Distribution Agreement

In presenting this thesis or dissertation as a partial fulfillment of the requirements for an advanced degree from Emory University, I hereby grant to Emory University and its agents the non-exclusive license to archive, make accessible, and display my thesis or dissertation in whole or in part in all forms of media, now or hereafter known, including display on the world wide web. I understand that I may select some access restrictions as part of the online submission of this thesis or dissertation. I retain all ownership rights to the copyright of the thesis or dissertation. I also retain the right to use in future works (such as articles or books) all or part of this thesis or dissertation.

Signature:

_____ _____

Chenchen Yu Date

Associations between Circulating Endothelial Progenitor Cells and Renal Function

By

Chenchen Yu

MPH

Emory University

Rollins School of Public Health

Department of Biostatistics

_____ [Chair's signature]

Yi-An Ko

_____ [Chair's signature]

Qi Long

_____ [Member's signature]

Michael H Kutner

Associations between Circulating Endothelial Progenitor Cells and Renal Function

By

Chenchen Yu

B.S., China Pharmaceutical University, 2008

MPH, Emory University

Rollins School of Public Health

2015

Thesis Committee Chair: Yi-An Ko, Ph.D.

Qi Long, Ph.D.

An abstract of

A thesis submitted to the Faculty of the

Rollins School of Public Health of Emory University

in partial fulfillment of the requirements for the degree of

Master of Public Health

in Biostatistics

2015

Abstract

Associations between Circulating Endothelial Progenitor Cells
and Renal Function

By Chenchen Yu

Objective: Decreased number and function of circulating endothelial progenitor cells (EPCs) have been reported in patients with chronic kidney disease (CKD). However, there is little information about the associations between circulating EPC levels and renal function as measured by Glomerular Filtration Rate (GFR) in patients with some type of coronary heart disease (CHD). We propose to obtain a summary measure of multiple circulating EPC types based on principal component analysis, and to investigate the associations between EPCs and GFR after adjusting for conventional risk factors.

Methods and Results: A total of 2,145 adult patients were enrolled prior to undergoing elective or emergent cardiac catheterization. Measurement of EPC counts by flow cytometry was based on the single or combination expression of surface markers on $CD45^{med}$ and CD34 cells. Kidney function was measured by GFR using the Cockcroft-Gault formula that adjusts for body weight and body mass index. The associations between circulating EPCs and GFR were studied using principal component scores in multiple linear regression analyses. Statistically significant associations between GFR and circulating EPCs were observed in both unadjusted and adjusted analyses, indicating improvement in the number of EPCs especially those with certain VEGF cell types contribute to an increase in GFR levels.

Conclusions: Age, race, hypertension and diabetes are significant risk factors for CKD. EPCs are shown to be significantly associated with the GFR levels in the patients with CHD after adjusting for conventional risk factors. EPCs might be considered to be a potential therapeutic target for CKD, although more investigations still need to be implemented.

Associations between Circulating Endothelial Progenitor Cells and Renal Function

By

Chenchen Yu

B.S., China Pharmaceutical University, 2008

MPH, Emory University

Rollins School of Public Health

2015

Thesis Committee Chair: Yi-An Ko, Ph.D.

Qi Long, Ph.D.

A thesis submitted to the Faculty of the
Rollins School of Public Health of Emory University
in partial fulfillment of the requirements for the degree of
Master of Public Health
in Biostatistics
2015

Acknowledgements

I would like to express my deepest gratitude to my advisors, Dr. Yi-An Ko and Professor Qi Long, who provide me the opportunity to experience the real research and practical issues beyond the textbooks. Also a thanks for their excellent guidance, patience and support to help me edit the thesis. Special thanks goes to Professor Michael H Kutner, who was willing to take the time to help me revise the thesis in the most professional way possible. I would never have been able to finish my thesis without their help.

I would also like to thank the faculty, advisors, and staff of the Biostatistics Department at Rollins School of Public Health for the dynamic two years of learning that I have had. This thesis is only a sample of the vast knowledge that was attained and applied through my two years here at Rollins.

Lastly I want to give a special thanks to my parents for their loving support and encouragement that motivated me to pursue a degree at the graduate level.

Table of Contents

1. Introduction

1.1 Chronic Kidney Disease (CKD)

CKD has increasingly been recognized as a global health problem affecting over 70 million people, with more than 1 million of whom receiving a kidney transplant [1]. The number of patients affected by CKD has increased substantially worldwide, mainly as a consequence of increasing prevalence of CKD risk factors, such as diabetes, hypertension, obesity, and cardiovascular disease [2]. Currently, roughly 19 million (11%) people in the United States are having some degree of CKD. The number of prevalent CKD people will continue to rise due partly to the increase in diabetes and hypertension. It is predicted that by 2030, more than 2 million people in the United States will need dialysis or transplantation for kidney failure [3]. Complications associated with progression of CKD include increased incidence of cardiovascular disease, cognitive decline, hyperlipidemia, anemia and metabolic bone disease [4]. Specifically according to the National Health and Nutrition Examination Surveys (NHANES) III database, the prevalence of the complications of CKD increases with disease progression [5].

In addition, the progressive nature of CKD and the consequent end-stage renal disease (ESRD), i.e., kidney failure, would lead to a considerable decline in patient health-related quality of life, a substantial burden on global health care resources as well as burgeoning costs due to its comorbid conditions. Renal replacement therapy (dialysis or transplantation) is necessary for patients with ESRD. The limited treatment options for ESRD, however, are costly and sometimes ineffective. As such, CKD imposes significant financial costs to the overall health care system together with significant burden on health care teams (patients, clinicians, and payers) [6].

1.2 Glomerular Filtration Rate (GFR)

Accurate measurement of kidney function is critical to the prevention, detection, evaluation, and treatment of CKD. General recommendations for evaluating kidney function for people at increased risk include measuring urine albumin (proteinuria) to assess kidney damage and estimating GFR based on the level of serum creatinine. Although it cannot be measured directly, GFR is widely accepted as the best indicator of overall kidney function and therefore it forms the basis of definition and classification system for CKD [7]. Normal values of GFR that need to be adjusted for age, sex and body size are considered to be above 130 milliliters per minute (mL/min). CKD is defined clinically as a GFR below 60 mL/min or the presence of kidney damage for three or more months [8]. The level of GFR and its decline rate over time are vital to the detection of kidney disease, understanding its severity, as well as making informed decisions about prevention, diagnosis, prognosis, and treatment.

At advanced stages of kidney disease, i.e., Stage 5 with a GFR of 15 mL/min or less, the kidneys have lost almost all their ability to function effectively, and eventually dialysis or a kidney transplant is required. Thus, accurate and early detection of CKD are critical in order to appropriately intervene to slow down or reverse the disease progression.

1.3 Current Treatment and Prevention Strategies for CKD

Treatment of CKD aims to slow the progression to ESRD and to reduce complications. Namely, treatment of CKD can slow its progression to ESRD. However, the therapies remain limited. Blood pressure control using angiotensin-converting

enzyme (ACE) inhibitors or angiotensin II receptor blockers (ARBs) has shown promising results [9]. Glycemic control in diabetes also has been shown to minimize the progression [10]. A number of metabolic disturbances of CKD (e.g., acidosis, hyperphosphatemia, vitamin D deficiency) may be useful therapeutic targets, but their effectiveness has not been determined [5, 11]. While drugs targeting treatment of CKD are at various stages of development [11], there are no proven effective therapies. In addition, the timing of when to begin existing treatments and properly educating patients about their disease remain major practical barriers to better outcomes.

