Supplementary Methods and Results for

Modulation of Red Blood Cell Population Dynamics is a Fundamental Homeostatic Response to Disease

Contents

Supplementary Methods

Reticulocyte and Complete Blood Counts

We measured CBCs and reticulocyte counts for more than 60,000 randomly selected blood samples. Blood samples were randomly selected from among those sent to the Massachusetts General Hospital (MGH) Clinical Hematology Laboratory. We measured an average of 150 CBCs and reticulocyte counts per day for 18 months. All reticulocyte and complete blood counts (CBCs) used in modeling were measured on an Abbott Cell-DYN Sapphire 4000 automated hematology analyzer (Abbott Hematology, Santa Clara, California). For each CBC and reticulocyte count, this instrument measures the volume and hemoglobin concentration of approximately 50,000 individual RBCs. Both volume and hemoglobin concentration are measured for each individual RBC, yielding the joint volume-hemoglobin distribution for the entire population of 50,000 RBCs. RBC population modeling has been validated on the Abbott Cell-DYN Sapphire 4000 as well as Siemens Advia 120, 2120, and 2120i instruments.

All CBC parameters used in hypothesis testing were measured on current instruments of record in the MGH Clinical Hematology Laboratory. Instruments included the Siemens Advia 2120 and the Sysmex XE-5000. All reference ranges used in this study were identical to those in place at MGH at the time of each test.

Table S1. Normal Ranges for Complete Blood Counts

TEST (units)	POPULATION	NORMAL RANGE
RBC (millions/mm^3)	ADULT MALES >18 YEARS	4.50-5.90
	ADULT FEMALES >18 YRS	4.00-5.20
	MALE 12-18 YEARS	4.50-5.30
	FEMALE 12-18 YEARS	4.10-5.10
HGB (g/dl)	ADULT MALES >18 YEARS	13.5-17.5
	ADULT FEMALES >18 YRS	12.0-16.0
	MALE 12-18 YEARS	13.0-16.0
	FEMALE 12-18 YEARS	12.0-16.0
HCT (%)	ADULT MALES >18 YEARS	41.0-53.0
	ADULT FEMALES >18 YRS	36.0-46.0
	MALE 12-18 YEARS	37.0-49.0
	FEMALE 12-18 YEARS	36.0-46.0
MCV (fl)	ALL ADULTS >18 YEARS	80-100
	MALE 12-18 YEARS	78-98
	FEMALE 12-18 YEARS	78-102
MCH (pg/RBC)	ADULT >18 YEARS	26.0-34.0
	12-18 YEARS	25.0-35.0
MCHC (g/dl)	ALL >2 YEARS	31.0-37.0
RDW (%)	ALL	11.5-14.5

Other Clinical Laboratory Testing

CBC and reticulocyte counts for each blood sample were modeled as described below. Blood samples with appropriate CBC parameters and clearance thresholds (v_c) were then selected for further measurement of ferritin and hemoglobin A1c. Ferritin levels were measured on a Roche

COBAS instrument (Roche Diagnostics, Indianapolis, Indiana). Glycated hemoglobin was measured on a BIO-RAD Variant II Turbo instrument (BIO-RAD, Hercules, California).

Table S2. Normal Ranges for Clinical Laboratory Tests

TEST (units)	POPULATION	NORMAL RANGE
Glucose (mg/dl)	ALL	70 - 110
Hemoglobin A1c (%)	ALL	3.8 - 6.4
Ferritin (ng/ml)	ADULT MALE	30-300
	ADULT FEMALE	10-200

Modeling RBC Population Dynamics

We used routine complete blood and reticulocyte count data and a mathematical model of *in vivo* RBC population dynamics to infer single-RBC rates of maturation and clearance for individual patients [1]. See Figures 1, S-1, and S-2. RBCs become smaller and lose hemoglobin as they age, with an initial fast phase of reduction followed by a slow phase [2-5]. The rates of the initial fast phase are quantified for each individual by model parameters β_v and β_h. The model quantifies the slower rate with the parameter α. Rates of volume and hemoglobin reduction vary for a single RBC over time and from one RBC to the next in the population. The model enables estimation of the magnitude of variation in rates of both hemoglobin (D_h) and volume reduction (D_v). The molecular trigger and mechanism for RBC clearance are not fully understood [6], but common proposed mechanisms and empirical measurements of RBC populations such as those in Figure 1 (red contours) show that the probability of clearance is correlated with an RBC's volume and hemoglobin, and the true clearance function can be approximated as a threshold function of RBC volume and hemoglobin. As shown in Figure S-2, the probability of clearance for a particular RBC increases rapidly to 100% as the cell nears the clearance threshold, a line perpendicular to and

intersecting the MCHC (mean corpuscular hemoglobin concentration) line at a volume equal to v_c. An RBC with a hemoglobin concentration equal to the MCHC will be most likely to be recycled when its volume reaches v_c. RBCs with higher hemoglobin concentrations will circulate until reaching slightly lower volumes, and RBCs with lower hemoglobin concentrations will be cleared at slightly higher volumes. Details of model derivation, validation, and parameter estimation have previously been published [1].

