

## Extreme erythrocyte macrocytic and microcytic percentages are highly predictive of morbidity and mortality

Benjamin D. Horne, ... , Jeffrey L. Anderson, Kirk U. Knowlton

*JCI Insight*. 2018;3(14):e120183. <https://doi.org/10.1172/jci.insight.120183>.

Clinical Medicine

Cell biology

Hematology

**BACKGROUND.** The red cell distribution width (RDW) is associated with health outcomes. Whether non-RDW risk information is contained in RBC sizes is unknown. This study evaluated the association of the percentage of extreme macrocytic RBCs (%Macro, RBC volume > 120 fl) and microcytic RBCs (%Micro, RBC volume < 60 fl) and the RDW–size distribution (RDW-sd) with mortality and morbidity.

**METHODS.** Patients (females,  $n = 165,770$ ; males,  $n = 100,210$ ) at Intermountain Healthcare were studied if they had a hematology panel between May 2014 and September 2016. Adjusted sex-specific associations of %Macro/%Micro and RDW-sd with mortality and 33 morbidities were evaluated.

**RESULTS.** Among females with fourth-quartile values of %Macro quartile and %Micro (referred to throughout as 4/4), there was an average of 7.2 morbidities versus 2.9 in the lowest risk (LR1) categories, 1/1, 1/2, 2/1, and 2/2 ( $P < 0.001$ ). Among males, those in the 4/4 category had 8.0 morbidities, while those in the LR1 had 3.4 ( $P < 0.001$ ). Cox regressions found %Macro/%Micro (4/4 vs. LR1, females: hazard ratio [HR] = 1.97 [95% CI = 1.53, 2.54]; males: HR = 2.17 [CI = 1.72, 2.73]), RDW-sd (quartile 4 vs. 1, [...])

Find the latest version:

<https://jci.me/120183/pdf>



# Extreme erythrocyte macrocytic and microcytic percentages are highly predictive of morbidity and mortality

Benjamin D. Horne,<sup>1,2</sup> Joseph B. Muhlestein,<sup>1,3</sup> Sterling T. Bennett,<sup>4,5</sup> Joseph Boone Muhlestein,<sup>3</sup> Kurt R. Jensen,<sup>1</sup> Diane Marshall,<sup>1</sup> Tami L. Bair,<sup>1</sup> Heidi T. May,<sup>1</sup> John F. Carlquist,<sup>1,3</sup> Matthew Hegewald,<sup>6</sup> Stacey Knight,<sup>1,7</sup> Viet T. Le,<sup>1</sup> T. Jared Bunch,<sup>1,8</sup> Donald L. Lappé,<sup>1,3</sup> Jeffrey L. Anderson,<sup>1,3</sup> and Kirk U. Knowlton<sup>1,9</sup>

<sup>1</sup>Intermountain Heart Institute, Intermountain Medical Center, Salt Lake City, Utah, USA. <sup>2</sup>Department of Biomedical Informatics and <sup>3</sup>Cardiology Division, Department of Internal Medicine, University of Utah, Salt Lake City, Utah, USA. <sup>4</sup>Intermountain Central Laboratory, Intermountain Medical Center, Salt Lake City, Utah, USA. <sup>5</sup>Department of Pathology, University of Utah, Salt Lake City, Utah, USA. <sup>6</sup>Pulmonary Division, Department of Internal Medicine, Intermountain Medical Center, Salt Lake City, Utah, USA. <sup>7</sup>Genetic Epidemiology Division, Department of Internal Medicine, University of Utah, Salt Lake City, Utah, USA. <sup>8</sup>Department of Internal Medicine, Stanford University, Palo Alto, California, USA. <sup>9</sup>Division of Cardiovascular Medicine, Department of Medicine, University of California San Diego, San Diego, California, USA.

**BACKGROUND.** The red cell distribution width (RDW) is associated with health outcomes. Whether non-RDW risk information is contained in RBC sizes is unknown. This study evaluated the association of the percentage of extreme macrocytic RBCs (%Macro, RBC volume > 120 fl) and microcytic RBCs (%Micro, RBC volume < 60 fl) and the RDW-size distribution (RDW-sd) with mortality and morbidity.

**METHODS.** Patients (females,  $n = 165,770$ ; males,  $n = 100,210$ ) at Intermountain Healthcare were studied if they had a hematology panel between May 2014 and September 2016. Adjusted sex-specific associations of %Macro/%Micro and RDW-sd with mortality and 33 morbidities were evaluated.

**RESULTS.** Among females with fourth-quartile values of %Macro quartile and %Micro (referred to throughout as 4/4), there was an average of 7.2 morbidities versus 2.9 in the lowest risk (LR1) categories, 1/1, 1/2, 2/1, and 2/2 ( $P < 0.001$ ). Among males, those in the 4/4 category had 8.0 morbidities, while those in the LR1 had 3.4 ( $P < 0.001$ ). Cox regressions found %Macro/%Micro (4/4 vs. LR1, females: hazard ratio [HR] = 1.97 [95% CI = 1.53, 2.54]; males: HR = 2.17 [CI = 1.72, 2.73]), RDW-sd (quartile 4 vs. 1, females: HR = 1.33 [CI = 1.04, 1.69]; males: HR = 1.41 [CI = 1.10, 1.80]), and RDW (quartile 4 vs. 1, females: HR = 1.59 [CI = 1.26, 2.00]; males: HR = 1.23 [CI = 0.99, 1.52]) independently predicted mortality. Limitations include that the observational design did not reveal causality and unknown confounders may be unmeasured.

**CONCLUSIONS.** Concomitantly elevated %Macro and %Micro predicted the highest mortality risk and the greatest number of morbidities, revealing predictive ability of RBC volume beyond what is measured clinically. Mechanistic investigations are needed to explain the biological basis of these observations.

**FUNDING.** This study was supported by internal Intermountain Heart Institute funds and in-kind support from Sysmex America Inc.

**Role of funding source:** The funding sources had no role in the design of the study, the data analysis, the interpretation of the findings, or the writing or publication of the study manuscript.

**Conflict of interest:** BDH, HTM, and JLA are inventors of clinical decision tools that are licensed to CareCentra. BDH is the principal investigator on grants funded by Intermountain Healthcare's Foundry innovation program, the Intermountain Research and Medical Foundation, CareCentra, GlaxoSmithKline, and AstraZeneca for the development and/or clinical implementation of clinical decision tools.

**Submitted:** January 29, 2018

**Accepted:** June 14, 2018

**Published:** July 25, 2018

**Reference information:**

JCI Insight. 2018;3(14):e120183.

<https://doi.org/10.1172/jci.insight.120183>.

insight.120183.

## Introduction

For more than 80 years, the degree of anisocytosis has been recognized as a useful characteristic for differentiating types of anemia (1). Measured as the coefficient of variation of RBC size, the addition of the red cell distribution width (RDW) to the complete blood count (CBC) around 1980 brought the evaluation of anemia types and the early diagnosis of iron deficiency anemia into routine practice (2, 3). The discovery

just 10 years ago of unexpected prognostic information in the RDW (4) expanded its possible uses and produced an active line of scientific investigation. Numerous studies of RDW and various outcomes were published in the last decade (5–30). Despite that wealth of data, why the RDW is higher among individuals with poor prognosis remains unclear, and the routine use of RDW as a clinical tool for personalizing care and ameliorating risk is only in its infancy.

The RDW is a statistical calculation of the dispersion in RBC sizes, not a directly measured physiological parameter. As a summary statistic, the RDW characterizes RBC size dispersion in a single value that can be used for consideration of diagnoses or health status. Inherently though, summary statistics like RDW fail to capture all available information, with their use based on the assumption that this loss of information is immaterial. If that assumption is incorrect and RDW does not capture all of the important information, it is possible that important predictive RBC size data remain overlooked. Further, almost all RDW studies evaluated what some call the RDW–coefficient of variation (RDW-CV; here referred to as RDW), while some hematology analyzers now provide an additional measure: the RDW–size distribution (RDW-sd). Hematology analyzers may also produce parameters that could be moved with relative ease into clinical production, including the percentage of RBCs that are macrocytic (%Macro, defined as RBC volume > 120 fl) and the percentage of microcytic RBCs (%Micro, RBC volume < 60 fl).

This study evaluated %Macro, %Micro, and RDW-sd as measures of RBC size among 165,770 females and 100,210 males of 18 years of age and older at Intermountain Healthcare who had CBCs and clinical outcomes. It was hypothesized that (a) additional risk prediction ability exists in the RBC size dispersion beyond that of the RDW and (b) mechanistic insights can be discovered from these additional measures of RBC size dispersion, when examined together with RDW, regarding the still undescribed biology involved in risk-associated differences in RBC sizes.

