Personalized Cytomic Assessment of Vascular Health: Evaluation of the Vascular Health Profile in Diabetes Mellitus

Nicholas Kurtzman, Lifeng Zhang, Benjamin French, Rebecca Jonas, Andrew Bantly, Wade T. Rogers, Jonni S. Moore, Michael R. Rickels, and Emile R. Mohler III*

1Department of Pathology, Perelman School of Medicine at the University of Pennsylvania, Philadelphia, Pennsylvania
2Department of Medicine, Division of Cardiovascular Disease, Section of Vascular Medicine, Perelman School of Medicine at the University of Pennsylvania, Philadelphia, Pennsylvania
3Center for Clinical Epidemiology and Biostatistics, Perelman School of Medicine at the University of Pennsylvania, Philadelphia, Pennsylvania
4Department of Medicine, Division of Endocrinology, Diabetes and Metabolism, Perelman School of Medicine at the University of Pennsylvania, Philadelphia, Pennsylvania

Background: An inexpensive and accurate blood test does not currently exist that can evaluate the cardiovascular health of a patient. This study evaluated a novel high dimensional flow cytometry approach in combination with cytometric fingerprinting (CF), to comprehensively enumerate differentially expressed subsets of pro-angiogenic circulating progenitor cells (CPCs), involved in the repair of vasculature, and microparticles (MPs), frequently involved in inflammation and thrombosis. CF enabled discovery of a unique pattern, involving both MPs and CPCs and generated a personalized signature of vascular health, the vascular health profile (VHP).

Methods: Levels of CPCs and MPs were measured with a broad panel of cell surface markers in a population with atherosclerosis and type 2 diabetes mellitus (DM) and age-similar Healthy controls (HC) using an unbiased computational approach, termed CF.

Results: Circulating hematopoietic stem and progenitor cell (CHSPCAng) levels were detected at significantly lower concentrations in DM (P < 0.001), whereas levels of seven phenotypically distinct MPs were present at significantly higher concentrations in DM patients and one MP subset was present at significantly lower concentration in DM patients. Collectively, the combination of CHSPCAng and MP levels was more informative than any one measure alone.

Conclusions: This work provides the basis for a personalized cytomic vascular health profile that may be useful for a variety of applications including drug development, clinical risk assessment and companion diagnostics. © 2013 International Clinical Cytometry Society

Key terms: circulating progenitor cell; microparticle; cytometric fingerprinting; type 2 diabetes mellitus; diagnostic testing

How to cite this article: Kurtzman N, Zhang L, French B, Jonas R, Bantly A, Rogers WT, Moore JS, Rickels MR, Mohler ER III. Personalized cytomic assessment of vascular health: Evaluation of the vascular health profile in diabetes mellitus. Cytometry Part B 2013; 84B: 255–266.
The prevalence of cardiovascular disease (CVD) in the United States is high (36.9% of adults or about 81 million people) and is projected to increase by about 10% over the next 20 years, and by 2050 it is estimated that over 40% of adults (approximately 116 million people) will have one or more forms of CVD (1). Symptomatic, clinical CVD events generally only occur once atherosclerosis progresses to a point where obstructed blood flow causes ischemia, or when a thrombus develops from an existing plaque due to rupture or erosion (1). Therefore, an important unmet clinical need is an early, cost-effective diagnostic test that provides a measure of cardiovascular health prior to overt CVD as well as an assessment of the efficacy of therapeutic interventions.

Cell-based measurements may provide a means to address these limitations. Endothelial progenitor cells (EPCs) are a heterogeneous population of bone marrow-derived progenitor cells that mobilize into circulation in response to endogenous (e.g., from ischemic tissue, tumor cells) or exogenous (e.g., statins) signals. Significant evidence indicates that these cells participate in angiogenesis and perhaps postnatal vasculogenesis by supporting vascular growth (2–4). Microparticles (MPs) are 0.1–1 \( \mu \)m particles shed from all eukaryotic cells, formed by exocytotic budding due to activation or apoptosis. The role of MPs in coagulation and inflammation and their presence in atherosclerotic plaques indicate that they are an important part of atherosclerotic pathophysiology and potential biomarkers of vascular health (5,6). Early studies combining independent measures of EPCs or EPC-like cells and MPs have been reported (7,8) but the present study aiming to integrate these measures using a systems biology approach is the first of its kind.