Therapy of CKD is usually conducted at an asymptomatic condition detected only by laboratory testing because of the slowly developing symptoms of chronically kidney failure, which is made more difficult as it usually stands for a late attempt at prevention. In other words, to some extent main risk factors associated with ESRD (i.e., hypertension, diabetes) could be prevented by primary preventative measures such as exercise, diet, and weight control. Also, after either hypertension or diabetes is diagnosed, renal complications can be alleviated by further prevention efforts targeted at blood pressure and glycemic control. Thus, treatment of CKD often represents an example of advanced prevention in populations who have failed the first line of prevention but who are still relatively asymptomatic. These features make CKD prevention and treatment a challenging task in practice.

1.4 Endothelial Progenitor Cells

Endothelial progenitor cells (EPCs) have attracted considerable attention due to their therapeutic potential for many diseases, such as myocardial infarction and cardiovascular

disease [2, 12]. Besides reproduction and migration, as immature primitive bone marrow (BW) cells, EPCs also have the capacity to undergo differentiation towards various mature endothelial cells [13]. EPCs were first identified in adult human peripheral blood in 1997 and were initially described by *Asahara* et al. –"human peripheral blood cells that differentiate into mature endothelial cells and form new blood vessels in vivo" [14]. Contrary to the theory of angiogenesis, the discovery of EPCs shows that vessel networks are not necessary for the formation of new blood vessels, because hematopoietic stem cells can differentiate into cells that are capable of vessel formation [15]. Since this discovery, intense experimental and clinical investigation has been conducted for the subject of pathophysiological role and therapeutic application of adult EPCs. Experimental studies have showed that EPCs are of great importance in maintaining the endothelial integrity and hemostasis. In majority of studies, strong associations have been found between circulating EPCs and cardiovascular risk [2]; the number of circulating EPCs has also been showed an inverse association with cardiovascular risk factors and vascular function [16].

Meanwhile, several studies have been conducted to explore the role of EPCs in CKD and their therapeutic prospective in regenerative medicine [17-20]. Patients with CKD are at a high risk for atherosclerotic cardiovascular disease (CVD) with varying degrees of endothelium dysfunction. It is showed in animal models of CKD that adult EPCs have the capacity to reduce proteinuria and improve GFR [20]. Various diseases including CKD are reported to be associated with decreased level and dysfunction of EPCs [21, 22], which suggests that EPCs may indicate the cardiovascular health condition including kidney function. However, in recent experiments and clinical trials, EPCs have been

shown to have limitations in reproducing ischemic tissues, and thus may be partially responsible for the increased morbidity of patients with cardiovascular risk factors [23]. It has been shown for patients with CKD Stage 5 receiving hemodialysis therapy that a low level of circulating CD34 cells predict independently both cardiovascular diseases and all-cause death [24]. Most available experimental and clinical data are consistent with the idea of decreased number and impaired functions of EPCs in patients with CKD, which may trigger reparative processes and put more pressure on the cardiovascular system [23].

1.5 Study Goals

The purpose of this study is to investigate the associations between circulating EPCs and renal function in totally 2,145 adult patients (aged 18 years or older) with or without active cardiovascular disease enrolled into the study prior to undergoing elective or emergent cardiac catheterization across three Emory Healthcare sites, between 2008 and 2014. We used data from the Emory Cardiovascular Biobank database [25], which was established to investigate the genetic basis of oxidative stress, vascular dysfunction, cardiovascular disease and stroke.

Using the Biobank data, we aim to obtain a summary measure of multiple circulating EPC types based on principal component analysis, and to investigate the associations between EPCs and GFR after adjusting for conventional risk factors.

1.6 Significance

The findings from the Emory Biobank database will contribute significantly to the limited existing literature concerning the role of EPC in CKD. Given that accurate and

early detection of CKD, along with appropriate aggressive interventions, may help retard the disease progression. Early targeted pharmacological therapy via improvement in EPC numbers could become a potential effective treatment option for CKD. In this study, we provide the insight on the emerging role of EPCs as biomarkers of CKD. We also explore the possibility in the field of EPC-based early detection, intervention and therapy, to halt and slow down the progression of kidney disease.

2. Methods

2.1 Study Population

Study patients were recruited from the Emory Cardiology Biobank [25]. A total of 2,145 adults (aged 18 years and older) were enrolled prior to undergoing elective or emergent cardiac catheterization across three Emory Healthcare sites, between 2008 and 2014. Twelve (0.6%) patients were excluded because of extreme values of EPC counts, leaving 2,133 patients for analysis. Patients were interviewed to collect information on demographic characteristic, medical history including family history of disease, medication usage, health behaviors, psychological factors, and neuropsychological functioning. Furthermore, medical record information was collected if patients were recruited from a clinical or hospital site. Risk factors for cardiovascular disease (e.g. hypertension, hyperlipidemia, and diabetes) were determined by physician diagnosis and/or treatment. Smoking was classified as 'non-smoker' or 'ever smoked' if there was a lifetime history of smoking at least 100 cigarettes [16].

Two cohorts were collected under the same protocol, with identical sampling strategies and collection methods but different enrollment times and EPC quantification methodology. Cohort 1 (n=1,286) was enrolled between 2008 and 2012, whereas cohort 2 (n=859) was enrolled between 2013 and 2014. Cohort 1 was our primary analysis population. After excluding outliers, 1,280 patients were used for analysis. Cohort 2 was used in this study as a validation dataset. The study was approved by the Institutional Review Board at Emory University, Atlanta, GA, USA. All patients provided written informed consent at the time of enrollment [17].

2.2 Measurement of Kidney Function

Glomerular filtration rate (GFR) refers to the filtration rate of the functioning nephrons in the kidney, which cannot be measured directly. The gold standard for the measurement of GFR is the urinary or plasma clearance of an ideal filtration marker, such as inulin, iothalamate or iohexol [18]. However, this method is complicated and is not used in clinical practice. Instead, serum levels of endogenous filtration markers, such as creatinine, along with urinary measurements in some cases, have traditionally been used to estimate GFR. However, serum creatinine alone is not an adequate marker of kidney function.

The estimation of GFR via creatinine clearance (Ccr), on the other hand, provides a quick and simple way to evaluate kidney function, which is useful for screening and detecting early kidney damage as well as monitoring kidney function. In this study, we calculated GFR using the Cockcroft-Gault formula:

$$C_{Cr}=\{((140\text{-age}) \times \text{weight})/(72\ S_{Cr})\} \times 0.85 \text{ if female}$$

$$C_{Cr}=\{((140\text{-age}) \times \text{weight})/(72\ S_{Cr})\} \text{ if male}$$

where C_{Cr} is expressed in milliliters per minute, age in years, weight in kilograms, and serum creatinine (S_{Cr}) in milligrams per deciliter. C_{Cr} may over-estimate GFR by 10-20% but remains the standard for drug dosing adjustments. The Cockcroft-Gault formula was developed in 1976 using data from 249 men with C_{Cr} from approximately 30 to 130 mL/min [19]. The original formula does not adjust for body surface area.

The Cockcroft-Gault equation remains the gold standard for eGFR after almost 40 years, despite inaccuracies that arise from variations in body composition among patients. In this study, we performed adjustments to the Cockcroft-Gault equation by body weight

and body mass index (BMI) as recommended by *Brown* et. al [26] and *Winter* et. al [27], since the original formula appears to become less accurate in weight extremes (underweight and particularly overweight/obesity). The algorithm for the adjusted Cockcroft-Gault equation is listed as follows:

- Underweight (BMI < 18.5): Weight uses actual body weight
- Normal Weight (BMI 18.5 - 22.9): Weight uses ideal body weight (IBW)
- Overweight/Obese (BMI ≥ 23): Weight uses adjusted body weight (ABW)

Here, the formula of IBW and ABW are:

- IBW for men = 50 + (2.3 * (Height in inches - 60))
- IBW for women = 45.5 + (2.3 * (Height in inches - 60))
- ABW = IBW + 0.4 * (Actual Body Weight - IBW)

Another recently developed formula for calculating eGFR – the CKD-Epidemiology Collaboration (CKD-EPI) equation [28]– has been commonly used in kidney studies. The CKD-EPI equation was developed based on subjects with normal renal function in 2009 [20]. Since all participants of our study were patients undergoing elective or emergent cardiac catheterization, it is likely that they had impaired kidney function. Given that the Cockcroft-Gault equation was developed based on men with and without CKD, it is appropriate to use this equation to calculate eGFR in our study.

