bayarri, and J. Vicente-Rodríguez, “Control group and maintenance treatment with bacillus Calmette-Guerin for carcinoma in situ and/or high grade bladder tumors.,” *J. Urol.*, vol. 165, no. 5, pp. 1488–1491, May 2001.

[21] D. L. Lamm *et al.*, “Maintenance bacillus Calmette-Guerin immunotherapy for recurrent TA, T1 and carcinoma in situ transitional cell carcinoma of the bladder: a randomized Southwest Oncology Group

Study.,” *J. Urol.*, vol. 163, no. 4, pp. 1124–1129, Apr. 2000.

[22] A. M. Kamat *et al.*, “Expert consensus document: Consensus statement on best practice management regarding the use of intravesical immunotherapy with BCG for bladder cancer,” *Nat. Rev. Urol.*, vol. 12, no. 4, pp. 225–235, 2015.

[23] A. M. Kamat *et al.*, “Bladder cancer,” *Lancet*, vol. 388, no. 10061, pp. 2796–2810, 2016.

[24] M. Moschini *et al.*, “Comparing long-term outcomes of primary and progressive carcinoma invading bladder muscle after radical cystectomy,” *BJU Int.*, vol. 117, no. 4, pp. 604–610, 2016.

[25] M. S. Soloway, “Diagnosis and management of superficial bladder cancer.,” *Semin. Surg. Oncol.*, vol. 5, no. 4, pp. 247–254, 1989.

[26] M. Babjuk, W. Oosterlinck, R. Sylvester, E. Kaasinen, A. Böhle, and J. Palou-Redorta, “EAU Guidelines on Non-Muscle-Invasive Urothelial Carcinoma of the Bladder,” *Eur. Urol.*, vol. 54, no. 2, pp. 303–314, 2008.

[27] F. A. Yafi, F. Brimo, J. Steinberg, A. G. Aprikian, S. Tanguay, and W. Kassouf, “Prospective analysis of sensitivity and specificity of urinary cytology and other urinary biomarkers for bladder cancer,” *Urol. Oncol. Semin. Orig. Investig.*, vol. 33, no. 2, pp. 66.e25-66.e31, 2015.

[28] C. Scheiermann, P. S. Frenette, and A. Hidalgo, “Regulation of leucocyte homeostasis in the circulation,” *Cardiovasc. Res.*, vol. 107, no. 3, pp. 340–351, 2015.

[29] T. S. Kapellos *et al.*, “Human monocyte subsets and phenotypes in major

chronic inflammatory diseases," *Front. Immunol.*, vol. 10, no. AUG, pp. 1–13, 2019.

[30] M. Y. Zanna *et al.*, "Review of dendritic cells, their role in clinical immunology, and distribution in various animal species," *Int. J. Mol. Sci.*, vol. 22, no. 15, 2021.

[31] P. J. Murray and T. A. Wynn, "Protective and pathogenic functions of macrophage subsets," *Nat. Rev. Immunol.*, vol. 11, no. 11, pp. 723–737, 2011.

[32] G. A. Duque and A. Descoteaux, "Macrophage cytokines: Involvement in immunity and infectious diseases," *Front. Immunol.*, vol. 5, no. OCT, pp. 1–12, 2014.

[33] E. Z. M. da Silva, M. C. Jamur, and C. Oliver, *Mast Cell Function: A New Vision of an Old Cell*, vol. 62, no. 10. 2014.

[34] B. Min, M. A. Brown, and G. Legros, "Understanding the roles of basophils: Breaking dawn," *Immunology*, vol. 135, no. 3, pp. 192–197, 2012.

[35] T. Wen and M. E. Rothenberg, "The regulatory function of eosinophils," *Myeloid Cells Heal. Dis. A Synth.*, vol. 4, no. 5, pp. 257–269, 2017.

[36] C. Rosales, "Neutrophil: A cell with many roles in inflammation or several cell types?," *Front. Physiol.*, vol. 9, no. FEB, pp. 1–17, 2018.

[37] K. Verhoeckx *et al.*, "Peripheral Blood Mononuclear Cells," in *The Impact of Food Bioactives on Health: In Vitro and Ex Vivo Models*, 2015.

[38] N. Zhang and M. J. Bevan, "CD8+ T Cells: Foot Soldiers of the Immune

System," *Immunity*, vol. 35, no. 2, pp. 161–168, 2011.

[39] T. Samji and K. M. Khanna, "Understanding Memory CD8+ T cells," *Immunol. Lett.*, vol. 185, pp. 32–39, 2017.

[40] R. V. Luckheeram, R. Zhou, A. D. Verma, and B. Xia, "CD4 +T cells: Differentiation and functions," *Clin. Dev. Immunol.*, vol. 2012, 2012.

[41] D. J. Gasper, M. M. Tejera, and M. Suresh, "CD4 T-cell memory generation and maintenance," *Crit. Rev. Immunol.*, vol. 34, no. 2, pp. 121–146, 2014.

[42] D. A. A. Vignali, L. W. Collison, and C. J. Workman, "How regulatory T cells work," *Nat. Rev. Immunol.*, vol. 8, no. 7, pp. 523–532, 2008.

[43] M. Akkaya, K. Kwak, and S. K. Pierce, "B cell memory: building two walls of protection against pathogens," *Nat. Rev. Immunol.*, vol. 20, no. 4, pp. 229–238, 2020.

[44] E. Vivier, E. Tomasello, M. Baratin, T. Walzer, and S. Ugolini, "Functions of natural killer cells," *Nat. Immunol.*, vol. 9, no. 5, pp. 503–510, 2008.

[45] N. M. Donin *et al.*, "Immunotherapy for the Treatment of Urothelial Carcinoma.," *J. Urol.*, vol. 197, no. 1, pp. 14–22, Jan. 2017.

[46] T. Kawasaki and T. Kawai, "Toll-like receptor signaling pathways.," *Front. Immunol.*, vol. 5, p. 461, 2014.

[47] T. Sonoda, K. Sugimura, S.-I. Ikemoto, H. Kawashima, and T. Nakatani, "Significance of target cell infection and natural killer cells in the anti-tumor effects of bacillus Calmette-Guerin in murine bladder cancer.,"

Oncol. Rep., vol. 17, no. 6, pp. 1469–1474, Jun. 2007.

[48] H. Yamada, S. Matsumoto, T. Matsumoto, T. Yamada, and U. Yamashita, "Enhancing effect of an inhibitor of nitric oxide synthesis on bacillus Calmette-Guerin-induced macrophage cytotoxicity against murine bladder cancer cell line MBT-2 in vitro.," *Jpn. J. Cancer Res.*, vol. 91, no. 5, pp. 534–542, May 2000.

[49] M. Naoe *et al.*, "Bacillus Calmette-Guérin-pulsed dendritic cells stimulate natural killer T cells and gammadeltaT cells.," *Int. J. Urol. Off. J. Japanese Urol. Assoc.*, vol. 14, no. 6, pp. 532–8; discussion 538, Jun. 2007.

[50] M. G. Netea *et al.*, "Defining trained immunity and its role in health and disease.," *Nat. Rev. Immunol.*, vol. 20, no. 6, pp. 375–388, Jun. 2020.

[51] B. A. Inman *et al.*, "PD-L1 (B7-H1) expression by urothelial carcinoma of the bladder and BCG-induced granulomata: associations with localized stage progression.," *Cancer*, vol. 109, no. 8, pp. 1499–1505, Apr. 2007.

[52] J. Hamanishi, M. Mandai, N. Matsumura, K. Abiko, T. Baba, and I. Konishi, "PD-1/PD-L1 blockade in cancer treatment: perspectives and issues.," *Int. J. Clin. Oncol.*, vol. 21, no. 3, pp. 462–473, Jun. 2016.

[53] M. E. Keir *et al.*, "Tissue expression of PD-L1 mediates peripheral T cell tolerance.," *J. Exp. Med.*, vol. 203, no. 4, pp. 883–895, Apr. 2006.

[54] B. T. Fife *et al.*, "Interactions between PD-1 and PD-L1 promote tolerance by blocking the TCR-induced stop signal.," *Nat. Immunol.*,

vol. 10, no. 11, pp. 1185–1192, Nov. 2009.

