

Errors of measurement for blood volume parameters: a meta-analysis

Christopher J. Gore, Will G. Hopkins and Caroline M. Burge

J Appl Physiol 99:1745-1758, 2005. First published 23 June 2005; doi:10.1152/jappphysiol.00505.2005

You might find this additional info useful...

This article cites 142 articles, 60 of which can be accessed free at:

<http://jap.physiology.org/content/99/5/1745.full.html#ref-list-1>

This article has been cited by 8 other HighWire hosted articles, the first 5 are:

Carbon monoxide uptake kinetics of arterial, venous and capillary blood during CO rebreathing

Laura A. Garvican, Caroline M. Burge, Amanda J. Cox, Sally A. Clark, David T. Martin and Christopher J. Gore

Exp Physiol, December 1, 2010; 95 (12): 1156-1166.

[\[Abstract\]](#) [\[Full Text\]](#) [\[PDF\]](#)

Alterations of systemic and muscle iron metabolism in human subjects treated with low-dose recombinant erythropoietin

Paul Robach, Stefania Recalcati, Domenico Girelli, Cecilia Gelfi, Niels J. Aachmann-Andersen, Jonas J. Thomsen, Anne M. Norgaard, Alessandra Alberghini, Natascia Campostrini, Annalisa Castagna, Agnese Viganò, Paolo Santambrogio, Tibor Kempf, Kai C. Wollert, Stéphane Moutereau, Carsten Lundby and Gaetano Cairo

Blood, June 25, 2009; 113 (26): 6707-6715.

[\[Abstract\]](#) [\[Full Text\]](#) [\[PDF\]](#)

Stability of hemoglobin mass over 100 days in active men

Annette Eastwood, Will G. Hopkins, Pitre C. Bourdon, Robert T. Withers and Christopher J. Gore

J Appl Physiol, April 1, 2008; 104 (4): 982-985.

[\[Abstract\]](#) [\[Full Text\]](#) [\[PDF\]](#)

Greater uterine artery blood flow during pregnancy in multigenerational (Andean) than shorter-term (European) high-altitude residents

Megan J. Wilson, Miriam Lopez, Marco Vargas, Colleen Julian, Wilma Tellez, Armando Rodriguez, Abigail Bigham, J. Fernando Armaza, Susan Niermeyer, Mark Shriver, Enrique Vargas and Lorna G. Moore

Am J Physiol Regul Integr Comp Physiol, September 1, 2007; 293 (3): R1313-R1324.

[\[Abstract\]](#) [\[Full Text\]](#) [\[PDF\]](#)

Determinants of blood oxygenation during pregnancy in Andean and European residents of high altitude

Marco Vargas, Enrique Vargas, Colleen G. Julian, J. Fernando Armaza, Armando Rodriguez, Wilma Tellez, Susan Niermeyer, Megan Wilson, Esteban Parra, Mark Shriver and Lorna G. Moore

Am J Physiol Regul Integr Comp Physiol, September 1, 2007; 293 (3): R1303-R1312.

[\[Abstract\]](#) [\[Full Text\]](#) [\[PDF\]](#)

Updated information and services including high resolution figures, can be found at:

<http://jap.physiology.org/content/99/5/1745.full.html>

Additional material and information about *Journal of Applied Physiology* can be found at:

<http://www.the-aps.org/publications/jappl>

This information is current as of July 26, 2011.

Errors of measurement for blood volume parameters: a meta-analysis

Christopher J. Gore,¹ Will G. Hopkins,² and Caroline M. Burge³

¹Department of Physiology, Australian Institute of Sport, Belconnen, Australian Capital Territory, Australia;

²Health Science/Sport and Recreation, Auckland University of Technology, Auckland, New Zealand;

and ³Royal Brisbane and Women's Hospital, Brisbane, Queensland, Australia

Submitted 2 May 2005; accepted in final form 21 June 2005

Gore, Christopher J., Will G. Hopkins, and Caroline M. Burge. Errors of measurement for blood volume parameters: a meta-analysis. *J Appl Physiol* 99: 1745–1758, 2005. First published June 23, 2005; doi:10.1152/jappphysiol.00505.2005.—The volume of red blood cells (V_{RBC}) is used routinely in the diagnostic workup of polycythemia, in assessing the efficacy of erythropoietin administration, and to study factors affecting oxygen transport. However, errors of various methods of measurement of V_{RBC} and related parameters are not well characterized. We meta-analyzed 346 estimates of error of measurement of V_{RBC} for techniques based on Evans blue ($V_{RBC,Evans}$), ⁵¹chromium-labeled red blood cells ($V_{RBC,^{51}Cr}$), and carbon monoxide (CO) rebreathing ($V_{RBC,CO}$), as well as hemoglobin mass with the carbon-monoxide method ($M_{Hb,CO}$), in athletes and active and inactive subjects undergoing various experimental and control treatments lasting minutes to months. Subject characteristics and experimental treatments had little effect on error of measurement, but measures with the smallest error showed some increase in error with increasing time between trials. Adjusted to 1 day between trials and expressed as coefficients of variation, mean errors for $M_{Hb,CO}$ (2.2%; 90% confidence interval 1.4–3.5%) and $V_{RBC,^{51}Cr}$ (2.8%; 2.4–3.2%) were much less than those for $V_{RBC,Evans}$ (6.7%; 4.9–9.4%) and $V_{RBC,CO}$ (6.7%; 3.4–14%). Most of the error of $V_{RBC,Evans}$ was due to error in measurement of volume of plasma via Evans blue dye (6.0%; 4.5–7.8%), which is the basis of $V_{RBC,Evans}$. Most of the error in $V_{RBC,CO}$ was due to estimates from laboratories with a relatively large error in $M_{Hb,CO}$, the basis of $V_{RBC,CO}$. $V_{RBC,^{51}Cr}$ and $M_{Hb,CO}$ are the best measures for research on blood-related changes in oxygen transport. With care, $V_{RBC,Evans}$ is suitable for clinical applications of blood-volume measurement.

reliability; hemoglobin mass; volume of red blood cells; Evans blue dye; carbon monoxide

BLOOD VOLUME (V_{Blood}) and its component volumes of red blood cells (V_{RBC}) and plasma (V_{Plasma}) have been studied for more than 100 years (58). As oxygen transport is closely associated with circulating mass of hemoglobin (M_{Hb}) (13, 49, 65), this component of the V_{RBC} has also been studied extensively. Early investigators of V_{Blood} parameters were concerned mostly with establishing the methodology and providing normal values for body size, age, and sex (133). In the last 50 yr, the techniques have been applied clinically in the diagnostic workup of polycythemia rubra vera and anemia (31, 32, 105, 106), assessing response to erythropoietin administration (103), and in red cell survival studies (141). Additionally, V_{Blood} and related measures have featured, in studies of relationships between exercise and aging (24), changes in V_{Blood} , V_{Plasma} , and V_{RBC} with bed rest (90, 104, 108) and spaceflight

(1), and, among chronic exercisers (136), changes in cardiac function with training-induced hypervolemia (134) and body fluid redistribution during exercise (97), as well as in investigations of thermoregulatory stress (124), the contributions of V_{Plasma} and hemoglobin to oxygen transport and performance (28, 45, 53, 125), mechanisms of blood doping (27, 103), and adaptation to altitude (48, 59, 88).