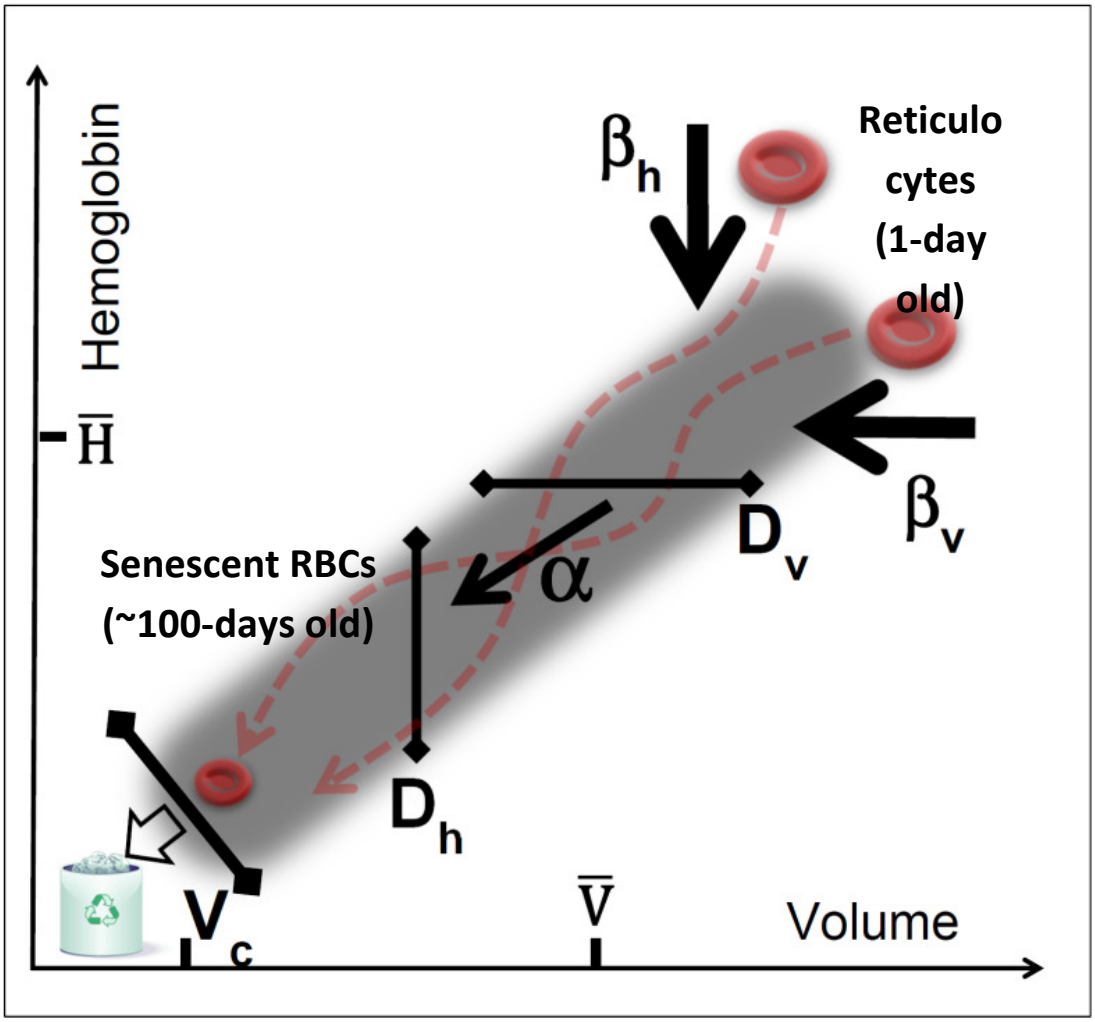

Figure S-1. Schematic of RBC maturation and clearance. Reticulocytes (top right) enter the circulation from the bone marrow. They circulate for ~100-120 days, losing ~30% of their volume and ~20% of their hemoglobin. There is an initial fast phase of reduction in both volume and hemoglobin, with rates quantified by β_v and β_h, followed by a slow phase, quantified by α. Rates of volume and hemoglobin reduction vary from cell to cell and for the same cell over time. D_v and D_h quantify the magnitude of this variation. After about 100 days in circulation, RBCs reach a clearance threshold (v_c) and are cleared from the circulation. The clearance threshold is a function of a cell's volume and hemoglobin as shown in Figure S-2.

The molecular and biophysical mechanisms for these maturation and clearance processes are unclear and likely involve multiple types of cells operating at distributed locations *in vivo* and consequently present significant experimental challenges for direct characterization. Because the probability of clearance for a particular cell correlates so strongly with its volume and hemoglobin,

we can use those quantities to estimate an RBC's probability of clearance even though we do not expect that either volume or hemoglobin *per se* is the actual trigger for clearance.

A mathematical model of RBC population dynamics [1] enables us to estimate typical rates for these *in vivo* maturation processes using measurements of single-RBC volume and hemoglobin concentration. The model (Equation 1) describes the volume and hemoglobin dynamics of a typical RBC ($\frac{dv}{dt}, \frac{dh}{dt}$) as a function ($\boldsymbol{f}$) of its current volume and hemoglobin and accounting for fluctuations in these rates of volume and hemoglobin loss over time ($\boldsymbol{\zeta}$). The dynamics of volume and hemoglobin in the full RBC population ($\frac{dP}{dt}$) include production of reticulocytes (b) and clearance of senescent cells (c).