[55] C. Ilcus *et al.*, “Immune checkpoint blockade: the role of PD-1-PD-L axis in lymphoid malignancies.,” *Onco. Targets. Ther.*, vol. 10, pp. 2349–2363, 2017.

[56] A. Lopez-Beltran *et al.*, “Immune Checkpoint Inhibitors for the Treatment of Bladder Cancer.,” *Cancers (Basel).*, vol. 13, no. 1, Jan. 2021.

[57] G. A. M. Povoleri, C. Scottà, E. A. Nova-Lamperti, S. John, G. Lombardi, and B. Afzali, “Thymic versus induced regulatory T cells - who regulates the regulators?,” *Front. Immunol.*, vol. 4, p. 169, 2013.

[58] Y.-N. Liu *et al.*, “Sphingosine 1 phosphate receptor-1 (S1P1) promotes tumor-associated regulatory T cell expansion: leading to poor survival in bladder cancer.,” *Cell Death Dis.*, vol. 10, no. 2, p. 50, Jan. 2019.

[59] A. Loskog, C. Ninalga, G. Paul-Wetterberg, M. de la Torre, P.-U. Malmström, and T. H. Tötterman, “Human bladder carcinoma is dominated by T-regulatory cells and Th1 inhibitory cytokines.,” *J. Urol.*, vol. 177, no. 1, pp. 353–358, Jan. 2007.

[60] R. Murai *et al.*, “Prediction of intravesical recurrence of non-muscle-invasive bladder cancer by evaluation of intratumoral Foxp3+ T cells in the primary transurethral resection of bladder tumor specimens.,” *PLoS One*, vol. 13, no. 9, p. e0204745, 2018.

[61] F. Veglia, M. Perego, and D. Gabrilovich, “Myeloid-derived suppressor cells coming of age.,” *Nat. Immunol.*, vol. 19, no. 2, pp. 108–119, Feb.

2018.

[62] A. Sica, C. Porta, A. Amadori, and A. Pastò, “Tumor-associated myeloid cells as guiding forces of cancer cell stemness.,” *Cancer Immunol. Immunother.*, vol. 66, no. 8, pp. 1025–1036, Aug. 2017.

[63] H. Zhang *et al.*, “CXCL2/MIF-CXCR2 signaling promotes the recruitment of myeloid-derived suppressor cells and is correlated with prognosis in bladder cancer.,” *Oncogene*, vol. 36, no. 15, pp. 2095–2104, Apr. 2017.

[64] M. Miyake *et al.*, “Regulatory T Cells and Tumor-Associated Macrophages in the Tumor Microenvironment in Non-Muscle Invasive Bladder Cancer Treated with Intravesical Bacille Calmette-Guérin: A Long-Term Follow-Up Study of a Japanese Cohort.,” *Int. J. Mol. Sci.*, vol. 18, no. 10, Oct. 2017.

[65] G. Yang *et al.*, “Accumulation of myeloid-derived suppressor cells (MDSCs) induced by low levels of IL-6 correlates with poor prognosis in bladder cancer.,” *Oncotarget*, vol. 8, no. 24, pp. 38378–38388, Jun. 2017.

[66] A. Mantovani, S. Sozzani, M. Locati, P. Allavena, and A. Sica, “Macrophage polarization: tumor-associated macrophages as a paradigm for polarized M2 mononuclear phagocytes.,” *Trends Immunol.*, vol. 23, no. 11, pp. 549–555, Nov. 2002.

[67] C. Chen *et al.*, “LNMAT1 promotes lymphatic metastasis of bladder cancer via CCL2 dependent macrophage recruitment.,” *Nat. Commun.*,

vol. 9, no. 1, p. 3826, Sep. 2018.

[68] H. Miyamoto, Y. Kubota, T. Shuin, S. Torigoe, Y. Dobashi, and M. Hosaka, “Expression of transforming growth factor-beta 1 in human bladder cancer.,” *Cancer*, vol. 75, no. 10, pp. 2565–2570, May 1995.

[69] X. Wang *et al.*, “Bladder cancer cells induce immunosuppression of T cells by supporting PD-L1 expression in tumour macrophages partially through interleukin 10.,” *Cell Biol. Int.*, vol. 41, no. 2, pp. 177–186, Feb. 2017.

[70] T. Hanada, M. Nakagawa, A. Emoto, T. Nomura, N. Nasu, and Y. Nomura, “Prognostic value of tumor-associated macrophage count in human bladder cancer.,” *Int. J. Urol. Off. J. Japanese Urol. Assoc.*, vol. 7, no. 7, pp. 263–269, Jul. 2000.

[71] B. Wang *et al.*, “High CD204+ tumor-infiltrating macrophage density predicts a poor prognosis in patients with urothelial cell carcinoma of the bladder.,” *Oncotarget*, vol. 6, no. 24, pp. 20204–20214, Aug. 2015.

[72] M. E. Shaul and Z. G. Fridlender, “Cancer-related circulating and tumor-associated neutrophils - subtypes, sources and function.,” *FEBS J.*, vol. 285, no. 23, pp. 4316–4342, Dec. 2018.

[73] K. Liu, K. Zhao, L. Wang, and E. Sun, “The prognostic values of tumor-infiltrating neutrophils, lymphocytes and neutrophil/lymphocyte rates in bladder urothelial cancer.,” *Pathol. Res. Pract.*, vol. 214, no. 8, pp. 1074–1080, Aug. 2018.

[74] M. Marchioni *et al.*, “The Clinical Use of the Neutrophil to Lymphocyte

Ratio (NLR) in Urothelial Cancer: A Systematic Review.," *Clin. Genitourin. Cancer*, vol. 14, no. 6, pp. 473–484, Dec. 2016.

[75] M. D. Vartolomei *et al.*, "Prognostic role of pretreatment neutrophil-to-lymphocyte ratio (NLR) in patients with non-muscle-invasive bladder cancer (NMIBC): A systematic review and meta-analysis.," *Urol. Oncol.*, vol. 36, no. 9, pp. 389–399, Sep. 2018.

[76] L. D. Moore, T. Le, and G. Fan, "DNA methylation and its basic function," *Neuropsychopharmacology*, vol. 38, no. 1, pp. 23–38, 2013.

[77] M. Ehrlich *et al.*, "Amount and distribution of 5-methylcytosine in human DNA from different types of tissues or cells," *Nucleic Acids Res.*, vol. 10, no. 8, pp. 2709–2721, 1982.

[78] A. M. Deaton and A. Bird, "CpG islands and the regulation of transcription," *Genes Dev.*, vol. 25, no. 10, pp. 1010–1022, 2011.

[79] J. Sandoval *et al.*, "Validation of a DNA methylation microarray for 450,000 CpG sites in the human genome," *Epigenetics*, vol. 6, no. 6, pp. 692–702, 2011.

[80] P. A. Wade, "Methyl CpG-binding proteins and transcriptional repression," *BioEssays*, vol. 23, no. 12, pp. 1131–1137, 2001.

[81] P. A. Jones, "Functions of DNA methylation: Islands, start sites, gene bodies and beyond," *Nat. Rev. Genet.*, vol. 13, no. 7, pp. 484–492, 2012.

[82] D. Aran, G. Toperoff, M. Rosenberg, and A. Hellman, "Replication timing-related and gene body-specific methylation of active human genes," *Hum. Mol. Genet.*, vol. 20, no. 4, pp. 670–680, 2011.

[83] B. Jin and K. D. Robertson, “DNA methyltransferases, DNA damage repair, and cancer,” *Adv. Exp. Med. Biol.*, vol. 754, pp. 3–29, 2013.

[84] B. C. Christensen *et al.*, “Aging and environmental exposures alter tissue-specific DNA methylation dependent upon CPG island context,” *PLoS Genet.*, vol. 5, no. 8, 2009.

[85] B. Kwabi-Addo *et al.*, “Age-related DNA methylation changes in normal human prostate tissues.,” *Clin. cancer Res. an Off. J. Am. Assoc. Cancer Res.*, vol. 13, no. 13, pp. 3796–3802, Jul. 2007.

[86] T. Bergsma and E. Rogaeva, “DNA Methylation Clocks and Their Predictive Capacity for Aging Phenotypes and Healthspan,” *Neurosci. Insights*, vol. 15, 2020.