The methods used to measure V_{Blood} values are all indirect and based on dilution of tracers injected into the circulation (31). The tracers are red blood cells labeled with radioactive chromium (⁵¹Cr) for measurement of V_{RBC} (51), albumin labeled with radioactive iodine (¹³¹I or ¹²⁵I) for measurement of V_{Plasma} (128), and the dye Evans blue, which delineates V_{Plasma} by staining plasma proteins (43). Measurement of M_{Hb} is also based on dilution of a tracer, inhaled carbon monoxide (CO), which binds to and changes the color of hemoglobin (58). The ⁵¹Cr method for estimating V_{RBC} ($V_{RBC,^{51}Cr}$) is regarded as the criterion method by the International Committee for Standardization in Haematology, on the “basis of reliability, reproducibility and ease of use in routine clinical use” (75). Iodine-labeled albumin is recommended for estimating V_{Plasma} (74, 75) in clinical settings, whereas researchers concerned about the health risks of radioactive iodine have used the Evans blue (12, 47, 90, 104, 108) or CO (25, 55, 99, 113, 151) methods extensively.

Application of a patient's V_{RBC} measurement to a clinical assessment requires an appreciation of population distribution of values (106) and of normal physiological within-person variation (31). Without supporting data, it is often assumed that the patient's V_{RBC} has been measured reliably; that is, if the test were repeated a few days later, a similar value would be obtained (32). However, meaningful interpretation of serial changes in blood parameters in myeloproliferative disease or experimentally induced erythropoiesis requires quantification of the errors of the measurement techniques (31). A reliable method to determine V_{RBC} or M_{Hb} will have small measurement error; moreover, reliability is a prerequisite of validity, the extent to which a test actually measures what it intends to measure. Measurement error is calculated as a coefficient of variation (CV) (109, 110) and is usually expressed as a percentage. The measurement error is also known as typical error (67) and includes random error (analytic error arising from using the method-specific apparatus and intraindividual biological variation) but not systematic error (bias). Measurement error can be estimated from studies that include interventions, where it will include the analytic error, the day-to-day biological variation, and the interindividual variation in response to

Address for reprint requests and other correspondence: C. J. Gore, Australian Institute of Sport, P.O. Box 176, Belconnen, ACT 2616, Australia (e-mail: chris.gore@ausport.gov.au).

The costs of publication of this article were defrayed in part by the payment of page charges. The article must therefore be hereby marked “advertisement” in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

the intervention. Measurement errors can arise in numerous ways for the different methods used to estimate V_{Blood} or M_{Hb} , as described in Table 1.

There has been no systematic review of errors in the various measures of V_{Blood} and related parameters, and it was apparent to us that the errors for V_{RBC} , V_{Plasma} , and M_{Hb} ranged widely,

between 1 and 10%. We have, therefore, performed a meta-analysis of the errors of measurement to characterize the contribution of sampling variation, differences between laboratories, effects of subject and study characteristics, and true differences between the errors of methods used to estimate V_{Blood} values and M_{Hb} .

Table 1. Comparison of sources of error for the common blood volume and hemoglobin mass methods

	Tracer			
	Evans blue	^{131}I , ^{125}I	^{51}Cr , $^{99\text{m}}\text{Tc}$	CO
Primary measurement	Plasma dye concentration, mg/ml (plasma)	Counts/ml (plasma)	Counts/ml (whole blood)	HbCO %
Primary circulatory space measured by tracer	Albumin space	Albumin space	V_{Blood}	$M_{\text{Hb,CO}}$
Administering the tracer	Reliability and accuracy of tracer dose administered			
Potential unquantifiable loss of tracer (different on a case-by-case basis. Usually remains unknown and contributes to overall method variability)	Proportion of dye unbound (if incubated) or proportion of dye directly injected becoming bound to extravascular albumin or other tissue	Proportion of dose as unbound counts	Proportion of dose as unbound counts (postincubation cell washing procedure adequacy)	Gas leak in rebreathing apparatus or at mouthpiece by volunteer
	Variable extravasation of the tracer	Variable extravasation of the tracer	Variable extravasation of the tracer	
Tracer mixing kinetics	2-compartment kinetics		1-compartment kinetics	
Estimating V_{D} of the tracer (or equivalent)	Plasma dye concentration at theoretical <i>time 0</i>	Count density at theoretical <i>time 0</i>	Δ Counts/ml (post – pre)	Δ HbCO% (post – pre)
Steps contributing variability to the V_{D} estimate	Assay reliability of serum dye concentration (single vs. dual-wavelength spectrophotometry) and reliability of %recovery if preextraction procedure used	Reliability and accuracy of pipetting		Accuracy of the CO-oximeter “span” calibration
	Construction of dye concentration (or counts/ml) vs. time curve. Determination of optimal number of time points and their interval OR Single specimens obtained pre- and ~10 min post-dye injection (52)	Scintillator detection sensitivity and error, background counts, reference material total counts		Reliability and precision of HbCO measurement (5 replicates or other criteria)
	Method selected to plot points (computer, hand) and interpolate curve (computer—spline function or other math function, eyeball)			
	Mathematical model selected for back-extrapolating plasma dye conc. or count density to time zero (including all assumptions about the early rapid mixing phase), if used.			
Primary method parameter	V_{Plasma}	V_{Plasma}	V_{Blood}	$M_{\text{Hb,CO}}$
Final derived parameter	V_{RBC}	V_{RBC}	V_{RBC}	$M_{\text{Hb,CO}}$
Additional parameter required	Hct	Hct	Hct	nil
Primary method equation	$V_{\text{Plasma}} = V_{\text{D}} = (\text{[dye]/ml dose/[dye]/ml plasma}) \times \text{dose volume}$	$V_{\text{Plasma}} = V_{\text{D}} = (\text{counts/ml dose/counts/ml plasma}) \times \text{dose volume}$	$V_{\text{Blood}} = (\text{counts/ml in dose/counts/ml in whole blood sample}) \times \text{dose volume}$	$M_{\text{Hb,CO}} = (\text{CO dose STPD}/1.34 \text{ ml CO/g Hb})/\Delta\% \text{HbCO}$
Secondary equations to derive blood measures	$V_{\text{Blood}} = V_{\text{Plasma}}/(1 - \text{Hct})$ or $V_{\text{RBC}} = V_{\text{Plasma}} \times (1 - \text{Hct})$ $V_{\text{RBC}} = V_{\text{Blood}} - V_{\text{Plasma}}$		$V_{\text{RBC}} = V_{\text{Blood}} \times \text{Hct}$	$V_{\text{Blood}} = M_{\text{Hb,CO}}/[\text{Hb}]$ $V_{\text{RBC}} = V_{\text{Blood}} \times \text{Hct}$ $V_{\text{Plasma}} = V_{\text{Blood}} - V_{\text{RBC}}$

