Development and validation of composite inflammaging metrics: Findings from two prospective cohort studies

Chenxi Li¹#, Yumeng Ge¹#, Zhenqing Yang¹, Gan Yang¹, Xingqi Cao¹, Jinjing Fu¹, Zuyun Liu¹*

Author affiliations:
¹School of Public Health and the Second Affiliated Hospital, The Key Laboratory of Intelligent Preventive Medicine of Zhejiang Province, Zhejiang University School of Medicine, Hangzhou 310058, Zhejiang, China

#These authors contributed equally to this work.

*Corresponding author:
Zuyun Liu, PhD, Professor, Center for Clinical Big Data and Analytics of the Second Affiliated Hospital and Department of Big Data in Health Science School of Public Health, The Key Laboratory of Intelligent Preventive Medicine of Zhejiang Province, Zhejiang University School of Medicine, 866 Yuhangtang Rd, Hangzhou, 310058, Zhejiang, China. Telephone: +86-0571-87077127. Email: Zuyun.liu@outlook.com or zuyunliu@zju.edu.cn
Abstract
Background
Aging-related inflammation is associated with chronic diseases and mortality. This study aimed to: 1) develop composite inflammaging metrics (CIMs) in UK biobank (UKB), and validate them in UKB and National Health and Nutrition Examination Survey (NHANES); 2) estimate mortality and CVD risk predictions of CIMs; 3) compare CIMs with single inflammatory blood biomarkers and conventional inflammatory indexes; 4) examine associations between lifestyles and CIMs.

Methods
We utilized algorithms including multiple linear regression, principal component analysis (PCA), allostatic load (AL), and Klemera and Doubal method (KDM), to develop four CIMs from five inflammatory blood biomarkers, using data of 438,321 adults (40-70 years) from UK Biobank (UKB). We validated these CIMs in UKB and 10,667 adults (20-84 years) from NHANES IV. We performed a parametric proportional hazard model based on Gompertz distribution to estimate CVD and mortality risk predictions of CIMs. Areas under receiver operating characteristic curves (AUCs) were calculated to compare the predictive abilities of CIMs. Multiple linear regression models were used to access associations between lifestyles and CIMs.

Results
With adjustment for age and sex, four CIMs were significantly associated with higher risks of all-cause mortality and incident CVD in UKB, among which CIM_{KDM} outperformed the others (all-cause mortality: hazard ratio [HR] = 1.48, 95% confidence interval [CI] = 1.46, 1.50; incident CVD: HR = 1.34, 95% CI = 1.33, 1.36). CIM_{KDM} had the best discriminative ability for predicting 10-year survival and incident CVD in UKB (all-cause mortality: AUC = 0.728; incident CVD: AUC = 0.712). CIMs were responsive to lifestyle variables. For example, in UKB, compared to never smokers, current smokers had a significant increment in CIM_{KDM} (coefficient = 0.30 SD, \( P < 0.001 \)). Similar results were well validated in NHANES IV.

Conclusions
We developed and validated four novel CIMs that were predictive of mortality and CVD risk. CIM_{KDM} outperformed the others and had the potential to be used in aging related preventive and intervention programs. Intervention programs targeting lifestyles could slow inflammaging and further reduce disease burden.

Keywords: Aging; Inflammation; Composite inflammaging metrics; Mortality; Cardiovascular Disease
1. Introduction
Inflammation has been recognized as one of the important factors driving aging\(^1\). Low-grade, chronic, and systemic inflammation has been termed as “inflammaging” in aging process, which can result in high risk of mortality and age-related chronic disease\(^2-4\), such as cardiovascular disease (CVD)\(^5\), and Alzheimer's disease\(^6\). Previous evidence has revealed that commonly used acute inflammation biomarkers, e.g., interleukin-6 (IL-6), C-reactive protein (CRP), and tumor necrosis factor-alpha (TNF-\(\alpha\)), can independently capture the effect of chronic inflammation on mortality\(^7,8\). There are still no canonical standard biomarkers for inflammaging\(^11\). Moreover, one certain inflammatory biomarker cannot comprehensively mirror the whole inflammatory status in the aging process, which emphasizes the importance of composite measures\(^12\).

Considering the necessity to identify the underlying inflammaging states, researchers have developed lots of composite inflammatory indexes as predictors of mortality and chronic diseases. Of conventional composite inflammatory indexes, systemic inflammation response index (SIRI)\(^14\), systemic immune-inflammation index (SII)\(^15\), neutrophil-to-lymphocyte ratio (NLR)\(^16\), platelet-to-lymphocyte ratio (PLR)\(^17\), and lymphocyte-to-monocyte ratio (LMR)\(^18\) have been widely applied in predicting mortality and pre-diagnostic age-related disease\(^20-22\). Besides, some researchers have used advanced algorithms, such as principal component analysis (PCA)\(^23\) and deep-learning\(^24\), trying to construct inflammaging metrics, but these metrics are too complex, and lack generalization in diverse population of wide age range. Therefore, there is a need to develop new inflammaging metric with practical value.