The potential impact at the systems level of high-dimensional, high-content flow cytometry, though recently enabled by advances in instrumentation and probe technology, has not been appreciated largely because development of data analysis approaches, optimized to deal with large and complex flow cytometry data sets, have lagged behind advances in the base technology (9). Cytometric fingerprinting (CF) (10,11) provides a means to rapidly analyze such complex data without investigator and system bias. Using CF, the multivariate probability distribution functions for multiple samples can be analyzed with conventional statistical analysis methods. For example, multivariate regions that are significantly up- (or down-) regulated in one group of samples as compared with another group of samples can be discovered using multiple hypothesis testing methods similar to those now routinely employed for the analysis of gene expression data (12). The use of CF not only enables a “data mining” approach to the analysis of flow cytometric data, whereby disease- or treatment-related alterations of the multivariate distributions are discovered directly from the data, but also builds a bridge to integrative analysis with other “-omics” technologies (13). Applying CF to high dimensional datasets empowers discovery of populations and relationships that are difficult or impossible to discern by conventional analysis methods.

The Vascular Health Profile (VHP) is a personalized cytomic assay, based on systems biology principles and enabled by CF that allows the integration of MP and pro-angiogenic circulating progenitor cell (CPC) measurements to provide a signature profile of the pathophysiological state of the vasculature of an individual. Diabetes mellitus (DM) is associated with high risk of cardiovascular complications including diseases of coronary, peripheral, and carotid arteries (14,15). Thus, we hypothesized that blood samples from patients at extremely high risk for cardiovascular events, those with long-term type 2 DM and with clinically apparent atherosclerosis, will display a unique VHP signature different from healthy controls (HC) (14).

**PATIENTS AND METHODS**

**Patients and Controls**

Patients diagnosed with type 2 DM for more than 5 years and with clinically apparent atherosclerosis, history of a myocardial infarction, stroke, claudication, or revascularization procedure were included in this study, excluding those with acute illness, recent myocardial infarction or stroke (within 3 months) and pregnancy. Age-similar ‘healthy’ controls, based on no prior or current history of diabetes or cardiovascular disease, and lack of major cardiovascular risk factors (smoking, hypertension or elevated LDL cholesterol) were included. Written informed consent was obtained from all study participants and study protocols were approved by the Institutional Review Board of the University of Pennsylvania.

Gender and race distribution differed slightly between the DM and HC with more females and fewer African Americans in HC group compared to the DM group and the mean age of the DM group was also slightly older (Table 1). In addition, there was a lower proportion reporting exercise, and higher average BMI in the DM group. As expected, HbA1c was elevated in the DM group, as was blood pressure. However, LDL levels were lower in DM compared to controls as 68% of the DM cohort for the CPC analysis and 73% of the DM cohort in the MP analysis were on statins as compared with none of the controls. While most of the DM patients were on antiplatelet and/or statin drugs, few of the controls were on preventative antiplatelet medications.

**Sample Collection**

After an overnight fast, blood was collected in a golden cap (Fisher Scientific) serum separator tubes for lipid analysis and a lavender cap tube with an EDTA additive using a 21-gauge needle (Fisher Scientific) for HbA1c and CBC analysis as previously described (7), as well as High-sensitivity C-reactive protein on an automated laser-based nephelometer (Siemens Healthcare Diagnostics, Model BNII) as per manufacturer’s recommended methods. Four sodium citrate vacutainer tubes...
## Table 1

### Participant Characteristics

<table>
<thead>
<tr>
<th>Demographic characteristics</th>
<th>CPC cohort</th>
<th>MP cohort</th>
<th>(P^a)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>n</strong> 62</td>
<td>n = 62</td>
<td>n = 51</td>
<td>0.019</td>
</tr>
<tr>
<td><strong>n</strong> 48</td>
<td>n = 48</td>
<td>n = 48</td>
<td>0.010</td>
</tr>
<tr>
<td><strong>Diabetics</strong></td>
<td><strong>Controls</strong></td>
<td><strong>Diabetics</strong></td>
<td><strong>Controls</strong></td>
</tr>
<tr>
<td><strong>Age, years</strong></td>
<td>67 (58, 73)</td>
<td>59 (55, 68)</td>
<td>0.019</td>
</tr>
<tr>
<td><strong>Female, (n)%</strong></td>
<td>26 (42)</td>
<td>32 (63)</td>
<td>0.037</td>
</tr>
<tr>
<td><strong>Black or African-American, (n)%</strong></td>
<td>33 (53)</td>
<td>12 (24)</td>
<td>0.002</td>
</tr>
<tr>
<td><strong>Smoking status, (n)%</strong></td>
<td>14 (23)</td>
<td>0 (0)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td><strong>Regular exercise, (n)%</strong></td>
<td>38 (61)</td>
<td>18 (35)</td>
<td>0.021</td>
</tr>
<tr>
<td><strong>Body mass index, (kg/m(^2))</strong></td>
<td>30 (27, 36)</td>
<td>24 (23, 26)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td><strong>Tonsillectomy, (n)%</strong></td>
<td>21 (34)</td>
<td>25 (49)</td>
<td>0.13</td>
</tr>
</tbody>
</table>