(1)

$$\begin{bmatrix} dv \\ dh \end{bmatrix} = \boldsymbol{f}dt + d\boldsymbol{\xi}(t)$$

$$\boldsymbol{f} = \begin{Bmatrix} -\alpha \cdot e^{\beta_v(v-h)} \\ -\alpha \cdot e^{\beta_h(h-v)} \end{Bmatrix}$$

$$\boldsymbol{\xi}(t) = \begin{Bmatrix} \sqrt{2D_v} \cdot B^1(t) \\ \sqrt{2D_h} \cdot B^2(t) \end{Bmatrix}$$

$B^i(t)$ *are independent Brownian motion.*

$$\frac{dP(v,h,t)}{dt} = -\nabla(P\boldsymbol{f}) + \nabla(\boldsymbol{D}\nabla \mathrm{P}) + \mathrm{b}(v,h) - \mathrm{c}(v,h) \cdot \mathrm{P}$$

$b(v,h)$ *is the scaled empirical reticulocyte distribution.*

$$c(v,h) = \frac{1}{1 + e^{\Delta(v,h)}}$$

$$\Delta(v,h) = 100\,\frac{\cos(\theta)\sqrt{v^2+h^2} - v_c\sqrt{\bar{v}^2+\bar{h}^2}}{v_c\sqrt{\bar{v}^2+\bar{h}^2}}$$

$$\theta = tan^{-1}\left(\frac{\bar{h}}{\bar{v}}\right) - tan^{-1}\left(\frac{h}{v}\right)$$