^{125}I , ^{125}I iodine; ^{131}I , ^{131}I iodine; ^{51}Cr , ^{51}Cr chromium; $^{99\text{m}}\text{Tc}$, $^{99\text{m}}\text{Tc}$ technetium; CO, carbon monoxide; HbCO, carboxyhemoglobin; V_{Blood} , blood volume; $M_{\text{Hb,CO}}$, hemoglobin mass via CO; V_{D} , volume of distribution; Δ HbCO, the difference in carboxyhemoglobin levels before and after CO rebreathing; V_{Plasma} , plasma volume; V_{RBC} , volume of red blood cells; Hct, hematocrit; CO dose, volume (STPD) of pure CO added to a rebreathing circuit; STPD, standard temperature and pressure dry; 1.34 is the binding capacity of hemoglobin for CO (57); [Hb], hemoglobin concentration.

METHODS

Data Sources, Techniques, and Method Variations

The data used in the meta-analysis were obtained from original publications that reported values of individual subjects and from published studies whose authors provided us with their de-identified raw data (1, 2, 4–6, 10–12, 15–17, 21, 24, 25, 33, 35–37, 39–44, 47–50, 55, 63, 65, 77–80, 83, 86–91, 93, 94, 98–102, 104, 107, 108, 111, 113, 114, 117–120, 122, 132, 137–139, 141, 143, 144, 146, 148–151). Three unpublished data sets have also been included in the analysis, courtesy of the respective senior investigator responsible for data collection (Gore CJ, Slater GJ and Schmidt W, personal communications). The first two of these studies were approved by the Australian Institute of Sport Human Ethics Committee and the third was approved by the Human Ethics Committee of the Friedrich-Alexander-Universität Erlangen-Nürnberg, Germany.

The V_{Blood} techniques selected for the meta-analysis are those most commonly used to estimate V_{RBC} and M_{Hb} (Table 1). We were able to obtain acceptable amounts of data for each of the following: CO rebreathing for M_{Hb} ($M_{\text{Hb,CO}}$); CO rebreathing for V_{RBC} ($V_{\text{RBC,CO}}$); $V_{\text{RBC,51Cr}}$; and Evans blue dye for V_{Plasma} , V_{Blood} , and V_{RBC} ($V_{\text{Plasma,Evans}}$, $V_{\text{Blood,Evans}}$, and $V_{\text{RBC,Evans}}$). Some data sets included values for hematocrit (Hct) and/or hemoglobin concentration ([Hb]). Only one study (30) provided sufficient data for V_{Plasma} , estimated using ^{131}I -labeled albumin, and none provided data for ^{125}I .

Modification of the CO-rebreathing technique may affect the magnitude of measurement error for $M_{\text{Hb,CO}}$ and its related volumes. We have previously recommended that, to minimize likely sources of error, researchers should use relatively large doses of CO, a small rebreathing volume, and at least four replicate measures of percent carboxyhemoglobin (%HbCO) (15). Others have generally used smaller doses of CO, and/or a larger rebreathing volume, and/or duplicate or single measures of %HbCO. We took this opportunity to compare the error of measurement of the “Burge and Skinner” (15) method (4–6, 15, 48–50, 122) with that of others (10, 39–41, 55, 63, 65, 117, 119, 120).

Methodological variations of the Evans blue technique could affect the magnitude of measurement error for $V_{\text{Plasma,Evans}}$ and its related volumes. These variations are in terms of extraction (column chromatography) or not of Evans blue dye from plasma. There are also variants that do or do not back-extrapolate a multitime point disappearance curve. Some authors do not back-extrapolate but use the simple change in plasma dye concentration obtained from one pre- and one postdye injection specimen (with post being obtained 10–20 min afterwards) (11, 24, 47, 55). The use of just one postinjection sample is justified by the statement of Greenleaf et al. (52) that a single 10-min post-dye-injection specimen gives the same V_{Plasma} value as using the more rigorous back-extrapolation method, of which there are also several mathematical approaches (43) (Table 1). The three common method variations for $V_{\text{Plasma,Evans}}$ in the data that we acquired were as follows: extracted/not back-extrapolated (11, 12, 24, 44, 47, 55), not extracted/not back-extrapolated (2, 16, 21, 33, 36, 37, 42, 43, 50, 77, 78, 86–91, 93, 94, 101, 102, 104, 107, 108, 111, 118, 132, 143, 149, 150), and not extracted/back-extrapolated (17, 146).

Classically, [Hb] has been determined by pipetting blood into Drabkin’s reagent, to form cyanmethemoglobin that is measured spectrophotometrically at 540 nm (76). We distinguished between those studies that used the traditional, manual cyanmethemoglobin method to determine [Hb] and those that used fully automated analyzers.

There are no substantial variants of the CO-based method for estimation of $V_{\text{RBC,CO}}$. Several altitude studies from which we obtained data (55, 98, 113, 119, 151) calculated the circulatory volumes by combining M_{Hb} with [Hb] or Hct (Table 1: secondary equations for

CO); other studies (5, 40, 64, 117, 122, 127) reported only $M_{\text{Hb,CO}}$ (15) (Table 1, primary method equation for CO).

Calculation of Error of Measurement

Studies were included in the meta-analysis if there were at least two assays of V_{RBC} , V_{Blood} , V_{Plasma} , or M_{Hb} in at least three individuals receiving the same experimental treatment. The error of measurement for each blood-related measure was calculated for pairs of assays as follows (67): after log-transformation, the standard deviation of the differences was calculated, and the result was divided by $\sqrt{2}$ and back-transformed to a CV, which was recorded along with its degrees of freedom ($df = \text{sample size} - 1$). These calculations were performed in a spreadsheet designed for analysis of crossovers (68). Each estimate of error was inflated by a factor $1 + 1/(df)$ to correct for small-sample bias (56).