2.3 Circulating EPC Counts

Measurements of the circulating EPCs by flow cytometry were performed by technicians masked to the study data. Before the quantification, circulating EPCs were identified based on the single or in combination expression of CD34+, CD133+,

CXCR4+ and VEGF2+ surface markers on arterial blood mononuclear cells that enumerated as $CD45^{med}$ and CD34 cells. We measured circulating numbers of CD34+, CD133+, VEGF2+ and CXCR4+, dual positive CD34+/VEGF2+, CD34+/CD133+, CD34+/CXCR4+, CD34+/VEGF2+, CD133+/VEGF2+, CD133+/CXCR4+ and CXCR4+/VEGF2+ and triple positive CD34+/CD133+/CXCR4+, CD34+/CD133+/VEGFR2+ and CD34+/VEGF2+/CXCR4+ cell populations [25].

2.4 Statistical Analysis

2.4.1 Descriptive Analysis

Patient characteristics are summarized as mean ± standard deviation (SD) for continuous variables if they follow an approximately normal distribution or as median (interquartile range) if they are highly skewed. Categorical variables are summarized as a number (percentage). The distributions of all EPC counts were right skewed. Log-transformations (natural log [cell counts + 0.5 * the minimum detectable value for EPC counts (0.005)]) was applied to reduce heteroskedasticity and to help stabilize estimation.

2.4.2 Univariate Analysis

Spearman correlation coefficients between GFR and the 14 EPC types, as well as the correlations among all 14 EPCs, were calculated in the target study population. We found that all EPC types were highly correlated with each other (Table S1 in the Supplementary Materials). In order to reduce the dimension of variables and to avoid repeated testing for the association between GFR and each EPC, we apply principal component analysis to the 13 highly correlated EPCs.

2.4.3 Principal Component Analysis

Principal component analysis (PCA) is a widely used multivariate statistical method [29, 30]. The objective of PCA is to reduce the dimension of variables and transform them into a new set of summary variables, which retain the maximum possible variance of the original set. The new variables, called principal components (PCs), are uncorrelated linear combinations of the original data and are ordered by the fraction of the total information each retains. Coefficients in the PCs, which indicate the percent contribution of corresponding variables, are called loadings. The general form for the formula to compute a score value (c_{ki}) for *i*th observation in *k*th PC using PCA is:

$$c_{ki} = t_{1k}x_{1i} + t_{2k}x_{2i} + \cdots + t_{pk}x_{pi}$$

where c_{ki} is the standardized score value of *i*th observation in *k*th PC; *i* is the number of observation; *k* is the number of selected PC number; *p* is the number of independent variables; x_{pi} is the value of *p*th variable of *i*th observation; and t_{pk} is the standardized weight of the *p*th variable in *k*th PCs.

Before performing PCA, we standardized the data to make all the log-transformed EPC counts with mean 0 and SD 1. PCA was performed on the 13 EPCs to explore their relative contributes to GFR. The first PC (PC1) and the second PC (PC2) were selected for multiple linear regression analysis. PC1 and PC2 explained 61% and 28% of all the variance, respectively. The scree plot (Figure 1) is also a useful visual aid for determining an appropriate number of PCs to select. The scree plot graphs the eigenvalue against the component number. The "elbow" in the scree plot suggests the appropriate number of components, which is also two. Therefore, we retained the first two PCs, which together explained 89% of the variability in the data.

2.4.4 Multiple Linear Regression Analysis

The potential confounding covariates that we considered were age, gender, race, BMI, low-density lipoprotein (LDL), triglycerides (TRIG), high-density lipoprotein (HDL), systolic blood pressure (SBP), smoking, hypertension (HTN), diabetes (DM) and hyperlipidemia (HLD). Multiple linear regression of GFR on the PC scores and all the covariates was performed using a backward variable section procedure with a specific p-value (0.05) to identify significant predictors of GFR. The first two PCs were retained in the model during the variable selection process. Non-significant variables were removed from the model by the backward elimination method. Finally we obtained model 1 as follows:

Model 1 $$\text{E (GFR)} = \beta_0 + \beta_1 * \text{PC1} + \beta_2 * \text{PC2} + \beta_3 * \text{age} + \beta_4 * \text{BMI} + \beta_5 * \text{LDL} + \beta_6 * Male + \beta_7 * \text{African American} + \beta_8 * \text{Smoking} + \beta_9 * \text{HTN} + \beta_{10} * \text{DM} + \beta_{11} * \text{HLD}$$

Next, interactions were explored between PC1, PC2 and age, gender, race, and diabetes, to test for heterogeneity. Using partial F-statistics test for interaction terms, first we added the interaction term *Age*PC2* into model 1 because of its smallest p-value (0.0037) among all the significant terms. After that, we continued the process until all significant terms were included. Finally, we added *DM*PC2* to the model. The final model is:

Model 2 $$\text{E (GFR)} = \beta_0 + \beta_1 * \text{PC1} + \beta_2 * \text{PC2} + \beta_3 * \text{age} + \beta_4 * \text{BMI} + \beta_5 * \text{LDL} + \beta_6 * Male + \beta_7 * \text{African American} + \beta_8 * \text{Smoking} + \beta_9 * \text{HTN} + \beta_{10} * \text{DM} + \beta_{11} * \text{HLD} + \beta_{12} * \text{Age} * \text{PC2} + \beta_{13} * \text{DM} * \text{PC2}$$

2.4.5 Validation of Results

We repeated the analysis using Model 2 in cohort 2 to evaluate whether the findings are consistent with cohort 1. Specifically, PC scores for PC1 and PC2 were calculated using the standardized weights from the PCA results of cohort 1.

2.4.6 Handling Missing Data

We assumed that our data were missing completely at random (MCAR) and used available case analysis [32]. As such, the sample size for analysis was not consistent across models. Table 1 lists the total number of available cases for each variable in both cohorts. The percentage of missing data in cohort 1 varied from 1% (age) to 27% (total cholesterol). A total of 1,001 patients were used in the multiple linear regression analyses (Table 3 and Table 4).

3. Results

3.1 Patient Characteristics

Baseline patient characteristics of both the 1,260 subjects in cohort 1 and the 853 subjects in cohort 2 are shown in Table 1. In cohort 1, 803 (63.5%) were males while 662 (36.5%) were female. Patient ages ranged from 18 to 91 years old with a mean age of 64.4 years. Race was divided into three categories, 891 (70.4%) whites, 297 (23.5%) blacks and 77 (6.1%) other patients. The average BMI was 28.9 km/m^2 ranging between 13.7 km/m^2 to 58.4 km/m^2. Among them, 148 (14%) patients were smokers, 496 (41%) had diabetes and 1053 (84%) were diagnosed with hypertension. The distribution of GFR appears to be slightly skewed, ranging from 3.5 to 251.5 mL/min with an average of 69.7 mL/min and a standard deviation of 32.1. Patients in cohort 2 were comparable to cohort 1.