**Laboratory values**

<table>
<thead>
<tr>
<th>Blood pressure (mm Hg)</th>
<th>CPC cohort</th>
<th>MP cohort</th>
<th>(P^a)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Systolic</strong></td>
<td>136 (120, 150)</td>
<td>122 (111, 129)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td><strong>Diastolic</strong></td>
<td>78 (70, 85)</td>
<td>77 (72, 85)</td>
<td>0.93</td>
</tr>
<tr>
<td>Cholesterol level (mg/dL)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>154 (126, 201)</td>
<td>211 (186, 226)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td><strong>Low-density lipoprotein</strong></td>
<td>80 (66, 106)</td>
<td>127 (117, 148)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td><strong>High-density lipoprotein</strong></td>
<td>38 (32, 48)</td>
<td>58 (49, 73)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td><strong>High-sensitivity C-reactive protein (mg/L)</strong></td>
<td>2.4 (1.0, 6.8)</td>
<td>0.9 (0.4, 1.5)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td><strong>Hemoglobin A1c (%)</strong></td>
<td>7.0 (6.5, 8.8)</td>
<td>5.6 (5.5, 5.7)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td><strong>Absolute lymphocyte count (10(^3) cells/microliter)</strong></td>
<td>1.8 (1.2, 2.4)</td>
<td>1.4 (1.2, 1.8)</td>
<td>0.009</td>
</tr>
</tbody>
</table>

**Medication use**

| Antiplatelet, \(n\)% | 45 (73) | 5 (10) | <0.001 | 37 (77) | 3 (6) | <0.001 |
| Statin, \(n\)% | 42 (68) | 0 (0) | <0.001 | 35 (73) | 0 (0) | <0.001 |

Summaries presented as median (inter-quartile range) unless otherwise noted as \(n\)%.

\(^a\)CPC and MP cohorts overlap but differ due to an analysis protocol revision during the study requiring analysis of fresh (rather than frozen) samples for MP assessment.

\(^b\)\(P\) values obtained from Wilcoxon rank-sum tests or Fisher’s exact tests, as appropriate.
were filled with 3 mL of peripheral blood for MP analysis and 30 mL of peripheral blood were also drawn into a 60 mL heparin-coated syringe for the CPC analysis.

**CPC Immunophenotyping**

Within one hour after sample collection, 30 mL of whole blood were lysed with ammonium chloride, washed twice with 3% FCS in PBS and resuspended in 10 mL of 3% FCS in PBS. After incubation with Mouse IgG (Sigma, Cat# I5381-10MG) for 10 min on ice, 8 × 10^6 cells were stained for 45 min on ice with the following antibodies in pretitrated optimized concentrations: FITC-CD31 (BD Cat# 555445, Clone WM59), PE-Cy7-CD4 (BD Cat#348791, Clone 8G12), Perccp-Cy5.5-CD3 (BD Cat# 340949, Clone SK7), Perccp-Cy5.5-CD33 (BD Cat# 341650, Clone P67.6), Perccp-Cy5.5-CD19 (BD Cat# 340951, Clone SJ25C1), V450-CD45 (BD Cat#560367 Clone HI30), PE-CD133 (Miltenyl Biotec Cat# 130-080-801, Clone AC133), APC-VEGFR2 (R&D Cat# FAB357A, Clone HI30), PE-CD144 (BD Bioscience Cat# 556570), PE-CD14 (BD Bioscience Cat# 552843), and were stained using the same reagents as were used in the sample tubes.

**Automated Gating and Data Analysis Process**

Data were analyzed using the R environment for statistical computing (version 2.13.1, R Development Core Team, Vienna, Austria) and the *flowFP* (18), *flowCore* (19), and *KernSmooth* (20) packages. For each subject, the list mode data were read and processed using the *flowCore* Biocoord package (19) in untransformed linear coordinates. Digital compensation was applied based on the spillover matrix determined by the FACSDiva acquisition software (Becton Dickinson, San Jose, CA) and stored in the FCS header, and data were mathematically normalized based on the brightest peak of reference beads (Spherotech, Cat# RCP-30-5A) run each day. The normalized fluorescence data were then biexponentially transformed and the scattering data were linearly transformed to put the fluorescence and scattering data on a similar scale. Fully automated gating strategies were developed for both the CPC and MP samples in order to eliminate possible operator bias. Fingerprinting analysis was carried out on the gated events using the *flowFP* package (18).