While Klemera and Doubal method (KDM) has been widely used in developing biological aging measure\(^25-27\), it has rarely been utilized in inflammatory estimation yet. Besides, allostatic load (AL) is usually regarded as an indicator of physiological dysregulation resulting from cumulative physiological stress that in relation to health outcomes along aging\(^28-30\), of which inflammation strain is a portion. In this study, four algorithms including multiple linear regression, PCA, AL, and KDM were used to develop composite inflammaging metrics (CIMs). Afterwards, it is vital to verify whether these CIMs have a better performance than single inflammatory blood biomarkers and conventional inflammatory indexes. Slowing inflammaging process through modifiable factors is a natural need for public health. Whether healthy lifestyles could influence inflammaging status reflected by CIMs is significant. Thus, this study aimed to: 1) develop four CIMs in UK biobank (UKB), and validate them in UKB and National Health and Nutrition Examination Survey (NHANES); 2) estimate mortality and CVD risk predictions of CIMs; 3) compare CIMs with single inflammatory blood biomarkers and conventional inflammatory indexes; 4) examine associations between lifestyles and CIMs.

2. Methods
2.1 Study participants
UKB is a large prospective cohort of national participants from across the UK, which recruited over 500,000 participants from 2006 to 2010 at 22 assessment centers.
Participants’ information on demographics, lifestyle, medical conditions, and other health-related aspects was provided through interviews and physical measurements at baseline. More detailed information on UKB can be seen elsewhere\textsuperscript{31}. We excluded participants with missing data on covariates (N = 23,243) or five inflammatory blood biomarkers (N = 40,882), leaving 438,321 participants aged 40-70 years to develop and validate CIMs. Details of the participants selection in UKB are shown in Figure S1. UKB study was approved by the North West Multi-centre Research Ethics Committee (ref. 11/NW/0382). All participants provided written informed consent.

NHANES is an ongoing programme aiming to evaluate US civilian’s health and nutritional status, carried out by the National Center for Health Statistics. NHANES collected data from nationally representative cross-sectional surveys. More detailed information on NHANES can be seen elsewhere (https://www.cdc.gov/nchs/nhanes/index.htm). We used data from NHANES IV (1999–2010, N = 13,823) as an independent validation sample. We excluded participants who did not accomplish 8 hours of fasting before blood sampling or those with missing data on biomarkers (N = 1,368), follow-up time (N = 15), survey weights (N = 1,008), or covariates (N = 765). We included 10,667 participants aged 20-84 years in final analytic sample to validate CIMs. Details of participants selection in NHANES IV are shown in Figure S1. NHANES was approved by the National Center for Health Statistics Research Ethics Review Board. All participants provided informed consent.

2.2 Inflammatory metrics

2.2.1 Inflammatory blood biomarkers
Considering the feasibility and accessibility of novel inflammaging metric, we included five blood biomarkers: platelet count (1000cell/uL), lymphocyte count (1000cell/uL), monocyte count (1000cell/uL), neutrophil count (1000cell/uL), and CRP (g/L). Lymphocyte count, monocyte count, neutrophil count, and CRP were logarithmically transformed to approximate normal distribution in regression.

2.2.2 Conventional inflammatory indexes
We calculated five conventional inflammatory indexes based on inflammatory blood biomarkers: SII\textsuperscript{15}, SIRI\textsuperscript{14}, NLR\textsuperscript{32,33}, PLR\textsuperscript{17,34}, and LMR\textsuperscript{18,35}. The formula is presented as follows, where P, N, L, and M are platelet, neutrophil, lymphocyte, and monocyte counts, respectively:

\[
\begin{align*}
\text{SII} & = P \times N / L \\
\text{SIRI} & = N \times M / L \\
\text{NLR} & = N / L \\
\text{PLR} & = P / L \\
\text{LMR} & = L / M
\end{align*}
\]

2.2.3 CIM\textsubscript{lin} (composite inflammaging metric using linear regression)
We used multiple linear regression model to regress five inflammatory blood
biomarkers on age, and defined prediction as CIM\textsubscript{lm}. The formula of CIM\textsubscript{lm} is presented as follows:

\[
\text{CIM}_{\text{lm}} = -0.01219 \times P - 0.66830 \times \ln(L) + 2.00501 \times \ln(M) - 0.14427 \times \ln(N) \\
+ 0.95186 \times \ln(CRP) + 62.02598
\]

2.2.4 CIM\textsubscript{PCA} (composite inflammaging metric using Principal Component Analysis)

We did PCA with five inflammatory blood biomarkers, and took four components out of five according to Scree Plot (Figure S2) and cumulative proportion of variance interpretation (over 80%, in Table S1). Then, we used multiple linear regression model to regress four selected components on age, and defined prediction as CIM\textsubscript{PCA}. The formula of CIM\textsubscript{PCA} is presented as follows:

\[
\text{CIM}_{\text{PCA}} = -0.00182 \times P - 0.01717 \times \ln(L) + 0.16207 \times \ln(M) + 0.06417 \times \ln(N) \\
+ 0.11384 \times \ln(CRP) + 0.47787
\]

2.2.5 CIM\textsubscript{AL} (composite inflammaging metric using allostatic load)

We used AL\textsuperscript{29} to calculate CIM\textsubscript{AL} based on five inflammatory blood biomarkers. The 75th percentile (25th percentile additionally for platelet count) was determined as the cut-off for five biomarkers. Every biomarker with values above the cut-off (below additionally for platelet count) was defined as ‘1’ otherwise ‘0’, and the sum of five biomarkers was calculated as CIM\textsubscript{AL}, ranging from 0 to 5. The higher CIM\textsubscript{AL}, the worse a participant’s inflammatory status. The normal range (defined as ‘0’) for five inflammatory blood biomarkers is presented as follows:

1) 213.4-287.0 1000cell/uL for P;
2) 0-2.30 1000cell/uL for L;
3) 0-0.57 1000cell/uL for M;
4) 0-4.95 1000cell/uL for N;
5) 0-2.72 g/L for CRP.