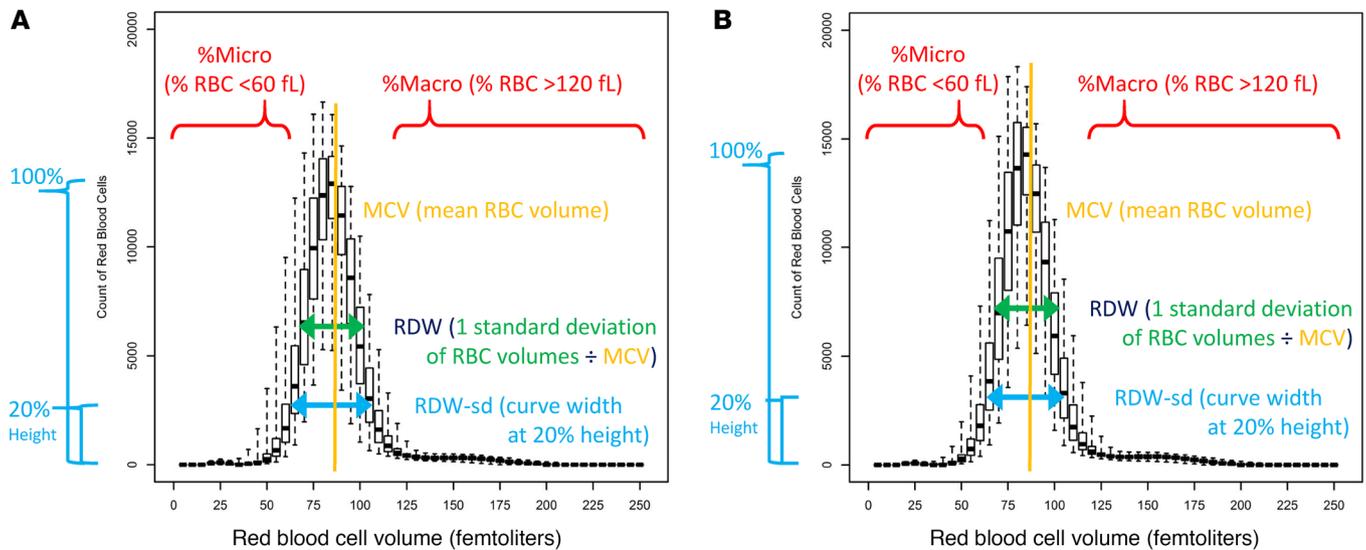
## Results

A graphical relationship among %Macro, %Micro, RDW-sd, and RDW is shown in Figure 1 using summary RBC size histogram data from this study. Baseline characteristics stratified by %Macro/%Micro are shown in Supplemental Table 1 for females and males (Supplemental material available online with this article; <https://doi.org/10.1172/jci.insight.120183DS1>). These characteristics stratified by RDW quartiles are provided in Supplemental Table 2. Essentially all characteristics differed significantly ( $P < 0.05$ ) across %Macro/%Micro categories and RDW quartiles. Figure 2 displays the patterns of association of %Macro/%Micro quartile categories with each of the 33 comorbidity diagnoses.

When examined as continuous variables, %Macro and %Micro had statistically significant correlations ( $P < 0.001$ ) with each other (females,  $r = -0.292$ ; males,  $r = -0.243$ ) and with age and other CBC components. However, the only moderate ( $0.4 \leq r < 0.7$ ) or strong ( $r \geq 0.7$ ) correlations were found with RDW (%Macro: females,  $r = 0.625$ ; males,  $r = 0.537$ ), RDW-sd (%Macro: females,  $r = 0.514$ ; males,  $r = 0.597$ ), mean corpuscular volume (MCV) (%Macro: females,  $r = 0.733$ ; males,  $r = 0.756$ ; %Micro: females,  $r = -0.704$ ; males,  $r = -0.605$ ), and mean corpuscular hemoglobin (MCH) (%Macro: females,  $r = 0.529$ ; males,  $r = 0.530$ ; %Micro: females,  $r = -0.690$ ; males,  $r = -0.622$ ). %Macro was not correlated with RDW (females,  $r = 0.022$ ; males,  $r = 0.185$ ), and %Micro was not correlated with RDW-sd (females,  $r = 0.076$ ; males,  $r = 0.057$ ). RDW and RDW-sd correlated among females ( $r = 0.739$ ) and males ( $r = 0.795$ ).

Scatterplots of %Macro and %Micro within each RDW quartile were similar for females (Figure 3) and males (Supplemental Figure 1), wherein some subjects with low %Macro had high %Micro in higher RDW quartiles, others with low %Micro had high %Macro, and some had elevated %Macro and %Micro. In females, for RDW quartiles 1, 2, 3, and 4, the %Macro interquartile ranges were 6.0%–8.1%, 6.0%–8.3%, 5.9%–8.6%, and 5.2%–8.8%, respectively, and among males were 6.3%–8.1%, 6.4%–8.5%, 6.5%–9.2%, and 6.3%–10.7%, respectively. The %Micro interquartile ranges for RDW quartiles 1, 2, 3, and 4 among females were 0.6%–1.1%, 0.7%–1.5%, 0.9%–2.1%, and 1.6%–5.9%, respectively, and for males were 0.7%–1.3%, 0.8%–1.6%, 0.9%–1.9%, and 1.2%–3.6%, respectively.

The number of morbidity diagnoses that each subject had at baseline was stratified (Figure 4) by %Macro/%Micro categories (Figure 4, C and F), RDW-sd quartiles (Figure 4, B and E), and — to a lesser extent — RDW quartiles (Figure 4, A and D). Among females, those in low-risk group 1 (LR1) had a mean of 2.9 morbidities, while females in the 4/4 category had a mean of 7.2 (Figure 4C). In multivariable analysis, including RDW, RDW-sd, and age, %Macro/%Micro provided greater ability to stratify the number of morbidities than RDW or RDW-sd (with RDW-sd being second and RDW third) (Supplemental Table 3).



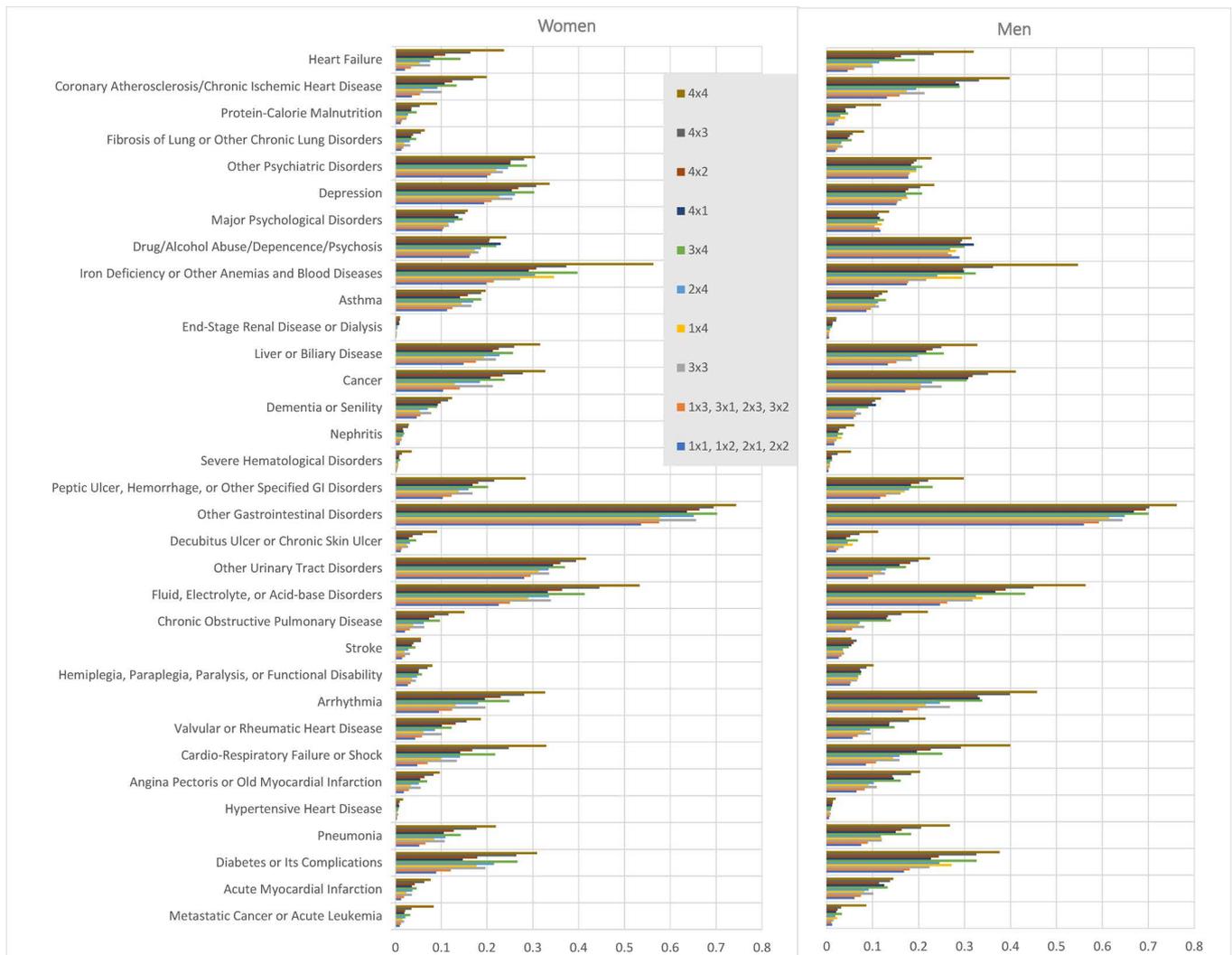
**Figure 1. Relationships of RDW, RDW-sd, %Macro, and %Micro.** Box-and-whisker plots of RBC counts in each of 50 bins of the RBC size histograms for (A) females and (B) males, along with representations of the measurement methods for %Macro, %Micro, RDW-size distribution (RDW-sd), and RDW. For each 5-fl-wide histogram bin, the boxes are 25th to 75th percentiles of RBC count, the black line is the median RBC count, and whiskers are 2.5th to 97.5th percentiles of RBC count. For displaying the RDW-sd at 20% height, the full height of the curve was the bin with the highest median RBC count.