3.2 Principal Component Analysis (PCA)

Figure 2 shows the standardized weights for all the 13 EPCs of both PC1 and PC2, indicating that all EPCs were positively correlated with PC1 but 7 of them were negatively correlated with PC2. Moreover, all 13 EPCs had a similar positive loading for PC1. However, in PC2 6 EPCs with VEGF cell type had similar positive loadings, whereas the other 7 without VEGF cell type had negative loadings. Since PC2 is significantly negative associated with GFR with consideration of interaction (Table 4), those EPCs without VEGF cell type contributed positively for GFR. Thus, CD34/CD133, CD34/CD133/CD45, CD34/CD45, CD34/CDCR4, CD34/CXCR4/CD133, CD34/CXCR4/CD133/CD45 and CD34/CXCR4/CD45 contribute positively to GFR,

indicating generally GFR level would increase with an increase in those 7 types of EPCs.

3.3 Associations between Circulating EPCs and GFR

Pairwise correlational analysis (Table 2) indicates that all 13 EPCs except CD34/CD146/CD45 are significantly associated with GFR ($p < 0.001$) and the two PCs are also significantly correlated with GFR ($p < 0.001$ for PC1 and $p = 0.03$ for PC2).

Without considering interactions, increased PC1 was found to be significantly associated with higher GFR ($p = 0.02$) after adjusting for all the other covariates. Age, race, hypertension, and diabetes were significant independent risk factors for GFR (Table 3). Specifically, GFR was estimated to decrease 1.16 (95% CI = [-1.30, -1.02]) mL/min annually holding all other variables constant. Compared with non-African Americans, African Americans had a significantly lower GFR level by 15.78 (95% CI = [11.76, 19.81]) mL/min on average. GFR of those who have hypertension or diabetes were significantly lower by 6.24 (95% CI = [0.92, 11.59]) mL/min and 4.06 (95% CI = [0.65, 7.47]) mL/min, respectively. On the other hand, a higher level of HDL was found to be significantly beneficial for GFR and GFR would be increased by 4.47 (95% CI = [0, 0.09]) mL/min with 1 mg/dl higher of HDL (Table 3).

Table 4 displays the results of the interaction analysis. PC1 was found to be significant and be positively associated with GFR ($p = 0.05$); PC2 was negatively associated with GFR ($p=0.01$) for patients without DM but positively correlated with GFR for those with DM. Risk factors and the parameter estimates are consistent with Model 1 (Table 4). The interaction terms are significant between age and PC2 ($p < 0.01$) as well as between DM and PC2 ($p = 0.02$), indicating that the effect of PC2 on GFR is

dependent on age and DM status.

Figure 3 shows plots of GFR and PC2 separately in age groups of <45, 45-60, and >60 years and among patients with and without diabetes. In patients without diabetes, patients younger than 45 years had more PC2-related decline in GFR compared to those in age group 45-60. Similar patterns were observed for the age group between 45 and 60 years compared to those older than 60, indicating that increased EPCs with VEGF cell type is associated with decline in GFR -- especially in young patients. Among patients with diabetes, an increase in PC2 is associated with a decrease in GFR in patients less than 45 years old, indicating that an elevated level of EPCs with VEGF cell type may contribute to impaired kidney function (or vice versa) in the younger diabetic population. In contrast, individuals with diabetes in the other two age groups displayed a modest PC2-related increase in GFR, showing that their renal function was not worse with a higher level of EPCs with VEGF cell type.

3.4 Validation Analysis

Table S4 shows the parameter estimates from model 2 using data from cohort 2. The results indicate that both PC1 and PC2 were not significantly associated with GFR after adjusting for potential covariates and the direction of association between GFR and PCs was opposite from what was obtained in cohort 1. Nevertheless, the risk factors from cohort 2 were consistent with the findings in cohort 1.

4. Discussion

4.1 Results

To the best of our knowledge, this is the first study using a large sample to show that decreased circulating EPC level is associated with impaired kidney function in patients undergoing cardiac catheterization. More importantly, we applied principal component analysis to EPC data metrics to reduce their dimensionality in order to bypass adjustment procedure associated with multiple testing for individual EPC. We found that all these 13 types of EPCs, as well as the first two principal components, were strongly associated with GFR, even after adjusting for conventional risk factors for CVD and other potential confounders. All the 13 EPCs contributed to PC1 with a similar positive loading, determining an approximately equal variance explained by PC1. On the other hand, 6 of these 13 EPCs with the VEGF+ cell type had a comparable positive loading for PC2, whereas the other 7 types without the VEGF+ cell type showed negative loadings for PC2. Moreover, higher levels of circulating EPCs without VEGF+ may contribute to improved kidney function in younger patients regardless of diabetic status, whereas EPCs with VEGF+ might be considered as potential therapeutic targets for older patients with diabetes.

Our study shows consistent findings with previous studies regarding the associations between EPCs and kidney function as well as significant risk factors for CKD [17-20]. A cross-sectional study enrolling 50 patients with varying degree of CKD indicated that EPC number and function decrease with advancing CKD [17]. Decreased circulating EPC level has also found in hypertensive patients with nephropathy through a cross-sectional study enrolling 125 patients with essential hypertension [18]. Additionally,

our recent studies had demonstrated patients who develop contract-induced nephropathy have decreased EPC level [19]. Compared to these studies, this study has several strengths including, but not limited to, a different study population, a considerably larger sample size, the use of EPC subtypes, the use of GFR as a measurement of renal function, and advanced statistical analysis methods.

In this study, we provide the insights on the emerging role of EPCs as biomarkers of CKD in patients with coronary heart disease (CHD). We demonstrate that the contributions of EPC profiles to kidney function may depend on patient's age and diabetic status. Further research can be done to validate these findings. Nevertheless, the findings from this study will provide the basis to explore the possibility of detection and intervention of kidney disease based on EPC profiles.

4.2 Limitations

The cross-sectional design of our study precludes definitive conclusions regarding the causal associations between EPCs and GFR. Including individuals only in the Atlanta area, the target population (cohort 1) may not be generalizable representative patients. It is of great importance to assess whether the results from this study could be replicated in a nationally diverse population. Furthermore, all the individuals in this study are patients with CKD with an average age of 65 ranging from 18 to 91 years old, so these findings may not be applicable to people with normal kidney function. The dataset included mostly white and black individuals (94%); therefore, these results may not be applicable to all ethnicities. Additionally, the final regression model, only 1,001 available patients are included automatically by SAS, excluding all patients with missing values under the

assumption of MCAR [32]. However, whether this assumption is fully satisfied in our population needs further investigation. Also, the potential confounding covariates used in the regression analyses are based on previous studies and might not account for some important variables that also effect the associations between GFR and EPCs. Finally, EPC counts were measured by different technicians between the two cohorts and hence may not be comparable without calibration. As a result, a direct validation using the second cohort as conducted in this work may not be ideal.

4.3 Future Research

It would be beneficial to conduct a longitudinal study to explore potential causal relationships between EPCs and GFR, i.e., to identify whether EPCs can predict changes in GFR. Furthermore, a clinical trial that provides CKD patients with EPC supplements could help answer important questions regarding how GFR changes as the levels of EPCs are boosted, namely, whether kidney repairs occur with a growth of EPC counts, and whether higher levels of EPCs could prevent further kidney impairment. To strengthen the generalizability of the results, future studies should target enrolling patients both nationally or worldwide with various health conditions (especially healthy people) from different racial groups. Lastly, given that the technology for measuring EPC counts is relatively new and has changed and matured over time, future research should address the issue of reproducibility for EPC measurements.

4.4 Conclusions

Our results suggest that EPCs were significantly associated with GFR in patients with CHD after adjusting for conventional risk factors for CVD and other potential

confounders. EPCs may be considered to be a new treatment target for CKD, although more investigations are still needed. In addition, age, race, hypertension and diabetes were shown to be significant risk factors for GFR.