2.2.6 CIM\textsubscript{KDM} (composite inflammaging metric using Klemera-Doubal Method)

We used KDM\textsuperscript{25,36} to calculate CIM\textsubscript{KDM} based on five inflammatory blood biomarkers. The formula of CIM\textsubscript{KDM} is presented as follows:

\[
\text{CIM}_{\text{KDM}} = -0.06911 \times P - 0.65665 \times \ln(L) + 15.47522 \times \ln(M) + 5.07968 \times \ln(N) \\
+ 6.92480 \times \ln(CRP) + 0.78113 \times \text{age} + 33.59987
\]

2.3 Outcomes

The primary outcome in UKB was all-cause mortality, which was obtained from the National Health Service Central Register Scotland (Scotland) and the National Health Service Information Centre (England and Wales). Detailed information is available from https://content.digital.nhs.uk/services. In this study, participants were followed up from baseline to the date of death or Aug 29, 2021 (censoring time), whichever
occurred first. The secondary outcome in UKB was incident CVD events. We defined CVD according to the International Statistical Classification of Diseases and Related Health Problems 9th (ICD-9) codes (410.x, 411.x, 412.x, 413.x, 414.x, 429.79, 430.x, 431.x, 432.x, 433.x, 434.x, 435.x, 436.x, 437.x, 438.x), and 10th edition (ICD-10) codes (I20.x, I21.x, I22.x, I23.x, I24.1, I25.x, I46, I60.x, I61.x, I63.x, I64.x)37. Participants who were diagnosed as CVD at baseline were excluded. Participants were followed up from baseline to the first diagnosis of CVD, the date of death, or Aug 29, 2021 (censoring time), whichever occurred first.

The primary outcome in NAHNES IV was all-cause mortality. NHANES IV data were linked to the National Death Index, provided by the Centers for Disease Control and Prevention (https://www.cdc.gov/nchs/data-linkage/mortality-public.htm). Participants were followed up from baseline to the date of death or December 31, 2019 (censoring time), whichever occurred first. The secondary outcome in NAHNES IV was CVD-specific mortality. We defined CVD-specific mortality using ICD-10 codes (I00-I09, I11, I13, I20-I51, or I60-I69)38.

### 2.4 Covariates and lifestyles

In UKB, covariates included age, sex (male and female), ethnic group (white and others), educational level (high, intermediate, and low), current employment status (working, retired, and others), Townsend deprivation index39, and disease counts (ranging 0-9, including dementia, CVD, hypertension, cancer, diabetes, cholelithiasis, osteoarthritis, respiratory disease, and depression). In NHANES IV, covariates included age, sex (male and female), ethnic group (non-Hispanic white, non-Hispanic black, Hispanic, and others), educational level (high, intermediate, and low), and disease counts (ranging 0-10, including congestive heart failure, stroke, cancer, chronic bronchitis, emphysema, cataracts, arthritis, type 2 diabetes mellitus, hypertension, and myocardial infarction).

In both UKB and NHANES IV, lifestyles included smoking status (never, previous, and current), alcohol intake frequency (“never or special occasions only”, “one to three times per month”, “one to four times per week”, and “daily or almost daily”), regular exercise (yes and no)39,40, healthy diet (yes and no)39,40, and body mass index (BMI) (<18.5, ≥18.5 and <25, ≥25 and <30, and ≥30 kg/m2).

### 2.5 Statistical analyses

Characteristics of the participants in two analytic samples were presented as numbers (percentages) for categorical variables, and mean ± standard deviation (SD) for continuous variables, respectively. All inflammatory metrics were scaled in analyses to assure them comparable.

First, we developed four CIMs in UKB. Second, we examined whether CIMs could stratify the risk of outcomes onset. We observed incident event rate plotted against CIMs percentiles. We plotted Kaplan-Meier curves of cumulative hazard of outcomes for persons in the lowest 20%, the middle 20% and the highest 20% of CIMs. Third, we estimated CIMs’ predictive validity in mortality and incident CVD with a basic model including age and sex. We used a parametric proportional hazard
model (Gompertz distribution) to evaluate the associations between CIMs and outcomes, and estimated hazard ratios (HRs). We estimated areas under receiver operating characteristic curves (AUCs) of 10-year outcomes incidence risk prediction. We used calibration curve to observe CIMs’ discriminative performance on 10-year outcomes incidence risk predictions. Fourth, we estimated HRs and AUCs for single inflammatory blood biomarkers and conventional inflammatory indexes to examine whether CIMs had a better performance than them. Sixth, we used multiple linear regression model to evaluate associations between lifestyles and inflammatory metrics with all covariates adjusted for, to observe whether healthy lifestyles can slow a person’s inflamming status. Seventh, we calculated four CIMs in NHANES IV and repeated all the analyses with outcomes as all-cause mortality and CVD-specific mortality. Eighth, we did two sensitivity analyses of estimation on HRs and AUCs: 1) in relatively healthy population (disease counts equal to 0); 2) with disease counts additionally adjusted for.