[87] S. Horvath, “DNA methylation age of human tissues and cell types,” *Genome Biol.*, vol. 16, no. 1, 2015.

[88] and K. Z. Gregory Hannum, Justin Guinney, Ling Zhao, Li Zhang, Guy Hughes, SriniVas Sadda, Brandy Klotzle, Marina Bibikova, Jian-Bing Fan, Yuan Gao, Rob Deconde, Menzies Chen, Indika Rajapakse, Stephen Friend, Trey Ideker, “Genome-wide Methylation Profiles Reveal Quantitative Views of Human Aging Rates,” *Mol Cell.*, vol. 49, no. 2, pp. 359–367, 2013.

[89] L. A. Salas *et al.*, “An optimized library for reference-based deconvolution of whole-blood biospecimens assayed using the Illumina HumanMethylationEPIC BeadArray.,” *Genome Biol.*, vol. 19, no. 1, p. 64, 2018.

[90] E. A. Houseman *et al.*, “DNA methylation arrays as surrogate measures of cell mixture distribution,” *BMC Bioinformatics*, vol. 13, no. 1, 2012.

[91] S. Moran, C. Arribas, and M. Esteller, “Validation of a DNA methylation microarray for 850,000 CpG sites of the human genome enriched in enhancer sequences,” *Epigenomics*, vol. 8, no. 3, pp. 389–399, 2016.

[92] R. Darst, C. Pardo, and Etc, “Bisulphite sequencing of DNA,” *Curr Protoc Mol Biol*, 2011.

[93] H. D. Woo and J. Kim, “Global DNA hypomethylation in peripheral blood leukocytes as a biomarker for cancer risk: A meta-analysis,” *PLoS One*, vol. 7, no. 4, 2012.

[94] R. Illingworth *et al.*, “A novel CpG island set identifies tissue-specific methylation at developmental gene loci,” *PLoS Biol.*, vol. 6, no. 1, pp. 0037–0051, 2008.

[95] E. Kitamura *et al.*, “Analysis of tissue-specific differentially methylated regions (TDMs) in humans,” *Genomics*, vol. 89, no. 3, pp. 326–337, 2007.

[96] L. E. Reinius *et al.*, “Differential DNA methylation in purified human blood cells: Implications for cell lineage and studies on disease susceptibility,” *PLoS One*, vol. 7, no. 7, 2012.

[97] V. K. Rakyan, T. A. Down, D. J. Balding, and S. Beck, “Epigenome-wide association studies for common human diseases,” *Nat. Rev. Genet.*, vol. 12, no. 8, pp. 529–541, 2011.

[98] A. J. Titus, R. M. Gallimore, L. A. Salas, and B. C. Christensen, “Cell-

type deconvolution from DNA methylation: A review of recent applications," *Hum. Mol. Genet.*, vol. 26, no. R2, pp. R216–R224, 2017.

[99] A. E. Teschendorff and S. C. Zheng, "Cell-type deconvolution in epigenome-wide association studies: A review and recommendations," *Epigenomics*, vol. 9, no. 5, pp. 757–768, 2017.

[100] J. K. Wiencke *et al.*, "Immunomethylomic approach to explore the blood neutrophil lymphocyte ratio (NLR) in glioma survival," *Clin. Epigenetics*, vol. 9, no. 1, pp. 1–11, 2017.

[101] Z. Zhang, J. K. Wiencke, K. T. Kelsey, D. C. Koestler, B. C. Christensen, and L. A. Salas, "HiTIMED: hierarchical tumor immune microenvironment epigenetic deconvolution for accurate cell type resolution in the tumor microenvironment using tumor-type-specific DNA methylation data," *J. Transl. Med.*, vol. 20, no. 1, pp. 1–17, 2022.

[102] L. A. Salas *et al.*, "Enhanced cell deconvolution of peripheral blood using DNA methylation for high-resolution immune profiling," *Nat. Commun.*, vol. 13, no. 1, 2022.

[103] C. López-Otín, M. A. Blasco, L. Partridge, M. Serrano, and G. Kroemer, "The Hallmarks of Aging," *Cell*, vol. 153, no. 6, pp. 1194–1217, 2013.

[104] J. Jylhävä, N. L. Pedersen, and S. Hägg, "Biological Age Predictors," *EBioMedicine*, vol. 21, pp. 29–36, 2017.

[105] C. E., V. S., K. J.K., and A. D., "Biomarkers related to aging in human populations.," *Advances in clinical chemistry*, vol. 46. pp. 161–216, 2008.

[106] S. Horvath and K. Raj, “DNA methylation-based biomarkers and the epigenetic clock theory of ageing,” *Nat. Rev. Genet.*, vol. 19, no. 6, pp. 371–384, 2018.

[107] A. E. Field, N. A. Robertson, T. Wang, A. Havas, T. Ideker, and P. D. Adams, “DNA Methylation Clocks in Aging: Categories, Causes, and Consequences,” *Molecular Cell*, vol. 71, no. 6. pp. 882–895, 2018.

[108] V. K. Rakyan *et al.*, “Human aging-associated DNA hypermethylation occurs preferentially at bivalent chromatin domains,” *Genome Res.*, vol. 20, no. 4, pp. 434–439, 2010.

[109] S. Bocklandt *et al.*, “Epigenetic Predictor of Age,” *PLoS One*, vol. 6, no. 6, 2011.

[110] C. M. Koch and W. Wagner, “Epigenetic-aging-signature to determine age in different tissues,” *Aging (Albany. NY).*, vol. 3, no. 10, pp. 1018–1027, 2011.

[111] B. H. Chen *et al.*, “DNA methylation-based measures of biological age: Meta-analysis predicting time to death,” *Aging (Albany. NY).*, vol. 8, no. 9, pp. 1844–1865, 2016.

[112] Q. Zhang *et al.*, “Improved precision of epigenetic clock estimates across tissues and its implication for biological ageing,” *Genome Med.*, vol. 11, no. 1, pp. 1–11, 2019.

[113] M. Levine *et al.*, “An epigenetic biomarker of aging for lifespan and healthspan,” *Aging (Albany NY)*, vol. 10, no. 4, pp. 573–591, 2018.

[114] Y. Zhang *et al.*, “DNA methylation signatures in peripheral blood

strongly predict all-cause mortality," *Nat. Commun.*, vol. 8, pp. 1–11, 2017.

[115] C. G. Bell *et al.*, "DNA methylation aging clocks: Challenges and recommendations," *Genome Biol.*, vol. 20, no. 1, pp. 1–24, 2019.

[116] Atlanta: American Cancer Society, "Cancer Facts & Figures 2020," *American Cancer Society*. 2020.

[117] R. H. Martinez Rodriguez, O. Buisan Rueda, and L. Ibarz, "Bladder cancer: Present and future," *Med. Clínica (English Ed.*, vol. 149, no. 10, pp. 449–455, 2017.

[118] R. S. Svatek and Y. Lotan, "Is there a rationale for bladder cancer screening?," *Curr. Urol. Rep.*, vol. 9, no. 5, pp. 339–341, 2008.

[119] S. Larré *et al.*, "Screening for bladder cancer: Rationale, limitations, whom to target, and perspectives," *Eur. Urol.*, vol. 63, no. 6, pp. 1049–1058, 2013.

[120] K. C. DeGeorge, H. R. Holt, and S. C. Hodges, "Bladder Cancer: Diagnosis and Treatment.," *Am. Fam. Physician*, vol. 96, no. 8, pp. 507–514, Oct. 2017.

[121] M. C. Hall *et al.*, "Guideline for the Management of Nonmuscle Invasive Bladder Cancer (Stages Ta, T1, and Tis): 2007 Update," *J. Urol.*, vol. 178, no. 6, pp. 2314–2330, 2007.

[122] M. Mossanen and J. L. Gore, "The burden of bladder cancer care: Direct and indirect costs," *Curr. Opin. Urol.*, vol. 24, no. 5, pp. 487–491, 2014.

[123] C. J. Lim *et al.*, "Immunological Hallmarks for Clinical Response to

BCG in Bladder Cancer," *Front. Immunol.*, vol. 11, no. January, pp. 1–13, 2021.