Table 1. Summary Characteristics of the Study Population

Characteristics	Cohort 1 (N=1,280)		Cohort 2 (N=853)	
	N	**Mean ± SD**	**N**	**Mean ± SD**
Age (year)	1266	64.4 ± 13.2	825	64.7 ± 13.2
Weight (kg)	1260	84.9 ± 19.9	823	87.0 ± 20.5
Body mass index (kg/m^2)	1260	28.9 ± 6.1	823	29.8 ± 6.3
High-density lipoprotein (mg/dL)	1196	43.1 ± 14.8	691	46.8 ± 14.2
Total cholesterol (mg/dl)	934	160.6 ± 44.1	588	163.0 ± 44.0
Systolic blood pressure (mm Hg)	1234	139.1 ± 22.7	577	140.0 ± 24.2
Glomerular filtration rate (mL/min)	1252	69.7 ± 32.1	762	68.5 ± 31.7
		Median (Q1, Q3)		**Median (Q1, Q3)**
Creatinine (mg/dL)	1257	1.0 (0.9, 1.3)	769	1.0 (0.9, 1.4)
Low-density lipoprotein (mg/dL)	1186	85.5 (65.0, 113.0)	675	83.0 (64.0, 107.0)
Triglycerides (mg/dL)	1194	118.0 (79.0, 177.0)	690	111.0 (76.0, 166.0)
		n (%)		**n (%)**
Sex - Male	1265	803 (63.5)	832	525 (63.1)
Race	1265		832	
White		891 (70.4)		618 (74.3)
Black		297 (23.5)		182 (21.9)
Other		77 (6.1)		32 (3.8)
Smoker	1057	148 (14.0)	309	38 (12.3)
Hypertension	1248	1053 (84.4)	342	237 (69.3)
Diabetes	1208	496 (41.1)	342	131 (38.3)
Hyperlipidemia	1247	903 (72.4)	342	237 (69.3)

Table 2. Correlations between Glomerular Filtration Rate and Endothelial Progenitor Cells in Cohort 1

Cell Type	Spearman Correlation Coefficient	P-value
CD34/CD45	0.197	**<.001**
CD34/CD133/CD45	0.193	**<.001**
CD34/VEGF/CD45	0.132	**<.001**
CD34/CXCR4/CD45	0.192	**<.001**
CD34/CD146/CD45	-0.052	0.06
CD34/VEGF/CD133/CD45	0.099	**<.001**
CD34/CXCR4/VEGF/CD45	0.134	**<.001**
CD34/CXCR4/CD133/CD45	0.190	**<.001**
CD34/CD133	0.203	**<.001**
CD34/VEGF	0.145	**<.001**
CD34/VEGF/CD133	0.113	**<.001**
CD34/CXCR4	0.197	**<.001**
CD34/CXCR4/VEGF	0.148	**<.001**
CD34/CXCR4/CD133	0.201	**<.001**
PC1	0.218	**<.001**
PC2	-0.061	**0.031**

PC1 = Principal Component 1; PC2 = Principal Component 2.

Figure 1. Scree Plot of Principal Component Analysis

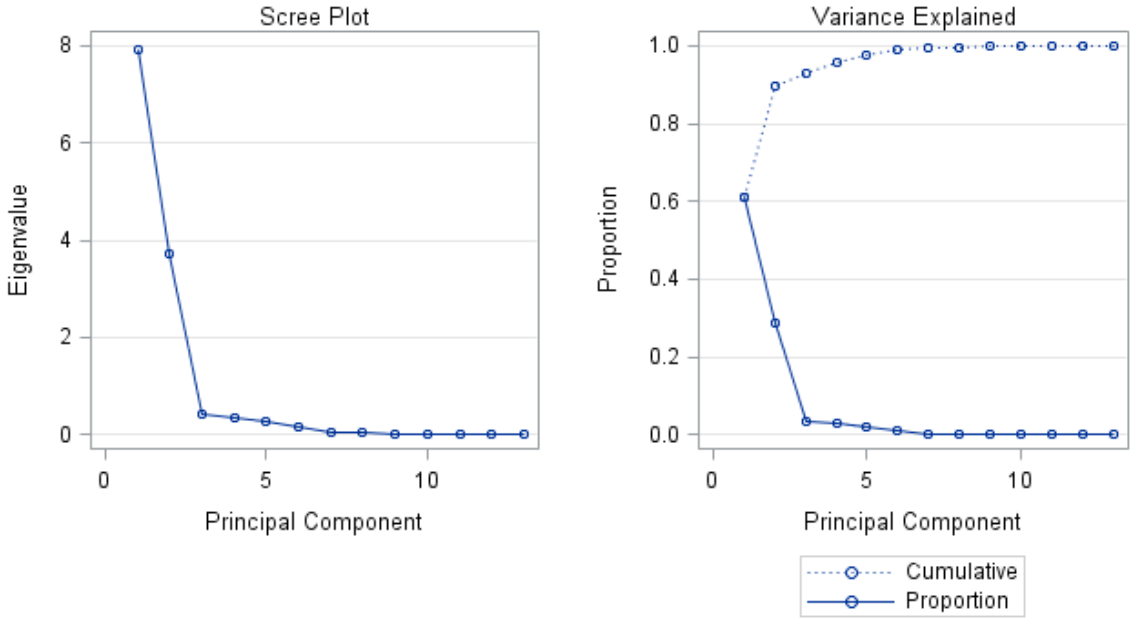


Figure 2. Eigenvectors for 13 Endothelial Progenitor Cells Corresponding to the First (upper panel) and the Second (lower panel) Principal Components