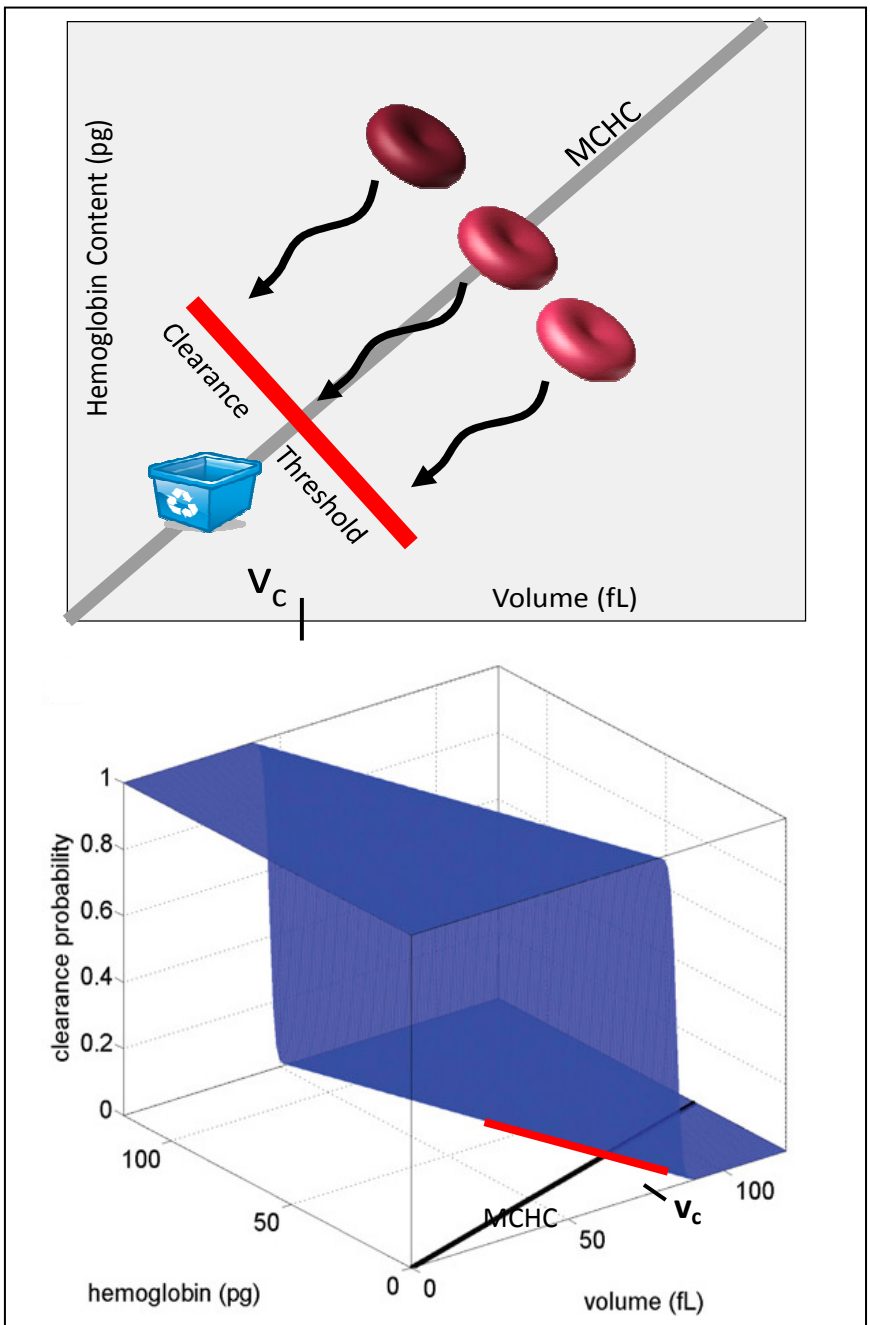

Figure S-2. Approximating RBC clearance as a threshold function of RBC hemoglobin and volume. The top panel shows a schematic of three RBCs maturing toward the clearance threshold (red line), moving roughly parallel to the line representing the population mean hemoglobin concentration (MCHC) in the volume-hemoglobin plane. v_c represents the volume at which an RBC with a hemoglobin concentration equal to MCHC is likely to be cleared. The bottom panel shows probability of clearance as a function of RBC volume and hemoglobin. As an RBC nears the clearance threshold (red line), its probability of clearance increases rapidly from 0% to 100% as shown. Delayed clearance is modeled as a shift of the blue surface toward the origin.

Parameter estimation is robust to the mild erythropoietic perturbations expected even for the wide range of conditions show to be associated with elevated RDW. Estimated parameters are less accurate in rapidly changing or severe conditions such as acute blood loss or significant hemolysis. In more mild cases of conditions associated with hemolysis like thalassemia trait or iron deficiency both from occult GI bleeding and nutritional deficiency, the model-based estimates of parameters are stable and show slight extensions of RBC lifespan, consistent with prior reports [7-10].

Random Selection of Complete Blood Counts ("Correlation Cohort")

To assess the correlation between the possible causes of elevated and RDW across a general inpatient and outpatient hospital patient population, we randomly selected 450 blood samples from the MGH Clinical Laboratory over the course of 4 weeks. At different times each day, we selected the most recently analyzed rack of tubes for the study. This cohort was 45% male, with a median age of 59 years and an interquartile range of 26 years. There was no overlap between this patient cohort and any of the others. The specific correlation coefficients calculated may not apply to other hospitals or clinics where the average level of patient severity and complexity differs.

Identification of Healthy Patients ("Healthy Patient Cohort")

We defined healthy patients as those who met the following criteria:

(1) The patient had at least three CBCs spanning at least 3 years.

(2) The patient had no recorded abnormal CBC index. See Table 1 for normal ranges.

It is possible that individuals with significant morbidity due to chronic illness may still meet these criteria. A manual review of randomly selected medical records suggests that the prevalence of chronic illness is below 5% and likely well below 2%. We use this healthy population to estimate a lower bound on the typical variation of clinical parameters in a truly healthy population. Rare

biological outliers will therefore have little impact. Also, we only compare measurements in this cohort against other measurements made in this cohort and therefore provide some control for the potential bias caused by the inadvertent inclusion of a small number of individuals with chronic illness. This cohort was 43% male, with a median age of 53 years and an interquartile range of 23 years. There was no overlap between this patient cohort and any of the others.

Case Control Study of RDW and v_c Association with All-Cause Mortality ("Mortality Rate Cohort")

We measured CBCs and reticulocyte counts for patients randomly selected from among those who had complete blood counts ordered as part of their medical care at MGH between June, 2012, and December 2013. We identified 900 of those patients ("cases") who had died by June, 2014, according to the Social Security Death Master File. For each case, we identified another patient ("controls") who

(1) had the same gender as the case

(2) had a date of birth within one year of the case

(3) had a CBC and reticulocyte measurement within 30 days of the case

(4) was not reported as deceased in the Social Security Death Master File.

We control for some potential biases in patient selection by selecting cases and controls from this same cohort and by matching them for age, gender, and timing of the CBC used for the comparison. There was no overlap between this patient cohort and any of the others. The patients in this cohort were 50% female and had a median age of 71 years with an interquartile range of 26 years.

Prospective Study of Dependence of Glycated Hemoglobin Fraction on v_c ("A1c Cohort")

We assessed the relationship between v_c and glycated hemoglobin fraction for non-diabetic individuals. "Non-diabetic" individuals were those who met the following criteria:

(1) Normal non-fasting glucose at the time of the complete blood count

(2) No diagnosis of diabetes in the individual's medical record

(3) No recorded abnormally high glucose measurements within the two years prior to the CBC

(4) No hemoglobin A1c measurement outside the reference range

There was no overlap between this patient cohort and any of the others. The patients in this cohort were 47% male and had a median age of 52 years with an interquartile range of 30 years. See Table 2 for normal ranges.

Case Control Study of v_c-based Prediction of All-Cause Anemia ("Anemia Prediction Cohort")

We determined the odds ratio for v_c and subsequent all-cause anemia. We were interested in assessing the predictive efficiency of v_c in identifying previously hematologically stable patients with a latent condition that might lead to future anemia.