We used R version 4.1 to conduct all statistical analyses. A P value < 0.05 (two-tailed) was defined as statistically significant.

3. Results
3.1 Characteristics of study participants
Characteristics of participants were presented in Table S2. Of 438,321 participants in UKB, the mean age was 57.0 ± 8.1 years, and 200,788 (45.8%) participants were male. A total of 29,218 (6.7%) participants died during a median follow-up time of 12.1 years (interquartile range, IQR 11.4–12.8 years), and a total of 31,812 (7.8%) participants out of 409,460 (excluding 28,861 participants diagnosed with CVD at baseline) developed CVD during a median follow-up time of 12.4 years (IQR 11.6–13.2 years).

Of 10,667 participants in the NHANES IV, the mean age was 48.9 ± 17.7 years, and 5,265 (49.4%) participants were male. A total of 2,047 (19.2%) participants died during a median follow-up time of 13.1 years (IQR 10.3–17.0 years), and a total of 608 (5.7%) participants developed CVD-specific death during the same follow-up time.

3.2 Characteristics of CIMs
In UKB, mean of CIM lm was 57.0 years (SD = 1.4). Mean of CIM PCA was 0.0 (SD = 0.2). Mean of CIM AL was 1.5 (SD = 1.2). Mean of CIM KDM was 56.5 years (SD = 13.3). In NHANES IV, mean of CIM lm was 57.6 years (SD = 1.5). Mean of CIM PCA was 0.0 (SD = 0.2). Mean of CIM AL was 1.8 (SD = 1.2). Mean of CIM KDM was 54.7 years (SD = 19.5). Figure S3 presents CIMs’ distribution. Figure S4 presents CIMs’ correlation with age. As expected, CIM KDM were highly correlated with age, because age was in CIMKDM measure. Other CIMs showed weak correlation with age.

3.3 CIMs stratify the risk of outcomes onset
As shown in Figure 1a, we observed increasing all-cause mortality and CVD incidence rates over CIM percentiles in UKB. As shown in Figure 1b, we observed
increasing all-cause mortality and CVD-specific mortality rates roughly over CIM percentiles in NHANES IV. CIMKDM outperformed the other.

We calculated the cumulative hazard for persons in the lowest 20%, the middle 20% and the highest 20% of CIMs (Figure 2), with 95% confidence intervals (CIs). As shown in Figure 2a, all CIMs stratified the different risk level of outcomes onset clearly in UKB, among which CIMKDM outperformed the other. As shown in Figure 2b, CIMlm and CIMPCA stratified the different risk level of outcomes onset clearly in NHANES IV. CIMAL barely stratify the risk level of persons in the middle 20% and the highest 20%, while CIMKDM barely stratify the risk level of persons in the middle 20% and the lowest 20%.

3.4 Associations between inflammatory metrics and outcomes

Table 1 shows the associations between inflammatory metrics and outcomes with age and sex adjusted for. We found that in UKB, each 1-SD increase in CIMlm increased the risk of all-cause mortality and CVD incidence by 32% (HR = 1.32, 95% CI = 1.30, 1.33) and 22% (HR = 1.22, 95% CI = 1.20, 1.23). Each 1-SD increase in CIMPCA increased the risk of all-cause mortality and CVD incidence by 33% (HR = 1.33, 95% CI = 1.31, 1.34) and 23% (HR = 1.23, 95% CI = 1.22, 1.25). Each 1-SD increase in CIMAL increased the risk of all-cause mortality and CVD incidence by 33% (HR = 1.33, 95% CI = 1.31, 1.34) and 23% (HR = 1.23, 95% CI = 1.22, 1.24). Each 1-SD increase in CIMKDM increased the risk of all-cause mortality and CVD incidence by 48% (HR = 1.48, 95% CI = 1.46, 1.50) and 34% (HR = 1.34, 95% CI = 1.33, 1.36). Among single inflammatory blood biomarkers, CRP performed the best. Each 1-SD increase in ln(CRP) increased the risk of all-cause mortality and CVD incidence by 34% (HR = 1.34, 95% CI = 1.33, 1.36) and 26% (HR = 1.26, 95% CI = 1.25, 1.28). Among conventional inflammatory indexes, SIRI performed the best. Each 1-SD increase in SIRI increased the risk of all-cause mortality and CVD incidence by 31% (HR = 1.31, 95% CI = 1.30, 1.33) and 13% (HR = 1.13, 95% CI = 1.12, 1.14). Overall, CIMs had a better performance than single inflammatory blood biomarkers and conventional inflammatory indexes, among which CIMKDM outperformed the other. Similar results were found in NHANES IV as outcomes were all-cause mortality and CVD-specific mortality.