For males, those in LR1 had a mean of 3.4 morbidities and those in the 4/4 category had 8.0 morbidities (Figure 4F), but in multivariable analysis RDW-sd had the greatest ability to stratify the number of morbidities (Supplemental Table 3).

Kaplan-Meier survival curves of %Macro/%Micro categories are displayed in Figure 5, along with survival curves for RDW and RDW-sd quartiles. Fully adjusted Cox regressions of %Macro/%Micro categories confirmed that %Macro/%Micro not only had strong univariable ability to predict mortality (females, Table 1; males, Supplemental Table 4), but that ability to predict mortality remained after multivariable adjustment for RDW, RDW-sd, and the 43 other variables (Table 1 and Supplemental Table 4). RDW and RDW-sd were also independent mortality predictors after full adjustment. Secondary analyses of %Macro quartiles alone and %Micro quartiles alone demonstrated a weaker predictive ability (Supplemental Figure 2) compared with being modeled jointly.

C-statistics (Supplemental Table 5) for the combination of %Macro/%Micro and RDW-sd were 0.002 (females) and 0.006 (males) lower than those for the combination of RDW plus these two variables and were substantially higher than c-statistics for any of the 3 variables alone, showing that a greater general level of prediction can be achieved with %Macro/%Micro and RDW-sd without loss of risk information contained in RDW. Net reclassification improvement (NRI) indices also revealed that %Macro/%Micro and RDW-sd provided risk information beyond RDW (Supplemental Table 5). Further, examination of the overall predictive ability of Cox models using the  $-2$  log likelihood demonstrated that while RDW-sd and %Macro/%Micro provided additional risk information beyond the RDW, models containing RDW-sd had substantial downgrading of the RDW-sd contribution to risk ( $-2$  log likelihood was greater) compared with that of RDW when adjusted for age and other covariables, while models including %Macro/%Micro had improved risk prediction ( $-2$  log likelihood was lower) compared with RDW or RDW-sd models when age and other factors were utilized in multivariable adjustments. The proportion of additional risk prediction ability was greater for %Macro/%Micro ( $-85$  for females,  $-71$  for males) added to a fully adjusted RDW model (i.e., with age and all other covariables) than for RDW-sd ( $-30$  for females,  $-14$  for males) added to the same RDW model, while fully adjusted models with RDW-sd and %Macro/%Micro were improved by only  $-17$  for females and  $-11$  for males when RDW was added.

As further evidence of their independent contribution to risk prediction, survival curves for RDW-sd and %Macro/%Micro categories were drawn for only those with a fourth-quartile RDW (Supplemental Figure 3). Cox regression was also performed in RDW quartile 4 (Supplemental Table 6), wherein both %Macro/%Micro categories and RDW-sd quartiles independently predicted mortality. Cox regression was used further in the fourth quartile of RDW-sd, showing that %Macro/%Micro (for category 4/4 vs. LR1:

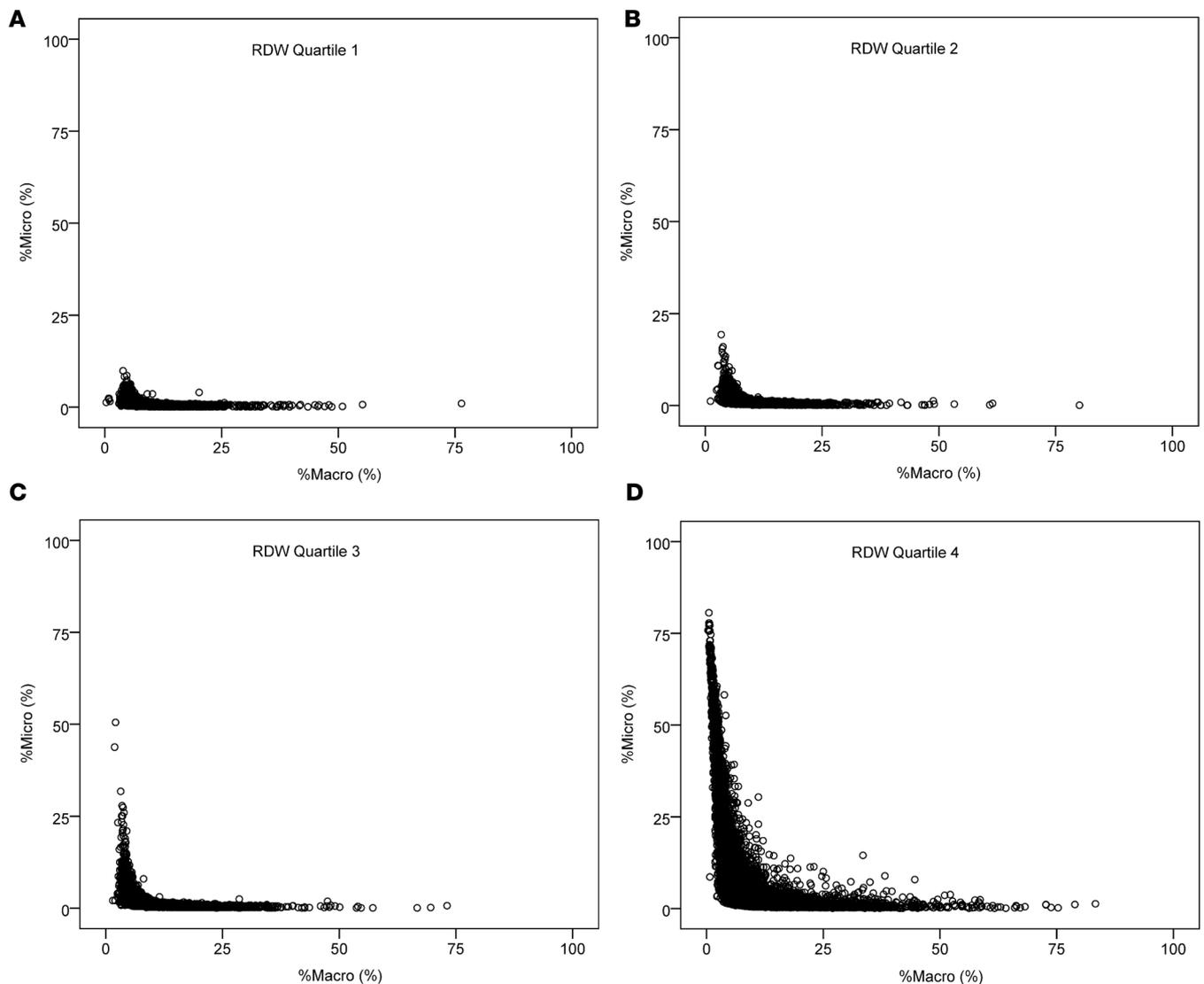


**Figure 2. Morbidity levels by %Macro/%Micro category.** Bar graphs of the percentage of individuals with each morbidity based on %Macro/%Micro quartile categories (females,  $n = 165,770$ ; males,  $n = 100,210$ ). Numeric data for these associations are provided in Supplemental Table 1, and comorbidity diagnosis definitions are provided in Supplemental Table 9. The distribution of each of the morbidities was significantly different ( $P < 0.001$  by the  $\chi^2$  test) across the %Macro/%Micro quartile categories (except for major psychological disorder among males, where  $P = 0.006$ ).