**Automated CPC Gating and Analysis**

The sequential gating strategy for CPCs is depicted in Figure 1. To gate the data, first a viability gate was applied in which all events were excluded above a constant threshold on the PE-A (Phycoerythrin) detector. A fixed concentration. Area, height and width FSC and SSC parameters were analyzed and side scatter width (SSC-W) was found to best resolve the beads (Fig. 2A).

The acquisition was stopped when a fixed number of 3.0 μm beads (20,000) were counted resulting in 82 thousand to 2.2 million MPs per sample. Compensation tubes were also run using PPP, BD CompBead (BD Bioscience Cat# 552843), and were stained using the same reagents as were used in the sample tubes.
Forward Scatter Area vs. Width. Finally a two-dimensional rectangular gate in CD45 vs. lineage cocktail (CD3 T-cells, CD19 B-cells, and CD33 monocytes) was created to eliminate cells already differentiated into the hematopoietic lineage and/or cells strongly expressing CD45, leaving only lineage-negative and CD45dim-negative events. The upper-right corner of the two-dimensional rectangular region was determined by first locating the CD45+ ⁄dump+ blob using the same 2D kernel density method described above, and then selecting the point along the blob boundary at its intersection with a ray drawn from the centroid of the blob horizontally to the left. The mean, median, range, minimum, maximum, and standard deviation of the resulting number of gated cells were 17,050, 15,050, 44,842, 4,348, 49,190, and 9,453, respectively.

For the gated CPC samples, a binning model was constructed using the method flowFPModel with default resolution based on the aggregate of all gated events from HC using the four measured fluorescence parameters not used in the gating (PE-Cy7/CD34, PE/CD133, APC/VEGF-R2 and FITC/CD31). The resulting binning model contained 512 bins. Fingerprints were then generated for all of the samples from both DM and HC. Relative event counts in each bin were computed by dividing the number of events in the bin by the number of events in the small cell gate (generally regarded as representing the number of lymphocytes measured in the flow cytometer). Absolute event counts were obtained by multiplying the relative event counts by the ALC laboratory result expressed as 1,000s of lymphocytes per μl of whole blood. Finally, bins were compared between DM and HC samples using the Wilcoxon test, and P-values were corrected for multiple comparisons using the Benjamini-Hochberg correction. Corrected P-values <0.05 were considered significant.

Automated MP Gating and Analysis Strategy

Data were first gated on Side Scatter Width (Fig. 2) as a relative measure of particle size to eliminate all events producing a SSC-W signal greater than that of 1μm reference beads collected each day. Thresholds for positive expression of markers were identified by visually examining the kernel density estimates of the univariate distributions for each parameter independently (Supporting Information Fig. 2). To determine sensitivity to the
choice of thresholds, the thresholds for fluorescence markers without clear separation in the kernel density estimate between the positive and negative populations were increased between about 1.4-to 2.5-fold and the results were shown to be stable, demonstrating that the choice of thresholds did not materially affect the results.

A selection of markers that could indicate cell types potentially involved in vascular health or disease were chosen for the MP panel and events that were not positive for any of these markers were gated out and not considered further in this study.

For the gated MP samples, analysis proceeded in three steps. First, a binning model was created using the method flowFPModel based on aggregated data from all HC samples using all fluorescence markers at default resolution, resulting in 8,192 bins. Fingerprints were then generated for each sample based on this model using flowFP (18). Fingerprint counts were normalized to the total number of gated events in the sample (e.g. each bin value represented the fraction of the total probability in that bin). Bins were compared between DM and HC samples using the Wilcoxon test, corrected for multiple comparisons using the Benjamini-Hochberg correction. Corrected $p$-values of $<0.05$ were considered significant. In the second step, bins judged to be significant were further grouped by phenotypic similarity. Each bin was determined to be positive with respect to each parameter if the mean minus the standard deviation of the aggregate events in that bin was greater than the threshold for positive expression for that parameter. Bins of like phenotype were further divided into separate groups based on the direction of variation of DM compared to HC. In the third and final step, the fraction of events in each of the eight phenotypic groups was computed by summing, for each sample, the number of events in the bins comprising the subset and then dividing by the total number of gated events. The results are reported in Table 2, with unadjusted $p$-values and multivariable linear regression coefficients with confidence intervals.