Figure 3 shows AUCs with a basic model including age and sex. All models of CIMs well calibrated in UKB and NHANES IV for 10-year event rate (Figure S5). For 10-year all-cause mortality and CVD incidence in UKB, basic model had an AUC of 0.713 (95% CI = 0.709, 0.716) and 0.700 (95% CI = 0.696, 0.703). CIMlm had an AUC of 0.725 (95% CI = 0.721, 0.728) and 0.708 (95% CI = 0.705, 0.711). CIMPCA had an AUC of 0.725 (95% CI = 0.722, 0.729) and 0.709 (95% CI = 0.706, 0.712). CIMAL had an AUC of 0.728 (95% CI = 0.724, 0.731) and 0.710 (95% CI = 0.707, 0.713). CIMKDM had an AUC of 0.728 (95% CI = 0.725, 0.732) and 0.712 (95% CI = 0.709, 0.715). Among single inflammatory blood biomarkers, CRP performed the best. ln(CRP) had an AUC of 0.728 (95% CI = 0.725, 0.731) and 0.712 (95% CI = 0.709, 0.715). Among conventional inflammatory indexes, SIRI performed the best. SIRI had an AUC of 0.724 (95% CI = 0.721, 0.728) and 0.703 (95% CI = 0.700, 0.706).
CIMs had a better performance than conventional inflammatory indexes. CIM\textsubscript{KDM} and CRP had similar great performance in UKB. However, in NHANES IV, CIM\textsubscript{KDM} outperformed CRP. For 10-year all-cause mortality and CVD-specific mortality, ln(CRP) had an AUC of 0.886 (95% CI = 0.876, 0.896) and 0.885 (95% CI = 0.868, 0.902). CIM\textsubscript{KDM} had an AUC of 0.888 (95% CI = 0.878, 0.898) and 0.889 (95% CI = 0.872, 0.906). Compared to single inflammatory blood biomarkers and conventional inflammatory indexes, CIMs had a better performance in NHANES IV.

3.5 Associations between lifestyles with inflammatory metrics

Figure 4 shows the associations between all lifestyle variables (including smoking status, alcohol intake frequency, regular exercise, healthy diet and BMI, simultaneously input) with inflammatory metrics.

Overall, inflammatory blood biomarkers, conventional inflammatory indexes, and CIMs were responsive to lifestyle variables. Unhealthy lifestyles were associated with a serious inflammatory status. For example, in UKB, compared to never smokers, current smokers had a significant increment in ln(CRP) (coefficient = 0.28 SD, \( P < 0.001 \)), SII (coefficient = 0.05 SD, \( P < 0.001 \)), and CIM\textsubscript{KDM} (coefficient = 0.30 SD, \( P < 0.001 \)). Interestingly, two relatively unhealthy lifestyle statuses (high alcohol intake frequency and BMI < 18.5) were associated with lower level of inflammatory blood biomarkers and CIMs. For example, in UKB, compared to participants with alcohol intake frequency of “never or special occasions only”, participants with alcohol intake frequency of “daily or almost daily” had a significant decrement in ln(CRP) (coefficient = -0.10 SD, \( P < 0.001 \)) and CIM\textsubscript{KDM} (coefficient = -0.04 SD, \( P < 0.001 \)). Compared to participants with normal weight (BMI \( \geq 18.5 \) and < 25), underweight participants (BMI < 18.5) had a significant decrement in ln(CRP) (coefficient = -0.50 SD, \( P < 0.001 \)) and CIM\textsubscript{KDM} (coefficient = -0.27 SD, \( P < 0.001 \)). The results in NHANES IV were similar to those in UKB.

3.6 Sensitivity analyses

We repeated analyses of estimation on HRs and AUCs twice under conditions below: 1) in relatively healthy population (disease counts equal to 0) (Table S3); 2) with disease counts additionally adjusted for (Table S4). We observed no substantial changes in results.

4. Discussion

In this large-scale population study, we developed and validated novel measures of inflammaging status—CIMs, including CIM\textsubscript{Im}, CIM\textsubscript{PCA}, CIM\textsubscript{AL} and CIM\textsubscript{KDM}. These metrics were demonstrated to be leveraged as predictors of all-cause mortality, CVD incidence, and CVD-specific mortality. It is observed that CIMs capturing risk more precisely beyond single inflammatory blood biomarkers and conventional inflammatory indexes, with exception of CRP. Furthermore, CIM\textsubscript{KDM} outperformed other CIMs in that it successfully stratified risk and manifested its preponderance of discrimination and calibration in 10-year outcomes predictions. The association
between some common unhealthy lifestyles and lower CIMs elucidated the feasibility of deferring inflammaging process by lifestyle intervention.

