

RESEARCH

Open Access



# Assessment of immune cell profiles among post-menopausal women in the Women's Health Initiative using DNA methylation-based methods

Emily Nissen<sup>1</sup>, Alexander Reiner<sup>2</sup>, Simin Liu<sup>3</sup>, Robert B. Wallace<sup>4</sup>, Annette M. Molinaro<sup>5</sup>, Lucas A. Salas<sup>6</sup>, Brock C. Christensen<sup>6,7,8</sup>, John K. Wiencke<sup>5,9</sup>, Devin C. Koestler<sup>1†</sup> and Karl T. Kelsey<sup>10\*†</sup>

## Abstract

**Background** Over the past decade, DNA methylation (DNAm)-based deconvolution methods that leverage cell-specific DNAm markers of immune cell types have been developed to provide accurate estimates of the proportions of leukocytes in peripheral blood. Immune cell phenotyping using DNAm markers, termed immunomethylomics or methylation cytometry, offers a solution for determining the body's immune cell landscape that does not require fresh blood and is scalable to large sample sizes. Despite significant advances in DNAm-based deconvolution, references at the population level are needed for clinical and research interpretation of these additional immune layers. Here we aim to provide some references for immune populations in a group of multi-ethnic post-menopausal American women.

**Results** We applied DNAm-based deconvolution to a large sample of post-menopausal women enrolled in the Women's Health Initiative (baseline,  $N = 58$ ) or the ancillary Long Life Study (WHI-LLS,  $N = 1237$ ) to determine the reference ranges of 58 immune parameters, including proportions and absolute counts for 19 leukocyte subsets and 20 derived cell ratios. Participants were 50–94 years old at the time of blood draw, and  $N = 898$  (69.3%) self-identified as White. Using linear regression models, we observed significant associations between age at blood draw and absolute counts and proportions of naïve B, memory CD4+, naïve CD4+, naïve CD8+, memory CD8+ memory, neutrophils, and natural killer cells. We also assessed the same immune profiles in a subset of paired longitudinal samples collected 14–18 years apart across  $N = 52$  participants. Our results demonstrate high inter-individual variability in rates of change of leukocyte subsets over this time. And, when conducting paired t tests to test the difference in counts and proportions between the baseline visit and LLS visit, there were significant changes in naïve B, memory CD4+, naïve CD4+, naïve CD8+, memory CD8+ cells and neutrophils, similar to the results seen when analyzing the association with age in the entire cohort.

**Conclusions** Here, we show that derived cell counts largely reflect the immune profile associated with proportions and that these novel methods replicate the known immune profiles associated with age. Further, we demonstrate the value this methylation cytometry approach can add as a potential application in epidemiological studies.

<sup>†</sup>Devin C. Koestler and Karl T. Kelsey have contributed equally

\*Correspondence:

Karl T. Kelsey

karl\_kelsey@brown.edu

Full list of author information is available at the end of the article



**Keywords** Population, WHI, Immune cell reference limits, Reference values, Immune cell phenotyping, Methylation cytometry, Immunomethylomics, DNA methylation, Deconvolution

## Background

Peripheral blood immune cell counts are the entry point for interrogation of the immune system either in response to an environmental insult or therapeutic intervention or in relation to disease states. However, it is important to have well-established reference values for cell types in healthy populations to interpret immune cell counts in clinically relevant populations. Reference values of various leukocyte immune parameters (i.e., cell-type subsets and cell ratios) are numerous in the literature for specific populations/cohorts of individuals across different age groups, races, ethnicities, nationalities, and sexes [1–19].

Some reference values are reported using a complete blood count (CBC), including differential assessed using hematology analyzers, which returns data on both red blood cells and white blood cells (WBC), enumerating five WBC subtypes (neutrophils, eosinophils, basophils, monocytes, and lymphocytes). For example, Cheng et al. [18] used NHANESIII data to develop reference values based on CBC and reported that the granulocyte fraction appears to increase with age, with the lymphocyte fraction showing an inverse association with age. Cheng et al. [18] also reported race-related differences, such as overall WBC count lower in both non-Hispanic black males and females, compared to non-Hispanic white and Mexican American males and females. This trend is a commonly accepted physiologic norm, driven by the inheritance of the Duffy antigen variant [20]. Mononuclear and lymphocyte percentages were increased, and granulocyte percentage was decreased in non-Hispanic African Americans compared to non-Hispanic whites and Mexican Americans. These findings were largely replicated by Coates et al. [7], which included CBC data collected on 7,157 healthy volunteers in the UK. While informative, CBC does not differentiate lymphocyte lineage subtypes, requiring flow cytometry approaches for some specific clinical applications. For instance, CD4+ counts are important in HIV monitoring [21, 22], CD4+ and CD8+ counts have been associated with COVID-19 severity and progression [23, 24] and natural killer cells have been implicated in various autoimmune diseases [25].

Numerous studies have assessed reference values for leukocyte percentages and leukocyte counts using flow cytometry ([1–17, 19] as above). Among the largest and most recent was that of Thyagarajan et al. [1], who studied isolated peripheral blood mononuclear cells (PBMCs) in 8848 participants in the Health and Retirement Study. This population is roughly representative

of the US population over 55 years of age. They found total T cells and CD4+ cells declined markedly with age, while CD8+ cells declined with age, though to a lesser degree than CD4+ cells. These findings are independent of cytomegalovirus (CMV) positivity or sex. CD4+ and CD8+ naïve cells were observed to be strongly inversely related to age, and these findings are also independent of CMV positivity. CMV positivity was associated with increases in total T cells and CD8+ cells. Of course, this approach cannot estimate granulocyte prevalence since these cells are lost in the preparation of the sample. One of the few studies to include granulocytes [3] studied 608 Germans and found the number of neutrophils in women to decrease with age, while the relationship between lymphocyte subtypes and age was largely consistent with Thyagarajan et al. [1].

All of the prior studies are beneficial; however, there is opportunity to apply newer research and knowledge to expand upon them. The CBC studies assess only five cell types, and most of the flow cytometry studies are quite small, limiting their power, and they also are conducted primarily on PBMCs where the isolation procedures can introduce additional variability.

DNA methylation (DNAm) deconvolution methods are used to estimate the proportions of leukocyte subsets from peripheral blood. Numerous deconvolution methods have been developed [26–29] with recent efforts focused on expanding the repertoire of immune cell types that can be accurately deconvolved. The application of DNAm deconvolution to estimate immune cell subsets, which we term immunomethylomics or methylation cytometry, is not typically used to establish population-specific reference ranges in large epidemiological studies. In an effort to show the potential use in population research of our recently developed extended DNAm deconvolution library [27], we applied deconvolution methods to Illumina EPIC array DNAm data collected on multi-ethnic participants in a subset of the original Women's Health Initiative (WHI) participants and a subset of participants in respectively. While Bas has a correlation) in samples that included concurrent CBC, allowing us to estimate both cell counts and cell proportions. This group provides an important population to study factors of immunosenescence and healthy aging, as it is a sample of women ranging in age from 50 to 94 years. With this study, we have implemented highly detailed immune cell phenotyping that includes

58 different immune measures for 1295 individuals to define population-specific reference ranges of DNAm deconvolution estimates and their associations with age and self-reported race.

## Materials/methods

### Study population

The Women's Health Initiative (WHI) enrolled post-menopausal women nationwide between the ages of 50 and 79 from 1993–1998 to study common causes of morbidity and mortality. There were ~160,000 enrolled in one of the clinical trials or an observational study [30]. The Long Life Study (LLS) is an ancillary study (WHI W64) that occurred from 2012 to 2014, about 12–14 years after initial enrollment in WHI. There were 7875 women who completed a one-time in-person visit in which a blood draw was taken, along with clinical and functional status assessments [31]. For a fraction of this LLS group ( $N \sim 1300$ ), blood samples were used for measuring DNA methylation (DNAm) using the Illumina EPIC array. For some women, blood samples from the original baseline screening for participation in the WHI also had measured DNAm on the EPIC array ( $N = 58$ ). After quality control, 37 samples were removed, yielding a total of 1237 samples from the LLS visit and 58 from the baseline visit. Of these baseline samples, 52 had a matching sample at the LLS visit, meaning there were 104 paired samples and 1191 unique samples. All covariate data was downloaded from the Women's Health Initiative website (<https://www.whi.org/datasets>).

### DNA methylation quantification

DNA was extracted from whole blood samples using Five Prime (5 Prime, Inc., Gaithersburg, MD) kits and bisulfite-converted. DNAm was then measured using the Illumina Infinium Human MethylationEPIC Bead-Chip on the Illumina iScan System as per the manufacturer's protocol (Illumina, Inc., San Diego, CA; <https://www.illumina.com/products/by-type/microarray-kits/infinium-methylation-epic.html>), which allows for interrogation of over 850,000 methylation sites. For a specific locus, DNAm was quantified as the proportion of the methylated signal to the sum of the methylation and unmethylated signal, commonly known as the beta( $\beta$ )-value. Beta-values range from 0 (unmethylated) to 1 (methylated).

### Data preprocessing and quality control

DNAm was measured using the Illumina Infinium MethylationEPIC platform; probe intensity data (IDATs) were obtained from the WHI. Data were processed using the *minfi* (v.1.36.0) [32] and *ENmix* (v.1.26.10) [33] Bioconductor packages using R version

4.0.3. Background correction and dye-bias normalization was conducted using Noob [34] via the preprocess-Noob function in the *minfi* Bioconductor package. To assess data quality, low-quality data points were defined as an out-of-band detection  $P$ -value threshold of  $> 0.05$  or a beads per probe threshold of  $> 3$ . Low-quality data points were set as missing values. Low-quality samples were then defined as samples with greater than 5% missing values across CpGs, bisulfite intensity below the mean minus 3 standard deviations, or outliers in the beta distribution. Low-quality CpGs were defined as CpGs with greater than 5% missing values across samples. After quality control, 37 low-quality samples and 25,816 low-quality CpGs were removed and excluded from subsequent analyses. Further, probes on sex chromosomes and CpH probes were removed, and the Zhou et al. [35] general masking was used to remove probes, which are cross-reactive, polymorphic, and associated with SNPs. In total, 1295 samples and 729,797 CpGs were retained for downstream analyses.

