1-1-2023

IgG N-glycosylation cardiovascular age tracks cardiovascular risk beyond calendar age

Zhiyuan Wu
Edith Cowan University

Zheng Guo
Edith Cowan University

Yulu Zheng
Edith Cowan University

Yutao Wang

Haiping Zhang

See next page for additional authors

10.1016/j.eng.2022.12.004
Wu, Z., Guo, Z., Zheng, Y., Wang, Y., Zhang, H., Pan, H., ... & Wang, W. (2023). IgG N-glycosylation cardiovascular age tracks cardiovascular risk beyond calendar age. Engineering, 26, 99-107. https://doi.org/10.1016/j.eng.2022.12.004

Engineering 26 (2023) 99–107

Contents lists available at ScienceDirect

Engineering

journal homepage: www.elsevier.com/locate/eng

Research
Glycomedicine—Article

IgG N-Glycosylation Cardiovascular Age Tracks Cardiovascular Risk Beyond Calendar Age

Zhiyuan Wu [a,b,#], Zheng Guo [a,#], Yulu Zheng [a,#], Yutao Wang [c], Haiping Zhang [b], Huiying Pan [b], Zhiwei Li [b], Lois Balmer [a], Xia Li [d], Lixin Tao [b,*], Xiuhua Guo [b,*], Wei Wang [a,b,*]

[a] Centre for Precision Health, Edith Cowan University, Joondalup, WA 6027, Australia
[b] Beijing Municipal Key Laboratory of Clinical Epidemiology, School of Public Health, Capital Medical University, Beijing 100069, China
[c] Shanghai Fufan Information Technology Co., Ltd., Shanghai 200433, China
[d] Department of Mathematics and Statistics, La Trobe University, Melbourne, VIC 3086, Australia

ARTICLE INFO

Article history:
Received 27 September 2022
Revised 18 November 2022
Accepted 8 December 2022
Available online 2 January 2023

Keywords:
IgG N-glycosylation cardiovascular age
Cardiovascular aging
Immunoglobulin G
Glycosylation
Inflammation
Feature selection
Machine learning

ABSTRACT

The use of an altered immunoglobulin G (IgG) *N*-glycan pattern as an inflammation metric has been reported in subclinical atherosclerosis and metabolic disorders, both of which are important risk factors in cardiovascular health. However, the usable capacity of IgG N-glycosylation profiles for the risk stratification of cardiovascular diseases (CVDs) remains unknown. This study aimed to develop a cardiovascular aging index for tracking cardiovascular risk using IgG *N*-glycans. This cross-sectional investigation enrolled 1465 individuals aged 40–70 years from the Busselton Healthy and Ageing Study. We stepwise selected the intersection of altered *N*-glycans using feature-selection methods in machine learning (recursive feature elimination and penalized regression algorithms) and developed an IgG N-glycosylation cardiovascular age (GlyCage) index to reflect the deviation from calendar age attributable to cardiovascular risk. The strongest contributors to GlyCage index were fucosylated *N*-glycans with bisecting *N*-acetylglucosamine (GlcNAc) (glycan peak 6 (GP6), FA2B,) and digalactosylated *N*-glycans with bisecting GlcNAc (GP13, A2BG2). A one-unit increase of GlyCage was significantly associated with a higher Framingham ten-year cardiovascular risk (odds ratio (OR), 1.09; 95% confidence interval (95% CI): 1.05–1.13) and probability of CVDs (OR, 1.07; 95% CI: 1.01–1.13) independent of calendar age. Individuals with excessive GlyCage (exceeding a calendar age > 3 years) had an increased cardiovascular risk and probability of CVDs, with adjusted ORs of 2.22 (95% CI: 1.41–3.53) and 2.71 (95% CI: 1.25–6.41), respectively. The area under curve (AUC) values of discriminating high cardiovascular risk and events were 0.73 and 0.65 for GlyCage index, and 0.65 and 0.63 for calendar age. The GlyCage index developed in this study can thus be used to track cardiovascular health using IgG N-glycosylation profiles. The distance between GlyCage and calendar age independently indicates the cardiovascular risk, suggesting that IgG N-glycosylation plays a role in the pathogenesis of CVDs. The generalization of the observed associations and the predictive capability of GlyCage index require external and longitudinal validation in other populations.

1. Introduction

The treatment and post-onset rehabilitation and care of cardiovascular diseases (CVDs) comprise a heavy economic burden [1], as CVDs rank as the leading cause of disability and mortality worldwide. According to the "Global Burden of Disease Study 2019" [2], there were 523 million prevalent cases of CVDs in 2019, and the number of deaths due to CVDs reached 18.6 million. Globally, the years lived with disability due to CVDs doubled from 17.7 million in 1990 to 34.4 million in 2019. Thus, it is essential to upscale the primary prevention of CVDs and identify novel biomarkers and tools to facilitate risk assessment and targeted intervention.

Glycosylation, the covalent attachment of sugar moieties to proteins, is the most common and diverse form of post-transcriptional modification, providing a novel dimension for early diagnosis, targeted prevention, and precision medicine by means of glycomedicine [3]. After nucleic acids and proteins, the first and

* Corresponding authors.
E-mail addresses: taolixin@ccmu.edu.cn (L. Tao), statguo@ccmu.edu.cn (X. Guo), wei.wang@ecu.edu.au (W. Wang).
# These authors contributed equally to this work.

https://doi.org/10.1016/j.eng.2022.12.004

second alphabets of the life code in the central dogma law, glycans could be the third alphabet of the cellular life under the governance of so-called paracentral dogma in cellular materialisty [3]. Glycans are regulated by glycosyltransferases and glycosidases [4], and participate in key biological processes (i.e., cell adhesion, molecular trafficking and clearance, receptor activation, signal transduction, and endocytosis). It is notable that the interaction between the N-glycosylation pattern and immunoglobulin G (IgG) has a direct impact on the regulation of inflammation, as the attached glycans can shift the role of IgG in pro- and anti-inflammatory functions [5]. The core of the glycan domain consists of *N*-acetylglucosamine (GlcNAc) and mannose residues, which may contain an added bisecting GlcNAc residue linked either to the first mannose or to terminal galactose or sialic acid residues [6–8].

Previous studies have reported the association of IgG N-glycosylation profiles with aging [9], inflammation [10], and cardiometabolic diseases including hypertension [11], diabetes [12], and dyslipidaemia [13], which are important traditional risk factors of CVDs. One study confirmed that the IgG N-glycosylation profile is associated with atherogenic lipoprotein and subclinical atherosclerosis carotid plaque [14]. However, the link between the IgG N-glycosylation profile and cardiovascular risk or diseases remains unknown. The concept of IgG glycan age has been proposed to explain biological age [15,16]. In addition, a recent study developed an aging clock to track cardiovascular aging using inflammatory biomarkers [17]. However, the capacity of the IgG N-glycosylation profile to indicate cardiovascular health has not been elucidated.