In studies in which more than two assays were performed per subject, one assay was chosen as the reference, and the error of measurement and change in the mean were calculated for each of the other assays paired against this reference. Where two or more baseline assays were obtained before an experimental treatment, the reference was the assay closest in time to the experimental treatment. Time between pairs of assays was recorded in days.

The values of individual subjects giving rise to an estimate of error of measurement were examined for outliers in a plot of the subjects’ change score between assays against the subjects’ value of the reference assay. The following outliers were excluded from the error of measurement calculation: a subject with a 40% decline in $V_{\text{Plasma,CO}}$ in Poulsen et al. (111), and a subject with $\sim 40\%$ changes in several Evans blue measures in Grover et al. (55). Individual values for the residuals and random effects in the meta-analyses (see below) were also displayed graphically for identification of other possible outliers (t values ~ 3 or more), but none was observed. These analyses, therefore, were performed with essentially no subjective filtering of data, and any apparently large errors of measurement obtained in our analysis are not attributable to the presence of outliers.

Coding of Predictor Variables

Experimental treatments. The treatment that occurred between each pair of assays was coded either as “altitude,” “none,” or “other.” Treatments were coded as altitude for assays separated by a period of real or simulated altitude of 600–7,000 m, for exposures of 0.5–24 h/day, for durations of 0.3–40 days. Altitude treatments that included any time spent at or near sea level before, during, and/or after the altitude exposure were coded as altitude. Treatments were coded as none for pairs of assays from reliability studies, pairs of assays during a baseline period in experimental studies, and pairs of assays in any control group receiving no intervention that would be expected to change V_{Blood} parameters. Treatments coded as other were as follows: training, ingestion of propranolol during altitude exposure, effects of menstrual cycle and/or oral contraceptive pill or placebo in combination with altitude, feeding before and after Evans blue injection, bed rest with head-down tilt and/or supine cycle ergometer training, acute exercise, living 12.8 m below the ocean surface, spaceflight, weight loss before rowing races, V_{Plasma} expansion, heat exposure, iron supplementation, and venesection.

Subject characteristics. Sex and fitness were included in some meta-analyses. Sex was coded as a covariate representing the proportion of men in the sample (examples: 0 for all women; 0.375 for 5 women and 3 men; 1 for all men). Two studies using the ^{51}Cr method (114, 144) and one study using the Evans blue method (143) did not state the sex of the subjects. For these studies, we assigned a sex covariate value of 0.6, which was the mean value of all of the studies. Fitness was coded as “inactive” (sedentary), “active” (but not a competitive athlete), and “athlete” (competitive). Hospital inpatients [2 studies (30, 143)] were coded as “inactive.” Four studies using the Evans blue method (24, 43, 98, 111), two using the ^{51}Cr method (114,

144), and three using $V_{RBC,CO}$ (25, 80, 98) did not indicate physical activity or fitness of the subjects and were coded as inactive.

Laboratories. The effect of individual laboratories (or research groups) on error of measurement was modeled as a random effect in the meta-analysis. If a laboratory changed a particular technique substantially, a new identifier was assigned to the subsequent estimates.

Meta-analysis

The details of our unique, novel meta-analytic approach and its interpretation are contained in the APPENDIX. In summary, we used a mixed-model meta-analysis in which the dependent variable was the log-transformed error of measurement, the fixed effects were the method of measurement and characteristics of the study, and the random effects were within- and between-study variation. An unbiased weighting factor for each estimate was derived from the estimate's df. Inferences about the substantiveness of true differences between two errors of measurement were made in accordance with extent of overlap of the confidence interval of their ratio with the thresholds for substantial ratios (0.9 and 1.1).

RESULTS

Raw Data

The numbers of estimates of error of measurement for each blood parameter were as follows: 69 $M_{Hb,CO}$, 32 $V_{Blood,Evans}$, 64 $V_{Plasma,Evans}$, 29 $V_{RBC,CO}$, 9 $V_{RBC,51Cr}$, 69 $V_{RBC,Evans}$, 83 [Hb], and 57 Hct. The mean df for each parameter ranged from 5 to 10. The number of estimates for each of the methods for measurement of $V_{Plasma,Evans}$ were as follows: 19 extracted/not back-extrapolated, 41 not extracted/back-extrapolated, 3 not extracted/not back-extrapolated, and 1 unclear. Thirty-six of the estimates of error for $M_{Hb,CO}$ used the method of Burge and Skinner, and 33 used the other method. Of the 83 estimates for [Hb], 52 used automated methods, and 31 used the cyanmethemoglobin method. The median time between measurements giving rise to the estimates of error was 22–30 days for $M_{Hb,CO}$, $V_{Blood,Evans}$, $V_{Plasma,Evans}$, $V_{RBC,CO}$, $V_{RBC,Evans}$, [Hb], and Hct, but only 1 day for $V_{RBC,51Cr}$. There were 159 estimates involving real or simulated altitude, 132 estimates involving other treatments, and 121 involving no treatment (from control groups, baseline pairs of measurements in treatment groups, or reliability studies). The median real or simulated altitude was 3,000 m (range 600–7,000 m), and median time spent at altitude was 12 days (0.3–40 days). The estimates were collected on 84 recreationally active, 165 athletic, and 161 inactive groups. For the entire data set, there was a greater proportion of men (60%) than women (40%).

Four Comparable Hematology Parameters for the Red Cell Compartment

The measurement errors for $M_{Hb,CO}$, $V_{RBC,CO}$, $V_{RBC,51Cr}$, and $V_{RBC,Evans}$ are displayed in Fig. 1 as a function of time between measurements. A 10-fold increase in time between measurements had a substantial effect on the error for $M_{Hb,CO}$ (by a factor of 1.5; 90% confidence limits 1.2–1.9), trivial small effects on the errors for $V_{RBC,51Cr}$ (by a factor of 1.1; 1.0–1.2) and $V_{RBC,Evans}$ (by a factor of 1.1; 0.9–1.2), and an unclear trivial effect on the error for $V_{RBC,CO}$ (by a factor of 1.0; 0.6–1.6). The mean error of measurement predicted for 1 day between measurements and averaged over treatments ranged from 2.2 to 6.7% for

the four methods (Table 2). $M_{Hb,CO}$ had substantially less error than $V_{RBC,51Cr}$ (ratio of errors of $M_{Hb,CO}$ to $V_{RBC,51Cr}$ = 0.8; 0.5–1.3), although the confidence limits indicate that the real difference was unclear. Errors for both methods were clearly less than (about one-third) those for $V_{RBC,CO}$ and $V_{RBC,Evans}$. Estimates of the typical between-laboratory variations in the errors of measurement for $M_{Hb,CO}$, $V_{RBC,CO}$, and $V_{RBC,Evans}$ had considerable uncertainty, but were substantial over the range of their confidence limits: observed values (and confidence limits) were $\times/\div 1.8$ (1.5–3.0), $\times/\div 1.5$ (1.3–4.2), and $\times/\div 1.6$ (1.3–2.8), respectively. Thus, for example, a new user of the $M_{Hb,CO}$ method could anticipate that, in their hands, the error of measurement could easily be consistently as high as 4.0% ($= 2.2 \times 1.8\%$) or as low as 1.2% ($= 2.2 \div 1.8\%$). The between-laboratory variation for $V_{RBC,51Cr}$ was not estimable, owing to a paucity of data.