females had HR = 2.96, CI = 1.47, 6.00,  $P = 0.003$ ; males had HR = 2.06, CI = 1.26, 3.37,  $P = 0.004$ ) and age, but not RDW, were associated with mortality. Cox regression among individuals with a %Macro/%Micro category with a fourth quartile (4/1, 4/2, 4/3, 1/4, 2/4, 3/4, 4/4) showed RDW (fourth vs. first quartile: females had HR = 2.05, CI = 1.60, 2.63,  $P < 0.001$ ; males had HR = 2.30, CI = 1.76, 3.00,  $P < 0.001$ ), RDW-sd (fourth vs. first quartile: females had HR = 2.58, CI = 1.98, 3.34,  $P < 0.001$ ; males had HR = 2.78, CI = 2.15, 3.60,  $P < 0.001$ ), and age were associated with mortality. Finally, Cox regressions among subgroups of only those subjects with each morbidity diagnosis were performed to evaluate the association of %Macro/%Micro categories and RDW with mortality (Supplemental Tables 7 and 8).

### Discussion

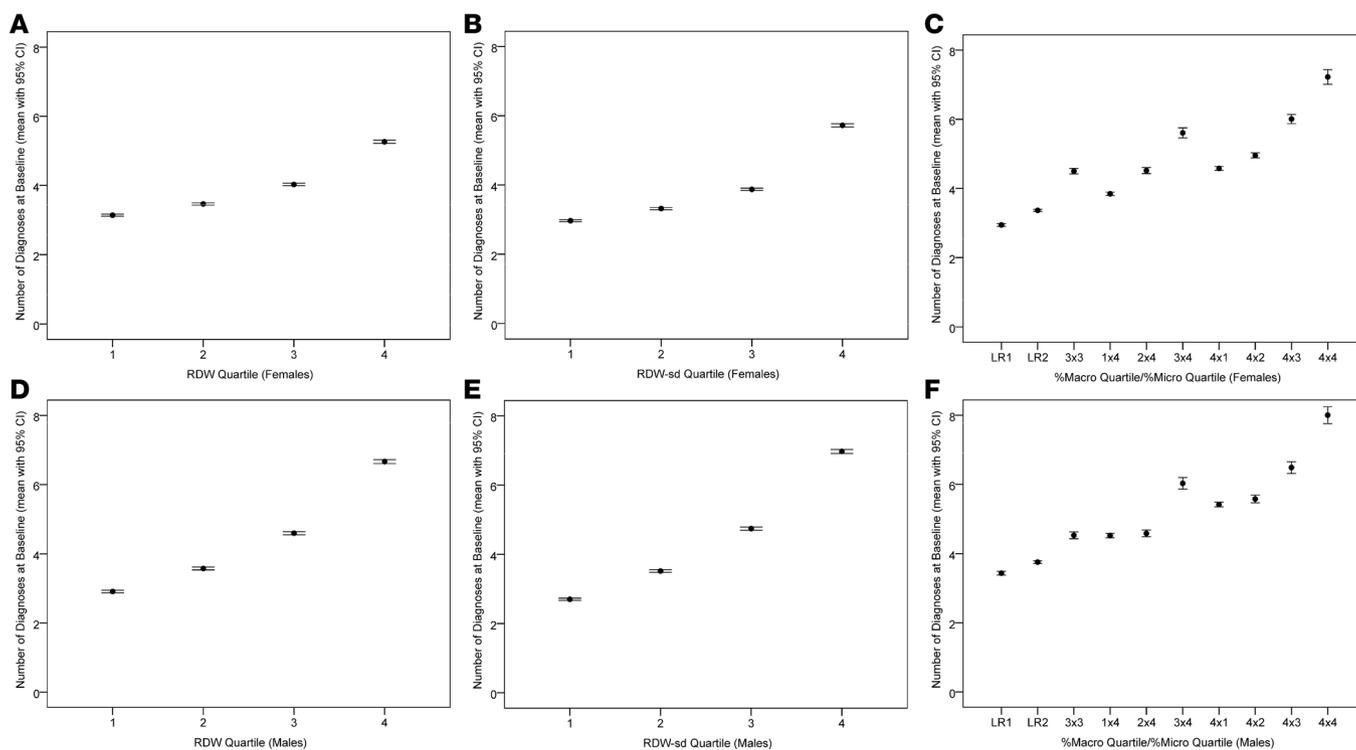
**Major findings.** Concomitantly elevated levels of %Macro (RBC volume > 120 fl) and %Micro (RBC volume < 60 fl) were the RBC size dispersion measures providing the greatest relative risk association with all-cause mortality in a very large population of general healthcare patients, and RDW-sd was also a strong independent risk marker beyond RDW. %Macro correlated best with RDW-sd and %Micro with RDW, but they were independent mortality predictors, despite adjustment for all of these and for many other variables. Measures of clinical utility, including the c-statistic and NRI index, and the Cox regression model predictive abilities suggested that %Macro/%Micro and RDW-sd together accounted for



**Figure 3. Distributions of %Macro and %Micro in RDW subgroups.** Scatterplots of the percentage of RBCs that are macrocytic (%Macro) and microcytic (%Micro) among females ( $n = 165,770$ ) in RDW quartile (A) 1 ( $n = 39,869$ ); (B) 2 ( $n = 43,942$ ); (C) 3 ( $n = 41,920$ ); and (D) 4 ( $n = 40,039$ ). Quartile thresholds are provided in the Methods (e.g., quartile 4 of %Macro was  $\geq 8.5\%$  and quartile 4 of %Micro was  $\geq 2.2\%$ ). For results among males, see Supplemental Figure 2.

almost all risk information in RDW, that RDW-sd was heavily confounded by age and not as relatively strong as it appeared in univariable analyses, and that %Macro/%Micro was the strongest contributor of the 3 RBC size factors when multivariable adjusted; additionally, it was found that %Macro/%Micro provided unique risk information not contained in the RDW, RDW-sd, age, or other variables. High combined %Macro/%Micro category and elevated RDW-sd were also strongly associated with a greater prevalence of each of 33 morbidities and were associated with high mean number of morbidities. Elevation of %Macro alone or %Micro alone was associated less strongly with mortality risk than when the two factors were combined.

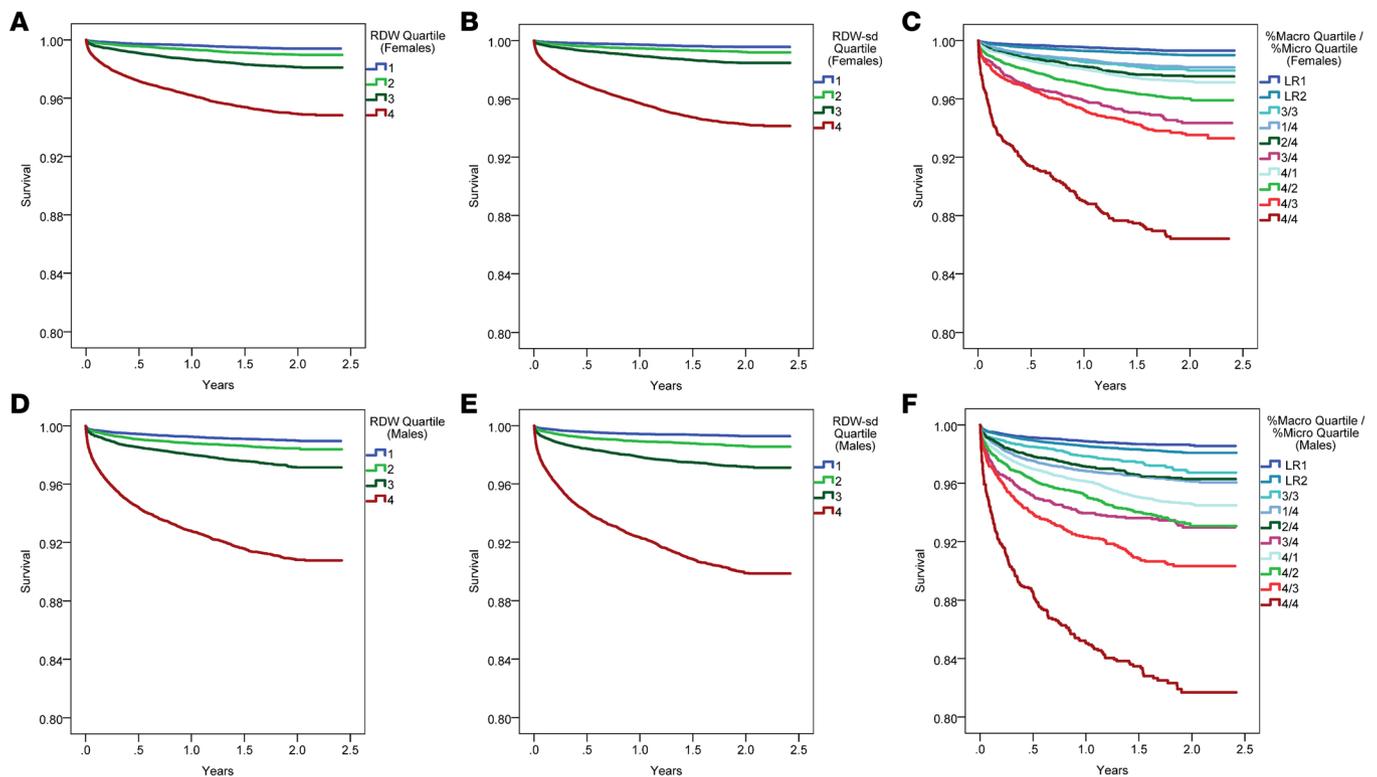
*Biological and clinical implications.* In the postgenomic era, the CBC panel is among the most common clinically measured parameters; however, it is not as thoroughly studied in clinical research. About a decade ago, the RDW was reported to be associated with mortality in cardiac populations (4, 5). Various etiologies have been hypothesized in the search for biological pathways involved in elevating RDW. Associations of RDW with mortality have been reported, though, despite adjustment for factors related to erythropoiesis, blood disorders, nutritional deficiencies, inflammation, and various other pathophysiologic pathways among populations defined by a plethora of distinct and seemingly unrelated diagnostic groups. Those groups include but are not limited to patients with coronary disease (4, 10, 18), heart failure (5, 7, 24), myocardial