Unlike conventional statistical models, which input limited variables and show relatively low efficiency for predictive performance, artificial intelligence (AI) approaches can incorporate high-dimensional and multivariate data to solve challenging issues [18]. Machine learning (ML) is a subdomain of AI involving the automated discovery of patterns within data [19,20]. For different ML-based models, supervised learning tools, such as random forest, the least absolute shrinkage and selection operator (LASSO), and recursive feature elimination (RFE), can learn complicated structures by incorporating numerous variables with multidimensional data [21]. Furthermore, due to their outstanding performance, ML techniques have been increasingly applied to clinically relevant domains, such as predictors selection, predictive diagnostics, targeted prevention, and personalized medical services [9,17,18].

In this study, we applied an ML approach for feature selection and further developed and validated an IgG N-glycosylation cardiovascular age (GlyCage) index to track cardiovascular risk, thereby providing a novel tool to evaluate cardiovascular aging beyond calendar age.

2. Methods

2.1. Study population

The data used in this analysis were collected from the Busselton Healthy and Ageing Study (BHAS). The study design and recruitment procedures of the BHAS have been described elsewhere [22]. A total of 5107 participants were recruited by the BHAS from May 2010 to December 2015 [23]. For the present cross-sectional study, 1465 participants—that is, those whose complete information was available, including the IgG N-glycome, medical history of CVDs, demographic data, anthropometric measurements, and biochemistry tests—were included in the final analysis. This study was conducted in accordance with the principles of the *Declaration of Helsinki* and approved by the Busselton Population Medical Research Institute (Amendment to BSN 22/10) and Edith Cowan University Human Research Ethics Committee (2019-00947-ZHENG and 2021-03164-WU). All participants gave their informed consent to participate.

2.2. Measurement and definition

The demographic characteristics, lifestyle factors, information of medication use, and medical history of diseases were assessed by the BHAS participant questionnaire [22]. Educational level was categorized as primary school or below, secondary school, or above secondary school. Marital status was classified as married and others. Smoking status was defined as currently smoking or not.

Body mass index (BMI) was measured by means of standard anthropometric techniques and calculated as weight in kilograms divided by height in meters squared. Systolic and diastolic blood pressures were measured using a vascular profiler (Omron VP1000, Japan) in all four limbs with the participant supine, after 5 min of rest. Tests for total cholesterol, high-density lipoprotein (HDL) cholesterol, fasting plasma glucose, glycated hemoglobin A1c (HbA1c), and C-reactive protein (CRP) were conducted on fresh specimens by PathWest laboratories (Busselton and QEII Medical Centre, Australia).

2.3. Definition of ten-year cardiovascular risk and CVDs

The ten-year cardiovascular risk was orthodoxly defined using eight clinical factors (i.e., age, sex, total cholesterol, HDL cholesterol, systolic blood pressure, use of anti-hypertensive medication, diabetes status, and smoking status) for individuals from 30 to 74 years without CVDs, with reference to the established Framingham Heart Study score [24]. In brief, the Framingham ten-year cardiovascular risk was calculated as $1 - 0.88936^{\exp(\Sigma\beta X - 23.9802)}$ for males and $1 - 0.95012^{\exp(\Sigma\beta X - 26.1931)}$ for females, where β is the regression coefficient and X is the level of each risk factor mentioned above. The detailed sex-specific formula is provided in Table S1 in Appendix A. Individuals with a ten-year cardiovascular risk $\geq$ 20% were defined as the high-risk group.

CVDs were defined as angina, myocardial infarction or heart attack, transient ischemic attack (TIA), stroke, experience of coronary angioplasty or stent, coronary bypass, and implant of cardiac pacemaker in this study. The composition of CVDs among the sample of this study is shown in Table S2 in Appendix A.

2.4. IgG isolation and N-glycan analysis

IgG was isolated from blood plasma utilizing 96-well protein G monolithic plates based on previously introduced methods [25]. The "in-solution" method was performed to release and label the *N*-glycans. IgG *N*-glycans were separated by means of hydrophilic interaction chromatography and measured on an ultra-performance liquid chromatography (UPLC) instrument (Waters Corporation, USA) into 24 glycan peaks (GPs) [26]. Detailed structures are shown in Section S2 in Appendix A.

2.5. Feature selection in machine learning

We used the LASSO and RFE ML algorithms to select the intersection of *N*-glycans. The features selected by those two methods enabled the models' stability and minimized the risks of overfitting. The lambda was set as "1se" and the area under curve (AUC) was chosen as the performance measure for LASSO. The bagged tree was embedded in RFE, and accuracy was chosen as the performance index to optimize. Based on the glycans selected by this ML approach, a multi-step logistics regression model was

used to determine the final glycans, along with calendar age and sex, for GlyCage index construction based on the Framingham ten-year cardiovascular risk.

2.6. Statistical analysis

Characteristics are presented as the mean (standard deviation (SD)) or number (proportions), as appropriate. Differences are compared using Student's *t*-test for continuous variables and the χ^2 test for categorical variables. The distribution of IgG glycans are presented using a box plot stratified by the ten-year cardiovascular risk and CVDs.

An adjusted logistics model was used to evaluate the association of the established GlyCage index with the Framingham ten-year high cardiovascular risk and CVDs in the validation stage. The calendar age, education level, marital status, and CRP were adjusted in the model. We then analyzed the effect of excessive GlyCage, defined as the distance between GlyCage and calendar age, on cardiovascular risk and diseases. A receiver operating characteristic (ROC) curve analysis was performed to compare the capacity to indicate cardiovascular health among GlyCage, calendar age, and the IgG glycosylation biological aging (GlycanAge) proposed in a previous study [27]. A comparison of the ROC curves was performed using the DeLong test. All the statistical analyses were conducted using *R* software, version 4.1.0 (R Foundation, Austria). The difference was considered to be statistically significant at a two-side significance level with a *P* value < 0.05.

3. Results

3.1. Population characteristics

A total of 1465 individuals were enrolled in the final analysis. The mean age (SD) was 59.58 (5.51) years, including 107 (7.3%) cases with a history of CVDs (Table 1). Among the 1358 individuals with no history of any CVDs, 376 cases (27.7%) had a high ten-year cardiovascular risk, as shown in Table S3 in Appendix A.

3.2. Multi-step IgG N-glycans selection in ML

Fig. S1 in Appendix A shows the distribution of 24 GPs stratified by ten-year cardiovascular risk (Fig. S1(a)) and CVDs (Fig. S1(b)). In the primary stage, a subset of five GPs (GP6, GP7, GP8, GP13, and GP20) was selected by the LASSO (Fig. S2 in Appendix A) and RFE algorithms (Fig. 1). Ultimately, fucosylated *N*-glycans with bisecting GlcNAc (GP6, FA2B) and digalactosylated *N*-glycans with bisecting GlcNAc (GP13, A2BG2) were retained in the logistical model ($P < 0.05$), as shown in Table S4 in Appendix A.

3.3. GlyCage index construction

GlyCage index was finally calculated (Table S5 in Appendix A) as follows:

$$\text{GlyCage index} = (0.29640 * (X_1 - 5.81422) - 5.88346 * (X_2 - 0.22277))/0.16601 + V \quad (1)$$

where X_1 is the expression level of GP6, X_2 is the expression level of GP13, and V is calendar age.