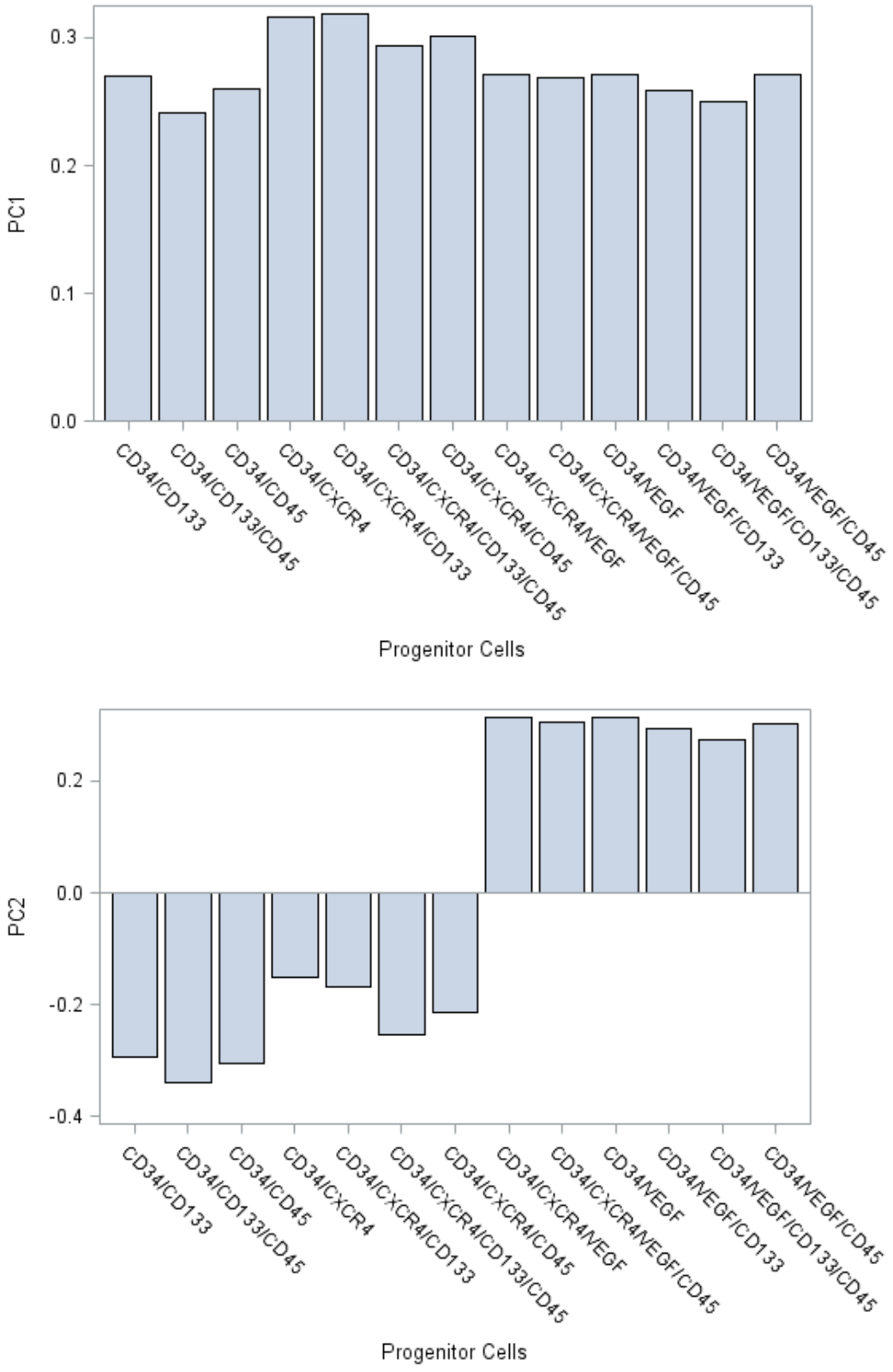


Table 3. Results of Model Fitting Without Interaction Terms for Cohort 1 (n=1,001)

Variable	Parameter Estimate	Standard Error	95% CI	p-value
PC1	0.72	0.32	(0.09, 1.34)	**0.02**
PC2	-0.20	0.46	(-1.10, 0.71)	0.67
Age	-1.16	0.07	(-1.30, -1.02)	**<.0001**
BMI	1.33	0.14	(1.05, 1.62)	**<.0001**
Low-density lipoprotein	0.04	0.02	(0, 0.09)	**0.04**
Male	9.78	1.76	(6.33, 13.23)	**<.0001**
African American	-15.78	2.05	(-19.81, -11.76)	**<.0001**
Smoking	7.11	2.44	(2.33, 11.90)	**<.01**
Hypertension	-6.24	2.71	(-11.59, -0.92)	**0.02**
Diabetes	-4.06	1.74	(-7.47, -0.65)	**0.02**
High-density lipoprotein	4.47	2.07	(0.41, 8.53)	**0.03**

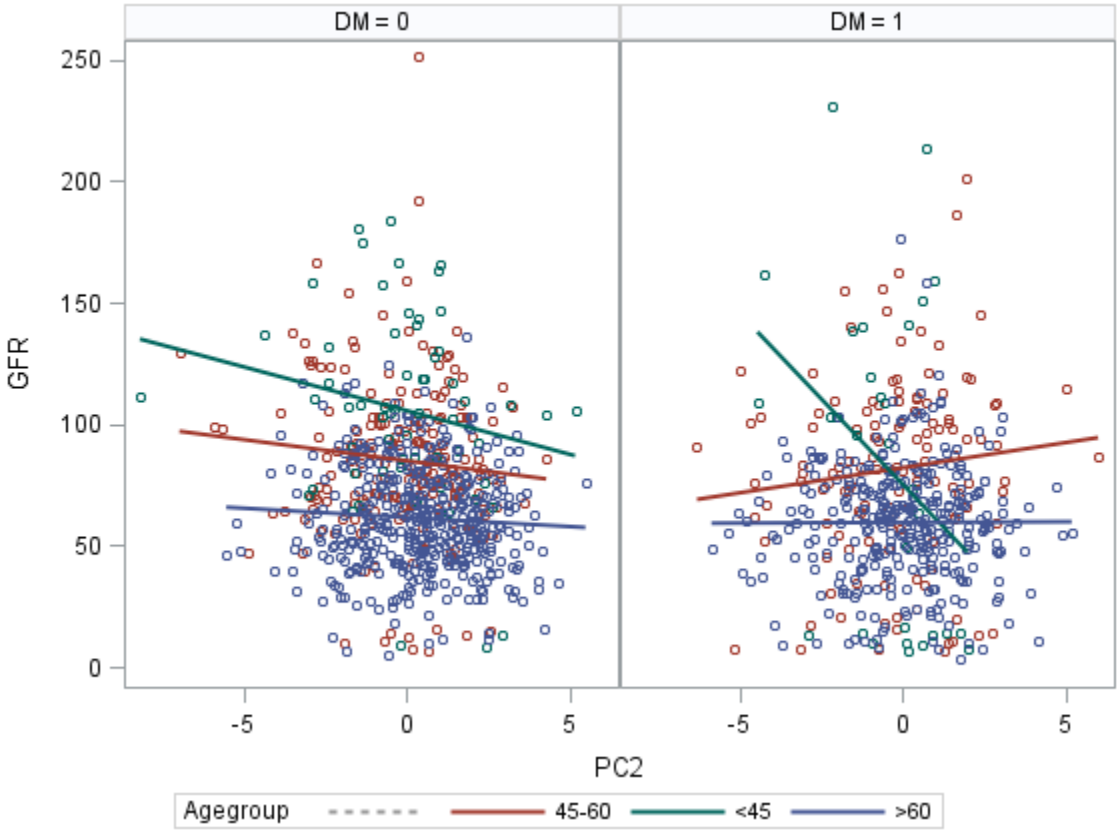
PC1 = Principal Component 1; PC2 = Principal Component 2; CI = Confidence Interval; BMI = Body Mass Index.

Table 4. Results of Model Fitting With Interaction Terms for Cohort 1 (n=1,001)

Variable	Parameter Estimate	Standard Error	95% CI	p-value
PC1	0.63	0.32	(0.01, 1.26)	**0.05**
PC2	-1.56	0.62	(-2.77, -0.35)	**0.01**
Age	-1.19	0.07	(-1.33, -1.05)	**<.0001**
BMI	1.34	0.14	(1.06, 1.62)	**<.0001**
Low-density lipoprotein	0.04	0.02	(0, 0.09)	**0.03**
Male	9.98	1.75	(6.55, 13.41)	**<.0001**
African American	-16.11	2.05	(-20.12, -12.11)	**<.0001**
Smoking	7.34	2.43	(2.58, 12.10)	**<.01**
Hypertension	-6.30	2.70	(-11.59, -1.01)	**0.02**
Diabetes	-4.53	1.75	(-7.96, -1.10)	**0.01**
High-density lipoprotein	4.73	2.06	(0.69, 8.77)	**0.02**
Age*PC2	0.10	0.04	(0.03, 0.17)	**<.01**
Diabetes*PC2	2.23	0.92	(0.42, 4.03)	**0.02**

PC1 = Principal Component 1; PC2 = Principal Component 2; CI = Confidence Interval; BMI = Body Mass Index.