[124] K. Ogihara *et al.*, "The Preoperative Neutrophil-to-lymphocyte Ratio is a Novel Biomarker for Predicting Worse Clinical Outcomes in Non-muscle Invasive Bladder Cancer Patients with a Previous History of Smoking," *Ann. Surg. Oncol.*, vol. 23, pp. 1039–1047, 2016.

[125] Y. G. Tan, E. Eu, W. Lau Kam On, and H. H. Huang, "Pretreatment neutrophil-to-lymphocyte ratio predicts worse survival outcomes and advanced tumor staging in patients undergoing radical cystectomy for bladder cancer," *Asian J. Urol.*, vol. 4, no. 4, pp. 239–246, 2017.

[126] M. Kang, C. W. Jeong, C. Kwak, H. H. Kim, and J. H. Ku, "Preoperative neutrophil-lymphocyte ratio can significantly predict mortality outcomes in patients with non-muscle invasive bladder cancer undergoing transurethral resection of bladder tumor," *Oncotarget*, vol. 8, no. 8, pp. 12891–12901, 2017.

[127] T. Yoshida *et al.*, "Prognostic impact of perioperative lymphocyte–monocyte ratio in patients with bladder cancer undergoing radical cystectomy," *Tumor Biol.*, vol. 37, no. 8, pp. 10067–10074, 2016.

[128] J. Y. Ma, G. Hu, and Q. Liu, "Prognostic significance of the lymphocyte-to-monocyte ratio in bladder cancer undergoing radical cystectomy: A meta-analysis of 5638 individuals," *Dis. Markers*, vol. 2019, 2019.

[129] B. Bhindi *et al.*, "Identification of the best complete blood count-based predictors for bladder cancer outcomes in patients undergoing radical

cystectomy," *Br. J. Cancer*, vol. 114, no. 2, pp. 207–212, 2016.

[130] E. Ojerholm *et al.*, "Neutrophil-to-lymphocyte ratio as a bladder cancer biomarker: assessing prognostic and predictive value in SWOG 8710," vol. 123, no. 5, pp. 794–801, 2018.

[131] L. R. Dixon, "The complete blood count: physiologic basis and clinical usage," *J Perinat Neonatal Nurs.*, vol. 11, no. 3, pp. 1–18, 1997.

[132] W. P. Accomando, J. K. Wiencke, E. A. Houseman, H. H. Nelson, and K. T. Kelsey, "Quantitative reconstruction of leukocyte subsets using DNA methylation," *Genome Biol.*, vol. 15, no. 3, pp. 1–12, 2014.

[133] M. Suelves, E. Carrió, Y. Núñez-Álvarez, and M. A. Peinado, "DNA methylation dynamics in cellular commitment and differentiation," *Brief. Funct. Genomics*, vol. 15, no. 6, pp. 443–453, 2016.

[134] L. A. Salas, J. K. Wiencke, D. C. Koestler, Z. Zhang, B. C. Christensen, and K. T. Kelsey, "Tracing human stem cell lineage during development using DNA methylation," *Genome Res.*, vol. 28, no. 9, pp. 1285–1295, 2018.

[135] U. Baron *et al.*, "DNA methylation analysis as a tool for cell typing," *Epigenetics*, vol. 1, no. 1, pp. 55–60, 2006.

[136] Y. Li *et al.*, "Stability of global methylation profiles of whole blood and extracted DNA under different storage durations and conditions," *Epigenomics*, vol. 10, no. 6, pp. 797–811, 2018.

[137] D. C. Koestler *et al.*, "DNA methylation-derived neutrophil-tolymphocyte ratio: An epigenetic tool to explore cancer inflammation

and outcomes," *Cancer Epidemiol. Biomarkers Prev.*, vol. 26, no. 3, pp. 328–338, 2017.

[138] D. Baris *et al.*, "A case-control study of smoking and bladder cancer risk: Emergent patterns over time," *J. Natl. Cancer Inst.*, vol. 101, no. 22, pp. 1553–1561, 2009.

[139] A. R. Schned, A. S. Andrew, C. J. Marsit, M. S. Zens, K. T. Kelsey, and M. R. Karagas, "Survival Following the Diagnosis of Noninvasive Bladder Cancer: WHO/International Society of Urological Pathology Versus WHO Classification Systems," *J. Urol.*, vol. 178, no. 4, pp. 1196–1200, 2007.

[140] K. T. Kelsey *et al.*, "A population-based study of immunohistochemical detection of p53 alteration in bladder cancer," *Br. J. Cancer*, vol. 90, no. 8, pp. 1572–1576, 2004.

[141] M. R. Karagas, T. D. Tosteson, J. Blum, J. S. Morris, J. A. Baron, and B. Klaue, "Design of an epidemiologic study of drinking water arsenic exposure and skin and bladder cancer risk in a U.S. population," *Environ. Health Perspect.*, vol. 106, no. SUPPL. 4, pp. 1047–1050, 1998.

[142] M. J. Aryee *et al.*, "Minfi: A flexible and comprehensive Bioconductor package for the analysis of Infinium DNA methylation microarrays," *Bioinformatics*, vol. 30, no. 10, pp. 1363–1369, 2014.

[143] Z. Xu, L. Niu, L. Li, and J. A. Taylor, "ENmix: a novel background correction method for Illumina HumanMethylation450 BeadChip.," *Nucleic Acids Res.*, vol. 44, no. 3, p. e20, Feb. 2016.

[144] W. E. Johnson, C. Li, and A. Rabinovic, “Adjusting batch effects in microarray expression data using empirical Bayes methods,” *Biostatistics*, vol. 8, no. 1, pp. 118–127, 2007.

[145] W. Zhou, P. W. Laird, and H. Shen, “Comprehensive characterization, annotation and innovative use of Infinium DNA methylation BeadChip probes,” *Nucleic Acids Res.*, vol. 45, no. 4, p. e22, 2017.

[146] H. KD, “IlluminaHumanMethylationEPICanno.ilm10b4.hg19: Annotation for Illumina’s EPIC methylation arrays.,” *R Packag. version 0.6.0*, 2017.

[147] N. C. Sheffield and C. Bock, “LOLA: Enrichment analysis for genomic region sets and regulatory elements in R and Bioconductor,” *Bioinformatics*, vol. 32, no. 4, pp. 587–589, 2016.

[148] B. Phipson, J. Maksimovic, and A. Oshlack, “MissMethyl: An R package for analyzing data from Illumina’s HumanMethylation450 platform,” *Bioinformatics*, vol. 32, no. 2, pp. 286–288, 2016.

[149] T. J. Peters *et al.*, “De novo identification of differentially methylated regions in the human genome,” *Epigenetics and Chromatin*, vol. 8, no. 1, pp. 1–16, 2015.

[150] W. Zhou, T. J. Triche, P. W. Laird, and H. Shen, “SeSAMe: Reducing artifactual detection of DNA methylation by Infinium BeadChips in genomic deletions,” *Nucleic Acids Res.*, vol. 46, no. 20, pp. 1–15, 2018.

[151] M. Aydın *et al.*, “Correlation of neutrophil-lymphocyte ratio and risk scores in non-muscle invasive bladder cancer,” *Actas Urológicas*

Españolas (English Ed., vol. 43, no. 9, pp. 503–508, 2019.

[152] H. D. Yuk, C. W. Jeong, C. Kwak, H. H. Kim, and J. H. Ku, "Elevated neutrophil to lymphocyte ratio predicts poor prognosis in non-muscle invasive bladder cancer patients: Initial intravesical bacillus Calmette-Guerin treatment after transurethral resection of bladder tumor setting," *Front. Oncol.*, vol. 9, no. JAN, pp. 1–8, 2019.

[153] I. Getzler, Z. Bahouth, O. Nativ, J. Rubinstein, and S. Halachmi, "Preoperative neutrophil to lymphocyte ratio improves recurrence prediction of non-muscle invasive bladder cancer," *BMC Urol.*, vol. 18, no. 1, pp. 1–10, 2018.

[154] Q. Zhang, Q. Lai, S. Wang, Q. Meng, and Z. Mo, "Clinical value of postoperative neutrophil-tolymphocyte ratio change as a detection marker of bladder cancer recurrence," *Cancer Manag. Res.*, vol. 13, pp. 849–860, 2021.

[155] F. Cantiello *et al.*, "Systemic Inflammatory Markers and Oncologic Outcomes in Patients with High-risk Non–muscle-invasive Urothelial Bladder Cancer," *Eur. Urol. Oncol.*, vol. 1, no. 5, pp. 403–410, 2018.