### Estimation of immune parameters

Leukocyte cell-type proportions were estimated via reference-based DNAm deconvolution using the 12-cell-type reference library, EPIC IDOL-Ext, contained in the *FlowSorted.BloodExtended.EPIC* Bioconductor package [27]. Three probes (cg12810503, cg11415852, cg08911152) that are part of the 12-cell-type reference library were excluded due to poor quality and thus were not used in deconvolution. The 12 cell types in this library include: neutrophils (Neu), eosinophils (Eos), basophils (Bas), monocytes (Mono), memory B cells (Bmem), naïve B cells (Bnv), naïve CD4+ cells (CD4nv), memory CD4+ cells (CD4mem), naïve CD8+ cells (CD8nv), memory CD8+ cells (CD8mem), T regulatory cells (Treg) and natural killer cells (NK). If a deconvolution estimate fell below the limit of detection for a cell type for a sample, the estimate was replaced with the limit of detection value [36] (Additional file 1: Table S1). Then, cell-type proportion estimates were multiplied by the respective total white blood cell (WBC) count provided in the WHI covariate data to obtain estimates of absolute counts (in cells per microliter). Counts and proportions for various cell types, including lymphocytes (Lymph), total T cell (Tcell), CD4+ T cells (CD4), CD8+ T cells (CD8), B cells (Bcell), myeloid (Mye), and granulocyte (Gran), were derived by summing their individual components (e.g., Bcell = Bmem + Bnv). We also derived values for multiple cell ratios, giving 58 immune parameters estimated in total (Additional file 2: Fig. S1). There was one WHI-LLS sample for which the WBC count was not available.

### Statistical analysis

Absolute cell counts and cell-type proportions were stratified by common demographic variables used in clinical settings and epidemiological studies: age and self-reported race. Age was grouped into deciles starting at age 50, and the self-reported race was stratified by White ( $N=898$ ) and Black ( $N=370$ ). Those who identified as other race ( $N=17$ ), American Indian/Alaska Native ( $N=1$ ), or those who did not report their race ( $N=9$ ) were excluded from race-specific stratifications due to small sample sizes. Such individuals were, however, included in the overall tables that look at the entire population ( $N=1294$  for cell counts and  $N=1295$  for cell proportions and cell ratios). Immune cell absolute counts (cells/ $\mu\text{l}$ ) and cell-type proportions (as the percent of the total WBC) were summarized by the mean, 95% confidence interval of the mean, median, and 2.5–97.5 percentile. To test the effect of age on immune cell parameters, linear regression models were fitted, modeling either the proportion (%) or absolute counts (cells/ $\mu\text{l}$ ) of an immune cell type as the dependent variable and age (continuous) as the independent variable. Models were also adjusted for Duffy antigen genotype, as this is well known to be associated with neutropenia [20], and thus could have an effect on other cell types. Linear regression was also used to test the effect of age on cell-type ratios (e.g., NLR) in the same way. Models were fitted separately for self-identified White women ( $N=895$ ) and self-identified Black women ( $N=367$ ) who had the Duffy antigen genotype available (Additional file 1: Table S8), totaling 88 regression models all together. For each model, interactions between the independent variables were assessed for statistical significance. Paired *t* tests were used to test differences for longitudinal samples. The comparison of the rate of change of longitudinal samples by age was analyzed using simple linear regression. A *p*-value  $< 0.05$  was considered significant for all tests.

### Software

R version 4.1.3 was used for all statistical analyses.

## Results

### Demographics

Table 1 summarizes the characteristics of the population in this study ( $N=1295$ ). The mean age at blood draw is 78.96 (SD  $\pm 7.5$ ) and ranges from 50 to 94 years of age; this includes women from both the Women's Health Initiative (WHI) baseline enrollment ( $N=58$ ) and the WHI Long Life Study (LLS) ( $N=1237$ ). Most of the population are White women (69.3%), 28.6% are Black women, and 7.1% of the participants report Spanish, Hispanic, or Latino ethnicity. The Duffy Antigen Receptor for

**Table 1** Demographics of WHI baseline and LLS individuals whose DNAm samples were used in this study

Variable name	Value†
<i>Samples</i>	
Paired	104 (8)
Unique	1191 (92)
Total	1295
<i>Visit</i>	
Baseline	58 (4.5)
LLS	1237 (95.5)
Age	78.96 (7.5)
<i>Age group</i>	
50–59	21 (1.6)
60–69	122 (9.4)
70–79	454 (35.1)
80–89	632 (48.8)
90+	66 (5.1)
<i>Race*</i>	
White	898 (69.3)
Black	370 (28.6)
Other‡	17 (1.3)
American Indian/Alaska Native	1 (0.08)
Missing	9 (0.7)
<i>Ethnicity*</i>	
Spanish/Hispanic/Latino	92 (7.1)
<i>Duffy Antigen Receptor for Chemokines genotype</i>	
FY–/–(DARC CC)	243 (18.8)
FY–/+(DARC CT)	128 (9.9)
FY+/+(DARC TT)	918 (70.9)
Missing	6 (0.5)
BMI (kg/m <sup>2</sup> )	28.43 (5.9)
<i>BMI categories</i>	
Underweight (< 18.5)	13 (1.0)
Normal (18.5–24.9)	381 (29.9)
Overweight (25.0–29.9)	457 (35.8)
Obesity I (30.0–34.9)	269 (21.1)
Obesity II (35.0–39.9)	99 (7.8)
Extreme obesity (> = 40)	57 (4.5)
<i>General health**</i>	
Excellent	84 (6.5)
Very good	444 (34.3)
Good	504 (38.9)
Fair	129 (9.9)
Poor	9 (0.7)
Missing	125 (9.7)
<i>Smoking</i>	
Never smoker	694 (53.6)
Ever smoker	494 (38.1)
Current smoker	35 (2.7)
Missing	72 (5.6)
<i>CBC availability</i>	
Total WBC count	1294 (99.9)
Differential	1237 (95.5)



**Table 1** (continued)

\*Self-identified race and ethnicity

\*\*Self-reported general health

†Values are *N*(%), except for age and BMI which is mean(*s.d.*)

‡“Other” was a category participants could identify with, and not a grouping of races

Chemokines (DARC) null genotype was found to be present in 18.8% of this population. The mean BMI is 28.43 (SD ± 5.9), with 29.9% of women categorized as normal weight, 35.8% as overweight, and 33.4% as obese. Only 2.7% of the women were current smokers at the time of blood draw, and over half (53.6%) of this population indicated that they had never been smokers. The vast majority of this population reported their general health as good, very good, or excellent (79.7%), while 10.6% reported it as fair or poor.

#### Estimates from DNAm deconvolution are comparable to estimates from CBC differentials

We first estimated the proportion of neutrophils (Neu), eosinophils (Eos), basophils (Bas), monocytes (Mono), memory B cells (Bmem), naïve B cells (Bnv), naïve CD4+ cells (CD4nv), memory CD4+ memory cells (CD4mem), naïve CD8+ cells (CD8nv), memory CD8+ cells (CD8mem), T regulatory cells (Treg) and natural killer cells (NK) for each sample via reference-based DNAm deconvolution via the 12 leukocyte subtype reference library described in Salas et al. [27]. Using total white blood cell (WBC) counts for each sample provided by the WHI, we also derived estimates of the absolute count of each cell type as cells per microliter (Additional file 2: Fig. S1). These estimates of proportions and counts were compared to the respective values for Bas, Eos, Mono, Neu, and Lymph obtained from a CBC differential, which was available for most of the samples (Table 1, *N* = 1237). The Pearson correlation between deconvolution estimates and CBC differential estimates are large for Eos ( $R = 0.82$ ,  $R = 0.84$ ), Mono ( $R = 0.78$ ,  $R = 0.86$ ), Neu ( $R = 0.89$ ,  $R = 0.98$ ), and Lymph ( $R = 0.92$ ,  $R = 0.95$ ) for proportions and absolute counts, respectively. While Bas has a correlation coefficient of 0.40 for absolute counts and 0.22 for proportions (Fig. 1A and B). The agreement between the two methods is fairly high, with a less than 1% absolute mean difference in deconvolution derived proportions and CBC estimated proportions for Bas, Eos, Mono and Neu and a 2.36% absolute mean difference for Lymph (Fig. 1C). When comparing counts derived from deconvolution and counts from the CBC, there is ≤ 50 cells/μl absolute difference for Bas, Eos, Mono, and Neu, and an absolute mean difference of 150 cells/μl for Lymph (Fig. 1D). Additionally, the majority of samples fall within the 95% confidence interval for the average