**Figure 4. RBC size dispersion and morbidities.** Association with the sum of morbidity diagnoses for the RBC size dispersion variables, with the mean and 95% CI provided among females ( $n = 165,770$ ), (A) RDW quartiles, (B) RDW-sd quartiles, and (C) %Macro/%Micro quartile categories, and among males ( $n = 100,210$ ), (D) RDW quartiles, (E) RDW-sd quartiles, and (F) %Macro/%Micro quartile categories. All comparisons to quartile 1 or LR1 were statistically significant at  $P < 0.001$  by ANOVA.

infarction (11, 13), stroke (8), cancer (26), trauma (20), chronic obstructive pulmonary disease (22), community-acquired pneumonia (15), acute kidney injury (19), pulmonary arterial hypertension (16), peripheral arterial disease (14), and critical illness (17) as well as unselected hospitalized patients (6, 21) and the general nonmedical nationally representative US population (9, 12). Associations of RDW with nonfatal events and outcomes have also been found, including risk of incident heart attack (10, 23), incident heart failure (10, 23), heart failure hospitalization (5, 7), 10-year coronary heart disease diagnosis (12), atrial fibrillation (10), stroke (10), cancer stage (26), complicated hospitalization (15), length of hospital stay (24), bloodstream infection (17), major bleeding (27), vascular complications (27), and many other outcomes. While the majority of studies have examined chronic RDW elevation, acute changes in RDW during or near the time of disease diagnosis or major health events are reportedly associated with differential outcomes (23, 24). Further search for causes of risk-associated RDW elevations has led to examinations of predictors of RDW level, and those that are described reveal a broad connection of red cell size differences with health in both systemic and physiology-specific processes (11, 18, 21, 24, 28–30). One key process that is implicated is inflammation; although many inflammatory factors are associated with RDW level (7, 18), they do not appreciably reduce RDW associations with mortality or other health outcomes (7, 18). With the systemic reach of inflammatory predictors and widely recognized role of inflammation in common diseases (31, 32), their relationship to RDW is important to understand since red cell size appears to integrate some of the risk information in inflammatory markers and because they are far from perfectly correlated.

The vast array of diagnostic groups in which RDW predicts health outcomes, the myriad of diagnoses that it predicts, and the plethora of other risk factors associated with RDW suggest that the dispersion of RBC size is a more general marker of health than any single traditional risk pathway can explain. Even if multiple pathways (e.g., inflammation, nutrition, blood disorders) are involved biologically in risk-associated RDW elevation, given that RDW's primary clinical use is in the realm of anemia and considering that nature tends toward parsimony, it is reasonable to conclude that a common bond between those biological pathways, such as oxygen demand and supply, is the underlying biological mechanism that RDW describes (23). Anemia, other blood disorders, inflammation, nutritional deficiencies, and other conditions may create



**Figure 5. RBC size dispersion and survival.** Kaplan-Meier survival curves illustrating the association with mortality among females ( $n = 165,770$ ) for (A) RDW quartiles, (B) RDW-sd quartiles, and (C) %Macro/%Micro quartile categories and among males ( $n = 100,210$ ) for (D) RDW quartiles, (E) RDW-sd quartiles, and (F) %Macro/%Micro quartile categories.

hypoxemia in various forms (even forms not described by RBC count or hemoglobin level, such as tissue perfusion) that stimulates the body's need for oxygen. The ability of the body's cells to communicate oxygen needs has been documented (33, 34), thus a common mechanistic pathway may exist regardless of what triggers it. Potentially, any pathophysiology that causes increased chronic or acute oxygen demand for which the body is not prepared or to which it struggles to respond, or normal demand with anemia or bone marrow insufficiency that may cause a supply/demand imbalance, could increase RBC size dispersion.

*Delay of eryptosis or early release from marrow?* With an epidemiologic design, this study cannot determine biological mechanisms but can provide some hints to inform biological investigations searching for the mechanisms. Variations in RBC sizes measured by the numerator of RDW and by the clearance rate of RBCs were previously reported to be associated with prognosis (21, 25). It was hypothesized that RDW elevation may result due to a greater %Micro (25), %Macro (23), or both (21). This study confirms that it is all 3 (see Figure 5), showing that having concomitantly higher %Macro and higher %Micro is the highest risk condition across all morbidities and for mortality, with elevation of one or the other having an intermediate level of risk. These data also demonstrate that an elevation in %Macro without greater %Micro is another condition connected with health risk (albeit an intermediate risk level). In a third patient subset, the study also confirms that a change in the set point for eryptosis (i.e., erythrocyte apoptosis) that reduces the RBC clearance rate and elevates %Micro is another process that elevates RBC size dispersion (25), although this is not the only or necessarily the most important process.

While it is unclear why the body would change the set point of eryptosis and develop an elevated %Micro (as previously described, ref. 25) in so many disparate diagnoses as found herein (see Figure 2), it is also unclear which larger erythrocytes are responsible for the elevated %Macro or why they are deposited early into the circulation in that same set of diagnoses regardless of whether %Micro is elevated or not. Some evidence suggests that those macrocytic RBCs may be immature macrocytic reticulocytes that are released from the bone marrow due to stimulation from increased erythropoietin (35–37). Presumably, the elevation in %Macro and %Micro (together or separately) provides more circulating RBCs and greater oxygen delivery capacity when an insufficient supply of oxygen is reaching target cells. But questions arise from these data.