Case patients were individuals presenting with mild anemia who had been hematologically stable within the prior 3 months. Mild anemia was defined as an HCT no more than 10% below the lower limit of the HCT reference range, or 33% <= HCT < 36% for females, and 37% <= HCT < 41% for males. Hematologic stability was defined by the following criteria:

(1) No anemic CBC in the prior 6 months

(2) No CBC within the prior 30 days

(3) No transfusion or surgery in the prior 6 months

Control patients were hematologically stable individuals with exactly two completely normal CBCs separated by between 10 and 14 months, where the second HCT was no less than the first HCT. There was no overlap between this patient cohort and any of the others. The patients in this cohort were 35% male and had a median age of 54 years with an interquartile range of 30 years. See Table 2 for normal ranges.

Prospective Study of Ferritin Levels ("Ferritin Cohort")

We measured ferritin levels in hematologically stable patients with completely normal CBCs and a range of v_c. We defined hematologically stable patients as those meeting the following criteria:

(1) No anemia in the previous 6 months

(2) No hospital visits in the prior 30 days

We define decreased iron stores as a ferritin in the bottom 5 percent of the normal range or below, which was equivalent to a ferritin < 20 ng/dl for females (normal range 10-200) and a ferritin < 44 ng/dl for males (normal range 30-300). There was no overlap between this patient cohort and any of the others. The patients in this cohort were 39% male and had a median age of 59 years with an interquartile range of 26 years.

Supplementary Results

HCT variability in a healthy population.

Figure S-3. Variation in hematocrit (%) in a healthy cohort of 600 individuals.

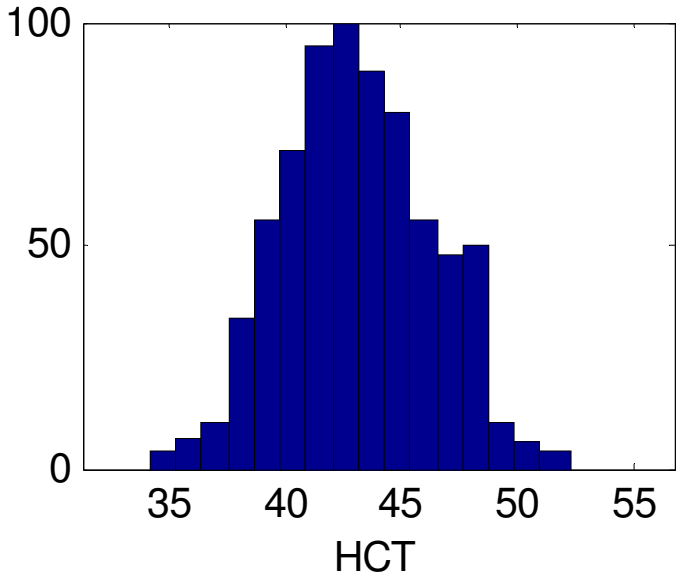

Figure S-3. Variation in hematocrit (%) in a healthy cohort of 600 individuals. The coefficient of variation for this distribution is 7.5% as compared to the much more tightly regulated v_c shown in Figure 1 with a CV of 1.1%. v_c is therefore likely to be a much more specific biomarker for the detection of perturbations to the hematologic system. See "Healthy Patient Cohort" in Methods for more details

rRDW variability in a healthy population.

Figure S-4. Reticulocyte volume variation (rRDW) for 600 healthy individuals.