The clearly smaller magnitude of measurement errors for $M_{Hb,CO}$ and $V_{RBC,51Cr}$ implies that these blood measures are better than $V_{RBC,CO}$ and $V_{RBC,Evans}$ for quantifying effects of subject characteristics and experimental treatment on error of measurement. Only $M_{Hb,CO}$ had enough estimates of error for such an analysis. The effect of a 10-fold increase in time between measurements on the error of measurement differed substantially among the three treatments (factors of 1.3, 1.1, and 1.5 for none, altitude, and other, respectively), but these differences were unclear, owing to large uncertainty in their ratios. The other effects on error for $M_{Hb,CO}$ were, therefore, estimated under the assumption that the effect of a 10-fold increase in time was the same for all three treatments; this effect was reasonably clear (a factor of 1.3; confidence limits 0.9–1.8), albeit somewhat less than that provided by the simpler model above. The observed error of $M_{Hb,CO}$ for women was slightly greater than that for men (ratio 1.1; 0.9–1.4), indicating that women might have substantially greater error than men but were very unlikely to have less error. Athletes were likely to have substantially greater error of measurement than active subjects (ratio 1.4; 1.0–1.9), but the greater error for athletes compared with inactive subjects was unclear, owing to a wide confidence interval for the comparison (ratio 1.5; 0.8–2.8). The active/inactive subjects comparison was also unclear (ratio 1.1; 0.6–2.0). The mean measurement error of the Burge and Skinner method of estimating $M_{Hb,CO}$ (1.7%; 1.0–2.8%) was less than one-half that of the other variant (3.9%; 2.4–6.4%); the difference was definitive (ratio of Burge and Skinner/other 0.4; 0.3–0.8). There was little observed effect of treatment on error (pairwise ratios for the three treatments all 1.0), but confidence limits for the ratios (all $\times/\div 1.2$) allowed for the possibility of small, true differences.

V_{Blood} Compartments with Evans Blue

The mean error of measurement for $V_{Blood,Evans}$, $V_{Plasma,Evans}$, and $V_{RBC,Evans}$ was ~5–7% (Table 3), and there were no clear differences between the ratios of their errors. However, the variation in error between laboratories for these measures was substantial: $\times/\div 1.6$ (1.4–2.0). A 10-fold increase in the time between measurements did not substantially increase the errors in $V_{Blood,Evans}$, $V_{Plasma,Evans}$, and $V_{RBC,Evans}$; the ratio for the pooled data of all three volumes was $\times/\div 1.0$ (0.9–1.1). There was little difference between errors for the pooled treatments (6.8%; 5.6–8.4%), altitude treatments (6.3%; 5.1–7.7%), and no treatment

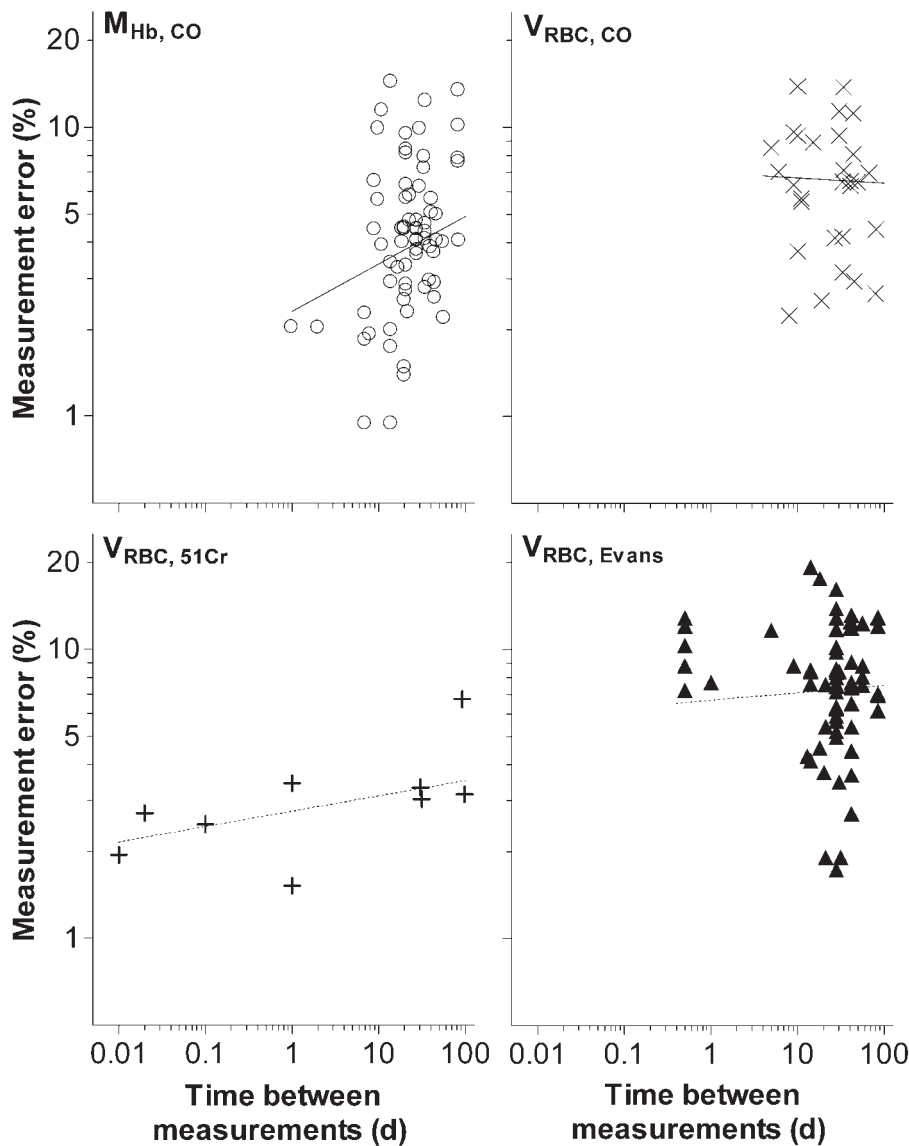


Fig. 1. Measurement error as a function of the time between measurements for 4 comparable hematological parameters: hemoglobin mass via CO re-breathing ($M_{Hb,CO}$; ○), and volume of red blood cells via CO re-breathing ($V_{RBC,CO}$; ×), 51 chromium ($V_{RBC,^{51}Cr}$; +), and Evans blue dye ($V_{RBC,Evans}$; ▲). Values of measurement error have been adjusted to remove sample-size bias. Lines of best fit are those derived from a meta-analysis, with time between measurements as the only fixed effect for each measure.