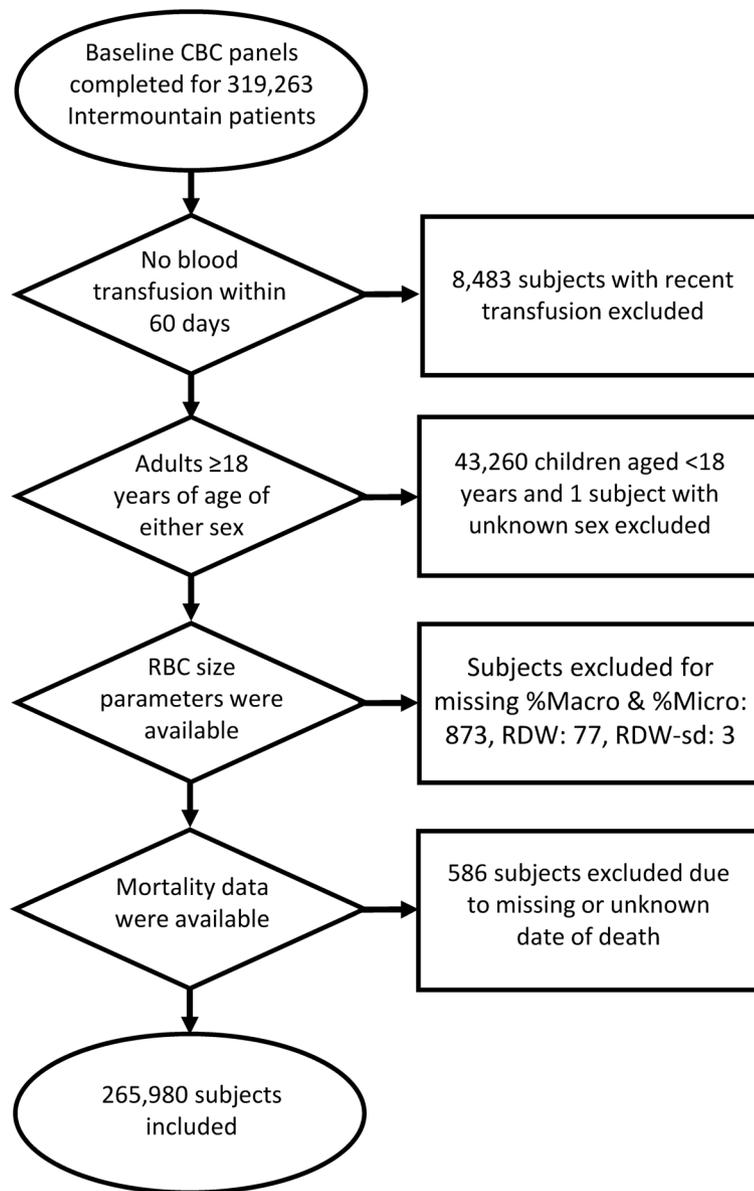
**Table 1. Hazard ratios and 95% CIs from Cox regression for the association with all-cause mortality of the RBC size dispersion variables among females (n = 165,770)**

Macrocytic/ microcytic	Absolute	Univariable	Age adjusted	Age, CBC <sup>A</sup>	Fully adjusted <sup>A</sup>	Sample
Quartile combination	Mortality risk	HR (95% CI)	HR (95% CI)	HR (95% CI)	HR (95% CI)	Size (n)
Quartile categories of %Macro/%Micro						
LR1 <sup>B</sup>	0.6%	1.0 (referent)	1.0 (referent)	1.0 (referent)	1.0 (referent)	23,235
LR2 <sup>C</sup>	0.8%	1.47 (1.20, 1.79) <sup>D</sup>	1.06 (0.87, 1.29)	0.97 (0.80, 1.19)	0.98 (0.79, 1.20)	53,486
3/3	1.7%	3.00 (2.37, 3.79) <sup>D</sup>	1.38 (1.09, 1.75) <sup>E</sup>	1.03 (0.80, 1.32)	0.96 (0.75, 1.24)	9,565
1/4	1.5%	2.66 (2.17, 3.25) <sup>D</sup>	2.23 (1.82, 2.73) <sup>D</sup>	1.67 (1.31, 2.13) <sup>D</sup>	1.46 (1.14, 1.86) <sup>E</sup>	25,910
2/4	2.0%	3.70 (2.95, 4.65) <sup>D</sup>	2.21 (1.76, 2.78) <sup>D</sup>	1.67 (1.29, 2.16) <sup>D</sup>	1.38 (1.07, 1.80) <sup>E</sup>	8,691
3/4	4.7%	8.47 (6.73, 10.66) <sup>D</sup>	3.64 (2.89, 4.59) <sup>D</sup>	2.34 (1.81, 3.04) <sup>D</sup>	1.77 (1.36, 2.30) <sup>D</sup>	3,751
4/1	2.3%	4.25 (3.50, 5.16) <sup>D</sup>	1.81 (1.49, 2.20) <sup>D</sup>	0.97 (0.78, 1.21)	0.90 (0.71, 1.13)	20,880
4/2	3.4%	6.17 (5.07, 7.51) <sup>D</sup>	2.30 (1.88, 2.80) <sup>D</sup>	1.26 (1.01, 1.57) <sup>E</sup>	1.14 (0.91, 1.43)	13,216
4/3	5.4%	9.62 (7.78, 11.89) <sup>D</sup>	3.07 (2.48, 3.80) <sup>D</sup>	1.39 (1.09, 1.77) <sup>E</sup>	1.18 (0.92, 1.51)	5,064
4/4	11.9%	21.72 (17.44, 27.05) <sup>D</sup>	7.06 (5.66, 8.81) <sup>D</sup>	2.55 (1.99, 3.26) <sup>D</sup>	1.97 (1.53, 2.54) <sup>D</sup>	1,972
RDW quartile						
1	0.5%	1.0 (referent)	1.0 (referent)	1.0 (referent)	1.0 (referent)	39,869
2	0.8%	1.72 (1.44, 2.06) <sup>D</sup>	1.21 (1.01, 1.45) <sup>E</sup>	1.18 (0.99, 1.41)	1.12 (0.92, 1.35)	43,942
3	1.5%	3.27 (2.78, 3.85) <sup>D</sup>	1.80 (1.53, 2.12) <sup>D</sup>	1.70 (1.44, 2.00) <sup>D</sup>	1.28 (1.04, 1.57) <sup>E</sup>	41,920
4	4.3%	9.13 (7.85, 10.62) <sup>D</sup>	4.17 (3.58, 4.86) <sup>D</sup>	3.43 (2.92, 4.02) <sup>D</sup>	1.59 (1.26, 2.00) <sup>D</sup>	40,039
RDW-sd quartile						
1	0.3%	1.0 (referent)	1.0 (referent)	1.0 (referent)	1.0 (referent)	42,305
2	0.6%	1.83 (1.50, 2.25) <sup>D</sup>	1.22 (0.99, 1.50)	1.04 (0.84, 1.29)	1.07 (0.86, 1.32)	41,223
3	1.2%	3.54 (2.95, 4.26) <sup>D</sup>	1.68 (1.39, 2.03) <sup>D</sup>	1.14 (0.93, 1.41)	1.12 (0.89, 1.39)	41,124
4	4.8%	14.04 (11.86, 16.63) <sup>D</sup>	4.37 (3.67, 5.20) <sup>D</sup>	1.71 (1.38, 2.12) <sup>D</sup>	1.33 (1.04, 1.69) <sup>E</sup>	41,117

<sup>A</sup>Adjustments in Cox regression for traditional CBC components included the RDW, RBC count, WBC count, platelet count, mean corpuscular volume, mean corpuscular hemoglobin concentration, and mean platelet volume (hemoglobin, hematocrit, and mean corpuscular hemoglobin were collinear with other CBC components and thus excluded from analyses), while full adjustment included age, all 3 RBC dispersion variables, other CBC components, and the 38 morbidities and health system encounter types listed in Supplemental Table 1. <sup>B</sup>LR1, combined group of %Macro/%Micro quartiles 1/1, 1/2, 2/1, and 2/2. <sup>C</sup>LR2, combined group of %Macro/%Micro quartiles 1/3, 3/1, 2/3, and 3/2. <sup>D</sup> $P < 0.001$ ; <sup>E</sup> $P < 0.05$ . For data from males (n = 100,210), please see Supplemental Table 4.

Why in some people does %Macro rise when %Micro does not? Why does %Micro become elevated in other people when %Macro does not? What is the very general but undescribed biological pathway that leads these distinct RBC processes to be potentially active in apparently any and all disease processes? Why do individuals with both elevated %Macro and %Micro have the greatest number of morbidities and risk of mortality?

*Limitations.* This study may be limited by its observational design, which used historical albeit prospectively collected data. Issues may include that unobserved variables or residual confounding may limit the generalizability of study findings. While patients receiving a blood transfusion were excluded, no information was available regarding included participants who were recent blood donors. In addition, data regarding menstruation status for females and medication information for this unselected healthcare population were not available. Further, the population studied here included individuals seen in the context of a healthcare encounter (the population was not selected from apparently healthy people), thus the severity of illness of those studied herein may be higher than that in the general population; longer follow up would be needed to study lower risk healthy populations. Interpretation of results should be performed with care, and additional studies should be conducted to confirm the associations, including in general nonhealthcare populations. The study did utilize a large number of covariables for multivariable adjustment, though, to remove confounding due to acuity (i.e., using data regarding the healthcare setting) and due to illness severity (i.e., using morbidity diagnoses for 33 broad disease classifications and other factors from the CBC panel), and the sex-specific study populations were very large (n > 100,000). Validation studies should also examine racially and geographically diverse populations. Whether measures of RBC size dispersion can be used to personalize care or target treatments for improving patient outcomes will require further studies in clinical practice settings.