Figure 3. Interactions Between Principal Component 2 and Diabetes/Age



References

1. Levey, A.S., et al., *Chronic kidney disease as a global public health problem: approaches and initiatives - a position statement from Kidney Disease Improving Global Outcomes.* Kidney Int, 2007. **72**(3): p. 247-59.
2. Xu, J.Y., et al., *Therapeutic application of endothelial progenitor cells for treatment of cardiovascular diseases.* Curr Stem Cell Res Ther, 2014. **9**(5): p. 401-14.
3. Coresh, J., et al., *Prevalence of chronic kidney disease in the United States.* JAMA, 2007. **298**(17): p. 2038-47.
4. Thomas, R., A. Kanso, and J.R. Sedor, *Chronic kidney disease and its complications.* Prim Care, 2008. **35**(2): p. 329-44, vii.
5. Murphree, D.D. and S.M. Thelen, *Chronic kidney disease in primary care.* J Am Board Fam Med, 2010. **23**(4): p. 542-50.
6. Couser, W.G., et al., *The contribution of chronic kidney disease to the global burden of major noncommunicable diseases.* Kidney Int, 2011. **80**(12): p. 1258-70.
7. Stevens, L.A. and A.S. Levey, *Measurement of kidney function.* Med Clin North Am, 2005. **89**(3): p. 457-73.
8. Stevens, L.A., et al., *Assessing kidney function--measured and estimated glomerular filtration rate.* N Engl J Med, 2006. **354**(23): p. 2473-83.
9. Wuhl, E. and F. Schaefer, *Therapeutic strategies to slow chronic kidney disease progression.* Pediatr Nephrol, 2008. **23**(5): p. 705-16.
10. Snively, C.S. and C. Gutierrez, *Chronic kidney disease: prevention and treatment of common complications.* Am Fam Physician, 2004. **70**(10): p. 1921-8.
11. Turner, J.M., et al., *Treatment of chronic kidney disease.* Kidney Int, 2012. **81**(4): p. 351-62.
12. Sun, Q., Z. Zhang, and Z. Sun, *The potential and challenges of using stem cells for cardiovascular repair and regeneration.* Genes Dis, 2014. **1**(1): p. 113-119.
13. Kawamoto, A. and D.W. Losordo, *Endothelial progenitor cells for cardiovascular regeneration.* Trends Cardiovasc Med, 2008. **18**(1): p. 33-7.
14. Asahara, T., et al., *Isolation of putative progenitor endothelial cells for angiogenesis.* Science, 1997. **275**(5302): p. 964-7.
15. Grisar, J.C., et al., *Endothelial progenitor cells in cardiovascular disease and chronic inflammation: from biomarker to therapeutic agent.* Biomark Med, 2011. **5**(6): p. 731-44.
16. Bakogiannis, C., et al., *Circulating endothelial progenitor cells as biomarkers for prediction of cardiovascular outcomes.* Curr Med Chem, 2012. **19**(16): p. 2597-604.
17. Krenning, G., et al., *Endothelial progenitor cell dysfunction in patients with progressive chronic kidney disease.* Am J Physiol Renal Physiol, 2009. **296**(6): p. F1314-22.
18. Huang, P.H., et al., *Increased circulating CD31+/annexin V+ apoptotic*

microparticles and decreased circulating endothelial progenitor cell levels in hypertensive patients with microalbuminuria. J Hypertens, 2010. **28**(8): p. 1655-65.

19. Chiang, C.H., et al., *Reduction of circulating endothelial progenitor cell level is associated with contrast-induced nephropathy in patients undergoing percutaneous coronary and peripheral interventions.* PLoS One, 2014. **9**(3): p. e89942.
20. Hsu, C.Y., et al., *Increased circulating endothelial apoptotic microparticle to endothelial progenitor cell ratio is associated with subsequent decline in glomerular filtration rate in hypertensive patients.* PLoS One, 2013. **8**(7): p. e68644.
21. Werner, N., et al., *Circulating endothelial progenitor cells and cardiovascular outcomes.* N Engl J Med, 2005. **353**(10): p. 999-1007.
22. Choi, J.H., et al., *Decreased number and impaired angiogenic function of endothelial progenitor cells in patients with chronic renal failure.* Arterioscler Thromb Vasc Biol, 2004. **24**(7): p. 1246-52.
23. Bahlmann, F.H., T. Speer, and D. Fliser, *Endothelial progenitor cells in chronic kidney disease.* Nephrol Dial Transplant, 2010. **25**(2): p. 341-6.
24. Humpert, P.M., et al., *Insulin stimulates the clonogenic potential of angiogenic endothelial progenitor cells by IGF-1 receptor-dependent signaling.* Mol Med, 2008. **14**(5-6): p. 301-8.
25. Patel, R.S., et al., *Circulating CD34+ progenitor cells and risk of mortality in a population with coronary artery disease.* Circ Res, 2015. **116**(2): p. 289-97.
26. Brown, D.L., A.J. Masselink, and C.D. Lalla, *Functional range of creatinine clearance for renal drug dosing: a practical solution to the controversy of which weight to use in the Cockcroft-Gault equation.* Ann Pharmacother, 2013. **47**(7-8): p. 1039-44.
27. Winter, M.A., K.N. Guhr, and G.M. Berg, *Impact of various body weights and serum creatinine concentrations on the bias and accuracy of the Cockcroft-Gault equation.* Pharmacotherapy, 2012. **32**(7): p. 604-12.
28. Florkowski, C.M. and J.S. Chew-Harris, *Methods of Estimating GFR - Different Equations Including CKD-EPI.* Clin Biochem Rev, 2011. **32**(2): p. 75-9.
29. Lengyel, A., et al., *Prediction of ozone concentration in ambient air using multivariate methods.* Chemosphere, 2004. **57**(8): p. 889-96.
30. Chia, K.S., H. Abdul Rahim, and R. Abdul Rahim, *Neural network and principal component regression in non-destructive soluble solids content assessment: a comparison.* J Zhejiang Univ Sci B, 2012. **13**(2): p. 145-51.
31. Weiner, D.E., et al., *Chronic kidney disease as a risk factor for cardiovascular disease and all-cause mortality: a pooled analysis of community-based studies.* J Am Soc Nephrol, 2004. **15**(5): p. 1307-15.
32. Little, R. J., & Rubin, D. B. (2014). *Statistical analysis with missing data*. John Wiley & Sons.

Table S1. Spearman Correlation Coefficients and P-values between All Progenitor Cells in Cohort 1 (N=1280)

	V1	V2	V3	V4	V5	V6	V7	V8	V9	V10	V11	V12	V13	V14
V1	1.00	0.90	0.26	0.86	-0.06	0.18	0.25	0.81	0.91	0.26	0.19	0.82	0.26	0.77
V2	p<.0001	1.00	0.13	0.73	-0.08	0.14	0.13	0.85	0.98	0.15	0.13	0.69	0.14	0.78
V3	p<.0001	p<.0001	1.00	0.45	-0.48	0.82	0.99	0.33	0.23	0.94	0.80	0.53	0.94	0.46
V4	p<.0001	p<.0001	p<.0001	1.00	-0.15	0.36	0.45	0.90	0.77	0.44	0.35	0.98	0.44	0.90
V5	p=0.02	P<.01	p<.0001	p<.0001	1.00	-0.45	-0.48	-0.18	-0.14	-0.54	-0.55	-0.22	-0.54	-0.26
V6	p<.0001	p<.0001	p<.0001	p<.0001	p<.0001	1.00	0.83	0.35	0.23	0.81	0.86	0.45	0.82	0.47
V7	p<.0001	p<.0001	p<.0001	p<.0001	p<.0001	p<.0001	1.00	0.33	0.22	0.94	0.80	0.53	0.94	0.46
V8	p<.0001	p<.0001	p<.0001	p<.0001	p<.0001	p<.0001	p<.0001	1.00	0.87	0.34	0.33	0.87	0.34	0.96
V9	p<.0001	p<.0001	p<.0001	p<.0001	p<.0001	p<.0001	p<.0001	p<.0001	1.00	0.27	0.27	0.75	0.26	0.85
V10	p<.0001	p<.0001	p<.0001	p<.0001	p<.0001	p<.0001	p<.0001	p<.0001	p<.0001	1.00	0.90	0.55	0.99	0.50
V11	p<.0001	p<.0001	p<.0001	p<.0001	p<.0001	p<.0001	p<.0001	p<.0001	p<.0001	p<.0001	1.00	0.47	0.91	0.52
V12	p<.0001	p<.0001	p<.0001	p<.0001	p<.0001	p<.0001	p<.0001	p<.0001	p<.0001	p<.0001	p<.0001	1.00	0.55	0.91
V13	p<.0001	p<.0001	p<.0001	p<.0001	p<.0001	p<.0001	p<.0001	p<.0001	p<.0001	p<.0001	p<.0001	p<.0001	1.00	0.50
V14	p<.0001	p<.0001	p<.0001	p<.0001	p<.0001	p<.0001	p<.0001	p<.0001	p<.0001	p<.0001	p<.0001	p<.0001	p<.0001	1.00