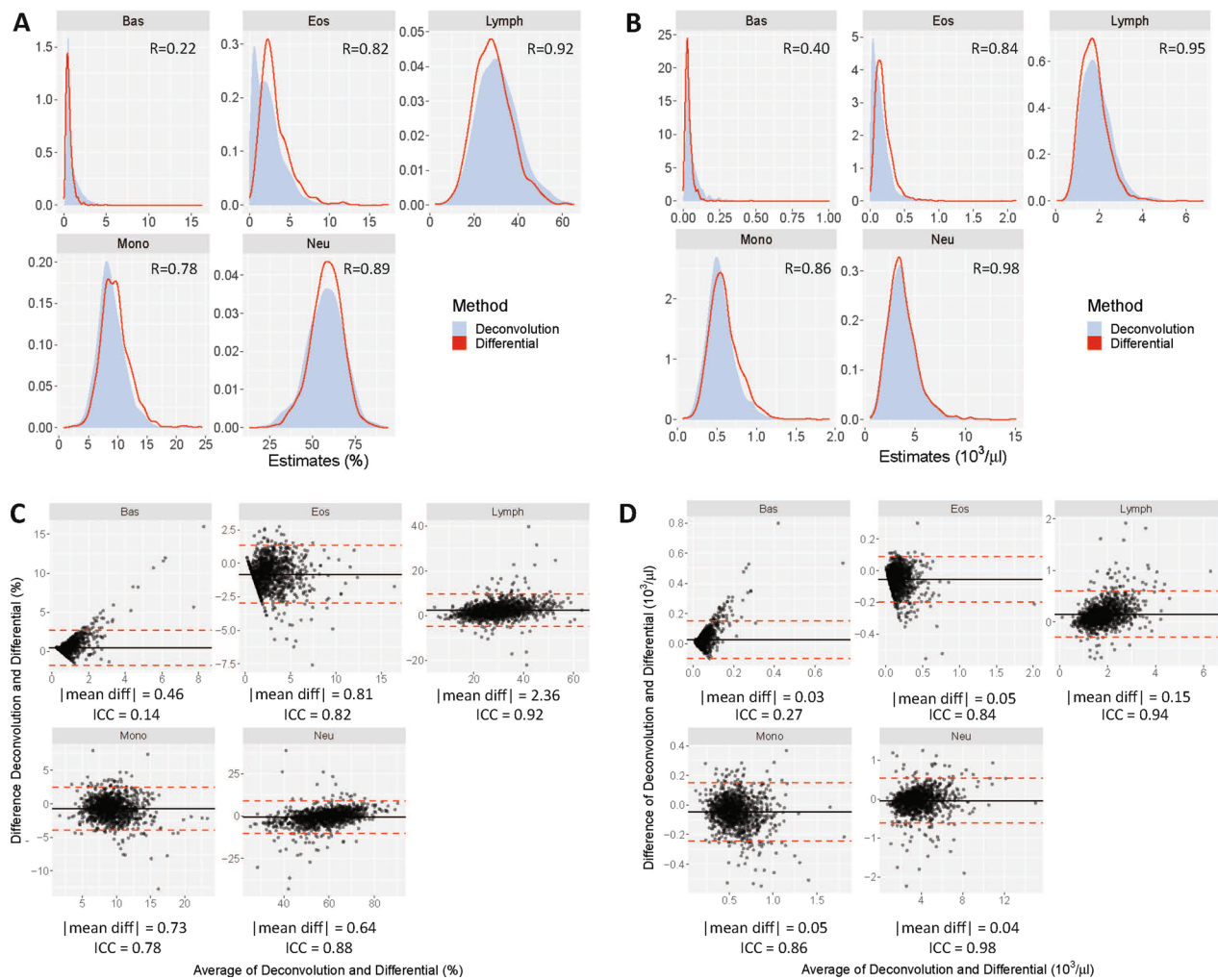
difference between the deconvolution estimates and the CBC differential estimates for all the cell types (Fig. 1C and D). The intraclass correlation coefficient (ICC) was also computed for the two methods. Basophils show poor reliability between the two methods (ICC = 0.14 cell proportions, ICC = 0.27 absolute counts). However, for proportions and counts of Eos (ICC = 0.82, ICC = 0.84), Lymph (ICC = 0.92, ICC = 0.94), Mono (ICC = 0.78, ICC = 0.86), and Neu (ICC = 0.88, ICC = 0.98), there is good to excellent reliability between the two methods.

#### Reference ranges established for 58 immune cell parameters

To establish reference ranges and characterize immune cell parameters derived using DNAm deconvolution estimates based on common demographic variables, data was grouped by age using deciles, starting from 50–59 years to 90+ years, and self-reported race, either Black or White. The immune cell types and ratios were summarized by their proportion (%) and absolute counts (cells/μl) as the mean and 95% confidence interval, median, and 2.5–97.5 percentile. Table 2 summarizes estimates for the entire population (*N* = 1295 for % or ratios and *N* = 1294 for cells/μl). There is a high correlation between derived proportions and their subsequent derived counts, with increasing variability as the deconvolution estimates increase (Additional file 3: Fig. S2). Additional file 1: Tables S2, S3, and S4 summarize estimates stratified by age and self-reported race. This totals 58 immune cell parameters (Additional file 2: Fig. S1) in which reference ranges are defined for this population of aging, post-menopausal women.

#### Immune cell parameters are associated with age and race regardless of Duffy antigen genotype

To test the effect of age on immune cell parameters, linear regression models were fitted modeling proportions and counts of immune cell parameters by age, controlling for Duffy antigen genotype. Models were stratified by self-reported race. For both Black and White women, there were significant associations with age for proportions and counts of Bnv, CD4mem, CD4nv, CD8nv, Neu and NK cells (Fig. 2, Additional file 1: Table S5). CD4mem and CD4nv cell proportions and counts decreased as age increased at similar rates between Black women and White women, and Neu and NK cells increased at similar rates (Fig. 2, Additional file 1: Table S5). There was a significant interaction between age and Duffy antigen genotype for CD4mem fraction only among White women, along with significant interaction for NK cells, but only among Black women (Additional file 1: Table S5). Compared to White women, Black women had an increased rate of cell decline as age increased for Bnv



**Fig. 1** Validity of DNA methylation deconvolution estimates. Density plots, correlation coefficients, and Bland–Altman plots comparing the estimates from deconvolution to the estimates from a CBC differential. **A, B** Density plot of estimates for proportions (**A**), represented as percentages, and absolute counts (**B**), represented as  $10^3/\mu\text{l}$ . The red line represents the distribution of estimates from the differential, and the light blue curve represents the distribution of estimates from using deconvolution. The value for  $R$  represents the Pearson correlation coefficient of the estimates for the two methods. **C, D** Bland–Altman plots to visualize the agreement between the two methods using the estimates for proportions (**C**) and absolute counts (**D**). The x-axis is the average of the estimate from the differential and deconvolution. The y-axis is the difference in the estimate from the deconvolution and differential. The horizontal black line is located at the average difference of the estimates, and the two dashed red lines are at the 95% confidence interval for the average difference. Each point is a sample. (ICC = intraclass correlation coefficient)

( $-5.351$  cells/year vs.  $-2.428$  cells/year and  $-0.103\%$ /year vs.  $-0.051\%$ /year) and CD8nv ( $-2.986$  cells/year vs.  $-0.860$  cells/year and  $-0.059\%$ /year vs.  $-0.019\%$ /year). Only White women saw a significant increase in CD8mem and Mono proportions and counts as age increased (Fig. 2, Additional file 1: Table S5).

The Neu/Lymph (NLR), Lymph/Mono, CD4nv/CD4, CD4nv/CD4mem, CD8nv/CD8, CD8nv/CD8mem, Treg/CD4, Bnv/B, Bnv/Bmem, and CD8/Bcell ratios all had significant linear associations with age in both Black women and White women, and even upon adjustment for Duffy antigen genotype (Fig. 3, Additional file 1: Table S5).

The NLR, Treg/CD4 and CD8/B ratios increased as age increased, while the others decreased as age increased (Fig. 3, Additional file 1: Table S5). The Eos/Neu, CD4/CD8, CD4nv/CD8nv, and CD4mem/CD8mem significantly decreased as age increased in White women, while the CD8mem/Treg and (CD8+ NK)/Mono significantly increased in White women (Additional file 1: Table S5, Additional file 4: Fig. S3). There were no significant associations between these ratios for Black women. The Bas/Lymph and Neu/Mono ratios significantly increased as age increased in Black women only (Additional file 1: Table S5, Additional file 4: Fig. S3).

**Table 2** Summary of immune cell parameters based on DNA methylation deconvolution estimates for the entire cohort

	Mean	(95% CI)	Median	Percentile range (2.5–97.5)
WBC*	6276.028	(6173.04, 6379.015)	6000	(3386.5, 10,680.25)
<i>Leukocyte subtypes</i>				
Myeloid**				
Cells/ $\mu$ l	4394.189	(4304.664, 4483.714)	4198.418	(1893.226, 8212.533)
%	69.161	(68.621, 69.701)	69.949	(46.791, 86.387)
Granulocytes**				
Cells/ $\mu$ l	3848.037	(3763.108, 3932.967)	3668.982	(1503.92, 7470.417)
%	60.231	(59.648, 60.815)	60.94	(36.342, 78.904)
Neutrophils				
Cells/ $\mu$ l	3642.355	(3558.256, 3726.454)	3469.95	(1334.334, 7201.745)
%	56.833	(56.207, 57.459)	57.457	(31.576, 76.845)
Basophils				
Cells/ $\mu$ l	65.7	(61.971, 69.43)	41.723	(18.371, 242.819)
%	1.08	(1.018, 1.141)	0.6	(0.478, 3.474)
Eosinophils				
Cells/ $\mu$ l	139.982	(133.382, 146.582)	112.492	(19.425, 444)
%	2.319	(2.219, 2.418)	2	(0.422, 6.539)
Monocytes				
Cells/ $\mu$ l	546.152	(536.271, 556.033)	524.352	(261.548, 961.111)
%	8.929	(8.801, 9.058)	8.66	(4.931, 14.337)
Lymphocytes**				
Cells/ $\mu$ l	1920.642	(1881.476, 1959.808)	1822.981	(842.129, 3620.853)
%	31.465	(30.92, 32.01)	30.507	(14.382, 54.122)
T cells**				
Cells/ $\mu$ l	1358.107	(1327.532, 1388.681)	1280.328	(525.1, 2645.941)
%	22.247	(21.809, 22.685)	21.155	(9.341, 40.643)
CD4+ total**				
Cells/ $\mu$ l	859.971	(838.43, 881.512)	794.503	(296.086, 1799.017)
%	14.139	(13.811, 14.467)	13.283	(5.096, 28.434)
CD4+ memory				
Cells/ $\mu$ l	609.586	(593.698, 625.473)	564.328	(170.535, 1328.026)
%	10.058	(9.81, 10.306)	9.378	(2.796, 21.22)
CD4+ naïve				
Cells/ $\mu$ l	200.559	(190.288, 210.829)	132.151	(34.738, 685.529)
%	3.283	(3.118, 3.447)	2.28	(0.878, 11.362)
T regulatory				
Cells/ $\mu$ l	49.826	(48.874, 50.779)	46.956	(26.777, 87.899)
%	0.798	(0.785, 0.81)	0.78	(0.78, 0.78)
CD8+ total**				
Cells/ $\mu$ l	498.136	(476.832, 519.44)	397.338	(84.81, 1473.431)
%	8.108	(7.793, 8.423)	6.715	(1.559, 22.443)
CD8+ memory				
Cells/ $\mu$ l	418.013	(396.375, 439.651)	306.012	(37.325, 1385.639)
%	6.803	(6.478, 7.128)	5.278	(0.801, 21.569)
CD8+ naïve				
Cells/ $\mu$ l	80.123	(76.878, 83.368)	58.101	(28.146, 246.615)
%	1.305	(1.253, 1.357)	0.758	(0.758, 4.127)
B cells**				
Cells/ $\mu$ l	227.216	(218.517, 235.914)	189.396	(50.299, 640.404)