[156] J. A. Witjes and K. Hendricksen, "Intravesical Pharmacotherapy for Non-Muscle-Invasive Bladder Cancer: A Critical Analysis of Currently Available Drugs, Treatment Schedules, and Long-Term Results," *Eur. Urol.*, vol. 53, no. 1, pp. 45–52, 2008.

[157] R. Pichler, J. Fritz, C. Zavadil, G. Schäfer, Z. Culig, and A. Brunner, "Tumor-infiltrating immune cell subpopulations influence the oncologic

outcome after intravesical bacillus calmette-guérin therapy in bladder cancer," *Oncotarget*, vol. 7, no. 26, pp. 39916–39930, 2016.

[158] F. Desgrandchamps *et al.*, "Prediction of non-muscle-invasive bladder cancer recurrence by measurement of checkpoint HLAG's receptor ILT2 on peripheral CD8+ T cells," *Oncotarget*, vol. 9, no. 69, pp. 33160–33169, 2018.

[159] S. Akan *et al.*, "Can the systemic immune inflammation index be a predictor of BCG response in patients with high-risk non-muscle invasive bladder cancer?," *Int. J. Clin. Pract.*, vol. 75, no. 4, pp. 1–9, 2021.

[160] R. G. Douglas, R. H. Alford, T. R. Cate, and R. B. Couch, "The leukocyte response during viral respiratory illness in man.," *Ann. Intern. Med.*, vol. 64, no. 3, pp. 521–530, 1966.

[161] J. E. Woodell-May and S. D. Sommerfeld, "Role of Inflammation and the Immune System in the Progression of Osteoarthritis," *J. Orthop. Res.*, vol. 38, no. 2, pp. 253–257, 2020.

[162] C. Bedel, M. Korkut, and H. H. Armag˘an, "NLR, d-NLR and PLR can be affected by many factors," *Int. Immunopharmacol.*, vol. 90, no. January, p. 107154, 2021.

[163] B. D. Lin *et al.*, "Causes of variation in the neutrophil-lymphocyte and platelet-lymphocyte ratios: A twin-family study," *Biomark. Med.*, vol. 10, no. 10, pp. 1061–1072, 2016.

[164] I. Elisia *et al.*, "Exploratory examination of inflammation state, immune

response and blood cell composition in a human obese cohort to identify potential markers predicting cancer risk," *PLoS One*, vol. 15, no. 2, pp. 1–21, 2020.

[165] F. J. Laso, J. M. Vaquero, J. Almeida, M. Marcos, and A. Orfao, "Chronic alcohol consumption is associated with changes in the distribution, immunophenotype, and the inflammatory cytokine secretion profile of circulating dendritic cells," *Alcohol. Clin. Exp. Res.*, vol. 31, no. 5, pp. 846–854, 2007.

[166] V. Grossmannm *et al.*, "Profile of the immune and inflammatory response in individuals with prediabetes and type 2 diabetes," *Diabetes Care*, vol. 38, no. 7, pp. 1356–1364, 2015.

[167] C. H. Graham *et al.*, "Innate immune memory is associated with increased disease-free survival in bladder cancer patients treated with bacillus Calmette-Guérin," *Can. Urol. Assoc. J.*, vol. 15, no. 8, pp. 1–13, 2021.

[168] Q. Luo *et al.*, "Slingshot homolog-1 expression is a poor prognostic factor of pT1 bladder urothelial carcinoma after transurethral resection," *World J. Urol.*, vol. 38, no. 11, pp. 2849–2856, 2020.

[169] E. Abdelzaher and A. F. Kotb, "High Coexpression of Runt-related Transcription Factor 2," vol. 24, no. 5, 2016.

[170] B. Liu, S. Pan, J. Liu, and C. Kong, "Cancer-associated fibroblasts and the related Runt-related transcription factor 2 (RUNX2) promote bladder cancer progression," *Gene*, vol. 775, no. August 2020, p. 145451, 2021.

[171] L. Li *et al.*, “A comprehensive enhancer screen identifies TRAM2 as a key and novel mediator of YAP oncogenesis,” *Genome Biol.*, vol. 22, no. 1, pp. 1–28, 2021.

[172] W. Deng *et al.*, “MICAL1 controls cell invasive phenotype via regulating oxidative stress in breast cancer cells,” *BMC Cancer*, vol. 16, no. 1, pp. 1–11, 2016.

[173] K. Chen, R. Zhao, G. Yao, Z. Liu, R. Shi, and J. Geng, “Overexpression of kin of IRRE-Like protein 1 (KIRREL) as a prognostic biomarker for breast cancer,” *Pathol. Res. Pract.*, vol. 216, no. 7, p. 153000, 2020.

[174] I. Lee, S. Y. Yeom, S. J. Lee, W. K. Kang, and C. Park, “A novel senescence-evasion mechanism involving Grap2 and cyclin D interacting protein inactivation by Ras associated with diabetes in cancer cells under doxorubicin treatment,” *Cancer Res.*, vol. 70, no. 11, pp. 4357–4365, 2010.

[175] M. A. Weniger *et al.*, “Gains of the proto-oncogene BCL11A and nuclear accumulation of BCL11AXL protein are frequent in primary mediastinal B-cell lymphoma [1],” *Leukemia*, vol. 20, no. 10, pp. 1880–1882, 2006.

[176] E. Satterwhite *et al.*, “The BCL11 gene family: Involvement of BCL11A in lymphoid malignancies,” *Blood*, vol. 98, no. 12, pp. 3413–3420, 2001.

[177] J. M. A. Moreira *et al.*, “Bladder cancer-associated protein, a potential prognostic biomarker in human bladder cancer,” *Mol. Cell. Proteomics*, vol. 9, no. 1, pp. 161–177, 2010.

[178] I. Gromova, S. Svensson, P. Gromov, and J. M. A. Moreira,

"Identification of BLCAP as a novel STAT3 interaction partner in bladder cancer," *PLoS One*, vol. 12, no. 11, pp. 1–18, 2017.

[179] D. I. Jacobs, Y. Mao, A. Fu, W. K. Kelly, and Y. Zhu, "Dysregulated methylation at imprinted genes in prostate tumor tissue detected by methylation microarray," *BMC Urol.*, vol. 13, no. 1, p. 1, 2013.

[180] R. L. Siegel, K. D. Miller, H. E. Fuchs, and A. Jemal, "Cancer statistics, 2022.," *CA. Cancer J. Clin.*, vol. 72, no. 1, pp. 7–33, Jan. 2022.

[181] K. Saginala, A. Barsouk, J. S. Aluru, P. Rawla, S. A. Padala, and A. Barsouk, "Epidemiology of Bladder Cancer," *Med. Sci.*, vol. 8, no. 1, p. 15, 2020.

[182] W. Lin, X. Pan, C. Zhang, B. Ye, and J. Song, "Impact of Age at Diagnosis of Bladder Cancer on Survival: A Surveillance, Epidemiology, and End Results-Based Study 2004-2015," *Cancer Control*, vol. 30, pp. 1–14, 2023.

[183] L. Hou *et al.*, "Association of smoking status with prognosis in bladder cancer: A meta-analysis," *Oncotarget*, vol. 8, no. 1, pp. 1278–1289, 2017.

[184] S. Antoni, J. Ferlay, I. Soerjomataram, A. Znaor, A. Jemal, and F. Bray, "Bladder Cancer Incidence and Mortality: A Global Overview and Recent Trends," *Eur. Urol.*, vol. 71, no. 1, pp. 96–108, 2017.

[185] J. A. Witjes, "Management of BCG Failures in Superficial Bladder Cancer: A Review," *Eur. Urol.*, vol. 49, no. 5, pp. 790–797, 2006.

[186] V. Soukup *et al.*, "Prognostic Performance and Reproducibility of the

1973 and 2004/2016 World Health Organization Grading Classification Systems in Non–muscle-invasive Bladder Cancer: A European Association of Urology Non-muscle Invasive Bladder Cancer Guidelines Panel Syst," *Eur. Urol.*, vol. 72, no. 5, pp. 801–813, 2017.