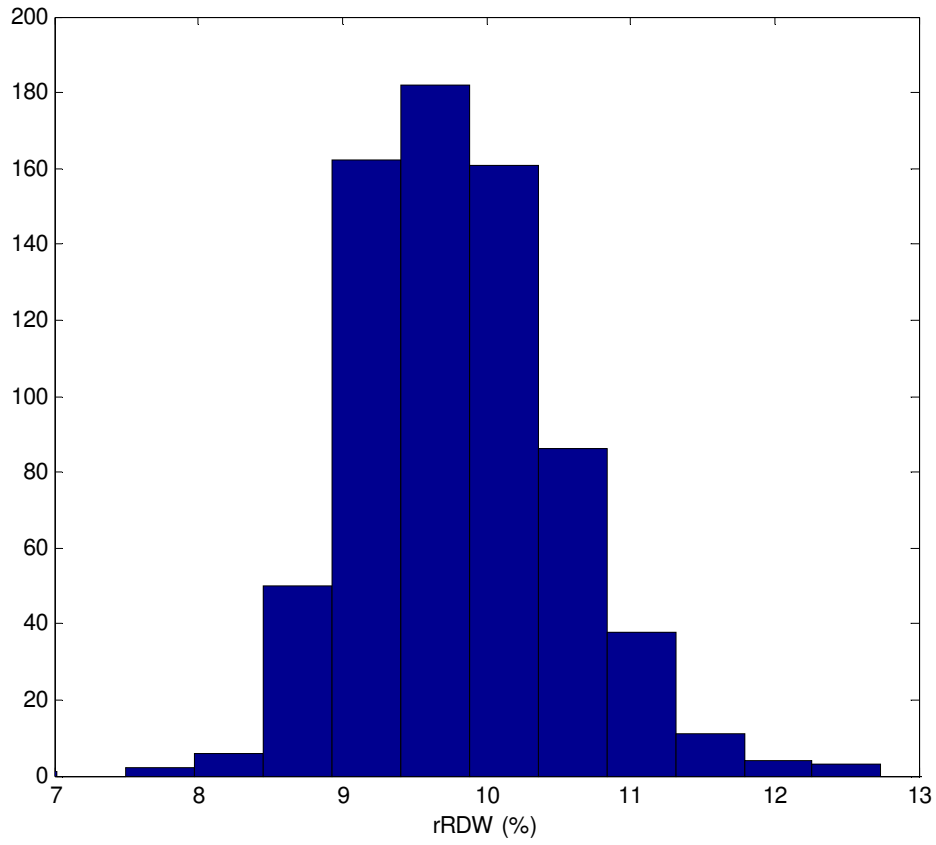

Figure S- 4. Reticulocyte volume variation (rRDW) for 600 healthy individuals. The coefficient of variation for this distribution is 8.2%. See Figure 1 and main text for comparison of this CV with the other three possible causes of an increased RDW. See "Healthy Patient Cohort" in Methods for more details.

RBC Population Dynamics in a Healthy Population

Figure S-5. Estimated Parameters for RBC Population Dynamics in a Healthy Population

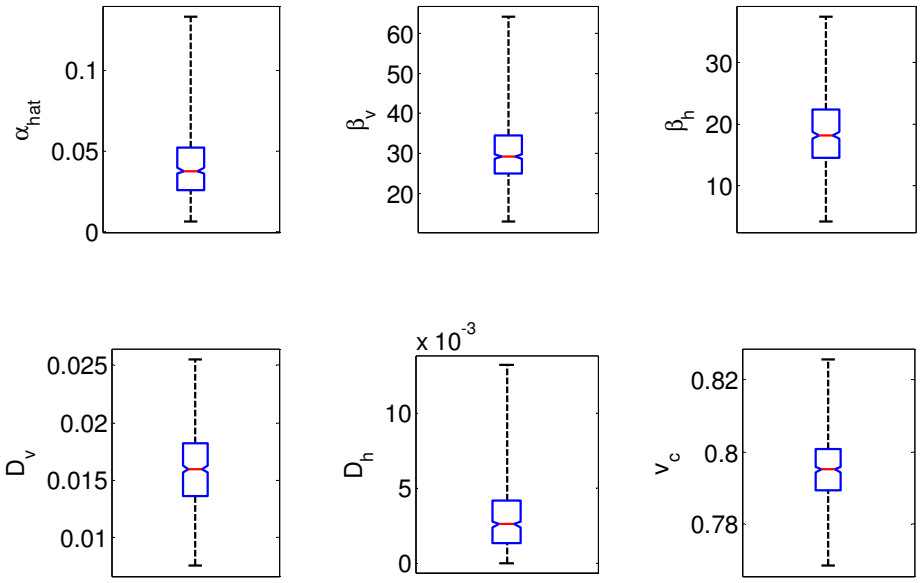

Figure S-5. Estimated Parameters for RBC Population Dynamics in a Healthy Population (n = 600). The panels show boxplots of parameters for the model of RBC population dynamics. See Figures 1, S-1, and S-2, and Materials and Methods "Modeling RBC Population Dynamics" and "Healthy Patient Cohort" for more detail.

Correlation Between RDW and Potential Causes of its Elevation.