**Table 2** (continued)

	Mean	(95% CI)	Median	Percentile range (2.5–97.5)
%	3.755	(3.617, 3.893)	3.183	(0.811, 9.76)
B memory				
Cells/ $\mu$ l	86.814	(81.799, 91.829)	67.363	(16.061, 255.53)
%	1.412	(1.335, 1.489)	1.172	(0.328, 4.097)
B naïve				
Cells/ $\mu$ l	140.402	(133.852, 146.952)	106.589	(23.055, 503.33)
%	2.343	(2.234, 2.452)	1.834	(0.483, 7.108)
Natural killer				
Cells/ $\mu$ l	335.32	(326.172, 344.467)	308.259	(94.861, 761.925)
%	5.463	(5.331, 5.596)	5.144	(1.587, 11.042)
<i>Cell ratios</i>				
Neutrophil/lymphocyte	2.162	(2.087, 2.236)	1.879	(0.597, 5.281)
Lymphocyte/monocyte	3.699	(3.625, 3.774)	3.502	(1.635, 6.829)
Eosinophil/lymphocyte	0.077	(0.073, 0.081)	0.063	(0.012, 0.253)
Basophil/lymphocyte	0.038	(0.035, 0.04)	0.024	(0.01, 0.134)
Eosinophil/neutrophil	0.046	(0.044, 0.048)	0.035	(0.006, 0.155)
(Eosinophil + Basophil)/neutrophil	0.067	(0.064, 0.071)	0.053	(0.012, 0.209)
Neutrophil/monocyte	7.126	(6.88, 7.372)	6.482	(2.855, 14.15)
CD4/CD8	3.012	(2.857, 3.166)	2.063	(0.35, 11.097)
CD4nv/CD8nv	3.014	(2.854, 3.174)	1.832	(0.419, 11.073)
CD4mem/CD8mem	3.806	(3.563, 4.049)	1.903	(0.275, 15.93)
CD4nv/CD4	0.218	(0.211, 0.226)	0.189	(0.047, 0.518)
CD4nv/CD4mem	0.365	(0.347, 0.383)	0.256	(0.052, 1.252)
CD8nv/CD8	0.266	(0.254, 0.279)	0.173	(0.034, 0.769)
CD8nv/CD8mem	0.61	(0.561, 0.659)	0.209	(0.035, 3.331)
Treg/CD4	0.069	(0.066, 0.071)	0.059	(0.027, 0.165)
Bnv/B	0.591	(0.579, 0.602)	0.607	(0.177, 0.919)
Bnv/Bmem	2.628	(2.44, 2.816)	1.544	(0.215, 11.359)
CD8/B	3.107	(2.918, 3.296)	1.999	(0.357, 12.643)
(CD8+ NK)/Mono	1.604	(1.555, 1.653)	1.427	(0.498, 3.788)
CD8mem/Treg	8.684	(8.269, 9.1)	6.752	(1.027, 27.648)

\*WBC = White blood cell count, from WHI data. All other cell count data is from multiplying cell deconvolution proportion estimates by WBC

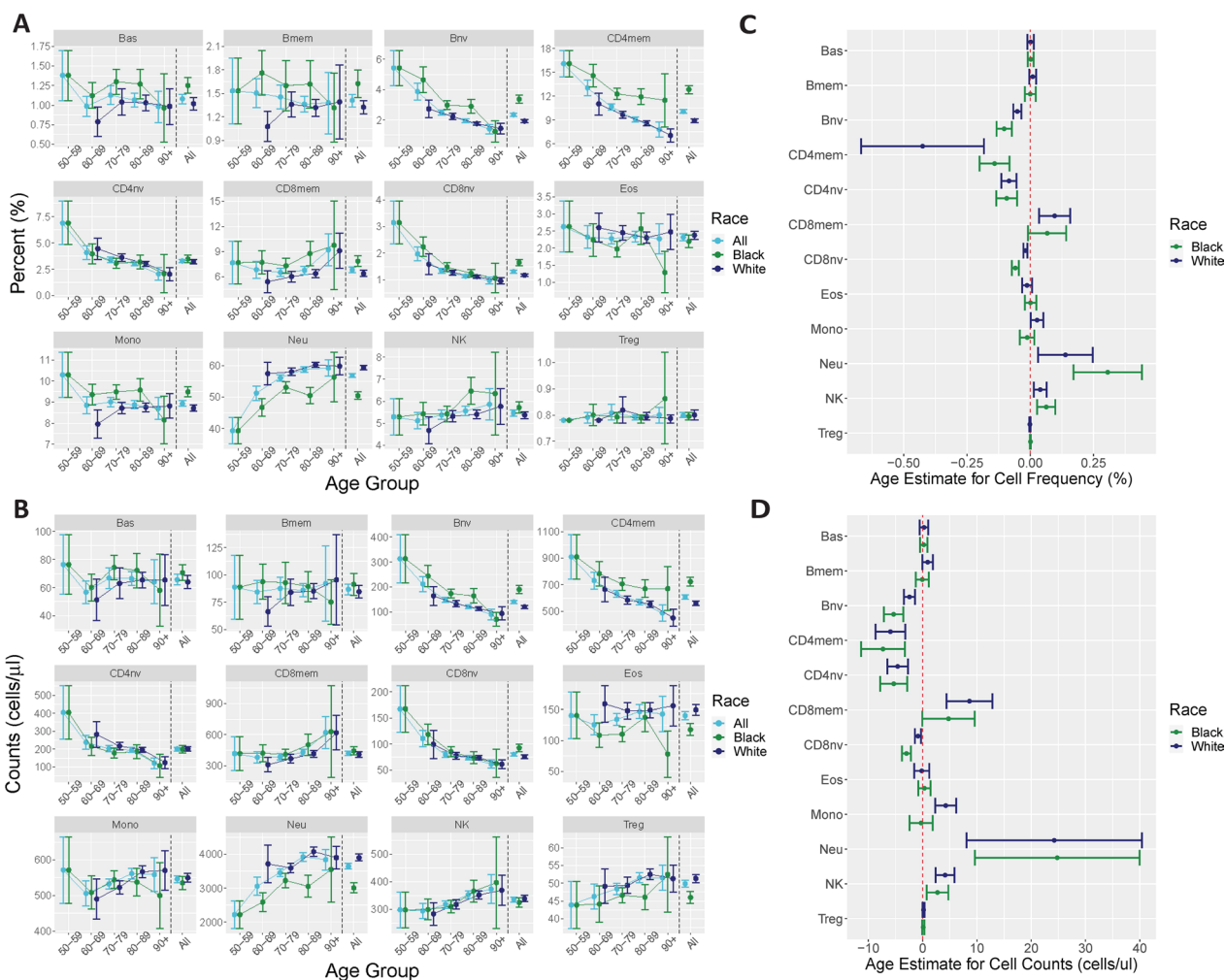
\*\*Derived cell counts: Lymph = CD4-total + CD8-total + B-total + NK + Treg; Tcell = CD4-total + CD8-total; CD4-total = CD4mem + CD4nv + Treg; CD8-total = CD8mem + CD8nv; B-total = Bmem + Bnv; Myeloid = Gran + Mono; Gran = Neu + Eos + Bas

### Longitudinal samples show similar associations with age with high variability between individuals

The results of the overall association of immune cell parameters and age were mirrored when examining the 52 individuals with longitudinal blood samples (Fig. 4A, B; Additional file 5: Fig. S4). Of the 104 paired samples, one sample came from the original baseline visit from enrollment in the WHI, and the second sample was from the WHI-LLS visit. There is a range of 14–18 years between two samples of the same individual; most of these samples come from Black women ( $N=48$ ). Using paired *t* tests, there are significant decreases from baseline to the LLS visit in cell proportions and counts for Bmem ( $|\Delta|=0.47\%$  and 28.8 cells/ $\mu$ l,

$p=0.007$  and 0.002), Bnv ( $|\Delta|=1.86\%$  and 101.3 cells/ $\mu$ l,  $p<0.0001$  and 0.0001), CD4mem ( $|\Delta|=2.96\%$  and 178.3 cells/ $\mu$ l,  $p<0.0001$  and  $p=0.0003$ ), CD4nv ( $|\Delta|=2.35\%$  and 131.3 cells/ $\mu$ l,  $p<0.0001$  and 0.0001), and CD8nv ( $|\Delta|=1.05\%$  and 54.2 cells/ $\mu$ l,  $p<0.0001$  and  $p=0.0002$ ). There is a significant increase in Neu proportions ( $|\Delta|=9.54\%$ ,  $p<0.0001$ ) and counts ( $|\Delta|=651.4$  cells/ $\mu$ l,  $p=0.0002$ ). However, NK cells did not show a significant increase in proportion or absolute counts in the longitudinal samples. The rates of change (calculated as the difference in proportion or absolute counts per year) across subjects is highly variable (Fig. 4C, D; Additional file 1: Table S6). For example, Neu range from  $-0.69$  to  $+2.11\%/year$  and  $-99.51$





**Fig. 2** Associations of immune cell profiles with age and race. **A, B** Line plots showing the proportion estimates as percentages from deconvolution (**A**) and absolute counts (**B**) in cells/ $\mu$ l, stratified by age and self-reported race. Data are represented by the mean values and 95% CIs. The y-axis is the percent value or the cells/ $\mu$ l value. The x-axis is age group, and self-identified race is indicated by color (light blue = all participants ( $N = 1295$ ), dark-green = Black women ( $N = 367$ ), dark-purple = White women ( $N = 895$ )). **C, D** Regression coefficients for age, after adjusting for Duffy antigen genotype, for proportions (%/year) (**C**) and absolute counts (cells/ $\mu$ l/year) (**D**) stratified by self-reported race. Data are represented as the regression coefficient for age and the 95% CI

to +207.68 cells/ $\mu$ l/year. However, using simple linear regression, the association of age and rate of change was not significant for most cell types. The one exception is for Bas ( $R = -0.42$ ,  $p = 0.0021$  for proportions,  $R = -0.33$ ,  $p = 0.017$  for counts, Fig. 4C, D).