GP6 was found to be significantly higher in individuals with a high ten-year cardiovascular risk and CVDs, while GP13 was found to be lower (Figs. 2(a)–(d)). The distributions of GP6 and GP13 among the low CVDs risk, high CVDs risk, and with CVDs groups are shown in Table S6 in Appendix A. Moreover, a correlation analysis revealed that GP6 was positively correlated and GP13 was negatively correlated with calendar age (Figs. 2(e) and (f)). The correlation coefficient was 0.899 between GlyCage and calendar age (Fig. 2(g)). The GlyCage and calendar age among the low ten-year cardiovascular risk group (57.71 vs 58.28) and without-CVDs group (59.40 vs 59.40) were similar, while GlyCage was higher in the high CVDs risk group (63.84 vs 62.34) and with-CVDs group (63.19 vs 61.75), as shown in Table S7 in Appendix A.

Table 1
Characteristics of the study population.

Characteristic	Without CVDs[a]	With CVDs	*P* value
Participants (*n* (proportion))	1358 (92.7%)	107 (7.3%)	
Sex (male; *n* (proportion))	612 (45.1%)	77 (72.0%)	< 0.001
Age (year, mean (SD))	59.40 (5.51)	61.75 (5.05)	< 0.001
Marital status (*n* (proportion))			
Married	1003 (73.9%)	94 (87.9%)	0.002
Others	355 (26.1%)	13 (12.1%)	
Education level (*n* (proportion))			
Primary school or below	20 (1.5%)	1 (0.9%)	0.886
Secondary school	707 (52.1%)	55 (51.4%)	
Above secondary school	630 (46.4%)	51 (47.7%)	
Current smoking (*n* (proportion))	111 (8.2%)	12 (11.2%)	0.362
Hypertension (*n* (proportion))	459 (33.8%)	63 (58.9%)	< 0.001
Diabetes (*n* (proportion))	267 (19.7%)	47 (43.9%)	< 0.001
BMI[b] ($kg \cdot m^{-2}$, mean (SD))	28.94 (5.25)	31.65(5.26)	< 0.001
SBP (mmHg, mean (SD))	133.99 (16.14)	137.57 (15.43)	0.028
DBP (mmHg, mean (SD))	78.04 (9.70)	79.41 (8.80)	0.161
Total cholesterol ($mmol \cdot L^{-1}$, mean (SD))	5.55 (1.11)	0.12 (0.03)	< 0.001
HDL cholesterol ($mmol \cdot L^{-1}$, mean (SD))	1.45 (0.36)	0.03 (0.01)	< 0.001
Fasting glucose ($mmol \cdot L^{-1}$, mean (SD))	5.48 (1.57)	2.58 (0.95)	< 0.001
HbA1c (mean (SD))	5.87% (0.82%)	6.34% (1.10%)	< 0.001
CRP ($mg \cdot L^{-1}$, mean (SD))	3.09 (4.44)	3.39 (4.77)	0.514

SI conversion factors: to convert fasting plasma glucose to $mg \cdot dL^{-1}$, multiply by 18.0; to convert triglyceride to $mg \cdot dL^{-1}$, multiply by 88.6; to convert cholesterol to $mg \cdot dL^{-1}$, multiply by 38.66.
SD: standard deviation; SBP: systolic blood pressure; DBP: diastolic blood pressure. 1 mmHg = 0.133 kPa.
[a] CVDs included angina, myocardial infarction or heart attack, TIA, stroke, experience of coronary angioplasty or stent, coronary bypass, and implant of cardiac pacemaker.
[b] Calculated as weight in kilograms divided by height in meters squared.

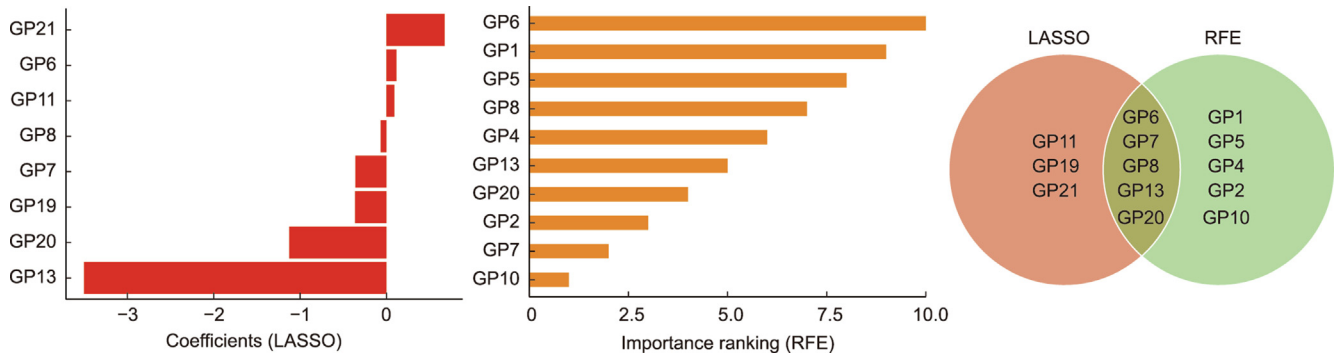

Fig. 1. Selected IgG *N*-glycans using LASSO and RFE algorithms.