[187] C. Te Lin *et al.*, "Prognostic relevance of preoperative circulating CD8-positive lymphocytes in the urinary bladder recurrence of urothelial carcinoma," *Urol. Oncol. Semin. Orig. Investig.*, vol. 30, no. 5, pp. 680–687, 2012.

[188] J. Q. Chen *et al.*, "Immune profiles and DNA methylation alterations related with non-muscle-invasive bladder cancer outcomes," *Clin. Epigenetics*, vol. 14, no. 1, pp. 1–14, 2022.

[189] D. Y. Oh *et al.*, "Intratumoral CD4+ T Cells Mediate Anti-tumor Cytotoxicity in Human Bladder Cancer," *Cell*, vol. 181, no. 7, pp. 1612-1625.e13, 2020.

[190] P. A. Dugue *et al.*, "Biological Aging Measures Based on Blood DNA Methylation and Risk of Cancer: A Prospective Study," *JNCI Cancer Spectr.*, vol. 5, no. 1, pp. 1–8, 2021.

[191] M. E. Levine, H. D. Hosgood, B. Chen, D. Absher, T. Assimes, and S. Horvath, "DNA methylation age of blood predicts future onset of lung cancer in the women's health initiative," *Aging (Albany. NY).*, vol. 7, no. 9, pp. 690–700, 2015.

[192] M. Chung *et al.*, "DNA methylation ageing clocks and pancreatic cancer risk: pooled analysis of three prospective nested case-control studies,"

Epigenetics, vol. 16, no. 12, pp. 1306–1316, 2021.

[193] C. Xiao *et al.*, "Association of Epigenetic Age Acceleration With Risk Factors, Survival, and Quality of Life in Patients With Head and Neck Cancer," *Int. J. Radiat. Oncol. Biol. Phys.*, vol. 111, no. 1, pp. 157–167, 2021.

[194] C. Xiao *et al.*, "Epigenetic age acceleration, fatigue, and inflammation in patients undergoing radiation therapy for head and neck cancer: A longitudinal study," *Cancer*, vol. 127, no. 18, pp. 3361–3371, 2021.

[195] P. A. Dugué *et al.*, "DNA methylation-based biological aging and cancer risk and survival: Pooled analysis of seven prospective studies," *Int. J. Cancer*, vol. 142, no. 8, pp. 1611–1619, 2018.

[196] J. Taylor and G. Kuchel, "Bladder cancer in the elderly: clinical outcomes, basic mechanisms, and future research direction," *Nat Clin Pr. Urol*, vol. 6, no. 3, pp. 135–144, 2009.

[197] A. M. Molinaro, K. Lostritto, and M. van der Laan, "partDSA: Deletion/substitution/addition algorithm for partitioning the covariate space in prediction," *Bioinformatics*, vol. 26, no. 10, pp. 1357–1363, 2010.

[198] D. C. Koestler *et al.*, "Semi-supervised recursively partitioned mixture models for identifying cancer subtypes," *Bioinformatics*, vol. 26, no. 20, pp. 2578–2585, 2010.

[199] N. A. Skopp, D. J. Smolenski, D. A. Schwesinger, C. J. Johnson, M. J. Metzger-Abamukong, and M. A. Reger, "Evaluation of a methodology to

validate National Death Index retrieval results among a cohort of U.S. service members.," *Ann. Epidemiol.*, vol. 27, no. 6, pp. 397–400, Jun. 2017.

[200] A. E. Teschendorff *et al.*, "A beta-mixture quantile normalization method for correcting probe design bias in Illumina Infinium 450 k DNA methylation data," *Bioinformatics*, vol. 29, no. 2, pp. 189–196, 2013.

[201] R. Pidsley, C. C. Y Wong, M. Volta, K. Lunnon, J. Mill, and L. C. Schalkwyk, "A data-driven approach to preprocessing Illumina 450K methylation array data," *BMC Genomics*, vol. 14, no. 1, 2013.

[202] K. Lostritto, R. L. Strawderman, and A. M. Molinaro, "A Partitioning Deletion/Substitution/Addition Algorithm for Creating Survival Risk Groups," *Biometrics*, vol. 68, no. 4, pp. 1146–1156, 2012.

[203] G. M. Poage *et al.*, "Genetic and epigenetic somatic alterations in head and neck squamous cell carcinomas are globally coordinated but not locally targeted," *PLoS One*, vol. 5, no. 3, 2010.

[204] B. C. Christensen *et al.*, "DNA methylation, isocitrate dehydrogenase mutation, and survival in glioma," *J. Natl. Cancer Inst.*, vol. 103, no. 2, pp. 143–153, 2011.

[205] M. Avissar-Whiting, D. C. Koestler, E. A. Houseman, B. C. Christensen, K. T. Kelsey, and C. J. Marsit, "Polycomb group genes are targets of aberrant DNA methylation in renal cell carcinoma," *Epigenetics*, vol. 6, no. 6, pp. 703–709, 2011.

[206] H. A. Yıldız, M. D. Değer, and G. Aslan, "Prognostic value of

preoperative inflammation markers in non-muscle invasive bladder cancer," *Int. J. Clin. Pract.*, vol. 75, no. 6, pp. 1–7, 2021.

[207] M. Ferro *et al.*, "Absolute basophil count is associated with time to recurrence in patients with high-grade T1 bladder cancer receiving bacillus Calmette–Guérin after transurethral resection of the bladder tumor," *World J. Urol.*, vol. 38, no. 1, pp. 143–150, 2020.

[208] R. Murai *et al.*, "Prediction of intravesical recurrence of nonmuscle-invasive bladder cancer by evaluation of intratumoral Foxp3 + T cells in the primary transurethral resection of bladder tumor specimens," *PLoS One*, vol. 13, no. 9, pp. 1–12, 2018.

[209] S. Oda *et al.*, "Expression of Spred2 in the urothelial tumorigenesis of the urinary bladder," *PLoS One*, vol. 16, no. 11 November, pp. 1–17, 2021.

[210] S. Lv, W. Wang, H. Wang, Y. Zhu, and C. Lei, "PPARγ activation serves as therapeutic strategy against bladder cancer via inhibiting PI3K-Akt signaling pathway," *BMC Cancer*, vol. 19, no. 1, pp. 1–13, 2019.

[211] M. Sun *et al.*, "Fibrous sheath interacting protein 1 overexpression is associated with unfavorable prognosis in bladder cancer: A potential therapeutic target," *Onco. Targets. Ther.*, vol. 10, pp. 3949–3956, 2017.

[212] M. Sun, Z. Chen, S. Tan, C. Liu, and W. Zhao, "Knockdown of fibrous sheath interacting protein 1 expression reduces bladder urothelial carcinoma cell proliferation and induces apoptosis via inhibition of the PI3K/AKT pathway," *Onco. Targets. Ther.*, vol. 11, pp. 1961–1971,

2018.

[213] J. Liu, X. Yu, B. Liu, H. Yu, and Z. Li, “Phosphorylated mapk14 promotes the proliferation and migration of bladder cancer cells by maintaining runx2 protein abundance,” *Cancer Manag. Res.*, vol. 12, pp. 11371–11382, 2020.

[214] J. H. Van Puffelen *et al.*, “Intravesical BCG in patients with non-muscle invasive bladder cancer induces trained immunity and decreases respiratory infections,” *J. Immunother. Cancer*, vol. 11, no. 1, pp. 1–19, 2023.

[215] J. Das, D. Verma, M. Gustafsson, and M. Lerm, “Identification of DNA methylation patterns predisposing for an efficient response to BCG vaccination in healthy BCG-naïve subjects,” *Epigenetics*, vol. 14, no. 6, pp. 589–601, 2019.

[216] K. Wu *et al.*, “Cisplatin inhibits the progression of bladder cancer by selectively depleting G-MDSCs: A novel chemoimmunomodulating strategy.,” *Clin. Immunol.*, vol. 193, pp. 60–69, Aug. 2018.

[217] E. Eruslanov *et al.*, “Circulating and tumor-infiltrating myeloid cell subsets in patients with bladder cancer.,” *Int. J. cancer*, vol. 130, no. 5, pp. 1109–1119, Mar. 2012.

[218] S. Jiang and G. Redelman-Sidi, “BCG in Bladder Cancer Immunotherapy.,” *Cancers (Basel).*, vol. 14, no. 13, Jun. 2022.