**Figure 6. Study inclusion criteria.** Flow chart of subject inclusion/exclusion and CBC panel selection.

*Conclusions.* Elevations in %Macro (the extreme erythrocyte macrocytic percentage or right tail of the RBC size histogram) or %Micro (the extreme erythrocyte microcytic percentage or left tail of the RBC size distribution) and in RDW-sd were associated with substantially higher risk of mortality and with a greater sum of morbidities that a patient carried, independent of the RDW. Despite that the RDW is one of the most predictive common laboratory tests for prognosis and mortality, use of additional variables that describe other components of the dispersion in RBC sizes provided additional risk prediction ability beyond what the RDW provides, and the new measures were more powerful than the RDW. New mechanistic investigations are needed to determine what the biological processes are that lead to elevated %Macro, %Micro, RDW-sd, and RDW; why %Macro or %Micro may be activated in some cases and not others; and why %Macro, %Micro, or both may be involved in the majority of disease states.

Measuring the dispersion of RBC size, including the %Macro that may result from early RBC release from the bone marrow and the %Micro that may represent the retention of older cells via a delay of eryptosis, appears to have broad applicability for the examination of human physiologic state and homeostatic adaptation to disease. Although presently only used as research parameters, implementation of

%Macro and %Micro along with RDW and RDW-sd in clinical practice may be feasible, since the CBC panel already generates the data to calculate them, some hematology platforms already provide them for research, and use of this information may lead to a number of clinical improvements. No other risk factor like the RBC size dispersion currently exists that may describe the underlying general health status of a patient, and, thus, further research is required to examine the combination of RDW, RDW-sd, %Macro, and %Micro or other metrics encapsulating the risk information in RBC sizes to determine whether they could be used to aid clinician decision making for diagnosis and treatment. Further, additional studies are needed to more fully explore the value and limits of %Macro/%Micro and RDW-sd, including whether they are more predictive of mortality in specific subgroups or if certain patient populations are at such low risk that the CBC panel does not need to be evaluated at a cost savings to the healthcare system. Other questions to be answered are whether the RBC size dispersion factors predict incident nonfatal disease or if they predict benefit from medication treatments and which treatments, medical or health interventions, or lifestyle changes normalize the RBC size dispersion.

## Methods

*Study aims.* The primary hypothesis of this study was that the combination of %Macro and %Micro predict mortality regardless of adjustment for RDW, RDW-sd, or other factors. These measures represent the right and left tails, respectively, of the RBC size histogram, and their thresholds of >120 fl and <60 fl were chosen as clearly indicative of extreme RBC size (i.e., volume). Secondary hypotheses included that the RDW-sd contains additional risk information and that %Macro, %Micro, and RDW-sd predict the presence of morbidities.

*Study population.* Study subjects included patients aged 18 years or older who were treated at an Intermountain Healthcare facility for any cause in an emergency, outpatient, or inpatient setting, with hematology panels evaluated at specific Intermountain clinical laboratories from May 2014 to early September 2016. The Intermountain integrated delivery system (21 hospitals, 185 clinics) serves about two-thirds of the population of Utah and southeastern Idaho. Subjects who received a blood transfusion within 60 days prior to laboratory testing or were missing %Macro, %Micro, RDW, or RDW-sd measurements were excluded from the study (Figure 6). For individuals with more than one CBC measurement, analyses used the first CBC test.

*Primary study outcome.* The primary study endpoint was all-cause mortality. Mortality outcomes were determined from a combination of Intermountain Healthcare electronic records, the US Social Security Administration death master file, and State of Utah death certificates, which provided complete follow up for mortality events. Subjects were followed until October 2016.

*Study variables.* Study characteristics included subject age, sex, type of healthcare encounter (i.e., inpatient hospitalization, outpatient exams or procedures, labor and delivery, or emergency visits), receipt of coronary revascularization, and 33 morbidities (Supplemental Table 1) defined based on International Classification of Diseases–9/10 groupings, using the Centers for Medicare and Medicaid Services Common Conditions methodology (Supplemental Table 9). These variables were collected from Intermountain's electronic data warehouse.

CBC testing was performed clinically on Sysmex hematology analyzers (Sysmex America), and in-kind support was provided by Sysmex to encode the software necessary to extract the RBC histogram data that are not clinically reportable. CBC parameters included the RDW, RDW-sd, RBC count, hemoglobin, hematocrit, WBC count, platelet count, MCV, MCH, MCH concentration (MCHC), and mean platelet volume (MPV). Finally, %Macro and %Micro were calculated from the histogram of RBC sizes, with %Macro defined as the percentage of RBCs with a volume >120 fl and %Micro defined as the percentage of RBCs with a volume <60 fl. CBC testing produces the histograms necessary to calculate %Macro and %Micro, but those data are typically discarded once the MCV, RDW, and — on some platforms — RDW-sd are calculated. Herein, assessments of individual RBC sizes are included in the %Macro and %Micro measures and are not the mean used to assess an MCV-based anemia.

*Statistics.* Baseline demographic, CBC component, morbidity, and health encounter characteristics are reported as percentages or means with standard deviations, as appropriate, with stratification by %Macro/%Micro quartile category or RDW quartile. Comparisons across those categories or quartiles were performed using the  $\chi^2$  test or ANOVA. Pearson's correlation coefficients ( $r$ ) were calculated for age, %Macro, %Micro, RDW, RDW-sd, and other CBC components.

Among females, quartiles of %Macro were  $\leq 5.8\%$ , 5.9%–6.8%, 6.9%–8.4%, and  $\geq 8.5\%$ ; quartiles of %Micro were  $\leq 0.7\%$ , 0.8%–1.2%, 1.3%–2.1%, and  $\geq 2.2\%$ ; quartiles of RDW were  $\leq 12.8\%$ , 12.9%–13.4%, 13.5%–14.2%, and  $\geq 14.3\%$ ; and quartiles of RDW-sd were  $\leq 41.5$  fl, 41.6–43.6 fl, 43.7–46.2 fl, and  $\geq 46.3$  fl. Among males, %Macro quartiles were  $\leq 6.3\%$ , 6.4%–7.3%, 7.4%–9.0%, and  $\geq 9.1\%$ ; %Micro quartiles were  $\leq 0.8\%$ , 0.9%–1.2%, 1.3%–1.9%, and  $\geq 2.0\%$ ; quartiles of RDW were  $\leq 12.7\%$ , 12.8%–13.2%, 13.3%–13.9%, and  $\geq 14.0\%$ ; and quartiles of RDW-sd were  $\leq 40.8$  fl, 40.9–43.1 fl, 43.2–46.2 fl, and  $\geq 46.3$  fl. Further, %Macro and %Micro quartiles were categorized jointly, with %Macro quartile listed first and %Micro second, with categories 1/1, 1/2, 2/1, and 2/2 combined into LR1 and categories 1/3, 3/1, 2/3, and 3/2 grouped into low-risk group 2 (LR2).

Sex-specific scatterplots of %Macro versus %Micro were constructed within each RDW quartile. Cross-sectional analyses using univariate ANOVA evaluated sex-specific associations of %Macro/%Micro categories, RDW-sd quartiles, and RDW quartiles with the 33 morbidities. ANOVA models analyzed individual variables, trivariable models, and age with the 3 variables.

Kaplan-Meier survival curves were drawn for %Macro/%Micro categories, RDW quartiles, and RDW-sd quartiles. Secondly, curves were also drawn for %Macro quartiles and %Micro quartiles individually. Cox regression calculated univariable HRs with 95% CIs, and multivariable models determined adjusted HRs and CIs. Sex-specific multivariable modeling used stepwise variable entry (forward and backward) adjusting for age, standard CBC components (RDW, RBC count, WBC count, platelet count, MCV, MCHC, and MPV), RDW-sd, 33 morbidities, receipt of coronary revascularization, and other healthcare encounter characteristics (inpatient hospitalization, outpatient exams or procedures, labor and delivery, or emergency visits).

Survival analyses (Kaplan-Meier and Cox regression) were also performed in RDW quartile 4 to evaluate associations of %Macro/%Micro categories and RDW-sd quartiles with mortality. Due to the reduced sample size in this high RDW population, LR1 and LR2 were combined and new RDW-sd quartiles were derived. Other survival analyses of %Macro/%Micro and RDW were performed within subgroups defined by each of the 33 morbidities. C-statistics and the NRI index were also computed for a selection of univariable, bivariable, and multiple variable combinations, with analyses of %Macro/%Micro, bivariable, and multivariable combinations using the predicted probabilities from logistic regression. For further model predictive ability comparisons, the  $-2$  log likelihood statistics were also computed in the Cox regressions described above and were compared with those from models only including RDW, RDW plus covariables other than age, and RDW plus age and covariables. All analyses used 2-sided tests of hypothesis, with  $P$  values of less than or equal to 0.05 considered significant. Statistical analyses used SPSS version 23.0 (IBM SPSS Statistics).