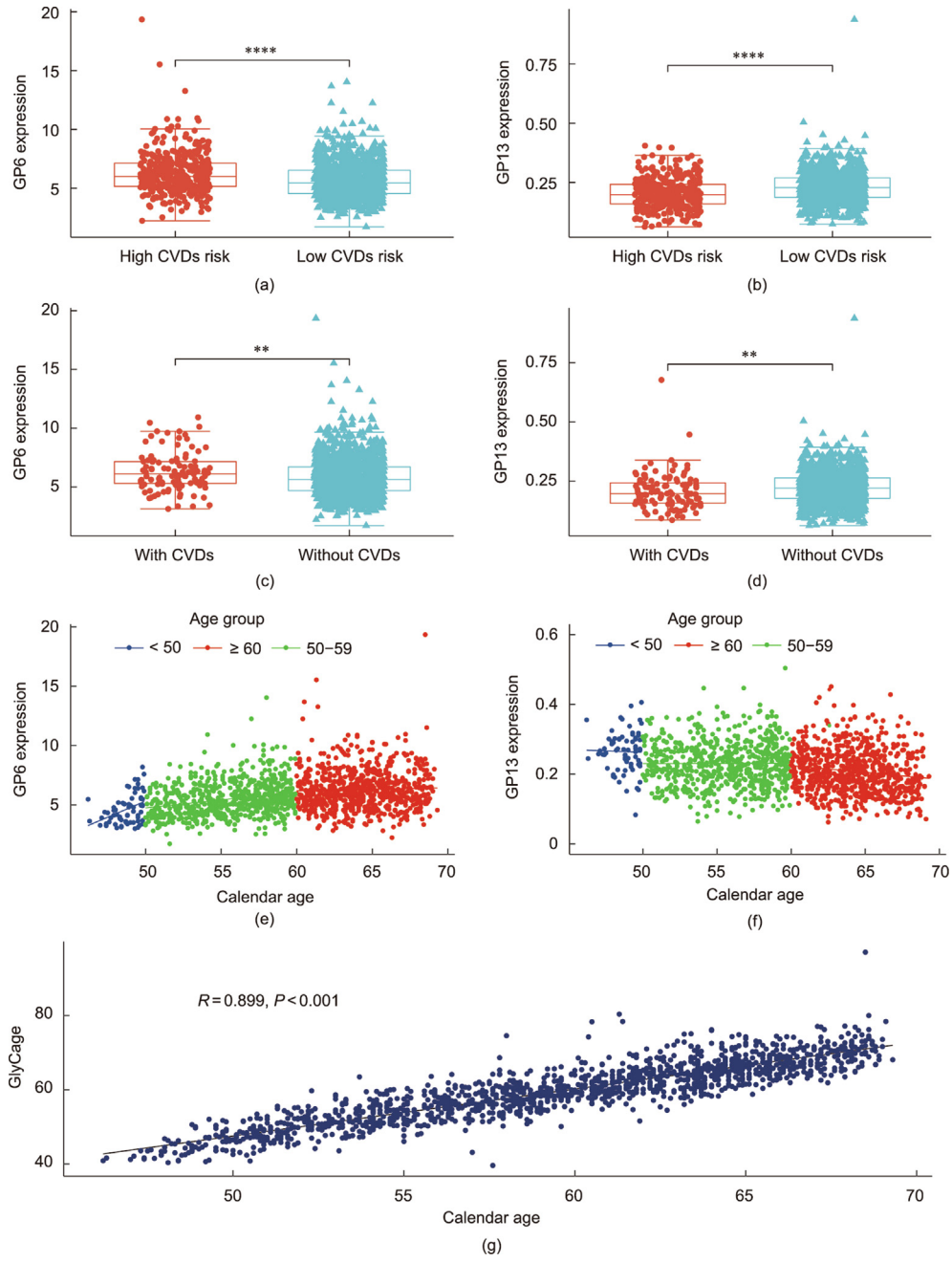

Fig. 2. Correlation of GlyCage and IgG *N*-glycans with calendar age. (a, b) Distribution of GP6 and GP13 among 1358 individuals without CVDs stratified by ten-year cardiovascular risk. (c, d) Distribution of GP6 and GP13 among 1465 individuals stratified by CVDs. (e, f) Correlation of GP6 and GP13 with calendar age. (g) Correlation between GlyCage and calendar age.

3.4. Association of GlyCage index with cardiovascular risk

In the adjusted logistics model, a one-year increase of GlyCage was significantly associated with a high ten-year cardiovascular risk and a high probability of existing CVDs; the ORs were 1.089 (95% confidence interval (95% CI): 1.050–1.131) and 1.068 (95% CI: 1.013–1.125), respectively, independent of calendar age (Table 2). People with a high ten-year cardiovascular risk or CVDs had an excessive GlyCage (Fig. S3 in Appendix A). Individuals were divided into three groups (excessive, equivalent, and regressive GlyCage), according to the distance between GlyCage and calendar age. People with excessive GlyCage (exceeding the calendar age by > 3 years) had the highest cardiovascular risk and probability of CVDs; the adjusted ORs were 2.220 (95% CI: 1.411–3.533) and 2.713 (95% CI: 1.253–6.406), respectively, compared with those with decreased GlyCage, as shown in Fig. 3 and Table 3.

In addition, the AUC value of the people identified to have high ten-year cardiovascular risks was 0.726 (95% CI: 0.698–0.754) for GlyCage index, which was significantly higher than those for calendar age (0.653 (95% CI: 0.621–0.685); $P < 0.001$ for the DeLong test) and GlycanAge (0.686 (95% CI: 0.658–0.715); $P = 0.041$ for the DeLong test), as presented in Fig. 4. The AUC value of the people identified to have CVDs was 0.645 (95% CI: 0.593–0.696) for GlyCage index; this was insignificantly higher than those for calendar age (0.626 (95% CI: 0.571–0.681); $P = 0.184$ for the DeLong test) and GlycanAge (0.627 (95% CI: 0.575–0.679); $P = 0.551$ for the DeLong test).

4. Discussion

Tools for the risk assessment and prediction of CVDs have been developed using clinical features. For example, the World Health Organization group derived and calibrated risk-prediction models to estimate CVDs risk in 21 Global Burden of Disease regions [28]. The pooled cohort equation was developed among US cohorts and is central in the development of prevention guidelines for CVDs [29]. The Systemic Coronary Risk Estimation project and Prediction for Atherosclerotic Cardiovascular Disease Risk project developed tools to predict the ten-year risk of CVDs for European [30] and Chinese [31] populations, respectively. In addition, genomic data and polygenic scores were recently applied to improve the accuracy of prediction tools [32–34], and inflammatory biomarkers were applied to reflect cardiovascular aging [17].

Against this backdrop, the present study is the first to develop a cardiovascular health index, GlyCage index, to track cardiovascular risk using IgG *N*-glycan profiles, which can identify cardiovascular age beyond calendar age. An excessive GlyCage was found to be significantly associated with higher cardiovascular risk; thus, GlyCage index offers a practical and comprehensible index based on glycomic analysis for cardiovascular risk stratification in a real-world setting. Compared with calendar age, GlyCage index achieved a significant improvement in AUC (0.073, $P < 0.001$) for ten-year cardiovascular risk and an insignificant improvement (0.019, $P = 0.184$) for CVDs—probably due to the insufficient statistical power. In parallel, the addition of the polygenic risk score to a baseline model including age and sex only improved the AUC by 0.024 for coronary artery disease prediction, which suggests the practical application of GlyCage index as an important supplement to cardiovascular risk stratification [35].

GlyCage index has great potential to promote cardiovascular risk stratification and inform clinical decision-making for primary prevention. A large proportion of residue cardiovascular risk still exists beyond the clinical-derived risk. Polygenic risk scores reflect the innate probability of CVDs, while GlyCage index may well inform people of the real-time risk based on a simple panel of two glycans [36]. Among individuals whose guideline-based recommendations are unknown, those with an excessive GlyCage can be recommended to initiate positive lifestyles changes. We hypothesize that the combination of GlyCage index, clinical risk, and polygenic risk could largely improve the precision prevention of CVDs and reduce the disease burden, considering the relatively low efficiency of
single-aspect variables and the high prevalence of unhealthy lifestyles. Individuals with both excessive GlyCage and high clinical risk may urgently need intensive prevention and management.

Through application of the LASSO and RFE ML algorithms, we selected the intersection of *N*-glycans among 24 IgG GPs detected via UPLC. Combining LASSO and RFE with multi-stage stepwise regression for *N*-glycans selection in this study made it possible to avoid the overfitting phenomenon and achieve a stable estimation for the glycan effect [37]. Identifying the most relevant IgG *N*-glycans played a crucial role in developing the accurate prediction tool GlyCage index for the assessment and management of cardiovascular risk. Thus, IgG *N*-glycan variables selected using advanced ML approaches are more likely to be effective predictors for establishing GlyCage index. On the other hand, generalization of the proposed GlyCage index requires further validation using different cohorts. The coefficients used in the construction of the GlyCage index tool by IgG *N*-glycans can be adjusted among different populations.