[219] J. W. Davis, S. I. Sheth, M. J. Doviak, and P. F. Schellhammer, “Superficial bladder carcinoma treated with bacillus Calmette-Guerin:

progression-free and disease specific survival with minimum 10-year followup.," *J. Urol.*, vol. 167, no. 2 Pt 1, pp. 494–500; discussion 501, Feb. 2002.

[220] D. L. Suzman *et al.*, "FDA Approval Summary: Atezolizumab or Pembrolizumab for the Treatment of Patients with Advanced Urothelial Carcinoma Ineligible for Cisplatin-Containing Chemotherapy.," *Oncologist*, vol. 24, no. 4, pp. 563–569, Apr. 2019.

[221] T. Zhang *et al.*, "A randomized phase 2 trial of pembrolizumab versus pembrolizumab and acalabrutinib in patients with platinum-resistant metastatic urothelial cancer.," *Cancer*, vol. 126, no. 20, pp. 4485–4497, Oct. 2020.

[222] K. C. Yuen *et al.*, "High systemic and tumor-associated IL-8 correlates with reduced clinical benefit of PD-L1 blockade.," *Nat. Med.*, vol. 26, no. 5, pp. 693–698, May 2020.

[223] M. Eckstein *et al.*, "Cytotoxic T-cell-related gene expression signature predicts improved survival in muscle-invasive urothelial bladder cancer patients after radical cystectomy and adjuvant chemotherapy.," *J. Immunother. cancer*, vol. 8, no. 1, May 2020.

[224] M. Peng, "Immune landscape of distinct subtypes in urothelial carcinoma based on immune gene profile.," *Front. Immunol.*, vol. 13, p. 970885, 2022.

[225] J. N. Kather *et al.*, "Topography of cancer-associated immune cells in human solid tumors.," *Elife*, vol. 7, Sep. 2018.

[226] J. K. Wiencke *et al.*, “DNA methylation as a pharmacodynamic marker of glucocorticoid response and glioma survival.,” *Nat. Commun.*, vol. 13, no. 1, p. 5505, Sep. 2022.

[227] M. E. Muse, C. D. Carroll, L. A. Salas, M. R. Karagas, and B. C. Christensen, “Application of Novel Breast Biospecimen Cell-Type Adjustment Identifies Shared DNA Methylation Alterations in Breast Tissue and Milk with Breast Cancer-Risk Factors.,” *Cancer Epidemiol. biomarkers Prev. a Publ. Am. Assoc. Cancer Res. cosponsored by Am. Soc. Prev. Oncol.*, vol. 32, no. 4, pp. 550–560, Apr. 2023.

[228] H. H. Nelson, J. K. Wiencke, L. Gunn, J. C. Wain, D. C. Christiani, and K. T. Kelsey, “Chromosome 3p14 alterations in lung cancer: evidence that FHIT exon deletion is a target of tobacco carcinogens and asbestos.,” *Cancer Res.*, vol. 58, no. 9, pp. 1804–1807, May 1998.

[229] S. Monti, P. Tamayo, J. Mesirov, and T. Golub, “Consensus Clustering A Resampling-Based Method for Class Discovery and Vi - Monti - Mach Learn,” *Mach. Learn.*, vol. 52, no. i, pp. 91–118, 2003.

[230] M. D. Wilkerson and D. N. Hayes, “ConsensusClusterPlus: a class discovery tool with confidence assessments and item tracking.,” *Bioinformatics*, vol. 26, no. 12, pp. 1572–1573, Jun. 2010.

[231] L. Vicente-Gonzalez and J. L. Vicente-Villardon, “PERMANOVA: Multivariate Analysis of Variance Based on Distances and Permutations.” 2021.

[232] M. Templ, K. Hron, and P. Filzmoser, “robCompositions: an R-package

for robust statistical analysis of compositional data," in *Compositional Data Analysis: Theory and Applications*, 2011, pp. 341–355.

[233] J. A. Joyce and D. T. Fearon, "T cell exclusion, immune privilege, and the tumor microenvironment.," *Science*, vol. 348, no. 6230, pp. 74–80, Apr. 2015.

[234] E. Lanitis, D. Dangaj, M. Irving, and G. Coukos, "Mechanisms regulating T-cell infiltration and activity in solid tumors.," *Ann. Oncol. Off. J. Eur. Soc. Med. Oncol.*, vol. 28, no. suppl_12, pp. xii18–xii32, Dec. 2017.

[235] S. TANGYE and F. MACKAY, "CHAPTER 11 - B Cells and Autoimmunity," *The Autoimmune Diseases (Fourth Edition)*. pp. 139–156, 2006.

[236] D. Allman and S. Pillai, "Peripheral B cell subsets.," *Curr. Opin. Immunol.*, vol. 20, no. 2, pp. 149–157, Apr. 2008.

[237] J. Sprent, "Circulating T and B lymphocytes of the mouse. I. Migratory properties.," *Cell. Immunol.*, vol. 7, no. 1, pp. 10–39, Apr. 1973.

[238] U. H. von Andrian and C. R. Mackay, "T-cell function and migration. Two sides of the same coin.," *N. Engl. J. Med.*, vol. 343, no. 14, pp. 1020–1034, Oct. 2000.

[239] L. Wu, S. Saxena, and R. K. Singh, "Neutrophils in the Tumor Microenvironment.," *Adv. Exp. Med. Biol.*, vol. 1224, pp. 1–20, 2020.

[240] R. Zahorec, "Neutrophil-to-lymphocyte ratio, past, present and future perspectives.," *Bratisl. Lek. Listy*, vol. 122, no. 7, pp. 474–488, 2021.

[241] M. D. Vartolomei *et al.*, “Prognostic role of pretreatment neutrophil-to-lymphocyte ratio (NLR) in patients with non–muscle-invasive bladder cancer (NMIBC): A systematic review and meta-analysis,” *Urol. Oncol. Semin. Orig. Investig.*, vol. 36, no. 9, pp. 389–399, 2018.

[242] I. Kerim, A. Ammad, H. Mahmood, and F. Q. J. Al-zayadi, “Comparative Analysis of Tumor Infiltrating T Cells and Serological Markers between MIBC and NMIBC Patients,” *Medico-Legal Updat.*, vol. 20, no. 4, pp. 770–775, 2020.

[243] J.-Q. Chen *et al.*, “Genome-scale methylation analysis identifies immune profiles and age acceleration associations with bladder cancer outcomes.,” *Cancer Epidemiol. biomarkers Prev. a Publ. Am. Assoc. Cancer Res. cosponsored by Am. Soc. Prev. Oncol.*, Aug. 2023.

[244] X. He *et al.*, “Clinical Outcomes and Immune Metrics in Intratumoral Basophil-Enriched Gastric Cancer Patients.,” *Ann. Surg. Oncol.*, vol. 28, no. 11, pp. 6439–6450, Oct. 2021.

[245] R. Poto *et al.*, “Basophils from allergy to cancer.,” *Front. Immunol.*, vol. 13, p. 1056838, 2022.

[246] J. Zhang *et al.*, “Basophils as a potential therapeutic target in cancer.,” *J. Zhejiang Univ. Sci. B*, vol. 22, no. 12, pp. 971–984, Dec. 2021.

[247] K.-C. Lin *et al.*, “Tumor Necrosis with Adjunction of Preoperative Monocyte-to-Lymphocyte Ratio as a New Risk Stratification Marker Can Independently Predict Poor Outcomes in Upper Tract Urothelial Carcinoma.,” *J. Clin. Med.*, vol. 10, no. 13, Jul. 2021.

[248] X. Zhu *et al.*, “The evaluation of monocyte lymphocyte ratio as a preoperative predictor in urothelial malignancies: a pooled analysis based on comparative studies.,” *Sci. Rep.*, vol. 9, no. 1, p. 6280, Apr. 2019.

[249] J. Liu *et al.*, “Preoperative Monocyte-to-lymphocyte Ratio Predicts for Intravesical Recurrence in Patients With Urothelial Carcinoma of the Upper Urinary Tract After Radical Nephroureterectomy Without a History of Bladder Cancer.,” *Clin. Genitourin. Cancer*, vol. 19, no. 3, pp. e156–e165, Jun. 2021.