**Statistical Analysis**

Participant characteristics were compared between DM and HC using Wilcoxon rank-sum tests or Fisher’s exact tests, as appropriate. CPC and MP counts were compared between DM and HC using Wilcoxon rank-sum tests. Multivariable linear regression models were used to estimate adjusted differences in CPC and MP counts between groups. Adjustment variables were selected using a stepwise model selection procedure based on the Akaike information criterion (AIC), for which a variable that reduced the AIC was retained. Variables evaluated were age, gender, race, current exercise, and body mass index. Because CPC and MP counts were positively skewed, a log transformation was applied such that the exponentiated regression coefficient for DM vs. HC quantified the ratio of the average count between DM and HC groups. In a post-hoc analysis, the MP and CPC counts that exhibited the strongest differences between DM and HC were each standardized by the median in the HC group. A plot of the standardized CPCs vs. the standardized MPs was used to graphically evaluate whether a combination of MPs and CPCs was useful in distinguishing DM from HC. More formally, receiver operating characteristic (ROC) curves were used to evaluate the ability of the MP and CPC counts, with and without hsCRP as well as with and without the confounding variables, to discriminate between DM and HC. The area under the ROC curve (AUC) was estimated from a logistic regression model for DM versus HC as the dependent variable. All analyses were completed using the R Statistical Programming Environment (21).

**RESULTS**

**Subjects**

Initially, 104 subjects were consecutively recruited (52 DM and 52 HC). Subsequently, due to a revision of our
Table 2
Comparison of CPC and MP Subsets Between DM and HC Groups

<table>
<thead>
<tr>
<th>Diabetics Median (IQR)</th>
<th>Controls Median (IQR)</th>
<th>Adjusted ratio (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Compared to CHSPC</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ang-Absc CHSPC</td>
<td>97 (61, 170)</td>
<td>1.005 (0.63, 1.2)</td>
</tr>
<tr>
<td>CHSPC</td>
<td>165 (100, 250)</td>
<td>1.005 (0.63, 1.2)</td>
</tr>
<tr>
<td></td>
<td>62 (71, 178)</td>
<td>1.005 (0.63, 1.2)</td>
</tr>
<tr>
<td></td>
<td>47 (175, 180)</td>
<td>1.005 (0.63, 1.2)</td>
</tr>
<tr>
<td></td>
<td>43 (175, 180)</td>
<td>1.005 (0.63, 1.2)</td>
</tr>
<tr>
<td></td>
<td>261</td>
<td>1.005 (0.63, 1.2)</td>
</tr>
<tr>
<td>CD31/CD41a⁺</td>
<td>9.4 (4.4, 19.9)</td>
<td>1.005 (0.63, 1.2)</td>
</tr>
<tr>
<td>CD31/CD41a⁻</td>
<td>6.1 (4.4, 19.9)</td>
<td>1.005 (0.63, 1.2)</td>
</tr>
<tr>
<td>Microparticle relative density</td>
<td>54 x 10⁻⁹ (2.9 x 10⁻³, 11 x 10⁻⁹)</td>
<td>1.001 (0.63, 1.2)</td>
</tr>
<tr>
<td>CD31⁺/CD41a⁺</td>
<td>15 x 10⁻⁹ (6.7 x 10⁻³, 2.8 x 10⁻⁹)</td>
<td>1.001 (0.63, 1.2)</td>
</tr>
<tr>
<td>CD31⁻/CD41a⁻</td>
<td>2.8 x 10⁻⁹ (1.8 x 10⁻², 4.8 x 10⁻⁹)</td>
<td>1.001 (0.63, 1.2)</td>
</tr>
<tr>
<td>Distinct micro particles</td>
<td>6.6 x 10⁻⁹ (4.5 x 10⁻², 1.2 x 10⁻⁸)</td>
<td>1.001 (0.63, 1.2)</td>
</tr>
<tr>
<td>CD105⁺</td>
<td>8.7 x 10⁻⁶ (4.8 x 10⁻³, 1.2 x 10⁻⁶)</td>
<td>1.001 (0.63, 1.2)</td>
</tr>
<tr>
<td>Annexin⁺</td>
<td>3.7 x 10⁻⁶ (2.8 x 10⁻³, 6.0 x 10⁻⁶)</td>
<td>1.001 (0.63, 1.2)</td>
</tr>
<tr>
<td>CD3⁺</td>
<td>1.6 x 10⁻⁹ (9.1 x 10⁻³, 3.4 x 10⁻⁹)</td>
<td>1.001 (0.63, 1.2)</td>
</tr>
<tr>
<td>CD4⁺</td>
<td>3.7 x 10⁻⁹ (2.8 x 10⁻³, 6.0 x 10⁻⁹)</td>
<td>1.001 (0.63, 1.2)</td>
</tr>
<tr>
<td>CD105⁻</td>
<td>1.3 x 10⁻⁶ (7.8 x 10⁻³, 3.2 x 10⁻⁶)</td>
<td>1.001 (0.63, 1.2)</td>
</tr>
<tr>
<td>CD3⁻</td>
<td>1.6 x 10⁻⁹ (9.1 x 10⁻³, 3.4 x 10⁻⁹)</td>
<td>1.001 (0.63, 1.2)</td>
</tr>
<tr>
<td>CD4⁻</td>
<td>3.7 x 10⁻⁹ (2.8 x 10⁻³, 6.0 x 10⁻⁹)</td>
<td>1.001 (0.63, 1.2)</td>
</tr>
<tr>
<td>CD105⁺/CD105⁻</td>
<td>8.7 x 10⁻⁹ (4.8 x 10⁻², 1.2 x 10⁻⁸)</td>
<td>1.001 (0.63, 1.2)</td>
</tr>
</tbody>
</table>