We determined that GP6 and GP13 were significantly associated with ten-year cardiovascular risk and CVDs. An individual's cardiovascular-risk-corrected age as calculated using GlyCage is equivalent to the age of a person with the same predicted risk and population average expression levels of GP6 and GP13. In brief, the proposed GlyCage index tracks the individual-level deviation from calendar age that is attributable to cardiovascular risk.

Table 2
Association of GlyCage index with high cardiovascular risk and CVDs beyond calendar age.

Variables	Model 1			Model 2[c]		
	OR	95% CI	*P* value	OR	95% CI	*P* value
High CVDs risk[a]						
Calendar age	1.049	0.997–1.104	0.067	1.051	0.997–1.107	0.062
GlyCage	1.092	1.053–1.132	< 0.001	1.089	1.050–1.131	< 0.001
CVDs[b]						
Calendar age	1.010	0.936–1.090	0.806	1.002	0.929–1.082	0.961
GlyCage	1.061	1.007–1.116	0.023	1.068	1.013–1.125	0.013

OR: odds ratio.
[a] Analyzed among individuals without CVDs ($n = 1358$); high risk refers to a Framingham ten-year cardiovascular risk ≥ 20%.
[b] Whole population ($n = 1465$); CVDs included angina, myocardial infarction or heart attack, TIA, stroke, experience of coronary angioplasty or stent, coronary bypass, and implant of cardiac pacemaker.
[c] Education level, marital status, and CRP adjusted in Model 2.

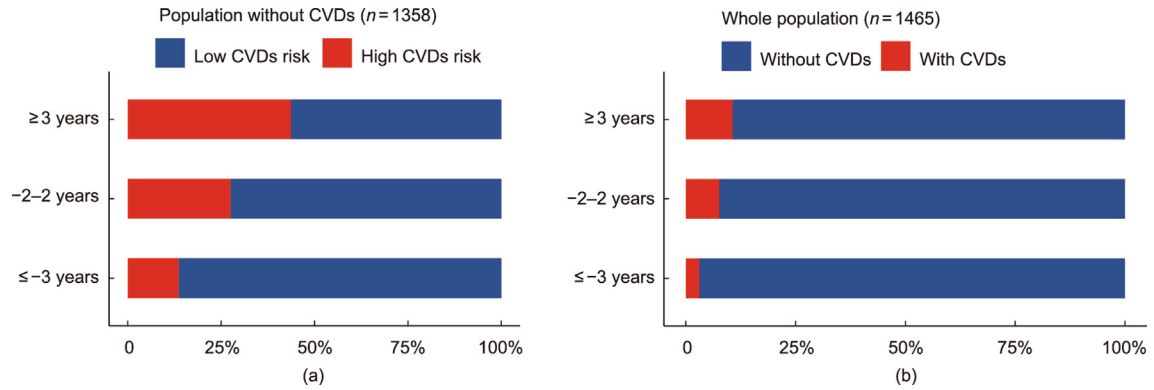

Fig. 3. Proportion of high cardiovascular risk and CVDs stratified by the distance between GlyCage and calendar age. The distance between GlyCage and calendar age was grouped into ≤ −3 years (group 1), −2 to 2 years (group 2), and ≥ 3 years (group 3). (a) Proportion of high cardiovascular risk among 1358 individuals without CVDs; high risk refers to a ten-year cardiovascular risk ≥ 20%; the proportions of high risk were 13.6%, 27.5%, and 43.7% for groups 1–3. (b) CVDs included angina, myocardial infarction or heart attack, TIA, stroke, experience of coronary angioplasty or stent, coronary bypass, and implant of cardiac pacemaker; the proportions of CVDs were 3.1%, 7.6%, and 10.6% for groups 1–3.

Table 3
Association of excessive GlyCage with cardiovascular risk and CVDs beyond calendar age.

Variables	β^{c}	OR	95% CI	*P* value
Population 1[a]				
Calendar age	0.139	1.149	1.118–1.182	0.001
Distance of GlyCage and calendar age (n_{hCVD}/n_{total}, proportion)				
≤ −3 years (39/286, 13.6%)	Reference	–	–	–
−2–2 years (223/811, 27.5%)	0.403	1.496	1.016–2.244	0.046
≥ 3 years (114/261, 43.7%)	0.797	2.220	1.411–3.533	0.001
Population 2[b]				
Calendar age	0.067	1.069	1.025–1.116	0.002
Distance of GlyCage and calendar age (n_{CVD}/n_{total}, proportion)				
≤ −3 years (9/295, 3.1%)	Reference	–	–	–
−2–2 years (67/878, 7.6%)	0.717	2.048	1.039–4.522	0.053
≥ 3 years (31/292, 10.6%)	0.998	2.713	1.253–6.406	0.015

n_{hCVD}: number of patients with high CVDs risk; n_{CVD}: number of patients with CVDs; n_{total}: number of corresponding total patients.
The distance between GlyCage and calendar age was grouped into ≤ −3 years (regressive), −2–2 years (equivalent) and ≥ 3 years (excessive).
[a] Analyzed among individuals without CVDs (*n* = 1358) for association between excessive GlyCage and high cardiovascular risk (ten-year cardiovascular risk ≥ 20%).
[b] Analyzed among whole population (*n* = 1465) for association between excessive GlyCage and CVDs including angina, myocardial infarction or heart attack, TIA, stroke, experience of coronary angioplasty or stent, coronary bypass, and implant of cardiac pacemaker.
[c] Adjusted for education level, marital status, and CRP level.