Our metrics took five inflammatory biomarkers into consideration, that is, platelet count, lymphocyte count, monocyte count, neutrophil count, and CRP, which are commonly tested in clinical practice. As we noticed, platelet count, lymphocyte count or monocyte count were not stable enough to be an indicator, which might be attributed to the complex mechanism of immune system. In the case of lymphocyte, T cells are deemed to be the main modulatory participants in atherosclerotic which accounts for the majority of CVD; yet the diverse differentiations and possible transformation of T lymphocytes determine their function to secrete pro- or anti-inflammatory cytokines, thus explaining why both the promotive and protective effects coexist. With that, some studies held the viewpoint that low lymphocyte count poses threat to CVD occurrence and progression. Though we assumed composite metrics would surpass individual biomarkers, the good performance of CRP could not be denied as its performance was close to CIMKDM. Of conventional inflammatory indexes, elevated SIRI, having the similar capability to assess the risk of mortality with CIMlm, seemed to be stronger predictors than others. Compared to these former indexes, in principle, CIMs had more edge over the prediction of onset across CVD than of mortality risk. On the one hand, conventional inflammatory indexes were calculated by multiplication and division among the platelet count, lymphocyte count, monocyte count, and neutrophil count, whereas the supposed functions of lymphocyte and monocyte were contradictory. According to our discussion above, it is untenable to posit that one certain inflammatory biomarker stay consistencies in its physiological roles of inflammaging. On the other hand, none of these indexes considered CRP while it is certainly an important inflammatory biomarker. Therefore, CIMs (CIMKDM in particular) are better than conventional inflammatory indexes from the perspective of algorithm. Moreover, conventional inflammatory indexes are regularly used, and might more suitable for linking systemic inflammation to prognosis or adverse outcomes in patients with cancer, CVD, and so on. For instance, our findings suggested the probability of the negative correlations between PLR/LMR and outcomes. Similarly, another prospective study based on UKB also reported inconsistent results that there are positive associations with risk of several cancers for PLR but negative associations for LMR.

To the best of our knowledge, we are the first to derive inflammaging metrics from platelet count, lymphocyte count, monocyte count, neutrophil count, and CRP, plus via four algorithms in such a nationwide prospective cohort. There are discrepancies between CIMs in predicting mortality and CVD risk. CIMKDM was the most pronounced one since it was robust in participants who were relatively healthy at baseline. Additionally, the strength of CIMKDM was conspicuous in drawing a distinction between high and low risk groups and higher accuracy of prediction. While the discriminatory power of CIMlm and CIMPCA were inferior to CIMKDM, they could also be valuable metrics for CVD risk prediction since they predicted with sufficient accuracy as well as CIMKDM. The hazard trajectory of all-cause mortality
and CVD-specific mortality in NHANES IV suggested that, when follow-up time became longer, CIM\textsubscript{KDM} was plausible of less predictive value for individual of moderate risk. Still, sample size was also an inevitable factor to this deficiency. Even though, as an excellent biological aging measures proved by many former investigations\textsuperscript{26,36}, KDM remains its intrinsic superiority of algorithm over multiple linear regression and PCA accordingly. Such a overlap of hazard trajectory emerged in different quantile of CIM\textsubscript{AL} likewise. CIM\textsubscript{AL} could be a fine indicator in terms of stability and discrimination, but its ability to identify people at risk was limited inasmuch as AL score is a discrete variable. More concretely, CIM\textsubscript{AL} was an integer ranged between 0 to 5, which meant the possible score of it was much smaller versus widespread AL score that combined more biomarkers from various systems\textsuperscript{28,49}.

Inflammaging has been perceived as a vital target for antiaging and curbing the onset or progression of age-related disease\textsuperscript{50}. Chronic exposure to exogenous risk factors like unhealthy lifestyles\textsuperscript{51}, augments inflammatory burden risk, thereby fueling inflammaging\textsuperscript{5}. Many researches on pathophysiology have already revealed that unhealthy lifestyles, such as smoking\textsuperscript{52,53}, alcohol intake\textsuperscript{54}, physical inactivity\textsuperscript{55,56}, unhealthy diet\textsuperscript{57,58}, and obesity\textsuperscript{59,60}, could trigger or increase chronic inflammation. At same time, associations of unhealthy lifestyles with CVD and mortality risk was widely illustrated in cohort studies\textsuperscript{61,62}. Therefore, with evaluation of the association of inflammaging indexes and unhealthy lifestyles, we found that those who kept away from smoking, did regular exercise, had healthy diet and kept fit, had lower CIMs. That gave us an insight into healthy lifestyles that may well attenuate inflammaging. However, alcohol intake and underweight were shown as positive factors to CIMs. Alcohol intake was assumed to have a J-shaped relationship between CVD, namely low-to-moderate drinkers have lower risk\textsuperscript{63} since some beverage types like red wine are likely to exert anti-inflammatory and antioxidant effects\textsuperscript{64}. Thus, while moderate amount of alcohol intake remains controversial, our study offered some evidence for the relationship between inflammatory biomarkers and alcohol intake. Obesity is accompanied by chronic, low-grade inflammation\textsuperscript{60}; but the concomitant low fat-free mass in the aging process has been linked to inflammation as well, for example, the cachexia in CVD\textsuperscript{65} or cancer\textsuperscript{66}. According to a previous study, middle-aged people’s BMI varied linearly with CRP while elderly people with low body mass showed chronic inflammation\textsuperscript{67}. So, a possible explanation was that most of the underweight participants were not being undernourished. Broadly speaking, positively alteration of unhealthy lifestyles would contribute to practical approaches to prevention and intervention of inflammaging and related diseases.