P-values: all P-values were calculated using the Wilcoxon rank-sum test except those for CHSPC which were calculated using the Student's t-test. The adjusted ratio is calculated using the formula: Adjusted ratio = Median (Diabetics) / Median (Controls) * (95% CI) = (Lower Limit, Upper Limit).

**CPC Subset Differences Between DM and HC**

Assessment of CPCs via a traditional manual gating analysis (using FlowJo, Treestar, Ashland, OR) demonstrated no statistically significant differences between DM and HC. However, when CF was applied, a distinct phenotypic subset was discerned in one fingerprint bin. This subset, for reasons explained in the discussion, termed pro-angiogenic Circulating Hematopoietic Stem and Progenitor Cell (CHSPC<sup>Ang</sup>), had the phenotype CD34⁺/CD31⁻/CD45⁻/CD133<sup>bright</sup>/CD41a<sup>dim</sup> (Table 2 and Fig. 3), and was lower on average in DM compared with HC. Notably, the distribution of VEGF-R2 for the events in this subset approximated the overall distribution for VEGF-R2 and ranged from negative to positive, indicating that VEGF-R2 did not contribute meaningfully to the phenotypic definition of this subset. The relative event count of this subset (CHSPC<sup>AngRel</sup>) was significantly different between DM and HC (P = 8 x 10⁻⁸), and remained significant after adjustment for covariates age, gender, race, current exercise, and body mass index (Table 2) as discussed above (CHSPC<sup>AngAbs</sup> was not significant after multiple comparisons).

**MP Subset Differences Between DM and HC**

Assessment of MPs via CF led to the discovery of 8 different phenotypic subsets of MPs that differed significantly between HC and DM groups (Table 2, Fig. 4 and Supporting Information Fig. 3). In all of the MP subsets except one (discussed below), concentrations were higher on average in DM compared with HC. Statistical significance of differential expression for all of the eight MP subsets remained after adjustment for the covariates age, gender, race, current exercise, and body mass index. Each MP population discovered via fingerprinting is actually a subset of the MPs positive for the indicated marker(s). For example, Figure 4G shows a subset of CD41a⁺ MPs. These are not all of the MPs that are positive for CD41a, but rather those events that are positive for CD41a while also being negative for all other markers in...
the panel and that fall into fingerprint bins that were differentially populated between the DM and HC cohorts. Subsets that were present at significantly higher concentrations in DM patients compared with HC include CD3⁺ T-Lymphocyte MPs, CD105⁺ Endothelial MP, Annexin V⁺ MPs and CD31⁺ MPs. CD41a⁺ and Annexin V⁺/CD31⁺/CD41a⁺ subsets as well as two CD31⁺/CD41a⁺ double-positive subsets of platelet MPs were also discovered by fingerprinting to significantly differ between DM and HC. One of these CD31⁺/CD41a⁺ double-positive subsets was upregulated in DM patients (CD31dim/CD41adim), while the other (CD31bright/CD41abright) was the only one of the differentially expressed subsets that was present at lower concentrations on average in DM compared with HC. Correlation of CD31 and CD41a with the side scatter signal suggested that the level of expression of these markers was directly proportional to MP size. Finally, the ratio of the CD31dim/CD41adim subset to the CD31bright/CD41abright subset was more strongly differentially expressed ($P \approx 10^{-7}$) than either the dim or the bright subsets individually, or any of the other differentially expressed MP subsets discovered by CF, and was consequently used in forming the combined measure of CPCs and MPs discussed below.

**Combination of CPCs and MPs into the VHP Provides High Discriminatory Accuracy**

Because pro-angiogenic subsets of CPCs play a role in vascular repair, and MPs reflect cellular damage, we hypothesized that a combination of the two would be
more clinically informative with respect to vascular status than either one alone. The combination of CHSPC<sup>Ang-Rel</sup> with the ratio of dim to bright CD31<sup>+</sup>/CD41a<sup>+</sup> MPs discriminated DM and HC (Fig. 5). The AUC for the combination of these two markers was 0.86, 95% CI: (0.79, 0.94), indicating high discrimination accuracy. For comparison, the AUC for hsCRP alone was 0.74, 95% CI: (0.64, 0.85), while the AUC for confounders alone (age, gender, race, current exercise, body mass index) was 0.91, 95% CI: (0.85, 0.97). When the joint MP/CHSPC<sup>Ang-Rel</sup> measure was combined with hsCRP, the AUC increased to 0.90, 95% CI: (0.83, 0.96), and when the joint MP/CHSPC<sup>Ang-Rel</sup> measure was combined with the confounders, the AUC increased to 0.96, 95% CI: (0.92, 0.99), which indicated almost perfect discrimination accuracy.