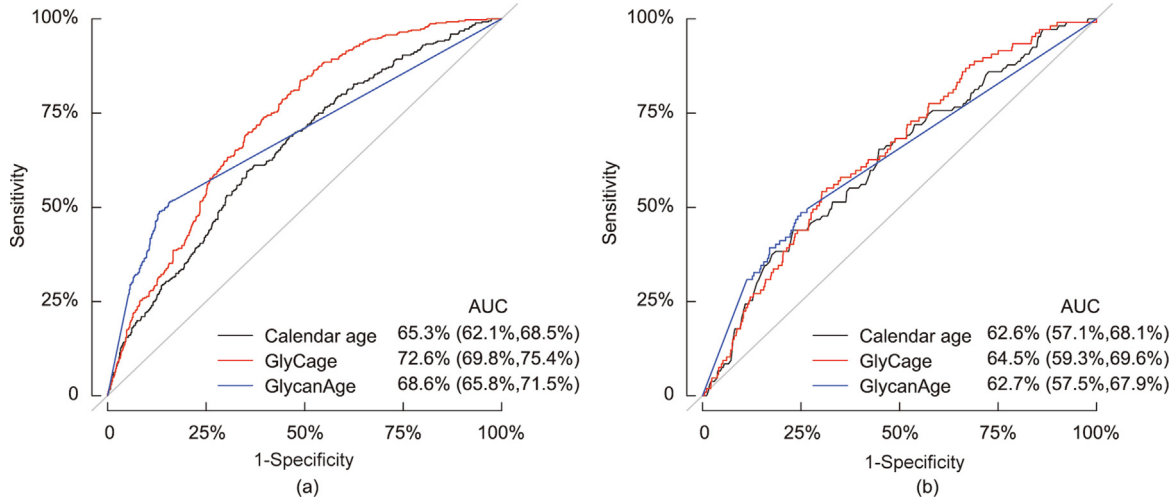

Fig. 4. Discrimination capacity of GlyCage, calendar age, and GlycanAge for cardiovascular risk and CVDs. (a) Capacity for discriminating high cardiovascular risk among 1358 individuals without CVDs, where high risk refers to a ten-year cardiovascular risk ≥ 20%. (b) Capacity for discriminating CVDs including angina, myocardial infarction or heart attack, TIA, stroke, experience of coronary angioplasty or stent, coronary bypass, and implant of cardiac pacemaker. GlyCage = (0.29640 ∗ (X_1 − 5.81422) − 5.88346 ∗ (X_2 − 0.22277))/0.16601 + *V*; GlycanAge = 139.9 + 85.1 ∗ X_1 − 5.2 ∗ X_1 ∗ X_1 − 34.6 ∗ X_3 + 11.8 ∗ X_4 (male)/= 110.0 + 164.5 ∗ X_1 − 46.7 ∗ X_1 ∗ X_1 − 22.4 ∗ X_3 − 1.9 ∗ X_4 (female); where X_1 is the expression level of GP6, X_2 is the expression level of GP13, X_3 is the expression level of GP14, X_4 is the expression level of GP15, and *V* is calendar age.

The two selected IgG *N*-glycans comprising GlyCage index (i.e., GP6 and GP13) have also been reported to be associated with the risk factors of CVDs in previous studies. Menni et al. [14] found that the presence of subclinical atherosclerosis pathology was predominantly explained by an increased relative abundance of GP6. Liu et al. [38] reported a lower abundance of GP13 in patients with ischemic stroke. Wang et al. [39] found that a high abundance of GP6 and low abundance of GP13 were significantly associated with hypertension. These reports are consistent with our findings, in which a higher abundance of GP6 and lower level of GP13 were associated with ten-year cardiovascular risk and CVDs.

There are several possible underlying mechanisms linking IgG N-glycosylation and cardiovascular risk. Chronic inflammation plays a crucial role in the development of CVDs by promoting the growth of plaques, loosening plaque in the arteries, and triggering blood clots, which are the main etiologies of coronary heart disease and stroke [40,41]. IgG is the most abundant immunoglobulin in the circulation system [5]. Glycosylation is one of the most common co- and post-translational modifications of proteins [42]. As described previously, IgG *N*-glycans have a complex-type biantennary structure, which can decisively influence the effector function via modulating the binding affinity of IgG to activate or inhibit Fcγ receptors (FcγRs) to induce the pro- or anti-inflammatory response [43].

Studies have shown that adding bisecting GlcNAc to fucosylated *N*-glycans can prevent fucosylation and then induce a pro-inflammatory response. This promotes the activity of antibody-dependent cell-mediated cytotoxicity (ADCC) through the enhanced binding affinity of IgG to FcγRIIIA on natural killer (NK) cells, macrophages, and neutrophils [44,45]. In addition, studies have reported that agalactosylation and asialylation can regulate the pro-inflammatory effect of IgG via complement system activation and ADCC [46,47]. In the current study, high levels of GP6 were observed in the high cardiovascular risk group. The presence of bisecting GlcNAc inhibits the process of fucosylation, leading to the enhancement of ADCC activity [3]. Moreover, GP6 is agalactosylated and asialylated in structure. The lack of a terminal galactose or sialic acid on GP6 has a pro-inflammatory effect by activating the complement cascade or the ADCC, respectively [48]. By extension, this effect contributes to the resultant inflammatory activity of IgG. Moreover, we observed low levels of galactosylated *N*-glycans (i.e., GP13) in the high cardiovascular risk group. A decrease in IgG galactosylation prevents the anti-inflammatory effect of circulating IgG through the activation of the lectin-initiated complement pathway. The presence of low levels of galactosylated *N*-glycans being associated with CVDs risk, as observed in the current study, supports the theory that the lectin-initiated complement pathway plays a crucial role in CVDs risk. Overall, the extent of IgG N-glycosylation (i.e., agalactosylated, asialylated, and fucosylated *N*-glycans with a bisecting GlcNAc) is altered in individuals with high CVDs risk, and the findings of this study and those of other CVDs-related studies have demonstrated that altered IgG N-glycosylation leads to an enhancement in the level of pro-inflammatory activity [38,49]. Taken together, these observations potentially explain why the detection of changes in IgG N-glycosylation even before the emergence of symptoms could be a promising predictive biomarker of CVDs onset.

Furthermore, the production of cytokines and inflammation-regulating molecules may potentially mediate the correlation between IgG N-glycosylation pattern and cardiovascular risk. An animal study demonstrated that the anti-inflammatory property of a high IgG galactosylation level is due to the activation of FcγRIIB and dectin-1, which prevents the function of pro-inflammatory effectors [50]. These findings indicate that individuals with lower levels of IgG galactosylation are expected to have more pro-inflammatory antibodies, leading to high levels of immune activation. Moreover, it has been reported that deficient IgG galactosylation may induce the activation of NK cells, and that such alterations are associated with higher levels of interferon-γ (IFN-γ) and tumor necrosis factor-α (TNF-α), thereby enhancing a pro-inflammatory effect [51]. An *in vivo* study revealed that cytokine IFN-γ triggers the pathogenicity of IgG by remodeling carbohydrate structure. Other cytokines, such as interleukin (IL)-21 and IL-17A, can switch the role of IgG between pro- and anti-inflammation in correlation with IgG glycan patterns [52]. Moreover, several random clinical trials (e.g., IL-6 inhibition with ziltivekimab in patients at high atherosclerotic risk (RESCUE), ASSessing the effect of Anti-IL-6 treatment in Myocardial Infarction (ASSAIL-MI), and Canakinumab Anti-Inflammatory Thrombosis Outcomes Study (CANTOS)) have targeted IL-1β or IL-6 to prevent cardiovascular risk and improve the outcome of CVDs [53–55]. In brief, the effect of an altered IgG N-glycosylation pattern on cardiovascular risk can be partially mediated by specific cytokines and molecules in the immune microenvironment, which warrants further research.

The results of this study should be interpreted in the context of the study's limitations. First, this is a cross-sectional study without intervention; thus, a causal association between IgG *N*-glycans and cardiovascular risk cannot be claimed. However, we analyzed the independent effect of GlyCage index on prospective ten-year cardiovascular risk, indicating the capacity of GlyCage index to predict the onset of CVDs. The underlining mechanisms linking IgG *N*-glycans and CVDs warrant further research, as they may provide potential molecular targets for the prevention and intervention of CVDs. Second, the history of CVDs was self-reported, which could cause recall bias. The effect of the selected IgG *N*-glycans and the capacity of GlyCage index to assess and predict cardiovascular health require further validation in other populations and cohort studies.