Figure S-6. Correlation of RDW with possible cause for its elevation

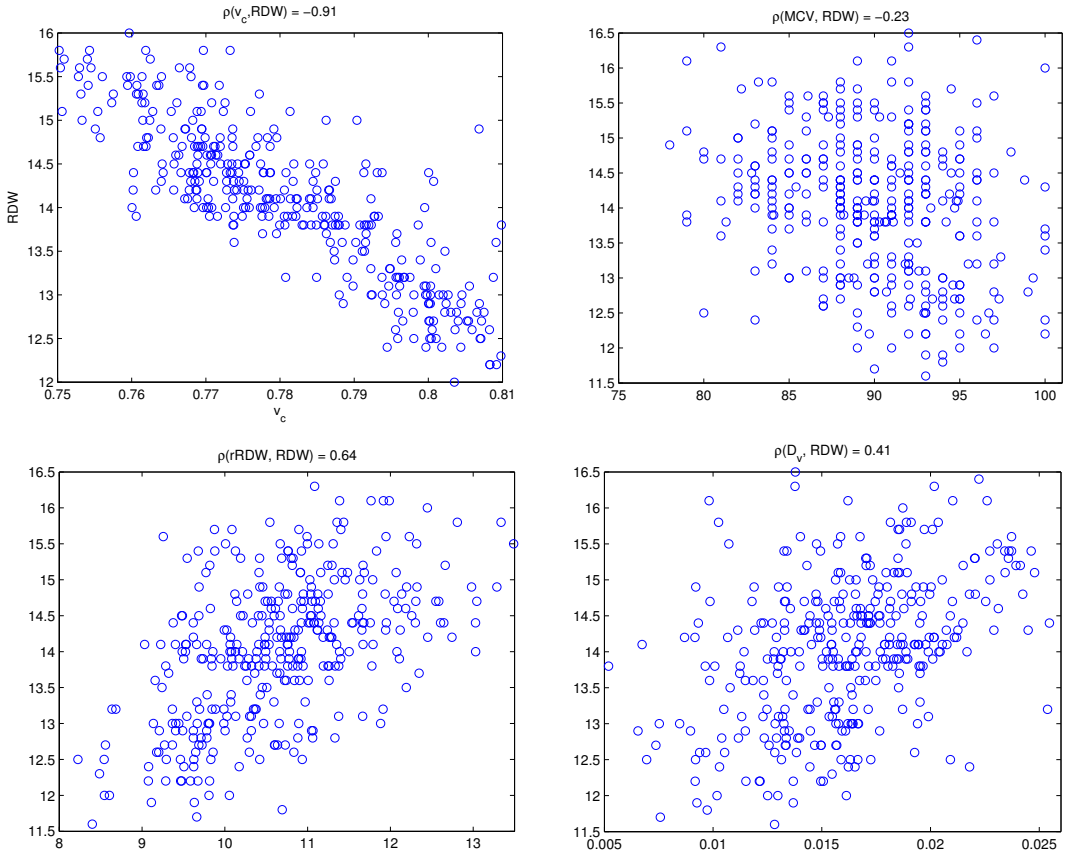

Figure S-6. Correlation of RDW with possible cause for its elevation. Panels show the correlation coefficient for RDW and each of the possible causes for its elevation: clearance threshold (v_c), mean cell volume (MCV), reticulocyte volume variation (rRDW), and magnitude of variation in the estimated rate of RBC volume reduction (D_v). v_c is much more strongly correlated with RDW than any of the other three possible causes, and RBC clearance delay as approximated by a decreasing v_c is therefore the most likely cause for the elevations in RDW found to be associated with a wide range of morbidity and mortality. Blood samples were taken from a randomly selected population of 400 patients receiving complete blood counts at Massachusetts General Hospital. See "Correlation Cohort" in Methods for more detail.

Odds Ratios of Future Anemia for Thresholds of CBC parameters and v_c

We assessed the predictive value of v_c and the traditional CBC indices for future anemia. See Materials and Methods for "Case Control Study of v_c-based Prediction of All-Cause Anemia." For each predictor, we let the diagnostic threshold vary across the measured range, as shown in Table S-1. Confidence intervals for odds ratios were calculated by assuming a log-normal distribution. Odds ratios statistically different from 1.0 are shown shaded. v_c has a statistically significant odds ratio across almost its entire measured range. HCT, HGB, MCH, MCHC, and MCV have no statistically significant odds ratios for any threshold value. RDW has thresholds with significant odds ratios, but the odds ratio (2.0) for the optimal threshold is less than ¼ that for v_c (8.5).

Table S3. Odds Ratios of Future Anemia for Thresholds of CBC parameters and v_c
(Odds ratios statistically different from 1.0 are shaded.)

Predictor	Threshold	Odds Ratio for Future Anemia	95% Confidence Interval	
v_c	**0.7750**	**8.5**	**1.1**	**63.4**
	0.7775	**6.2**	**1.5**	**26.1**
	0.7800	**5.6**	**1.7**	**18.5**
	0.7825	**3.7**	**1.5**	**8.8**
	0.7850	**3.1**	**1.5**	**6.4**
	0.7875	**2.1**	**1.2**	**3.7**
	0.7900	**1.8**	**1.1**	**3.1**
	0.7925	**1.7**	**1.0**	**2.7**
	0.7950	1.5	0.9	2.3
RDW	12.25	0.9	0.3	2.7
	12.50	1.1	0.6	2.2
	12.75	1.3	0.8	2.3
	13.00	**1.9**	**1.2**	**3.0**
	13.25	**2.0**	**1.2**	**3.2**
	13.50	**1.9**	**1.1**	**3.3**
	13.75	**2.0**	**1.1**	**3.8**
	14.00	2.3	0.9	5.2
	14.25	2.1	0.7	6.1
HGB	12.5	0.8	0.3	2.1
	13.0	1.2	0.7	2.1
	13.5	0.9	0.6	1.4
	14.0	0.9	0.6	1.4
	14.5	0.7	0.4	1.1
	15.0	0.5	0.3	1.0
MCHC	32.5	1.2	0.5	3.0
	33.0	1.4	0.8	2.4
	33.5	1.1	0.7	1.7
	34.0	0.9	0.5	1.5
	34.5	1.0	0.6	1.9
	35.0	0.9	0.4	2.2
MCH	28.0	2.7	0.8	9.1
	29.0	1.2	0.6	2.3
	30.0	1.3	0.8	2.1
	31.0	1.2	0.8	2.0
	32.0	0.7	0.4	1.3
	33.0	0.6	0.2	1.5
MCV	80.0	0.3	0.0	5.3
	83.0	0.7	0.2	2.4
	86.0	1.4	0.6	3.1
	89.0	1.7	0.9	3.0
	92.0	1.1	0.7	1.8
	95.0	1.2	0.7	2.1
HCT	38.0	1.0	0.5	1.9
	40.0	0.9	0.5	1.4
	42.0	0.8	0.5	1.2
	44.0	0.8	0.5	1.4
	46.0	0.5	0.2	1.5
	48.0	0.6	0.1	5.2

Glycated Hemoglobin is Elevated in Non-Diabetics with Normal CBCs and Decreased Ferritin