**DISCUSSION**

This proof of principle study aimed to determine a cytomic signature of high-risk individuals using the model system of DM with atherosclerosis, a population known to be at extremely high risk for cardiovascular events. A novel method was employed to identify this abnormal vascular health profile, combining a high dimensional cell surface marker panel with an unbiased analysis scheme using CF. The present results identified one sub-population of CD31<sup>+</sup>/CD34<sup>-</sup>/CD45<sup>dim</sup>...
CD133^bright CHSPC^{Ang} that was lower in DM compared to HC. In addition, the method determined that there were 8 subpopulations of differentially expressed MPs corresponding to Platelet, T-Lymphocyte, Annexin V^+, CD31^+, and Endothelial MPs. Therefore, this study (a) confirms the hypothesis that a signature vascular health profile based on an cytometric informatics approach exists, and (b) suggests the possibility that this personalized approach to risk assessment could be useful in the evaluation of vascular health. The determination of putative pathophysiological roles for cell and MP subsets that were discovered to be expressed differently in a model of vascular disease (DM) compared with healthy individuals (HC) would be of great interest, but is beyond the scope of the present study.

**Circulating Progenitor Cells and Vascular Repair**

The physiological function of pro-angiogenic CPCs contributes to vascular homeostasis, which is crucial to prevent the pathogenesis of various vascular diseases as reviewed previously (2). Correlating specific immunophenotypic subsets with physiological function and biological role remains challenging. Unlike the present study, many studies have presumed *a priori* a progenitor phenotype, and then sought to determine if they were present or differentially expressed in study populations leading to some inconsistencies in the literature (22–24).

A recent study by Estes et al. defined a population of cells with the phenotype CD31^+, CD34^{bright}, CD133^+ and CD45^{dim}, with *in vitro* hematopoietic colony forming activity, multilineage engraftment in NOD/SCID mice, and promotion of *in vivo* angiogenesis in NOD/SCID mice tumors (25), to which they refer as circulating hematopoietic stem and progenitor cells (CHSPCs) with proangiogenic capacity. In another study, these pro-angiogenic CHSPCs were found to be elevated in pediatric subjects with malignant solid tumors as compared to healthy controls (26). In the present study, there was one population with a phenotype of CD31^+/CD34^{bright}/CD45^{dim}/CD133^{bright} that was downregulated in DM. This population, shown in Figure 3, is phenotypically nearly identical to the CHSPCs with proangiogenic capacity described by Estes et al. Consequently, this subset will be referred to as CHSPC^{Ang}. The present finding of a reduction of CHSPC^{Ang} in DM suggests that this cell type plays a significant role in the impairment of peripheral collateral vessel development in DM (27).

**Microparticles and Vascular Damage**

MPs contain miRNA and proteins from their parent cell and are often pro-coagulative and pro-inflammatory (5,28,29). The role of MPs in coagulation and inflammation and their presence in atherosclerotic plaques suggest that they are an important part of atherosclerotic pathophysiology and especially attractive as potential biomarkers of vascular health (30). Indeed, many studies have demonstrated elevated cell-specific MPs in conditions of vascular dysfunction (5). Furthermore, MPs are significantly elevated in patients with acute coronary syndromes compared to patients with stable angina symptoms (31), and are a robust predictor of secondary myocardial infarction or death (32). They are also elevated following acute ischemic stroke/cerebrovascular accident (33). In a prospective observational study of patients with DM, elevated MPs robustly predicted the presence of coronary lesions, and proved to be a more significant independent risk factor than length of diabetic disease, lipid concentrations, or the presence of hypertension (34). Collectively, these studies suggest that MP presence, number, and type could be a component of a sensitive and specific VHP assay with clinical utility in predicting atherosclerotic risk in asymptomatic patients.
Platelet MPs have both beneficial and detrimental effects on vascular health (5). These claims are supported by the present study as four distinct platelet MPs populations were present at significantly different levels between DM and HC, three of which were higher and one was lower in DM. Omoto et al. (35) found that platelet MPs are significantly higher in DM patients with nephropathy compared to DM patients without complications suggesting that platelet MPs are involved in pathological activity leading to the kidney dysfunction. Furthermore, Tan et al. (36) discovered that DM patients with clinically apparent atherosclerosis had a significantly higher level of platelet MPs than DM patients without clinically apparent atherosclerosis and HC. The present results support the conclusion that the majority of MP subsets are markers of poor vascular health.