(6.0%; 4.9–7.3%), although confidence limits allowed for small differences between some treatment groups (range of confidence limits for ratios: 0.8–1.2).

The mean errors of measurement for $V_{Plasma,Evans}$, determined with or without extraction and/or back-extrapolation, differed substantially but not conclusively, owing to wide

confidence limits. The errors (and confidence limits) were 4.8% (2.3–10%) for not extracted/not back-extrapolated, 5.7% (3.4–9.6%) for extracted/not back-extrapolated, and 6.7% (4.5–10%) for not extracted/back-extrapolated. The estimate of the typical between-laboratory variation in the errors of measurement was substantial: $\times/\div 1.6$ (1.4–2.9).

Table 2. Measurement error of four comparable hematological parameters

Measure	CV for Error, %	Ratio of CV Between Measures			
		$M_{Hb,CO}$	$V_{RBC,^{51}Cr}$	$V_{RBC,CO}$	$V_{RBC,Evans}$
$M_{Hb,CO}$	2.2 (1.4–3.5)	1	0.8 (0.5–1.3)	0.3 (0.2–0.8)	0.3 (0.2–0.6)
$V_{RBC,^{51}Cr}$	2.8 (2.4–3.2)	1.2 (0.8–2.0)	1	0.4 (0.2–0.8)	0.4 (0.3–0.6)
$V_{RBC,CO}$	6.7 (3.4–14)	3.0 (1.3–6.9)	2.4 (1.2–4.9)	1	1.0 (0.5–2.1)
$V_{RBC,Evans}$	6.7 (4.9–9.4)	3.0 (1.8–5.3)	2.4 (1.7–3.5)	1.0 (0.5–2.1)	1

Data are means for both levels of treatment predicted for 1 day between measurements. Numbers in parentheses are 90% confidence limits. Errors [expressed as coefficients of variation (CV)] and their ratios were derived from a meta-analytic model in which the covariates were the parameter [4 levels: $M_{Hb,CO}$, V_{RBC} via ^{51}Cr ($V_{RBC,^{51}Cr}$), V_{RBC} via CO ($V_{RBC,CO}$), V_{RBC} via Evans blue dye ($V_{RBC,Evans}$)], its interaction with time between measurements (logarithmic metric), and its interaction with treatment between measurements (2 levels: none, other).

Table 3. Measurement error of the volume of blood, plasma, and red blood cells using Evans blue dye

Measure	CV for Error, %	Ratio of CV Between Measures		
		V _{Blood,Evans}	V _{Plasma,Evans}	V _{RBC,Evans}
V _{Blood,Evans}	5.4 (3.6–8.3)	1	0.9 (0.6–1.5)	0.8 (0.5–1.3)
V _{Plasma,Evans}	6.0 (4.5–7.8)	1.1 (0.7–1.8)	1	0.9 (0.6–1.3)
V _{RBC,Evans}	6.7 (5.0–9.0)	1.2 (0.8–2.0)	1.1 (0.8–1.6)	1

Data are means for the 2 levels of treatment (other and none) predicted for 1 day between measurements. Numbers in parentheses are 90% confidence limits. Errors (expressed as CV) and their ratios were derived from a meta-analytic model, the same as that described in Table 2 legend. V_{Blood,Evans}, V_{Plasma,Evans}, and V_{RBC,Evans}: V_{Blood}, V_{Plasma}, and V_{RBC} via Evans blue dye, respectively.

V_{Plasma} with ¹³¹I

In the only study providing usable data for the ¹³¹I method, the subjects were five female and eight male patients, and there was no treatment between the two trials, which were separated by 150 min. The error of measurement for V_{Plasma} was 4.9% (confidence limits 3.7–7.0%).

[Hb] and Hct

The mean errors of measurement for [Hb] were 2.5% (confidence limits 2.3–2.7%) and 2.1% (1.8–2.5%) for the automated and cyanmethemoglobin methods, respectively, whereas that for Hct was 3.2% (2.6–4.0%). The variation in error between laboratories for [Hb] was trivial ($\times/\div 1.1$), and its upper confidence limit was only marginally substantial ($\times/\div 1.2$). On the other hand, the variation in error between laboratories for Hct was substantial, although small: $\times/\div 1.2$ (1.1–2.8). There was little difference between men and women for the error in [Hb] (ratio men/women 1.1; 1.0–1.3) and Hct (ratio 1.0; 0.8–1.2). A 10-fold increase in time between measurements was associated with a decisive substantial increase in the error for [Hb] (ratio 1.3; 1.3–1.4); in contrast, the error in Hct was, at most, only marginally larger (ratio 1.1; 1.0–1.2) for a 10-fold increase in time.

DISCUSSION

The most important results in the present study are the meta-analytic estimates of error of measurement of V_{RBC} and M_{Hb}, the blood parameters directly related to oxygen transport. The short-term errors for V_{RBC,51Cr} and M_{Hb,CO} were ~2.5%, whereas those for V_{RBC,Evans} or V_{RBC,CO} were about threefold greater. Over a period of 1 mo, the errors for V_{RBC,51Cr} and M_{Hb,CO} were ~3.5%, about one-half of those for V_{RBC,Evans} or V_{RBC,CO}. The errors of measurement for M_{Hb,CO}, V_{RBC,51Cr}, and V_{RBC,Evans} also showed wide variation between laboratories, typically by a factor of approximately $\times/\div 1.6$. Thus a poor laboratory assessing M_{Hb,CO} and a good laboratory assessing V_{RBC,Evans} could have similar errors of measurement (~4%) and obtain similar precision in the estimates of effects on M_{Hb} and red cell volume with a given sample size, but an even greater disparity between the two methods is also possible. Unfortunately, we were unable to estimate whether V_{RBC,51Cr} shows substantial variation from laboratory to laboratory, owing to a paucity of data.