The major strength of our study is that we derived CIMs via four well-performed algorithms in a large-scale prospective cohort (i.e., UKB) and achieved external validation in another independent population-based cohort (i.e., NHANES). The generalization of results in sensitivity analyses were basically recognized. Besides, five primary inflammatory biomarkers used by CIMs are of quick availability, easy detectability and low-priced examination cost. Meanwhile, there are several limitations in our study. Firstly, given the small sample size of the validation sample (i.e., NHANES), general transferability of CIMs needs to be validated in larger
cohort. Secondly, to enhance the prediction ability, the coverage of biomarkers calls for extension, and more inflamming biomarkers with practical application value need to be identified. Finally, in sight of complicated mechanisms on pathophysiology, advanced methods that incorporating two opposite aspects of inflammation and immune may be favorable for representing inflamming. Of course, more aetiological researches are expected to offer deeper interpretation to inflamming.

5. Conclusions
In summary, applying four algorithms—multiple linear regression, PCA, AL, and KDM, we developed four novel CIMs in UKB and validated them externally in US adults. Furthermore, we elucidated their value in mortality and incident CVD risk prediction, and verified their response to modifiable lifestyle factors. The findings support the implementation of preventive strategies and intervention programs targeting lifestyles to slow inflamming and further reduce disease burden.
Funding
This work was supported by grants from the Natural Science Foundation of Zhejiang Province (LQ21H260003), the National Natural Science Foundation of China (82171584), the 2020 Milstein Medical Asian American Partnership Foundation Irma and Paul Milstein Program for Senior Health project award (ZL), fundings from Key Laboratory of Intelligent Preventive Medicine of Zhejiang Province (2020E10004), and Zhejiang University Global Partnership Fund (188170-11103). The funders had no role in the study design; data collection, analysis, or interpretation; in the writing of the report; or in the decision to submit the article for publication.

Acknowledgments
This study has been conducted using the UK Biobank resource under application number 61856. We wish to thank the UK Biobank participants providing the sample that made the data available. We also thank all participants who took part in the National Health and Nutrition Examination Survey.

Author contributions
Z.L. designed and supervised the study. C.L. and Y.G. analyzed the data. C.L., Y.G., and Z.L. interpreted the results. C.L. and Y.G. drafted the manuscript. All authors revised the manuscript. Z.L. took responsibility for the content of the article. All authors read and approved the final version of the manuscript.

Competing interests
The authors declare no competing interests.
Reference


Table 1. Associations between inflammatory metrics and outcomes.

<table>
<thead>
<tr>
<th></th>
<th>All-cause mortality in UKB</th>
<th>CVD incidence in UKB</th>
<th>All-cause mortality in NHANES</th>
<th>CVD-specific mortality in NHANES</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>HR (95% CI)</td>
<td>P value</td>
<td>HR (95% CI)</td>
<td>P value</td>
</tr>
<tr>
<td>Platelet count</td>
<td>1.05 (1.03, 1.06)</td>
<td>&lt;0.001</td>
<td>1.05 (1.04, 1.06)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Lymphocyte count</td>
<td>0.98 (0.97, 0.99)</td>
<td>0.004</td>
<td>1.11 (1.10, 1.12)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Monocyte count</td>
<td>1.31 (1.29, 1.33)</td>
<td>&lt;0.001</td>
<td>1.19 (1.18, 1.21)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Neutrophil count</td>
<td>1.17 (1.15, 1.18)</td>
<td>&lt;0.001</td>
<td>1.13 (1.12, 1.14)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>C-reactive protein</td>
<td>1.34 (1.33, 1.36)</td>
<td>&lt;0.001</td>
<td>1.26 (1.25, 1.28)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>SII</td>
<td>1.20 (1.19, 1.21)</td>
<td>&lt;0.001</td>
<td>1.06 (1.05, 1.07)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>SIRI</td>
<td>1.31 (1.30, 1.33)</td>
<td>&lt;0.001</td>
<td>1.13 (1.12, 1.14)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>NLR</td>
<td>1.23 (1.22, 1.25)</td>
<td>&lt;0.001</td>
<td>1.05 (1.04, 1.07)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>PLR</td>
<td>1.02 (1.01, 1.03)</td>
<td>0.001</td>
<td>0.93 (0.92, 0.94)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>LMR</td>
<td>0.86 (0.85, 0.87)</td>
<td>&lt;0.001</td>
<td>0.98 (0.97, 0.99)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>CIMlm</td>
<td>1.32 (1.30, 1.33)</td>
<td>&lt;0.001</td>
<td>1.22 (1.20, 1.23)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>CIMPCA</td>
<td>1.33 (1.31, 1.34)</td>
<td>&lt;0.001</td>
<td>1.23 (1.22, 1.25)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>CIMAL</td>
<td>1.33 (1.31, 1.34)</td>
<td>&lt;0.001</td>
<td>1.23 (1.22, 1.24)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>CIMKDM</td>
<td>1.48 (1.46, 1.50)</td>
<td>&lt;0.001</td>
<td>1.34 (1.33, 1.36)</td>
<td>&lt;0.001</td>
</tr>
</tbody>
</table>

Abbreviations: UKB, UK Biobank; NHANES, National Health and Nutrition Examination Survey; HR, hazard ratio; CI, confidence interval; CVD, cardiovascular disease; SII, Systemic immune-inflammation index; SIRI, Systemic Inflammation Response Index; NLR, Neutrophil to lymphocyte ratio; PLR, Platelet to Lymphocyte Ratio; LMR, Lymphocyte to monocyte Ratio; CIM, Composite inflammaging metric; CIMlm, Composite inflammaging metric using linear regression; CIMPCA, Composite inflammaging metric using Principal Component Analysis (PCA); CIMAL, Composite inflammaging metric using allostatic load; CIMKDM, Composite inflammaging metric using Klemera-Doubal Method (KDM).