*Study approval.* This study was approved by the Intermountain Institutional Review Board, with a waiver of consent, as this is a historical data-only study that posed minimal risk to subjects. The study was done in accordance with the Declaration of Helsinki.

## Author contributions

BDH had full access to study data, takes responsibility for the integrity of the data and the accuracy of the data analysis, and had authority over manuscript preparation and the decision to submit the manuscript for publication. BDH, Joseph B. Muhlestein, STB, Joseph Boone Muhlestein, MH, VTL, DLL, JLA, and KUK provided study concept and design. BDH, Joseph B. Muhlestein, STB, KRJ, DM, TLB, HTM, JFC, SK, and TJB acquired, analyzed, or interpreted data. BDH, Joseph B. Muhlestein, STB, Joseph Boone Muhlestein, and KUK drafted the manuscript. KRJ, DM, TLB, HTM, JFC, MH, SK, VTL, TJB, DLL, and JLA critically revised the manuscript for important intellectual content. BDH provided statistical analysis. BDH, Joseph B. Muhlestein, STB, DM, DLL, and KUK provided study supervision.

## Acknowledgments

This study was supported by internal Intermountain Heart Institute funds and by in-kind support from Sysmex America Inc.

Address correspondence to: Benjamin D. Horne, Intermountain Heart Institute, 5121 S. Cottonwood Street, Salt Lake City, Utah 84107, USA. Phone: 801.507.4708; Email: benjamin.horne@imail.org.

1. Daland GA, Heath CW, Minot GR. Differentiation of pernicious anemia and certain other macrocytic anemias by the distribution of red blood cell diameters. *Blood*. 1946;1:67–75.
2. Bessman JD, Gilmer PR, Gardner FH. Improved classification of anemias by MCV and RDW. *Am J Clin Pathol*. 1983;80(3):322–326.
3. McClure S, Custer E, Bessman JD. Improved detection of early iron deficiency in nonanemic subjects. *JAMA*. 1985;253(7):1021–1023.
4. Anderson JL, et al. Usefulness of a complete blood count-derived risk score to predict incident mortality in patients with suspected cardiovascular disease. *Am J Cardiol*. 2007;99(2):169–174.
5. Felker GM, et al. Red cell distribution width as a novel prognostic marker in heart failure: data from the CHARM Program and the Duke Databank. *J Am Coll Cardiol*. 2007;50(1):40–47.
6. Horne BD, et al. Exceptional mortality prediction by risk scores from common laboratory tests. *Am J Med*. 2009;122(6):550–558.
7. Föhrhéc Z, Gombos T, Borgulya G, Pozsonyi Z, Prohászka Z, Jánoskúti L. Red cell distribution width in heart failure: prediction of clinical events and relationship with markers of ineffective erythropoiesis, inflammation, renal function, and nutritional state. *Am Heart J*. 2009;158(4):659–666.
8. Ani C, Ovbiagele B. Elevated red blood cell distribution width predicts mortality in persons with known stroke. *J Neurol Sci*. 2009;277(1-2):103–108.
9. Patel KV, Ferrucci L, Ershler WB, Longo DL, Guralnik JM. Red blood cell distribution width and the risk of death in middle-aged and older adults. *Arch Intern Med*. 2009;169(5):515–523.
10. Horne BD, et al. The Intermountain Risk Score (including the red cell distribution width) predicts heart failure and other morbidity endpoints. *Eur J Heart Fail*. 2010;12(11):1203–1213.
11. Dabbah S, Hammerman H, Markiewicz W, Aronson D. Relation between red cell distribution width and clinical outcomes after acute myocardial infarction. *Am J Cardiol*. 2010;105(3):312–317.
12. Zalawadiya SK, Veeranna V, Niraj A, Pradhan J, Afonso L. Red cell distribution width and risk of coronary heart disease events. *Am J Cardiol*. 2010;106(7):988–993.
13. Azab B, et al. Usefulness of red cell distribution width in predicting all-cause long-term mortality after non-ST-elevation myocardial infarction. *Cardiology*. 2011;119(2):72–80.
14. Ye Z, Smith C, Kullo IJ. Usefulness of red cell distribution width to predict mortality in patients with peripheral artery disease. *Am J Cardiol*. 2011;107(8):1241–1245.
15. Braun E, Domany E, Kenig Y, Mazor Y, Makhoul BF, Azzam ZS. Elevated red cell distribution width predicts poor outcome in young patients with community acquired pneumonia. *Crit Care*. 2011;15(4):R194.
16. Rhodes CJ, Wharton J, Howard LS, Gibbs JS, Wilkins MR. Red cell distribution width outperforms other potential circulating biomarkers in predicting survival in idiopathic pulmonary arterial hypertension. *Heart*. 2011;97(13):1054–1060.
17. Bazick HS, Chang D, Mahadevappa K, Gibbons FK, Christopher KB. Red cell distribution width and all-cause mortality in critically ill patients. *Crit Care Med*. 2011;39(8):1913–1921.
18. Lappé JM, et al. Red cell distribution width, C-reactive protein, the complete blood count, and mortality in patients with coronary disease and a normal comparison population. *Clin Chim Acta*. 2011;412(23-24):2094–2099.
19. Oh HJ, et al. Red blood cell distribution width is an independent predictor of mortality in acute kidney injury patients treated with continuous renal replacement therapy. *Nephrol Dial Transplant*. 2012;27(2):589–594.
20. Majercik S, Fox J, Knight S, Horne BD. Red cell distribution width is predictive of mortality in trauma patients. *J Trauma Acute Care Surg*. 2013;74(4):1021–1026.
21. Horne BD, et al. Association of the dispersion in red blood cell volume with mortality. *Eur J Clin Invest*. 2015;45(6):541–549.
22. Horne BD, et al. Pulmonary-specific intermountain risk score predicts all-cause mortality via spirometry, the red cell distribution width, and other laboratory parameters. *Respir Care*. 2015;60(9):1314–1323.
23. Yčas JW, Horrow JC, Horne BD. Persistent increase in red cell size distribution width after acute diseases: A biomarker of hypoxemia? *Clin Chim Acta*. 2015;448:107–117.
24. Muhlestein JB, et al. Both initial red cell distribution width (RDW) and change in RDW during heart failure hospitalization are associated with length of hospital stay and 30-day outcomes. *Int J Lab Hematol*. 2016;38(3):328–337.
25. Patel HH, Patel HR, Higgins JM. Modulation of red blood cell population dynamics is a fundamental homeostatic response to disease. *Am J Hematol*. 2015;90(5):422–428.
26. Koma Y, et al. Increased red blood cell distribution width associates with cancer stage and prognosis in patients with lung cancer. *PLoS One*. 2013;8(11):e80240.
27. Fatemi O, et al. Red cell distribution width as a bleeding predictor after percutaneous coronary intervention. *Am Heart J*. 2013;166(1):104–109.
28. Guimarães PO, et al. Association of standard clinical and laboratory variables with red blood cell distribution width. *Am Heart J*. 2016;174:22–28.
29. Malandrino N, Wu WC, Taveira TH, Whitlatch HB, Smith RJ. Association between red blood cell distribution width and macrovascular and microvascular complications in diabetes. *Diabetologia*. 2012;55(1):226–235.
30. Loprinzi PD, Hall ME. Physical activity and dietary behavior with red blood cell distribution width. *Physiol Behav*. 2015;149:35–38.
31. Ross R. Atherosclerosis—an inflammatory disease. *N Engl J Med*. 1999;340(2):115–126.
32. Anker SD, von Haehling S. Inflammatory mediators in chronic heart failure: an overview. *Heart*. 2004;90(4):464–470.
33. Semalessza GL, Wang GL. A nuclear factor induced by hypoxia via de novo protein synthesis binds to the human erythropoietin gene enhancer at a site required for transcriptional activation. *Mol Cell Biol*. 1992;12(12):5447–5454.
34. Maxwell PH, Pugh CW, Ratcliffe PJ. Inducible operation of the erythropoietin 3' enhancer in multiple cell lines: evidence for a widespread oxygen-sensing mechanism. *Proc Natl Acad Sci USA*. 1993;90(6):2423–2427.
35. Brecher G, Stohlman F. Reticulocyte size and erythropoietic stimulation. *Proc Soc Exp Biol Med*. 1961;107:887–891.
36. Van Dilla MA, Spalding JF. Erythrocyte volume distribution during recovery from bone marrow arrest. *Nature*. 1967;213(5077):708–709.
37. Kario K, Matsuo T, Kodama K, Asada R. Clinical significance of red blood cell distribution width in the elderly: a potential indicator of bone marrow stimulation by erythropoietin. *Clin Lab Haematol*. 1993;15:185–193.