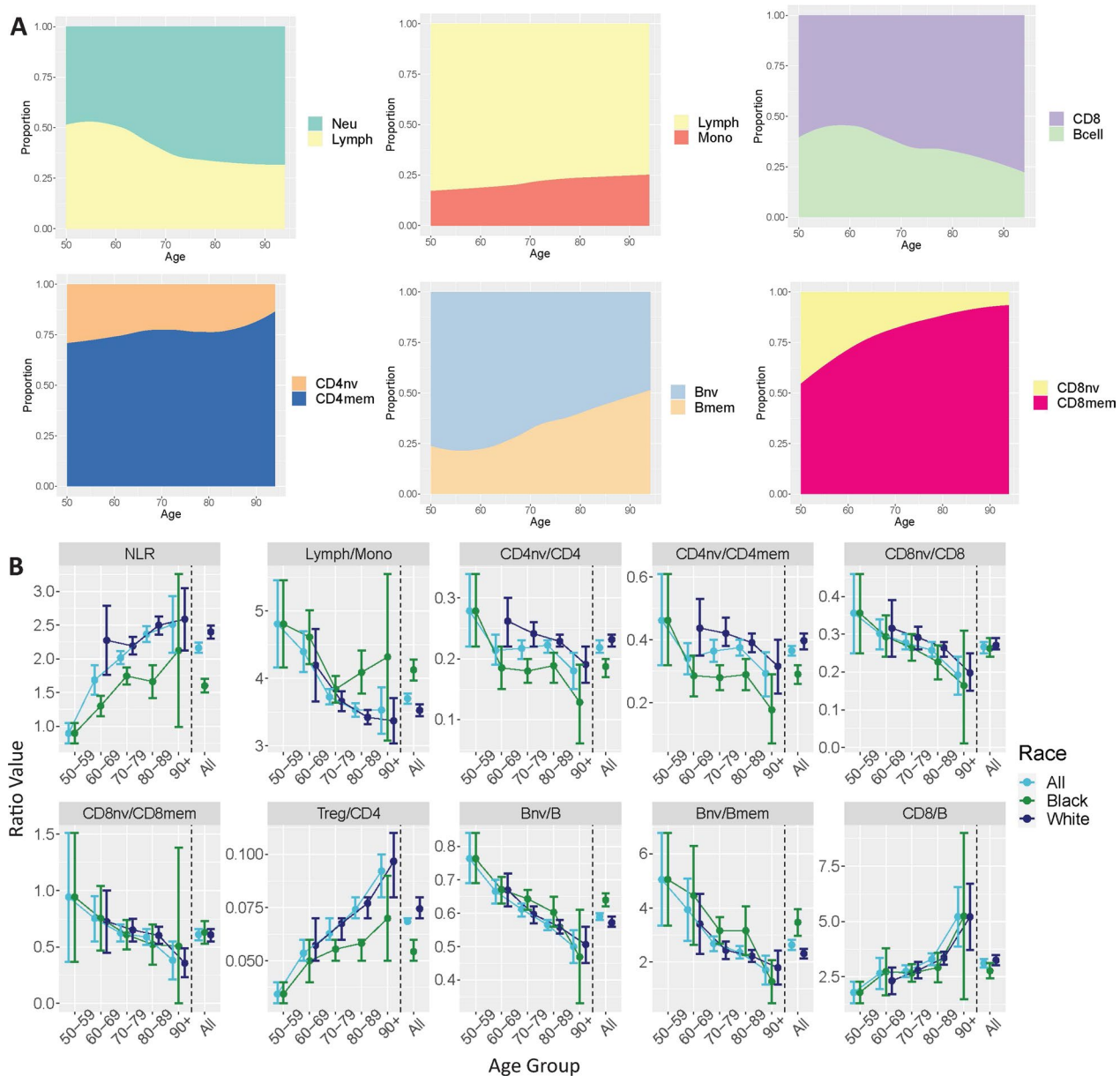
**Comparison immune cell counts to those in the literature**

We conducted a literature review of studies that include reference ranges derived from flow cytometry methods and CBC differentials for similar populations to this study. Table 3 shows a selected comparison from the full literature review of our estimates of 15 immune cell parameters to 4 other studies with the most similar

populations. Additional file 1: Table S7 includes all studies reviewed.

**Discussion**

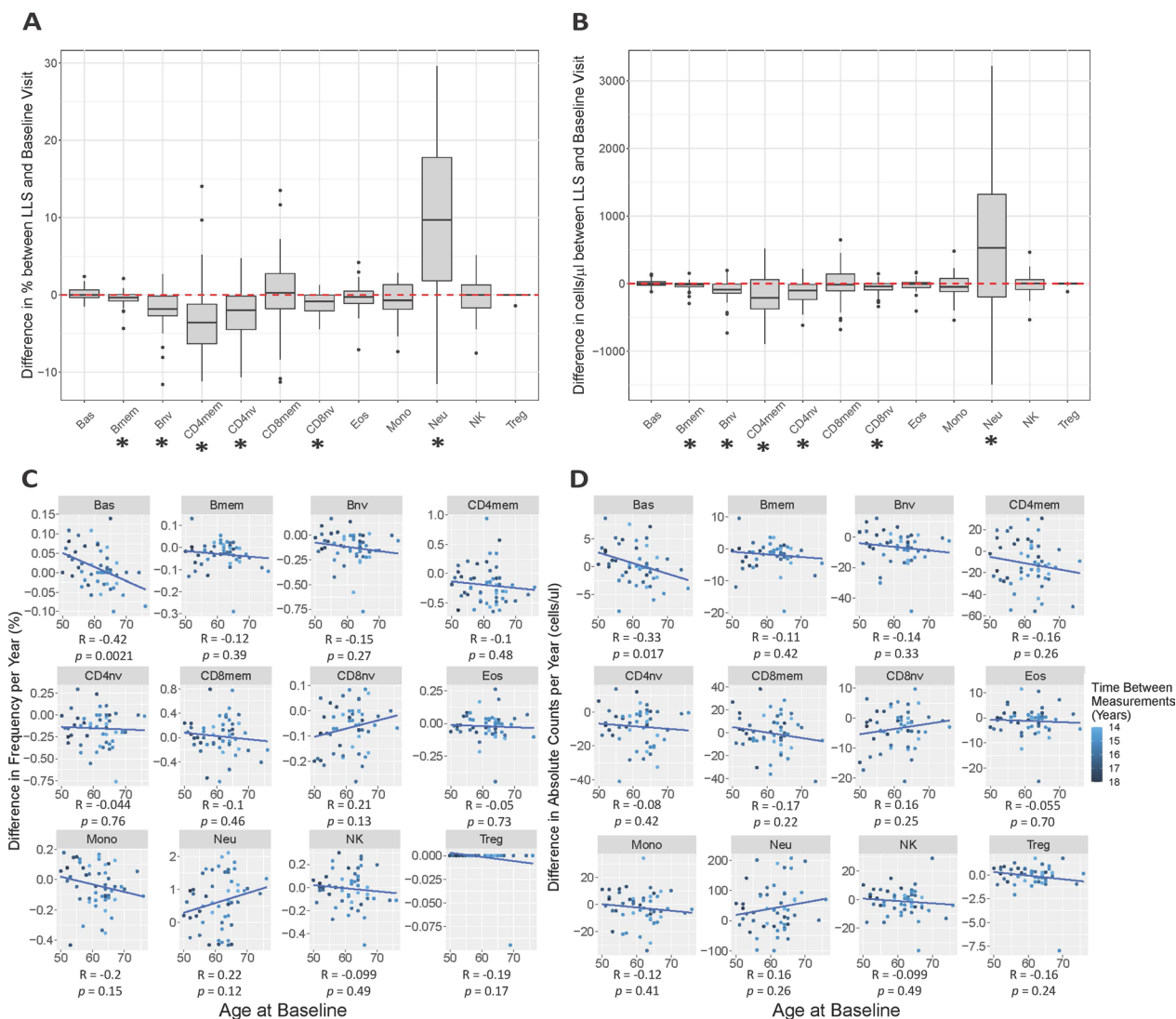
We applied our novel approach to peripheral blood methylation data in a subset of the WHI to establish reference ranges across 58 immune cell parameters. The deconvolution method used was recently developed by Salas et al. [27] where libraries for this high resolution immune phenotyping were created and validated against the gold standard of flow cytometry. This technique provides the opportunity for estimating proportions and counts for leukocyte subtypes without



**Fig. 3** Changes in cell ratios across age and race. **A** Area plots showing the average and smoothed cell proportion per compartment across age. **B** Line plots showing the cell ratios stratified by age and self-reported race. Data are represented by the mean values and 95% CIs. The y-axis is the ratio value. The x-axis is age group, and self-reported race is indicated by color (light blue = all participants ( $N = 1295$ ), dark green = Black women ( $N = 367$ ), dark purple = White women ( $N = 895$ ))

the need for fresh blood and in a cheaper and more standardized fashion compared to flow cytometry, thus making it scalable to large sample sizes. In these data, we showed deconvolution estimates were similar to complete blood cell count (including differential) data in the WHI, showing it to be comparable to conventional assessment of peripheral blood immune profile in the setting of blood collected as part of a large epidemiologic study. Our results showed a high correlation

between the proportions and count data, and the deconvolution estimates were almost all within the 95% confidence limits of the differential estimates from the initial blood draw. Further, the magnitude of difference between the two methods falls within the range seen in other studies [37, 38]. The overall leukocyte count data allows us to convert the relative prevalence to a count, making comparisons to prior published data enumerating immune profiles in blood using flow cytometry.