The substantial increase in the error for M_{Hb,CO} with increasing time between measurements can be accounted for partly by an increase in the contribution of biological variation. The error of measurement consists of biological variation and analytic error that are independent and combine as variances: (error of measurement)² = (biological variation)² + (analytic error)². If we assume that the 1-day error for M_{Hb,CO} (2.2%) comprises minimal biological variation, the 30-day error (~4%) indicates additional error of 3.3% during this interval, irrespective of treatment. The sources of technical error (Table 1) should be independent of time between measurements, so this additional error appears to be entirely biological. An intriguing possibility is that it arises from cyclical variation in hematopoiesis with a period of approximately weeks, similar to but of lesser magnitude than that described in some hematological disorders (62). On the other hand, the additional error could be an artifact of studies of longer duration being conducted by laboratories with poorer measurement error. Frequent serial measurements of M_{Hb} and other blood parameters over several months should resolve this issue.

The measurement errors of V_{RBC,CO}, V_{Blood,Evans}, V_{Plasma,Evans}, and V_{RBC,Evans} did not increase substantially with increasing time between measures, in contrast to that of M_{Hb,CO}. This finding is likely attributable to large measurement error swamping any biological variation. For example, the measurement error for V_{Plasma,Evans} was 6% after 1 day, which, combined with biological variation of $\leq 3\%$ (as described above), results in a total error of 6.2% [$\sqrt{6^2 + 3^2}$] for 30 days between measures.

Sources of Error

Sources of measurement error of the common blood measurement methods are summarized in Table 1. Several of these sources of method-specific measurement error are discussed below in relation to the results obtained in the meta-analysis.

M_{Hb,CO}. The major contributions to test-retest random analytic error for the CO rebreathing method include gas leaks in the mouthpiece, noseclip, and rebreathing system, inadequate CO dose, and using a rebreathing bag with excessive volume (15). The importance of careful attention to these sources of error is reinforced by the result that, in our hands, the Burge and Skinner method has less than one-half the error of the equivalent method used by others. An adequate dose of CO becomes particularly important if estimating %HbCO levels with commercial CO-oximeters that display readings only to a single decimal place (usually $\pm 0.1\%$) (15). With progressively lower doses of CO, a 0.1% difference in the %HbCO is associated with a substantial increase in the measurement error of M_{Hb,CO} (15); for example, doses of 75 and 25 ml produce errors of 1.3 and 4.1%, respectively, for a woman with 600 g of hemoglobin. Investigators can perform replicate measurements with each specimen to improve the single-replicate precision (15, 18). We have successfully used a CO dose of 1.25–1.5 ml/kg in athletes with low body mass index to induce a pre- to postdose change in %HbCO of ~6.5% and performed at least five replicate measures of %HbCO (OSM-3, Radiometer, Copenhagen, Denmark) on each blood sample (50). In clinical situations, a CO dose of 1 ml/kg at sea level (aiming to increase %HbCO by ~6%) to a maximum dose of 100 ml is adequate, with appropriate dose reduction for patients with significant anemia and/or morbid obesity. Use of high-prec-

sion gas chromatography to measure %HbCO (23) allows much smaller doses of CO to be administered but requires well-trained technical staff. On the other hand, well-maintained CO-oximeters require minimal staff training and are routinely found in major hospital emergency departments and intensive care units and now increasingly in exercise physiology laboratories. The increased CO dose required when using CO-oximetry to obtain adequate precision makes no substantial difference to the safety of the method (15).

Our meta-analysis indicates that the measurement error of $M_{\text{Hb,CO}}$ is small in most laboratories. Nevertheless, the CO method has sustained criticism of its validity, particularly the claim that inspired CO leaves the vascular space and binds to extravascular porphyrin moieties, such as myoglobin (73, 115, 126), leading to overestimation of $M_{\text{Hb,CO}}$ (84, 100, 123, 126). However, at relatively low HbCO saturation (<15% HbCO) and at the elevated blood oxygen tension induced in the CO method (15, 48), HbCO remains stable (9), and oxygen rather than CO is taken up by myoglobin (19, 95). Observations on the effects of ischemia on deoxymyoglobin in resting muscle in the presence of 20% HbCO using magnetic resonance spectroscopy (116) support our earlier findings (15) that measures of $M_{\text{Hb,CO}}$ do not need to be corrected for loss of CO to myoglobin. Mathematical modeling of CO uptake during 40 min of rebreathing (14) indicates that loss of CO to extravascular sites is, at most, a negligible ~1% of the CO dose over 10 min. Thus concerns of poor validity of M_{Hb} with CO rebreathing appear to be unwarranted. Additionally, the low rate of extravascular loss of CO appears to be consistent within an individual and, therefore, will have little or no substantial effect on the random error of ~2%.

$V_{\text{RBC},51\text{Cr}}$. $V_{\text{RBC},51\text{Cr}}$ is the defined criterion method of the International Committee for Standardization in Haematology (75). Our meta-analysis confirms the low measurement error of $V_{\text{RBC},51\text{Cr}}$, as concluded 25 years ago (75) and more recently (31, 32). A common misconception, however, is that the ^{51}Cr method primarily measures V_{RBC} . The primary estimate is actually $V_{\text{Blood},51\text{Cr}}$ (32), because counts in a whole blood specimen are compared with the assay reference counts. $V_{\text{RBC},51\text{Cr}}$ must then be secondarily derived from $V_{\text{Blood},51\text{Cr}}$ using the Hct (Table 1). It is difficult, and potentially unethical, to perform multiple ^{51}Cr studies on individuals simply for the sake of establishing the method measurement error. With multiple estimations in the same subject, residual radioactivity from previous estimations increases measurement error, unless compensated by progressively larger doses of ^{51}Cr (0.5–4 MBq for one estimate, 4–20 MBq after three consecutive estimates within ~28 days). The radiation exposure is thus a concern, especially if the studies are conducted in healthy people for nondiagnostic purposes. Hence, use of the ^{51}Cr method is rare. A variation of the chromium-labeling of red blood cells that may warrant further investigation is the use of the nonradioactive isotope ^{53}Cr (142).