In summary, this study developed the GlyCage index to track cardiovascular risk beyond calendar age, thereby providing an index for assessing and predicting cardiovascular health based on glycomic analysis.

Acknowledgments

This study was funded by National Natural Science Foundation of China (8177120753) and China–Australia International Collaborative Grant (NHMRC APP1112767, NSFC 81561128020) to Wei Wang. Zhiyuan Wu was supported by the China Scholarship Council (201908110447). Yulu Zheng and Zheng Guo were supported by the Edith Cowan University Higher Degree by Research Scholarship (ECU-HDR ST10469322 and ST10468211).

Authors' contribution

Study conception and design: Xiuhua Guo, Wei Wang, and Zhiyuan Wu; data collection: Zheng Guo and Yulu Zheng; data analysis and interpretation: Zhiyuan Wu, Yutao Wang, Huiying Pan, Zhiwei Li, and Xia Li; manuscript writing and reviewing: Zhiyuan Wu, Zheng Guo, Yulu Zheng, Lois Balmer, and Haiping Zhang; study supervision: Wei Wang, Xiuhua Guo, Lixin Tao, and Lois Balmer.

Compliance with ethics guidelines

Zhiyuan Wu, Zheng Guo, Yulu Zheng, Yutao Wang, Haiping Zhang, Huiying Pan, Zhiwei Li, Lois Balmer, Xia Li, Lixin Tao, Xiuhua Guo, and Wei Wang declare that they have no competing interests.

Appendix A. Supplementary material

Supplementary data to this article can be found online at https://doi.org/10.1016/j.eng.2022.12.004.

References

[1] Riegel B, Moser DK, Buck HG, Dickson VV, Dunbar SB, Lee CS, et al. Self-care for the prevention and management of cardiovascular disease and stroke: a scientific statement for healthcare professionals from the American Heart Association. J Am Heart Assoc 2017;6(9):e006997.

[2] Roth GA, Mensah GA, Johnson CO, Addolorato G, Ammirati E, Baddour LM, et al. Global burden of cardiovascular diseases and risk factors, 1990–2019: update from the GBD 2019 Study. J Am Coll Cardiol 2020;76 (25):2982–3021.

[3] Wang W. Glycomedicine: the current state of the art. Engineering. In press.

[4] Ohtsubo K, Marth JD. Glycosylation in cellular mechanisms of health and disease. Cell 2006;126(5):855–67.

[5] Chan AC, Carter PJ. Therapeutic antibodies for autoimmunity and inflammation. Nat Rev Immunol 2010;10(5):301–16.

[6] Alter G, Ottenhoff THM, Joosten SA. Antibody glycosylation in inflammation, disease and vaccination. Semin Immunol 2018;39:102–10.

[7] Wang H, Li X, Wang X, Liu D, Zhang X, Cao W, et al. Next-generation (glycomic) biomarkers for cardiometabolic health: a community-based study of immunoglobulin G *N*-glycans in a Chinese Han population. OMICS 2019;23 (12):649–59.

[8] Seeling M, Brückner C, Nimmerjahn F. Differential antibody glycosylation in autoimmunity: sweet biomarker or modulator of disease activity? Nat Rev Rheumatol 2017;13(10):621–30.

[9] Greto VL, Cvetko A, Štambuk T, Dempster NJ, Kifer D, Deriš H, et al. Extensive weight loss reduces glycan age by altering IgG N-glycosylation. Int J Obes 2021;45(7):1521–31.

[10] Dall'Olio F, Vanhooren V, Chen CC, Slagboom PE, Wuhrer M, Franceschi C. N-glycomic biomarkers of biological aging and longevity: a link with inflammaging. Ageing Res Rev 2013;12(2):685–98.

[11] Kifer D, Louca P, Cvetko A, Deriš H, Cindrić A, Grallert H, et al. N-glycosylation of immunoglobulin G predicts incident hypertension. J Hypertens 2021;39 (12):2527–33.

[12] Wu Z, Li H, Liu D, Tao L, Zhang J, Liang B, et al. IgG glycosylation profile and the glycan score are associated with type 2 diabetes in independent Chinese populations: a case-control study. J Diabetes Res 2020;2020: 5041346.

[13] Liu D, Chu X, Wang H, Dong J, Ge SQ, Zhao ZY, et al. The changes of immunoglobulin G N-glycosylation in blood lipids and dyslipidaemia. J Transl Med 2018;16(1):235.

[14] Menni C, Gudelj I, Macdonald-Dunlop E, Mangino M, Zierer J, Bešić E, et al. Glycosylation profile of immunoglobulin G is cross-sectionally associated with cardiovascular disease risk score and subclinical atherosclerosis in two independent cohorts. Circ Res 2018;122(11):1555–64.

[15] Yu X, Wang Y, Kristic J, Dong J, Chu X, Ge S, et al. Profiling IgG *N*-glycans as potential biomarker of chronological and biological ages: a community-based study in a Han Chinese population. Medicine 2016;95(28):e4112.

[16] Jurić J, Kohrt WM, Kifer D, Gavin KM, Pezer M, Nigrovic PA, et al. Effects of estradiol on biological age measured using the glycan age index. Aging 2020;12(19):19756–65.

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

[18] Heo J, Yoon JG, Park H, Kim YD, Nam HS, Heo JH. Machine learning-based model for prediction of outcomes in acute stroke. Stroke 2019;50 (5):1263–5.

[19] Bruffaerts R. Machine learning in neurology: what neurologists can learn from machines and vice versa. J Neurol 2018;265(11):2745–8.

[20] Wu Z, Li L, Jin R, Liang L, Hu Z, Tao L, et al. Texture feature-based machine learning classifier could assist in the diagnosis of COVID-19. Eur J Radiol 2021;137:109602.

[21] Abedi V, Goyal N, Tsivgoulis G, Hosseinichimeh N, Hontecillas R, Bassaganya-Riera J, et al. Novel screening tool for stroke using artificial neural network. Stroke 2017;48(6):1678–81.

[22] James A, Hunter M, Straker L, Beilby J, Bucks R, Davis T, et al. Rationale, design and methods for a community-based study of clustering and cumulative effects of chronic disease processes and their effects on ageing: the Busselton healthy ageing study. BMC Public Health 2013;13:936.

[23] Tan HE, Lan NSR, Knuiman MW, Divitini ML, Swanepoel DW, Hunter M, et al. Associations between cardiovascular disease and its risk factors with hearing loss-A cross-sectional analysis. Clin Otolaryngol 2018;43 (1):172–81.

[24] D'Agostino Sr RB, Vasan RS, Pencina MJ, Wolf PA, Cobain M, Massaro JM, et al. General cardiovascular risk profile for use in primary care: the Framingham Heart Study. Circulation 2008;117(6):743–53.

[25] Liu D, Xu X, Li Y, Zhang J, Zhang X, Li Q, et al. Immunoglobulin G *N*-glycan analysis by ultra-performance liquid chromatography. J Vis Exp 2020;155: e60104.

[26] Zhang X, Yuan H, Lyu J, Meng X, Tian Q, Li Y, et al. Association of dementia with immunoglobulin G *N*-glycans in a Chinese Han population. NPJ Aging Mech Dis 2021;7(1):3.