**Fig. 4** Immune cell profiles in paired, longitudinal samples. **A, B** Box plots showing the change in (A) proportion estimates and (B) absolute counts for paired samples. Change was calculated as the difference from the baseline visit to the LLS visit for a pair of samples (LLS estimate—baseline estimate). Red dashed line is at a change of 0. Starred cell types indicate  $p$ -value < 0.05 when conducting a paired t-test. **C, D** Scatter plots of the rate of change for longitudinal samples by age of the subject. Each point is the difference in proportion per year (C) or difference in absolute counts per year (D) from the baseline visit to the LLS visit for a pair of samples for a given subject. The points are colored in by the number of years between the pair of samples. Pearson correlation coefficients and  $p$ -values for associations between rate of change and age at baseline are below each cell type plot

This further validates this method and demonstrates its potential for application in other large epidemiologic studies.

There was little disagreement when we compared our values for immune cell counts to those in the literature derived from flow cytometry methods and CBC differentials (Table 3, Additional file 1: Table S7). There were no studies that established reference ranges for all the immune parameters defined in this study and few studies that share a similar population as compared to the

present study, mostly due to the age range. The studies with the most similar populations are Provinciali et al. [17], Thyagarajan et al. [1], and Valiathan et al. [11]; however, they only established reference ranges for various lymphocyte subsets using PBMCs (which may introduce variability in counts). The estimates derived for CD4+ and CD8+ cells and their naïve and memory subsets in Provinciali et al. [17] and Thyagarajan et al. [1] compared to the estimates derived in this study for the WHI cohort are fairly similar (Table 3), except for CD8mem

**Table 3** Comparing immune cell counts to those in the literature (selected version)

Study	WHI (this study)				Provinciali et al. [17] <sup>a</sup>		Thyagarajan et al. [1] <sup>b</sup>		Valiathan et al. [11] <sup>c</sup>		Melzer et al. [3] <sup>d</sup>	
	Geographic area	Sex	Age range	Sample size	United States Women	Italy Women	United States Women	United States Women	United States Women	Germany Women	Flow cytometry	Flow cytometry
Method	DNAm deconvolution				Flow cytometry		Flow cytometry		Flow cytometry		Flow cytometry	
	Type	Percentile (2.5–97.5)	Mean	Median	Weighted mean	Weighted mean	Weighted mean	Mean	Median	Median	Percentile (2.5–97.5)	
Neutrophils	Cells/μl	(1334, 7202)	3642	3470						3122	(1801, 5579)	
Eosinophils	Cells/μl	(19, 444)	140	112						101	(30, 383)	
Monocytes	Cells/μl	(262, 961)	546	524						230	(108, 422)	
Lymphocytes	Cells/μl	(842, 3621)	1921	1823						1624	(677, 2916)	
T cells	Cells/μl	(525, 2646)	1358	1280			1434	1424	1351	1098	(369, 2086)	
CD4+ total	Cells/μl	(296, 1799)	860	795	710		999	892	867	818	(263, 1707)	
CD4+ memory	Cells/μl	(171, 1328)	610	564	324		384					
CD4+ naive	Cells/μl	(35, 686)	201	132	176		486					
T regulatory	Cells/μl	(27, 88)	50	47						60	(22, 130)	
CD8+ total	Cells/μl	(85, 1473)	498	397	410		337	525	453	143	(32, 437)	
CD8+ memory	Cells/μl	(37, 1386)	418	306	178		200					
CD8+ naive	Cells/μl	(28, 247)	80	58	59		73					
B-total	Cells/μl	(50, 640)	227	189				229	164	216	(62, 513)	
Natural Killer	Cells/μl	(95, 762)	335	308				339	250	276	(85, 624)	
CD4/CD8		(0.35, 1.1)	3.01	2.07				2.4	1.7			

<sup>a</sup>Weighted means were calculated from data in Table 1 of Provinciali et al.

<sup>b</sup>Weighted means were copied from Additional file 1: Table S3 of Thyagarajan et al. and converted to cells/μl; CD4mem and CD8mem were added from their respective TCM + Effector + TEM compartments

<sup>c</sup>Mean and median values were copied from Additional file 1: Table S2 of Valiathan et al.

<sup>d</sup>Median and percentile values were copied from Additional file 1: Table S4 of Melzer et al. and converted to cells/μl; CD4 were taken from T helper (Th) cells, CD8 were taken from cytotoxic T (Tc) cell

mean absolute counts being much higher in the WHI compared to both Provinciali et al. [17] and Thyagarajan et al. [1] (Table 3; WHI = 418 cells/ $\mu$ l, Provinciali = 178 cells/ $\mu$ l, Thyagarajan = 200 cells/ $\mu$ l). Valiathan et al. [11] had a much smaller sample size ( $N=33$ ) of women in the same age range compared to the WHI. However, mean and median absolute counts of the total CD4+, CD8+, B, and NK cells were most comparable to those derived in the WHI (Table 3). Finally, Melzer et al. [3] included cell-type estimates of Neu, Eos, Mono, Lymph, total T, CD4+ total, Treg, CD8+ total, total B, and NK cells for a group of women ranging from 60 to 79 years of age, which is a similar age group to about 45% of the WHI cohort. The study's most similar median absolute count estimates were Eos, CD4+, Treg, total B, and NK cells (Table 3). As noted, across all other studies, most cell type estimates were in agreement to those derived in our study (Additional file 1: Table S7). One exception is CD8mem, where, similar to the above studies, we observed much larger absolute counts in the WHI than all other studies; this may be the result of small sample sizes of the other studies, systematic loss of CD8+ cells in PBMC preparation or perhaps chance.

Further evidence of the utility of this approach is found in the data showing that immune profiles derived from deconvolution demonstrate well-known relationships of age with immune subsets. For example, we show the expected diminution of the naïve compartments of CD4+ [1, 2, 17, 19, 39–42] and CD8+ [1, 17, 39–44] lymphocyte lineages, as well as the expected growth of CD8mem [1, 39, 42]. Also, similar to other published studies, we see a decrease in Bv [44] and an increase in Neu [11] and NK cells [8, 11, 12] with age. However, in contrast to other studies [1, 2, 17, 19, 39, 42], we show a significant decrease in CD4mem with age. We also show that the well-established Duffy antigen variant genotype predicts the expected neutropenia [20] assessed by both percentages and counts.

We have also presented novel reference ranges for numerous additional immune parameters that may find considerable utility in epidemiologic studies going forward. For example, the NLR has been widely used in studies of cancer survival [45–47]. This is now being applied to cancer risk as well as the risk of additional inflammatory-associated diseases (e.g., diabetes [48, 49], cardiovascular disease [50, 51], and stroke [52, 53]). We find many of the novel immune parameters that are generated easily using this method also are associated with age and self-reported race. Our data may be quite valuable for comparison with other epidemiologic studies.

Finally, we made use of methylation data from blood draws repeated over a number of years in the same

subject to assess individual changes in immune profile over an average of some 14–18 years. This is very unusual data and shows that changes in immune profiles in elderly women are highly variable. Other studies have reported this high inter-individual variability for similar cell types, including Lin et al. [44], which had paired longitudinal samples 5 years apart, and Alpert et al. [43], which had longitudinal samples collected over 9 years. This may hold important information and thus suggests that larger studies nested in ongoing large cohorts where multiple blood draws have occurred may hold novel findings of changes in immune profiles predicting subsequent types of chronic inflammatory disease.

This study is not without limitations. We only compared our deconvolution estimates within the WHI samples using CBC derived estimates, which lack the granularity to distinguish various lymphocyte subsets. Although we could not validate the deconvolution estimates with flow cytometry on the same samples, reference-based deconvolution using DNAm has been shown to be accurate across numerous studies [26, 27, 29]. We also did not have access to CMV status and thus could not test this association with the immune profile. Although we use self-reported race to describe immune cell profiles, we are limited in interpretation by not addressing other factors that could be relevant (e.g., environmental factors). However, interpretation for why there might be differences between groups is beyond the scope of this work but does provide the opportunity for future studies. Further, our study population included only women aged 50–94 years old at the time of blood draw, thus limiting generalizability to younger women or men.

## Conclusions

We have outlined a method to provide highly detailed immune phenotyping using stored peripheral blood samples and validated this method using complete blood count differential estimates and comparisons to estimates from flow cytometry from other published studies. Although this work was conducted in one particular population, it demonstrates the validity of the methylation cytometry approach and should stimulate additional investigation of immune profiles and chronic disease associations in prospect cohorts.

## Abbreviations

CBC	Complete blood count
WBC	White blood cell
CMV	Cytomegalovirus
DNAm	DNA methylation
WHI	Women's Health Initiative
LLS	Long Life Study



DARC	Duffy Antigen Receptor for Chemokines
Neu	Neutrophil
Eos	Eosinophil
Bas	Basophil
Mono	Monocyte
Bmem	B memory
Bnv	B naïve
CD4nv	CD4+ naïve
CD4mem	CD4+ memory
CD8nv	CD8+ naïve
CD8mem	CD8+ memory
Treg	T regulatory
NK	Natural killer
Lymph	Lymphocyte
Tcell	Total T cell
CD4	CD4+ T cell
CD8	CD8+ T cell
Bcell	B cell
Mye	Myeloid
Gran	Granulocyte
NLR	Neutrophil-to-lymphocyte ratio

## Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s13148-023-01488-8>.