CF was also able to discover a population of platelet MPs that were lower in DM, which is consistent with other findings that platelet MPs can have beneficial effects. For example, platelet MPs have been shown to aid in the angiogenic activity of human umbilical vein endothelial cells and augmented EPC differentiation in peripheral blood mononuclear cells (37). Mause et al. showed that platelet MPs can boost the potential for angiogenic early outgrowth cells to restore endothelial integrity after vascular injury (38). Based on the size correlation observed in the present study, it is possible that the larger platelet MPs are primarily responsible for the beneficial effects and the smaller platelet MPs for the detrimental effects. However, this speculation remains to be fully demonstrated. Importantly, the computational approach in this study distinguished four separate platelet MP phenotypes that have previously been shown to play a role in vascular health.

Endothelial MPs, like platelet MPs, are elevated in DM patients and are associated with vascular dysfunction, are a sign of cellular apoptosis, and therefore reflect vascular wall damage (39). In one study, endothelial MP levels were negatively correlated with flow-mediated dilation indicating that endothelial MPs are associated with endothelial dysfunction (40). Additionally, another study showed that endothelial MPs were significantly higher in patients with coronary artery disease than in HC (31), Singh et al. (41) found that Annexin V+ endothelial MPs were elevated in patients with cardiac allograft vasculopathy (CAV+), as compared with CAV− patients, supporting the association of these particles with vascular damage. Interestingly, they found that combining cellular and MP measures was more informative with respect to CAV than either measure alone, consistent with the present findings.

Furthermore, the present results support previous work that has shown T-Lymphocyte MPs as pro-inflammatory, decreasing NO (nitric oxide) production by reducing the level of endothelial nitric oxide synthase expression and increasing oxidative stress in endothelial cells (42). Additionally, T-Lymphocyte MPs induce endothelial dysfunction in both conductance and resistance arteries by alteration of NO prostacyclin pathways. Furthermore, T-Lymphocyte MPs impaired acetylcholine-induced relaxation of aortic rings at concentrations similar to those found in human blood attesting to their role in endothelial dysfunction (42).

Lastly, the present study supports the previous findings, as the Annexin V+ subset was significantly higher in DM patients. Annexin V binds to phosphatidylserine exposed on MPs due to inversion of the lipid membrane during apoptosis. A study of human atherosclerotic plaques showed that apoptotic MPs found in plaques account for almost all of the TF (tissue factor) activity of the plaque extracts (43). That study indicates that the MPs may play a role in the initiation of the coagulation cascade. Furthermore, MPs positive for Annexin V are higher in patients with acute coronary syndrome compared to patients with stable angina (44), indicating that an increase in Annexin V+ MPs reflects a worsening of atherosclerosis.

**Study Summary**

The results of this study indicate that CHSPCsAng are lower, and most MP subsets are higher in an atherosclerotic DM population compared to HC. Importantly, these results were obtained with CF, a novel unbiased method of data analysis capable of discovering distributional patterns that may remain hidden when using conventional methods of analysis. Several subsets were significantly differentially expressed between the two populations, some of which are supported in the literature and others are novel findings. CF is an objective, comprehensive and labor-saving method with general utility in the analysis of complex, multivariate distributions that may contain hidden data patterns.

Although there are published studies showing the prognostic value of measuring pro-angiogenic CPCs and MPs, further study is needed to determine the prognostic value of the VHP. In particular, this study demonstrated highly statistically significant differences between DM and HC cohorts of multiple cell and MP subsets, but independent prospective validation in consecutively recruited subjects will be required to validate the predictive/prognostic accuracy of the VHP. Beyond that, further study is warranted in patients with history of cardiovascular events, those with subclinical atherosclerotic disease and in healthy individuals prior to disease onset. Furthermore, although no functional assays were conducted, there is increasing evidence that surface protein expression is highly indicative of the function of the cells and subcellular particles. In conclusion, this study provides the basis for a personalized VHP, which may be useful for a number of applications including drug development, clinical risk assessment and companion diagnostics.

**ACKNOWLEDGMENTS**

The authors thank Theodore Mifflin for his assistance in gathering the hsCRP data. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Center for Research Resources or the National Institutes of Health.
Three of the authors (WTR, JSM, and ERM) declare financial interest in a company (Cytovas LLC) engaged in developing cell-based biomarkers for vascular health.

LITERATURE CITED


Cytometry Part B: Clinical Cytometry