[250] H. Shi *et al.*, “A high monocyte-to-lymphocyte ratio predicts poor prognosis in patients with radical cystectomy for bladder cancer.,” *Transl. Cancer Res.*, vol. 9, no. 9, pp. 5255–5267, Sep. 2020.

[251] T. Chanmee, P. Ontong, K. Konno, and N. Itano, “Tumor-associated macrophages as major players in the tumor microenvironment.,” *Cancers (Basel).*, vol. 6, no. 3, pp. 1670–1690, Aug. 2014.

[252] J. W. Pollard, “Tumour-educated macrophages promote tumour progression and metastasis.,” *Nature reviews. Cancer*, vol. 4, no. 1. England, pp. 71–78, Jan-2004.

[253] Q. Zhang *et al.*, “Prognostic significance of tumor-associated macrophages in solid tumor: a meta-analysis of the literature.,” *PLoS One*, vol. 7, no. 12, p. e50946, 2012.

[254] B.-Z. Qian *et al.*, “CCL2 recruits inflammatory monocytes to facilitate breast-tumour metastasis.,” *Nature*, vol. 475, no. 7355, pp. 222–225, Jun.

2011.

[255] J. Han, X. Gu, Y. Li, and Q. Wu, “Mechanisms of BCG in the treatment of bladder cancer-current understanding and the prospect.,” *Biomed. Pharmacother.*, vol. 129, p. 110393, Sep. 2020.

[256] J. H. van Puffelen *et al.*, “Intravesical BCG in patients with non-muscle invasive bladder cancer induces trained immunity and decreases respiratory infections.,” *J. Immunother. cancer*, vol. 11, no. 1, Jan. 2023.

[257] K. Taniguchi *et al.*, “Systemic immune response after intravesical instillation of bacille Calmette-Guérin (BCG) for superficial bladder cancer.,” *Clin. Exp. Immunol.*, vol. 115, no. 1, pp. 131–135, Jan. 1999.

[258] A. Ratna and P. Mandrekar, “Alcohol and Cancer: Mechanisms and Therapies.,” *Biomolecules*, vol. 7, no. 3, Aug. 2017.

[259] L. Dyck *et al.*, “Suppressive effects of the obese tumor microenvironment on CD8 T cell infiltration and effector function.,” *J. Exp. Med.*, vol. 219, no. 3, Mar. 2022.

[260] K. D. Sievert *et al.*, “Economic aspects of bladder cancer: what are the benefits and costs?,” *World J. Urol.*, vol. 27, no. 3, pp. 295–300, Jun. 2009.

[261] J. Daza, A. Charap, P. N. Wiklund, and J. P. Sfakianos, “Role of the Innate Immune System in the Development, Progression, and Therapeutic Response of Bladder Cancer.,” *Eur. Urol. Focus*, vol. 6, no. 4, pp. 650–652, Jul. 2020.

[262] M. Joseph and D. Enting, “Immune Responses in Bladder Cancer-Role

of Immune Cell Populations, Prognostic Factors and Therapeutic Implications.," *Front. Oncol.*, vol. 9, p. 1270, 2019.

[263] S. Sallami, H. Khouni, M. Ben Atta, S. Abou El Makarim, M. B. Zouari, and S. Ben Rhouma, "Long-term outcome of intravesical bacillus Calmette-Guérin therapy with maintenance for urinary bladder carcinoma in situ: About 47 cases.," *Tunis. Med.*, vol. 94, no. 12, p. 844, Dec. 2016.

[264] S. Mugiya *et al.*, "Long-term outcome of a low-dose intravesical bacillus Calmette-Guerin therapy for carcinoma in situ of the bladder: results after six successive instillations of 40 mg BCG.," *Jpn. J. Clin. Oncol.*, vol. 35, no. 7, pp. 395–399, Jul. 2005.

[265] L. A. Emens, "Chemotherapy and tumor immunity: an unexpected collaboration.," *Front. Biosci.*, vol. 13, pp. 249–257, Jan. 2008.

[266] J. M. Hamilton-Reeves *et al.*, "Effects of Immunonutrition for Cystectomy on Immune Response and Infection Rates: A Pilot Randomized Controlled Clinical Trial.," *Eur. Urol.*, vol. 69, no. 3, pp. 389–392, Mar. 2016.

[267] D. Sarkar, M. K. Jung, and H. J. Wang, "Alcohol and the Immune System.," *Alcohol Research : Current Reviews*, vol. 37, no. 2. pp. 153–155, 2015.

[268] Y. Pan *et al.*, "Incidence of Bladder Cancer in Type 2 Diabetes Mellitus Patients: A Population-Based Cohort Study.," *Medicina (Kaunas).*, vol. 56, no. 9, Aug. 2020.

[269] G. Mowatt *et al.*, “Systematic review of the clinical effectiveness and cost-effectiveness of photodynamic diagnosis and urine biomarkers (FISH, ImmunoCyt, NMP22) and cytology for the detection and follow-up of bladder cancer.,” *Health Technol. Assess.*, vol. 14, no. 4, pp. 1–331, iii–iv, Jan. 2010.

[270] X. Chen *et al.*, “Urine DNA methylation assay enables early detection and recurrence monitoring for bladder cancer.,” *J. Clin. Invest.*, vol. 130, no. 12, pp. 6278–6289, Dec. 2020.

[271] Q. Fang, X. Zhang, Q. Nie, J. Hu, S. Zhou, and C. Wang, “Improved urine DNA methylation panel for early bladder cancer detection.,” *BMC Cancer*, vol. 22, no. 1, p. 237, Mar. 2022.

[272] H. D. Pharo *et al.*, “BladMetrix: a novel urine DNA methylation test with high accuracy for detection of bladder cancer in hematuria patients.,” *Clin. Epigenetics*, vol. 14, no. 1, p. 115, Sep. 2022.

[273] S. N. Abraham and Y. Miao, “The nature of immune responses to urinary tract infections.,” *Nat. Rev. Immunol.*, vol. 15, no. 10, pp. 655–663, Oct. 2015.

[274] E. Castellano *et al.*, “CyTOF analysis identifies unusual immune cells in urine of BCG-treated bladder cancer patients.,” *Front. Immunol.*, vol. 13, p. 970931, 2022.

[275] H. Jiang, Y. Zhou, Y. Lin, R. C. K. Chan, J. Liu, and H. Chen, “Deep learning for computational cytology: A survey.,” *Med. Image Anal.*, vol. 84, p. 102691, Feb. 2023.

[276] M. Tsuneki, M. Abe, and F. Kanavati, “Deep Learning-Based Screening of Urothelial Carcinoma in Whole Slide Images of Liquid-Based Cytology Urine Specimens.,” *Cancers (Basel).*, vol. 15, no. 1, Dec. 2022.

[277] J. J. Levy *et al.*, “Large-scale validation study of an improved semiautonomous urine cytology assessment tool: AutoParis-X.,” *Cancer Cytopathol.*, Jun. 2023.

[278] J. J. Levy *et al.*, “Examining longitudinal markers of bladder cancer recurrence through a semiautonomous machine learning system for quantifying specimen atypia from urine cytology.,” *Cancer Cytopathol.*, Jun. 2023.

[279] M. I. Jordan and T. M. Mitchell, “Machine learning: Trends, perspectives, and prospects.,” *Science*, vol. 349, no. 6245, pp. 255–260, Jul. 2015.

[280] I. H. Sarker, “Deep Learning: A Comprehensive Overview on Techniques, Taxonomy, Applications and Research Directions.,” *SN Comput. Sci.*, vol. 2, no. 6, p. 420, 2021.

[281] A. Echle, N. T. Rindtorff, T. J. Brinker, T. Luedde, A. T. Pearson, and J. N. Kather, “Deep learning in cancer pathology: a new generation of clinical biomarkers.,” *Br. J. Cancer*, vol. 124, no. 4, pp. 686–696, Feb. 2021.

[282] W. Barrios *et al.*, “Bladder cancer prognosis using deep neural networks and histopathology images.,” *J. Pathol. Inform.*, vol. 13, p. 100135, 2022.

[283] J. Pan *et al.*, “An artificial intelligence model for the pathological diagnosis of invasion depth and histologic grade in bladder cancer.,” *J. Transl. Med.*, vol. 21, no. 1, p. 42, Jan. 2023.