**Additional file 1: Table S1.** Limit of detection for DNAm deconvolution estimates. **Table S2:** Summary of absolute counts of immune cells stratified by age and race. **Table S3:** Summary of proportions of immune cells stratified by age and race. **Table S4:** Summary of immune cell ratios stratified by age and race. **Table S5:** Results from linear regression analysis. **Table S6:** Rate of change of immune cell parameters in paired longitudinal samples. **Table S7:** Comparing immune cell counts to those found in literature. **Table S8:** Distribution of the Duffy antigen genotype.

**Additional file 2. Fig. S1:** Study design for deriving all 58 immune parameters. (1) First, DNAm deconvolution is performed using the DNAm measured from peripheral blood samples. The proportions of 12 leukocyte subsets are given. (2) Derive other leukocyte subsets by adding respective components together. Also, derive absolute counts of all the subsets by multiplying the proportion by the total WBC. (3) Derive cell ratios. This gives 58 immune parameters in total.

**Additional file 3. Fig. S2:** Correlation between deconvoluted proportions and derived cell counts. Scatter plots showing the correlation between DNAm deconvoluted proportions and their respective derived cell counts for each of the 12 subtypes. The blue line represents the linear best fit line and R represents the Pearson correlation coefficient.

**Additional file 4. Fig. S3:** Changes in cell ratios across age and race. Area plots showing the average and smoothed cell proportion per compartment across age. Area plots are separated by self-identified race. Line plots showing the cell ratio values stratified by age and self-identified race. Data are represented by the mean values and 95% CIs. The y-axis is the ratio value. The x-axis is age-group and self-identified race is indicated by color, (light-blue = all participants (N = 1295), dark green = Black women (N = 367), dark purple = White women (N = 895).

**Additional file 5. Fig. S4:** Change in immune cell parameters for paired, longitudinal samples. **A, B** Spaghetti plots showing the **A** proportion estimates and **B** absolute counts for paired samples. Points indicate the estimate and either baseline or LLS, and lines connect paired samples. The absolute mean difference and p-values from a paired t-test are below each cell type plot.

## Acknowledgements

The WHI program is supported by contracts from the National Heart, Lung and Blood Institute, NIH. The authors thank the WHI investigators and staff for their dedication, and the study participants for making the program possible. A

listing of WHI investigators can be found at <https://s3-us-west-2.amazonaws.com/www-whi-org/wp-content/uploads/WHI-Investigator-Long-List.pdf>.

## Author contributions

EN carried out the data preprocessing, quality control, and statistical analyses. EN wrote the manuscript with support from KTK and DCK, who also supervised the project. KTK, DCK, AR, SL, RBW, AMM, LAS, BCC, JKW, and EN helped design the study. KTK, DCK, AR, SL, RBW, AMM, LAS, BCC, JKW, and EN contributed to the interpretation of the results. All authors provided critical feedback and helped shape the research, analysis, and manuscript. All authors read and approved the final manuscript.

## Funding

The WHI program is funded by the National Heart, Lung, and Blood Institute, National Institutes of Health, U.S. Department of Health and Human Services through 75N92021D00001, 75N92021D00002, 75N92021D00003, 75N92021D00004, 75N92021D00005. Research was supported by the National Cancer Institute (NCI) Cancer Center Support Grant P30 CA168524; the Kansas IDeA Network of Biomedical Research Excellence Bioinformatics Core, supported by the National Institute of General Medical Science award P20 GM103418; the Kansas Institute for Precision Medicine COBRE, supported by the National Institute of General Medical Science award P20 GM130423; CA253976; the CDMRP/Department of Defense (W81XWH-20-1-0778); NIGMS (P20 GM104416); X01HL139376.

## Availability of data and materials

The data generated and/or analyzed during the current study are not publicly available. Authorization from the WHI is needed for access to the data.

## Declarations

### Ethics approval and consent to participate

The WHI project was reviewed and approved by the Fred Hutchinson Cancer Research Center (Fred Hutch) IRB in accordance with the U.S. Department of Health and Human Services regulations at 45 CFR 46 (approval number: IR# 3467-EXT). Participants provided written informed consent to participate. Additional consent to review medical records was obtained through signed written consent. Fred Hutch has an approved FWA on file with the Office for Human Research Protections (OHRP) under assurance number 0001920.

### Consent for publication

Not applicable.

### Competing interests

Drs. Kelsey and Wiencke are founders of Cellintec, which had no role in this research.

### Author details

<sup>1</sup>Department of Biostatistics and Data Science, University of Kansas Medical Center, Kansas City, KS, USA. <sup>2</sup>Division of Public Health Science, Fred Hutchinson Cancer Center, Seattle, WA, USA. <sup>3</sup>Departments of Epidemiology, Medicine, and Surgery, Brown University, Providence, RI, USA. <sup>4</sup>Departments of Epidemiology and Internal Medicine, School of Public Health, University of Iowa, Iowa City, IA, USA. <sup>5</sup>Department of Neurological Surgery, University of California San Francisco, San Francisco, CA, USA. <sup>6</sup>Department of Epidemiology, Geisel School of Medicine, Dartmouth College, Lebanon, NH, USA. <sup>7</sup>Department of Molecular and Systems Biology, Geisel School of Medicine, Dartmouth College, Lebanon, NH, USA. <sup>8</sup>Department of Community and Family Medicine, Geisel School of Medicine, Dartmouth College, Lebanon, NH, USA. <sup>9</sup>UCSF Weill Institute for Neurosciences, University of California San Francisco, San Francisco, CA, USA. <sup>10</sup>Departments of Epidemiology and Pathology and Laboratory Medicine, Brown University, 70 Ship St, Providence, RI 02903, USA.