All models were adjusted for age and sex. All biomarkers were taken the logarithm before being scaled except for Platelet count. All biomarkers, conventional inflammatory metrics and CIMs were scaled. HR represented increased/decreased risk per 1-SD increment.

...
Figure legend:

**Figure 1.** Associations between CIMs and event rate in UK Biobank and NHANES IV.

a. All-cause mortality rate and CVD incidence rate based on CIM percentile in UK Biobank. b. All-cause mortality rate and CVD-specific mortality rate based on CIM percentile in NHANES IV.

Abbreviation: NHANES, National Health and Nutrition Examination Survey; CIM, Composite inflammaging metric; CIMlm, Composite inflammaging metric using linear regression; CIMPCA, Composite inflammaging metric using Principal Component Analysis (PCA); CIM_AL, Composite inflammaging metric using allostatic load; CIM_KDM, Composite inflammaging metric using Klemera-Doubal Method (KDM).

**Figure 2.** Cumulative hazard curves of event rate in UK Biobank and NHANES IV.

a. Cumulative hazard of all-cause mortality and CVD incidence in UK Biobank stratified by CIMs quantiles. b. Cumulative hazard of all-cause mortality and CVD-specific mortality in NHANES IV stratified by CIMs quantiles.

CIMs quantiles are indicated by light green (bottom 20%), green (median 20%), dark green (top 20%) with 95% confidence intervals.

Abbreviations: NHANES, National Health and Nutrition Examination Survey; CIM, Composite inflammaging metric; CIMlm, Composite inflammaging metric using linear regression; CIMPCA, Composite inflammaging metric using Principal Component Analysis (PCA); CIM_AL, Composite inflammaging metric using allostatic load; CIM_KDM, Composite inflammaging metric using Klemera-Doubal Method (KDM).

**Figure 3.** The discrimination performance of biomarkers, conventional inflammatory metrics and CIMs in UK Biobank and NHANES IV.

This figure shows the predictive ability of inflammatory metrics for 10-year all-cause mortality, 10-year CVD incidence in UKB and NHANES IV. The dotted line indicates a basic model, Age+Sex.

All models were adjusted for age and sex. Data is shown as mean AUC (center dot of error bar) and 95% confidence interval (line of error bar).

Abbreviations: NHANES, National Health and Nutrition Examination Survey; AUC, area under the curve; SII, Systemic immune-inflammation index; SIRI, Systemic Inflammation Response Index; NLR, Neutrophil to lymphocyte ratio; PLR, Platelet to Lymphocyte Ratio; LMR, Lymphocyte to monocyte Ratio; CIM, Composite inflammaging metric; CIMlm, Composite inflammaging metric using linear regression; CIMPCA, Composite inflammaging metric using Principal Component Analysis (PCA); CIM_AL, Composite inflammaging metric using allostatic load; CIM_KDM, Composite inflammaging metric using Klemera-Doubal Method (KDM).

**Figure 4.** Associations between inflammatory metrics and lifestyles in UKB and NHANES IV.
a. Models were adjusted for age, sex, ethnic group, education, current employment status, Townsend deprivation index, and disease counts to examine the associations of indicators with lifestyles in UKB. b. Models were adjusted for age, sex, ethnic group, education, and disease counts to examine the associations of indicators with lifestyles in NHANES IV.

Abbreviations: NHANES, National Health and Nutrition Examination Survey; SII, Systemic immune-inflammation index; SIRI, Systemic Inflammation Response Index; NLR, Neutrophil to lymphocyte ratio; PLR, Platelet to Lymphocyte Ratio; LMR, Lymphocyte to monocyte Ratio; CIM, Composite inflammaging metric; CIMLin; Composite inflammaging metric using linear regression; CIMPCA, Composite inflammaging metric using Principal Component Analysis (PCA); CIMAl, Composite inflammaging metric using allostatic load; CIMKDM, Composite inflammaging metric using Klemera-Doubal Method (KDM).
<table>
<thead>
<tr>
<th></th>
<th>Platelet count</th>
<th>Coefficient (95% CI)</th>
<th>P-value</th>
<th>Lymphocyte count</th>
<th>Coefficient (95% CI)</th>
<th>P-value</th>
<th>Neutrophil count</th>
<th>Coefficient (95% CI)</th>
<th>P-value</th>
<th>Monocyte count</th>
<th>Coefficient (95% CI)</th>
<th>P-value</th>
<th>C-reactive protein</th>
<th>Coefficient (95% CI)</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Smoking status</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Never</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Current smoker</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Alcohol intake</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Never or special occasions only</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>One to three times per week</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Regular exercise</td>
<td>No</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Body mass index (kg/m²)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Normal (18.5 to &lt;25)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Overweight (25 to &lt;30)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Obese (≥30)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Notes:**
- The table above represents statistical data for various health outcomes, including smoking status, alcohol intake, and body mass index, associated with platelet, lymphocyte, neutrophil, monocyte counts, and C-reactive protein levels. The coefficients and P-values indicate the strength and significance of the associations. For a detailed interpretation, please refer to the original study or publication for further context.