[27] Krištić J, Vučković F, Menni C, Klarić L, Keser T, Beceheli I, et al. Glycans are a novel biomarker of chronological and biological ages. J Gerontol A Biol Sci Med Sci 2014;69(7):779–89.

[28] Lam TH. World Health Organization cardiovascular disease risk charts: revised models to estimate risk in 21 global regions. Lancet Glob Health 2019;7(10): e1332–45.

[29] Yadlowsky S, Hayward RA, Sussman JB, McClelland RL, Min YI, Basu S. Clinical implications of revised pooled cohort equations for estimating atherosclerotic cardiovascular disease risk. Ann Intern Med 2018;169(1):20–9.

[30] Conroy RM, Pyörälä K, Fitzgerald AP, Sans S, Menotti A, De Backer G, et al. Estimation of ten-year risk of fatal cardiovascular disease in Europe: the SCORE project. Eur Heart J 2003;24(11):987–1003.

[31] Yang X, Li J, Hu D, Chen J, Li Y, Huang J, et al. Predicting the 10-year risks of atherosclerotic cardiovascular disease in Chinese population: the China–PAR Project (prediction for ASCVD risk in China). Circulation 2016;134 (19):1430–40.

[32] Inouye M, Abraham G, Nelson CP, Wood AM, Sweeting MJ, Dudbridge F, et al. Genomic risk prediction of coronary artery disease in 480,000 adults: implications for primary prevention. J Am Coll Cardiol 2018;72(16):1883–93.

[33] Aragam KG, Natarajan P. Polygenic scores to assess atherosclerotic cardiovascular disease risk: clinical perspectives and basic implications. Circ Res 2020;126(9):1159–77.

[34] Sun L, Pennells L, Kaptoge S, Nelson CP, Ritchie SC, Abraham G, et al. Polygenic risk scores in cardiovascular risk prediction: a cohort study and modelling analyses. PLoS Med 2021;18(1):e1003498.

[35] Lu X, Liu Z, Cui Q, Liu F, Li J, Niu X, et al. A polygenic risk score improves risk stratification of coronary artery disease: a large-scale prospective Chinese cohort study. Eur Heart J 2022;43(18):1702–11.

[36] Štambuk T, Klasić M, Zoldoš V, Lauc G. *N*-glycans as functional effectors of genetic and epigenetic disease risk. Mol Aspects Med 2021;79 100891.

[37] Zheng Y, Guo Z, Zhang Y, Shang J, Yu L, Fu P, et al. Rapid triage for ischemic stroke: a machine learning-driven approach in the context of predictive, preventive and personalised medicine. Epma J 2022;13(2):285–98.

[38] Liu D, Zhao Z, Wang A, Ge S, Wang H, Zhang X, et al. Ischemic stroke is associated with the pro-inflammatory potential of N-glycosylated immunoglobulin G. J Neuroinflammation 2018;15(1):123.

[39] Wang Y, Klarić L, Yu X, Thaqi K, Dong J, Novokmet M, et al. The association between glycosylation of immunoglobulin G and hypertension: a multiple ethnic cross-sectional study. Medicine 2016;95(17):e3379.

[40] Costantino S, Paneni F, Cosentino F. Ageing, metabolism and cardiovascular disease. J Physiol 2016;594(8):2061–73.

[41] Willerson JT, Ridker PM. Inflammation as a cardiovascular risk factor. Circulation 2004;109(21 Suppl 1):Ii2–10.

[42] Li X, Wang H, Russell A, Cao W, Wang X, Ge S, et al. Type 2 diabetes mellitus is associated with the immunoglobulin G N-glycome through putative proinflammatory mechanisms in an Australian population. OMICS 2019;23 (12):631–9.

[43] Liu D, Li Q, Zhang X, Wang H, Cao W, Li D, et al. Systematic review: immunoglobulin G *N*-glycans as next-generation diagnostic biomarkers for common chronic diseases. OMICS 2019;23(12):607–14.

[44] Vučković F, Krištić J, Gudelj I, Teruel M, Keser T, Pezer M, et al. Association of systemic lupus erythematosus with decreased immunosuppressive potential of the IgG glycome. Arthritis Rheumatol 2015;67(11):2978–89.

[45] Takahashi M, Kuroki Y, Ohtsubo K, Taniguchi N. Core fucose and bisecting GlcNAc, the direct modifiers of the *N*-glycan core: their functions and target proteins. Carbohydr Res 2009;344(12):1387–90.

[46] Petrović T, Vijay A, Vučković F, Trbojević-Akmačić I, Ollivere BJ, Marjanović D, et al. IgG N-glycome changes during the course of severe COVID-19: an observational study. EBioMedicine 2022;81:104101.

[47] Russell A, Adua E, Ugrina I, Laws S, Wang W. Unravelling immunoglobulin G Fc N-glycosylation: a dynamic marker potentiating predictive, preventive and personalised medicine. Int J Mol Sci 2018;19(2):390.

[48] Russell AC, Šimurina M, Garcia MT, Novokmet M, Wang Y, Rudan I, et al. The N-glycosylation of immunoglobulin G as a novel biomarker of Parkinson's disease. Glycobiology 2017;27(5):501–10.

[49] Meng X, Song M, Vilaj M, Štambuk J, Dolikun M, Zhang J, et al. Glycosylation of IgG associates with hypertension and type 2 diabetes mellitus comorbidity in the Chinese Muslim ethnic minorities and the Han Chinese. J Pers Med 2021;11(7):614.

[50] Karsten CM, Pandey MK, Figge J, Kilchenstein R, Taylor PR, Rosas M, et al. Anti-inflammatory activity of IgG1 mediated by Fc galactosylation and association of FcγRIIB and dectin-1. Nat Med 2012;18(9):1401–6.

[51] Vicente MM, Alves I, Gaifem J, Rodrigues CS, Fernandes Â, Dias AM, et al. Altered IgG glycosylation at COVID-19 diagnosis predicts disease severity. Eur J Immunol 2022;52(6):946–57.

[52] Cao Y, Song Z, Guo Z, Zhao X, Gong Y, Zhao K, et al. Cytokines in the immune microenvironment change the glycosylation of IgG by regulating intracellular glycosyltransferases. Front Immunol 2021;12:724379.

[53] Ridker PM, Devalaraja M, Baeres FMM, Engelmann MDM, Hovingh GK, Ivkovic M, et al. IL-6 inhibition with ziltivekimab in patients at high atherosclerotic risk (RESCUE): a double-blind, randomised, placebo-controlled, phase 2 trial. Lancet 2021;397(10289):2060–9.

[54] Broch K, Anstensrud AK, Woxholt S, Sharma K, Tøllefsen IM, Bendz B, et al. Randomized trial of interleukin-6 receptor inhibition in patients with acute ST-segment elevation myocardial infarction. J Am Coll Cardiol 2021;77(15):1845–55.

[55] Ridker PM, Everett BM, Thuren T, MacFadyen JG, Chang WH, Ballantyne C, et al. Antiinflammatory therapy with canakinumab for atherosclerotic disease. N Engl J Med 2017;377(12):1119–31.