Received: 2 September 2022 Accepted: 19 April 2023

Published online: 28 April 2023

## References

- Thyagarajan B, Faul J, Vivek S, Kim JK, Nikolich-Zugich J, Weir D, et al. Age-related differences in T-cell subsets in a nationally representative sample of people older than age 55: findings from the health and retirement study. *J Gerontol Ser A*. 2021. <https://doi.org/10.1093/gerona/glab300>.
- Schultze-Florey CR, Chukhno E, Goudeva L, Blasczyk R, Ganser A, Prinz I, et al. Distribution of major lymphocyte subsets and memory T-cell subpopulations in healthy adults employing GLP-conforming multicolor flow cytometry. *Leukemia*. 2021;35(10):3021–5.
- Melzer S, Zachariae S, Bocsi J, Engel C, Löffler M, Tárnok A. Reference intervals for leukocyte subsets in adults: results from a population-based study using 10-color flow cytometry. *Cytometry B Clin Cytom*. 2015;88(4):270–81.
- Böhm I. Quantification of absolute peripheral white blood cells and their subsets in patients with lupus erythematosus: comparison with other inflammatory diseases with and without autoimmune background. *Biomed Pharmacother*. 2006;60(2):92–5.
- Kokuina E, Breff-Fonseca MC, Villegas-Valverde CA, Mora-Díaz I. Normal values of T, B and NK lymphocyte subpopulations in peripheral blood of healthy Cuban adults. *MEDICC Rev*. 2019;11(21):16–21.
- Andreu-Ballester JC, García-Ballesteros C, Benet-Campos C, Amigó V, Almela-Quilis A, Mayans J, et al. Values for  $\alpha\beta$  and  $\gamma\delta$  T-lymphocytes and CD4+, CD8+, and CD56+ subsets in healthy adult subjects: assessment by age and gender. *Cytometry B Clin Cytom*. 2012;82B(4):238–44.
- Coates S, Wang D, Pierscionek T, Fernandes S, Djumanov D, Lorch U, et al. Time- and race-specific haematological reference intervals for healthy volunteer trials: a retrospective analysis of pooled data from multiple phase I trials. *Front Pharmacol*. 2020;11:314.
- Jentsch-Ullrich K, Koenigsman M, Mohren M, Franke A. Lymphocyte subsets' reference ranges in an age- and gender-balanced population of 100 healthy adults—a monocentric German study. *Clin Immunol*. 2005;116(2):192–7.
- Torres AJL, Angelo ALD, Silva MO, Bastos MDC, Souza DFD, Inocencio LA, et al. Establishing the reference range for T lymphocytes subpopulations in adults and children from Brazil. *Rev Inst Med Trop Sao Paulo*. 2013;55(5):323–8.
- Saathoff E, Schneider P, Kleinfeldt V, Geis S, Haule D, Maboko L, et al. Laboratory reference values for healthy adults from southern Tanzania. *Tropical Med Int Health*. 2008;13(5):612–25.
- Valiathan R, Ashman M, Asthana D. Effects of ageing on the immune system: infants to elderly. *Scand J Immunol*. 2016;83(4):255–66.
- Choi J, Lee SJ, Lee YA, Maeng HG, Lee JK, Kang YW. Reference values for peripheral blood lymphocyte subsets in a healthy Korean population. *Immune Netw*. 2014;14(6):289.
- Amatya R, Vajpayee M, Kaushik S, Kanswal S, Pandey RM, Seth P. Lymphocyte immunophenotype reference ranges in healthy Indian adults: implications for management of HIV/AIDS in India. *Clin Immunol*. 2004;112(3):290–5.
- Zhang K, Wang F, Zhang M, Cao X, Yang S, Jia S, et al. Reference ranges of lymphocyte subsets balanced for age and gender from a population of healthy adults in Chongqing District of China. *Cytometry B Clin Cytom*. 2016;90(6):538–42.
- Shahal-Zimra Y, Rotem Z, Chezar J, Shochat T, Ross L, Pickholtz I, et al. Lymphocyte subset reference ranges in healthy Israeli adults. *Isr Med Assoc J*. 2016;18(12):739–43.
- Fest J, Rüter R, Ikram MA, Voortman T, Van Eijck CHJ, Stricker BH. Reference values for white blood-cell-based inflammatory markers in the Rotterdam study: a population-based prospective cohort study. *Sci Rep*. 2018. <https://doi.org/10.1038/s41598-018-28646-w>.
- Provinciali M, Moresi R, Donnini A, Lisa RM. Reference values for CD4+ and CD8+ T lymphocytes with Naive or memory phenotype and their association with mortality in the elderly. *Gerontology*. 2009;55(3):314–21.
- Cheng C, Chan J, Cembrowski G, Assendelft O. Complete blood count reference interval diagrams derived from NHANES III: stratification by age, sex, and race. *Lab Hematol Off Publ Int Soc Lab Hematol*. 2004;10:42–53.
- Qin L, Jing X, Qiu Z, Cao W, Jiao Y, Routy J-P, et al. Aging of immune system: immune signature from peripheral blood lymphocyte subsets in 1068 healthy adults. *Aging*. 2016;8(5):848–59.
- Reich D, Nalls MA, Kao WHL, Akyzbekova EL, Tandon A, Patterson N, et al. Reduced neutrophil count in people of African descent is due to a regulatory variant in the Duffy antigen receptor for chemokines gene. *PLoS Genet*. 2009;5(1):e1000360.
- Ford N, Meintjes G, Vitoria M, Greene G, Chiller T. The evolving role of CD4 cell counts in HIV care. *Curr Opin HIV AIDS*. 2017. <https://doi.org/10.1097/COH.0000000000000348>.
- Hogg RS, Yip B, Chan KJ, Wood E, Craib KJ, O'Shaughnessy MV, et al. Rates of disease progression by baseline CD4 cell count and viral load after initiating triple-drug therapy. *JAMA*. 2001;286(20):2568–77.
- De Zuani M, Lazničková P, Tomašková V, Dvončová M, Forte G, Stokin GB, et al. High CD4-to-CD8 ratio identifies an at-risk population susceptible to lethal COVID-19. *Scand J Immunol*. 2022;95(3):e13125.
- Huang W, Berube J, McNamara M, Saksena S, Hartman M, Arshad T, et al. Lymphocyte subset counts in COVID-19 patients: a meta-analysis. *Cytometry A*. 2020;97(8):772–6.
- Kucuksezer UC, Aktas Cetin E, Esen F, Tahrali I, Akdeniz N, Gelmez MY, et al. The role of natural killer cells in autoimmune diseases. *Front Immunol*. 2021. <https://doi.org/10.1097/COH.0000000000000348>.
- Salas LA, Koestler DC, Butler RA, Hansen HM, Wiencke JK, Kelsey KT, et al. An optimized library for reference-based deconvolution of whole-blood biospecimens assayed using the Illumina HumanMethylationEPIC Bead-Array. *Genome Biol*. 2018;19(1):64.
- Salas LA, Zhang Z, Koestler DC, Butler RA, Hansen HM, Molinaro AM, et al. Enhanced cell deconvolution of peripheral blood using DNA methylation for high-resolution immune profiling. *Nat Commun*. 2022. <https://doi.org/10.1038/s41467-021-27864-7>.
- Houseman EA, Accomando WP, Koestler DC, Christensen BC, Marsit CJ, Nelson HH, et al. DNA methylation arrays as surrogate measures of cell mixture distribution. *BMC Bioinformatics*. 2012;13(1):86.
- Koestler DC, Jones MJ, Usset J, Christensen BC, Butler RA, Kobor MS, et al. Improving cell mixture deconvolution by identifying optimal DNA methylation libraries (IDOL). *BMC Bioinformatics*. 2016. <https://doi.org/10.1186/s12859-016-0943-7>.
- Women's Health Initiative Study Group. Design of the women's health initiative clinical trial and observational study. *Control Clin Trials*. 1998;19(1):61–109.
- LaCroix AZ, LaCroix AZ. THE women's health initiative (WHI): still learning from 161,808 Postmenopausal women. *Innov Aging*. 2019;3(Suppl 1):S355.
- Aryee MJ, Jaffe AE, Corrada-Bravo H, Ladd-Acosta C, Feinberg AP, Hansen KD, et al. Minfi: a flexible and comprehensive bioconductor package for the analysis of Infinium DNA methylation microarrays. *Bioinformatics*. 2014;30(10):1363–9.
- Xu Z, Niu L, Li L, Taylor JA. ENmix: a novel background correction method for Illumina HumanMethylation450 BeadChip. *Nucl Acids Res*. 2016;44(3):e20.
- Triche TJ, Weisenberger DJ, Van Den Berg D, Laird PW, Siegmund KD. Low-level processing of illumina Infinium DNA methylation BeadArrays. *Nucl Acids Res*. 2013;41(7):e90.
- Zhou W, Laird PW, Shen H. Comprehensive characterization, annotation and innovative use of Infinium DNA methylation BeadChip probes. *Nucl Acids Res*. 2016. <https://doi.org/10.1093/nar/gkw967>.
- Bell-Glenn S, Salas LA, Molinaro AM, Butler RA, Christensen BC, Kelsey KT, et al. Calculating detection limits and uncertainty of reference-based deconvolution of whole-blood DAN methylation data [under review]. *Epigenomics*. 2023.
- Koestler DC, Christensen B, Karagas MR, Marsit CJ, Langevin SM, Kelsey KT, et al. Blood-based profiles of DNA methylation predict the underlying distribution of cell types: a validation analysis. *Epigenetics*. 2013;8(8):816–26.
- Gervin K, Salas LA, Bakulski KM, van Zelm MC, Koestler DC, Wiencke JK, et al. Systematic evaluation and validation of reference and library selection methods for deconvolution of cord blood DNA methylation data. *Clin Epigenetics*. 2019;11(1):125.
- Li M, Yao D, Zeng X, Kasakovski D, Zhang Y, Chen S, et al. Age related human T cell subset evolution and senescence. *Immun Ageing*. 2019. <https://doi.org/10.1186/s12979-019-0165-8>.

40. Márquez EJ, Chung C-H, Marches R, Rossi RJ, Nehar-Belaid D, Eroglu A, et al. Sexual-dimorphism in human immune system aging. *Nat Commun*. 2020. <https://doi.org/10.1038/s41467-020-14396-9>.
41. Patin E, Hasan M, Bergstedt J, Rouilly V, Libri V, Urrutia A, et al. Natural variation in the parameters of innate immune cells is preferentially driven by genetic factors. *Nat Immunol*. 2018;19(3):302–14.
42. Saule P, Trauet J, Dutriez V, Lekeux V, Dessaint J-P, Labalette M. Accumulation of memory T cells from childhood to old age: central and effector memory cells in CD4+ versus effector memory and terminally differentiated memory cells in CD8+ compartment. *Mech Ageing Dev*. 2006;127(3):274–81.
43. Alpert A, Pickman Y, Leipold M, Rosenberg-Hasson Y, Ji X, Gaujoux R, et al. A clinically meaningful metric of immune age derived from high-dimensional longitudinal monitoring. *Nat Med*. 2019;25(3):487–95.
44. Lin Y, Kim J, Metter EJ, Nguyen H, Truong T, Lustig A, et al. Changes in blood lymphocyte numbers with age in vivo and their association with the levels of cytokines/cytokine receptors. *Immun Ageing*. 2016. <https://doi.org/10.1186/s12979-016-0079-7>.
45. Wiencke JK, Koestler DC, Salas LA, Wiemels JL, Roy RP, Hansen HM, et al. Immunomethylomic approach to explore the blood neutrophil lymphocyte ratio (NLR) in glioma survival. *Clin Epigenet*. 2017. <https://doi.org/10.1186/s13148-017-0316-8>.
46. Yin X, Wu L, Yang H, Yang H. Prognostic significance of neutrophil-lymphocyte ratio (NLR) in patients with ovarian cancer: a systematic review and meta-analysis. *Medicine (Baltimore)*. 2019;98(45):e17475.
47. Mjaess G, Chebel R, Karam A, Moussa I, Pretot D, Abi Tayeh G, et al. Prognostic role of neutrophil-to-lymphocyte ratio (NLR) in urological tumors: an umbrella review of evidence from systematic reviews and meta-analyses. *Acta Oncol*. 2021;60(6):704–13.
48. Dudani S, Poodury S, Mangalesh S. Study of neutrophil-lymphocyte ratio (NLR) in recent onset type 2 diabetes mellitus. *Bali Medical J*. 2021;10(1):11.
49. Hessami K, Tabrizi R, Homayoon N, Hashemi A, Heydari ST, Pourhoseini SA. Gestational diabetes mellitus and inflammatory biomarkers of neutrophil-lymphocyte ratio and platelet-lymphocyte ratio: a systematic review and meta-analysis. *Biomarkers*. 2021;26(6):491–8.
50. Angkananard T, Anothaisintawee T, McEvoy M, Attia J, Thakkinstian A. Neutrophil lymphocyte ratio and cardiovascular disease risk: a systematic review and meta-analysis. *Biomed Res Int*. 2018;2018:2703518.
51. Haybar H, Pezeshki SMS, Saki N. Evaluation of complete blood count parameters in cardiovascular diseases: an early indicator of prognosis? *Exp Mol Pathol*. 2019;110:104267.
52. Tokgoz S, Kayrak M, Akpınar Z, Seyithanoğlu A, Güney F, Yürüten B. Neutrophil lymphocyte ratio as a predictor of Stroke. *J Stroke Cerebrovasc Dis*. 2013;22(7):1169–74.
53. Song S-Y, Zhao X-X, Rajah G, Hua C, Kang R-j, Han Y-p, et al. Clinical significance of baseline neutrophil-to-lymphocyte ratio in patients with ischemic stroke or hemorrhagic stroke: an updated meta-analysis. *Front Neurol*. 2019. <https://doi.org/10.3389/fneur.2019.01032>.

## Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Ready to submit your research? Choose BMC and benefit from:

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

Learn more [biomedcentral.com/submissions](https://biomedcentral.com/submissions)