Figure S-7. Decreased ferritin with normal CBC and low v_c is associated with longer RBC lifespan

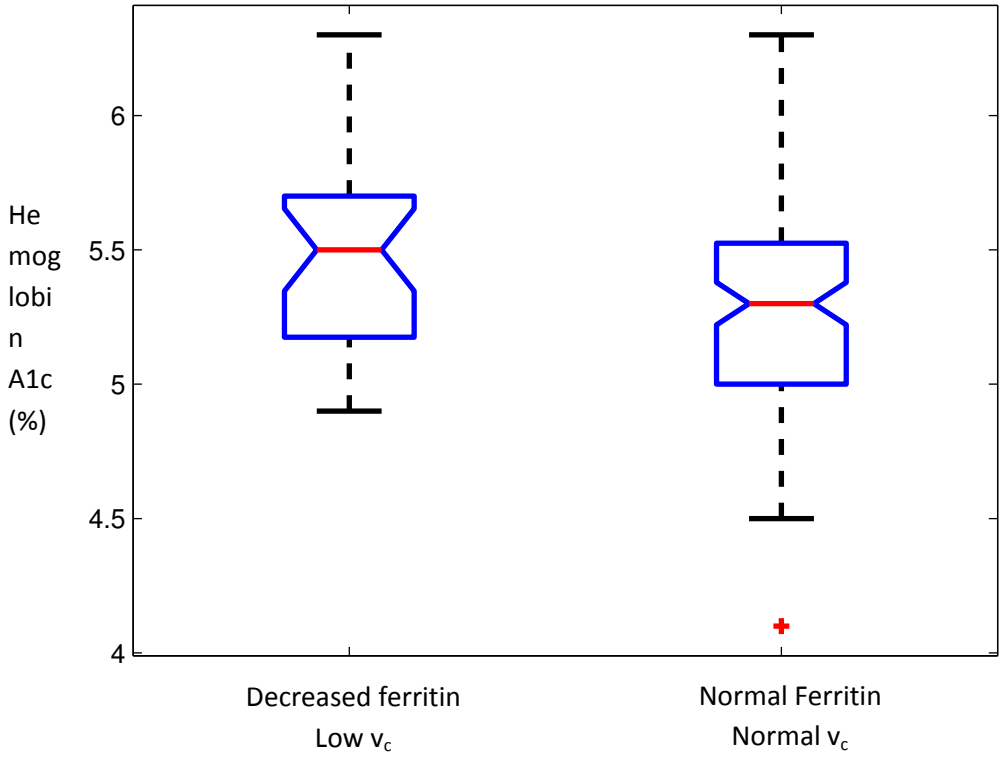

Figure S- 7. Decreased ferritin in patients with normal complete blood counts and low v_c is associated with longer RBC lifespan. Boxplots of glycated hemoglobin in non-diabetics with normal CBCs and either (1) ferritin below the bottom 5% of the normal range and a low v_c or (2) ferritin above the bottom 5% of the reference range and normal v_c. Group 1 had 29 individuals with median A1c of 5.5% and median ferritin of 17 ng/ml. Group 2 had 109 individuals with median A1c of 5.3% and median ferritin of 107 ng/ml. See "Ferritin Cohort" in Methods for more detail.

1. Higgins JM, Mahadevan L. Physiological and pathological population dynamics of circulating human red blood cells. Proc Natl Acad Sci U S A 2010;107:20587-20592.

2. Franco RS, Puchulu-Campanella ME, Barber LA, et al. Changes in the properties of normal human red blood cells during in vivo aging. Am J Hematol 2013;88:44-51.

3. Willekens FLA, Roerdinkholder-Stoelwinder B, Groenen-Dopp YAM, et al. Hemoglobin loss from erythrocytes in vivo results from spleen-facilitated vesiculation. Blood 2003;101:747-751.

4. Gifford SC, Derganc J, Shevkoplyas SS, et al. A detailed study of time-dependent changes in human red blood cells: from reticulocyte maturation to erythrocyte senescence. Br J Haematol 2006;135:395-404.

5. Willekens FLA, Werre JM, Groenen-Dop YAM, et al. Erythrocyte vesiculation: a self-protective mechanism? Br J Haematol 2008;141:549-556.

6. Franco RS. The measurement and importance of red cell survival. Am J Hematol 2009;84:109-114.

7. Saitta G, Catalano C, Artenisio AC, et al. GLYCOSYLATED HEMOGLOBIN IN CARRIERS OF BETA-THALASSEMIA TRAIT. British Medical Journal 1984;289:1382-1383.

8. Ellis G, Diamandis EP, Giesbrecht EE, et al. AN AUTOMATED HIGH-PRESSURE LIQUID-CHROMATOGRAPHIC ASSAY FOR HEMOGLOBIN-A1C. Clinical Chemistry 1984;30:1746-1752.

9. Coban E, Ozdogan M, Timuragaoglu A. Effect of iron deficiency anemia on the levels of hemoglobin A1c in nondiabetic patients. Acta Haematologica 2004;112:126-128.

10. Polage C, Little RR, Rohlfing CL, et al. Effects of beta thalassemia minor on results of six glycated hemoglobin methods. Clin Chim Acta 2004;350:123-128.