V1 = CD34/CD45; V2 = CD34/CD133/CD45; V3 = CD34/VEGF/CD45; V4 = CD34/CXCR4/CD45; V5 = CD34/CD146/CD45; V6 = CD34/VEGF/CD133/CD45; V7 = CD34/CXCR4/VEGF/CD45; V8 = CD34/CXCR4/CD133/CD45; V9 = CD34/CD133; V10 = CD34/VEGF; V11 = CD34/VEGF/CD133; V12 = CD34/CXCR4; V13 = CD34/CXCR4/VEGF; V14 = CD34/CXCR4/CD133

Table S2. Spearman Correlation Coefficients and P-values between All Progenitor Cells in Cohort 2 (N=853)

	V1	V2	V3	V4	V5	V6	V7	V8	V9	V10	V11	V12	V13	V14
V1	1.00	0.92	0.16	0.82	0.15	0.09	0.14	0.76	0.91	0.16	0.07	0.82	0.14	0.75
V2	p<.0001	1.00	0.15	0.75	0.12	0.15	0.14	0.83	0.99	0.16	0.12	0.75	0.14	0.82
V3	p<.0001	p<.0001	1.00	0.19	0.71	0.76	0.97	0.19	0.16	0.86	0.68	0.19	0.85	0.19
V4	p<.0001	p<.0001	p<.0001	1.00	0.19	0.15	0.19	0.92	0.75	0.21	0.15	0.99	0.20	0.92
V5	p<.0001	P<.001	p<.0001	p<.0001	1.00	0.58	0.71	0.15	0.13	0.61	0.52	0.19	0.61	0.16
V6	P<.01	p<.0001	p<.0001	p<.0001	p<.0001	1.00	0.74	0.22	0.16	0.67	0.87	0.16	0.66	0.22
V7	p<.0001	p<.0001	p<.0001	p<.0001	p<.0001	p<.0001	1.00	0.18	0.14	0.83	0.67	0.19	0.86	0.19
V8	p<.0001	p<.0001	p<.0001	p<.0001	p<.0001	p<.0001	p<.0001	1.00	0.83	0.21	0.21	0.92	0.21	0.99
V9	p<.0001	p<.0001	p<.0001	p<.0001	P<.001	p<.0001	p<.0001	p<.0001	1.00	0.17	0.13	0.76	0.15	0.83
V10	p<.0001	p<.0001	p<.0001	p<.0001	p<.0001	p<.0001	p<.0001	p<.0001	p<.0001	1.00	0.77	0.23	0.98	0.23
V11	P=.05	P<.001	p<.0001	p<.0001	p<.0001	p<.0001	p<.0001	p<.0001	P<.001	p<.0001	1.00	0.17	0.77	0.23
V12	p<.0001	p<.0001	p<.0001	p<.0001	p<.0001	p<.0001	p<.0001	p<.0001	p<.0001	p<.0001	p<.0001	1.00	0.22	0.92
V13	p<.0001	p<.0001	p<.0001	p<.0001	p<.0001	p<.0001	p<.0001	p<.0001	p<.0001	p<.0001	p<.0001	p<.0001	1.00	0.22
V14	p<.0001	p<.0001	p<.0001	p<.0001	p<.0001	p<.0001	p<.0001	p<.0001	p<.0001	p<.0001	p<.0001	p<.0001	p<.0001	1.00

V1 = CD34/CD45; V2 = CD34/CD133/CD45; V3 = CD34/VEGF/CD45; V4 = CD34/CXCR4/CD45; V5 = CD34/CD146/CD45; V6 = CD34/VEGF/CD133/CD45; V7 = CD34/CXCR4/VEGF/CD45; V8 = CD34/CXCR4/CD133/CD45; V9 = CD34/CD133; V10 = CD34/VEGF; V11 = CD34/VEGF/CD133; V12 = CD34/CXCR4; V13 = CD34/CXCR4/VEGF; V14 = CD34/CXCR4/CD133

Table S3. Spearman Correlation Coefficients and P-values between Some Potential Confounding Covariates in Cohort 1 (N=1280)

	PC1	PC2	Age	Weight	BMI	LDL	TRIG	TOTCHOL	HDL	SBP	Creatinine
PC1	1.00	0.024	-0.19	0.17	0.15	0.04	0.11	0.01	-0.17	-0.01	-0.06
PC2	p=0.39	1.00	0.13	-0.07	-0.09	-0.02	-0.11	-0.10	0.01	-0.02	-0.01
Age	p<.0001	p<.0001	1.00	-0.18	-0.16	-0.08	-0.07	-0.13	0.08	0.17	0.06
Weight	p<.0001	P=0.01	p<.0001	1.00	0.84	0.04	0.18	0.01	-0.26	0.04	0.15
BMI	p<.0001	P<.01	p<.0001	p<.0001	1.00	0.06	0.19	0.06	-0.16	0.06	0.03
LDL	P=0.14	P=0.57	P<.01	P=0.21	P=0.05	1.00	0.09	0.88	0.07	0.02	-0.11
TRIG	P<.001	P<.001	P=0.02	p<.0001	p<.0001	P<.01	1.00	0.32	-0.35	0.06	0.05
TOTCHOL	P=0.83	P<.01	p<.0001	P=0.81	P=0.06	p<.0001	p<.0001	1.00	0.24	0.07	-0.15
HDL	p<.0001	P=0.91	P<.01	p<.0001	p<.0001	P<0.02	p<.0001	p<.0001	1.00	0.07	-0.15
SBP	P=0.88	P=0.53	p<.0001	P=0.12	P=0.04	P=0.54	P=0.04	P=0.04	p=0.01	1.00	0.06
Creatinine	P=0.05	P=0.74	P=0.03	p<.0001	P=0.35	p<.0001	P=0.08	p<.0001	p<.001	p=0.03	1.00

PC1 = Principal Component 1; PC2 = Principal Component2; BMI = Body Mass Index; LDL = Low-density Lipoprotein; TRIG = Triglycerides; TOTCHOL = Total Cholesterol; HDL = High-density Lipoprotein; SBP = Systolic Blood Pressure

Table S4. Results of Model Fitting With Interaction Terms for Cohort 2 (n=263)

Variable	Parameter Estimate	Standard Error	95% CI	p-value
PC1	-0.29	0.69	(-1.64, 1.06)	0.68
PC2	0.04	1.02	(-1.96, 2.04)	0.97
Age	-1.76	0.50	(-2.74, -0.78)	**<.001**
BMI	1.12	0.29	(0.55, 1.69)	**<.001**
Low-density lipoprotein	-0.01	0.04	(-0.09, 0.07)	0.85
Male	9.28	3.44	(2.54, 16.02)	**<.01**
African American	-13.36	4.57	(-22.32, -4.40)	**<.01**
Smoking	13.85	5.44	(3.19, 24.51)	**0.01**
Hypertension	-4.54	3.79	(-11.97, 2.89)	0.23
Diabetes	-1.46	13.96	(-28.82, 25.90)	0.92
High-density lipoprotein	0.70	3.85	(-6.85, 8.25)	0.86
Age*PC2	-0.07	0.06	(-0.19, 0.05)	0.27
Diabetes*PC2	1.10	1.76	(-2.35, 4.55)	0.53

PC1 = Principal Component 1; PC2 = Principal Component 2; CI = Confidence Interval; BMI = Body Mass Index.