Fairbanks and colleagues (32) simulated best case scenario sources of variability affecting the measurement error of $V_{\text{RBC},51\text{Cr}}$ as follows: whole blood ^{51}Cr dose (~8,000 counts: pipetting error 1%, blood resuspension and reinjection errors 2.0%), whole blood specimen (pipetting error 1.0%, mixing error 1.0%), scintillation count error (1.1%), and Hct (1.7% biological within-person variability, 1.0% analytic variability),

giving an overall estimate of the measurement error of $V_{\text{RBC},51\text{Cr}}$ of 3.4%. Scintillation count errors of 1.5% (100) and 1.6% (148) have also been reported. The measurement error obtained in our meta-analysis for $V_{\text{RBC},51\text{Cr}}$ was of similar magnitude to that from Fairbanks' modeling (Table 2).

In our meta-analysis, the mean measurement error for $V_{\text{RBC},51\text{Cr}}$ was substantially more than that for $M_{\text{Hb,CO}}$, but the confidence limits of the ratio of errors indicate that the true difference between the methods is unclear (Table 2). If the real difference is substantial, it could be due to the fact that $V_{\text{RBC},51\text{Cr}}$, unlike $M_{\text{Hb,CO}}$, requires Hct for its estimation (Table 1). An interesting anomaly obtained in our results, however, was that the measurement error for $V_{\text{RBC},51\text{Cr}}$ is somewhat less than our estimate of the measurement error for Hct, which is not possible in reality (31). Our result can be explained as follows: the studies that we meta-analyzed for $V_{\text{RBC},51\text{Cr}}$ all had measurement errors for Hct < 3%, which is near the lower 90% confidence limit for Hct. Our results are consistent with the model of Fairbanks and associates (32) that the random, biological variation in Hct from day to day is the major source of error in $V_{\text{RBC},51\text{Cr}}$ (31).

$V_{\text{RBC,Evans}}$. The primary estimate of the Evans blue method is $V_{\text{Plasma,Evans}}$, and $V_{\text{RBC,Evans}}$ must be derived subsequently with the Hct (Table 1). It follows that the measurement errors for $V_{\text{RBC,Evans}}$ consist of those for $V_{\text{Plasma,Evans}}$ (6.0%) and Hct (3.2%) that combine as variances: $V_{\text{RBC,Evans}}^2 = \text{error in } V_{\text{Plasma,Evans}}^2 + \text{error in Hct}^2$. Thus error in $V_{\text{RBC,Evans}} = \sqrt{(6.0^2 + 3.2^2)} = 6.8\%$, which is consistent with our meta-analytic estimate of 6.7%. It is, therefore, clear that most of the error in $V_{\text{RBC,Evans}}$ derives from the measurement of $V_{\text{Plasma,Evans}}$.

The existence of a wide variety of Evans blue methods is suggestive of attempts by researchers to improve method reliability. However, errors of $V_{\text{Plasma,Evans}}$ were not decisively different, whether a dye-extraction procedure was used (11, 12, 47, 55) or not used (43, 90, 94, 104, 108, 118), and whether (43, 88, 93, 111, 118, 143) or not (11, 12, 17, 24, 55) multiple data points were back-extrapolated to obtain the volume of distribution of Evans blue. Consequently, our analysis does not support use of one method variant over another. Analysis of multiple replicates of specimens improves analytic precision, and it is noteworthy that Gordon et al. (47), who achieved one of the lowest measurement errors for $V_{\text{Plasma,Evans}}$, routinely performed triplicate spectrophotometric assays of Evans blue dye concentration. Among the other studies from which we acquired data for meta-analysis, only Branch et al. (11, 12), Chien et al. (17), and Levine et al. (personal communication) reported at least duplicate spectrophotometric assays. Different numbers of assayed replicates likely contribute to some of the variation in measurement error between laboratories, but it is unclear whether the use of multiple assays would be sufficient to make the error for $V_{\text{Plasma,Evans}}$ comparable to that of $M_{\text{Hb,CO}}$.

The Evans blue method has sustained criticism about its validity, particularly that variable loss of the dye from the vascular space causes overestimation of V_{Plasma} as well as a relatively high measurement error (8, 29, 54). Evans blue is primarily an estimate of the albumin space (112), but the dye also binds to globulins (85), fibrinogen (8), and connective tissue (129). The albumin space includes variable flux among

the vascular, interstitial, and lymphatic spaces (8, 34); postinjection disappearance curves are consistent with a rapid mixing phase and a slow disappearance phase (143), as initially described by Gibson and Evans in 1937 (43). The rates of decline in the two phases vary between individuals (96), and substantial change within an individual between two trials would increase measurement error, especially if there were changes in extravascular flux due to increased vascular permeability. Capillary permeability to albumin increases during acute exposure to high altitude (60, 61, 92, 152) but appears to be an inconsistent response (81, 82) at the moderate altitudes used by athletes. Serum vascular endothelial growth factor [VEGF; also known as vascular permeability factor (20, 130, 131)] increases substantially at moderate altitude (3), secondary to increased oxygen-regulated gene expression (147) mediated by hypoxia inducible factor-1 (26). An increase in VEGF provides a possible means by which vascular permeability is increased at moderate altitude, and its mechanism of action appears related to disruption of endothelial tight junctions (145). The chronic increase in vascular permeability resulting from an increase in VEGF persists for at least 24 h, but not 72 h (7). Consequently, Evans blue estimates of V_{Plasma} conducted at altitude, or 1 or 2 days after return to sea level, may be affected by an enhanced rate of loss of dye. Our analysis indicates little difference in measurement error for Evans blue volumes with various treatments, but additional error arising from a small and variable change in vascular permeability at or after altitude may have been overwhelmed by the noise of the Evans blue method.

$V_{\text{RBC,CO}}$. The meta-analysis demonstrated that there was threefold greater 1-day error for $V_{\text{RBC,CO}}$ than for $M_{\text{Hb,CO}}$. Part of this difference arises from the contributions that [Hb] and Hct make, along with the contribution of M_{Hb} , to the error in $V_{\text{RBC,CO}}$ (Table 1). If we assume little or no biological variation in M_{Hb} , [Hb], and Hct over 1 day, the errors in these three measures are methodological and, therefore, independent. Thus $(V_{\text{RBC,CO}} \text{ error})^2 = (M_{\text{Hb,CO}} \text{ error})^2 + (\text{Hct error})^2 + ([\text{Hb}] \text{ error})^2$. In the meta-analysis, the 1-day error was 2.2% for $M_{\text{Hb,CO}}$, 3.2% for Hct, and 2.5% for [Hb] via the cyanmethemoglobin method. The expected error in $V_{\text{RBC,CO}}$ is, therefore, $\sqrt{(2.2^2 + 3.2^2 + 2.5^2)} = 4.6\%$, which is still substantially less than the meta-analyzed value of 6.7%. The simple explanation for this difference is that errors in $V_{\text{RBC,CO}}$ came from studies with a higher than average error for M_{Hb} , which is certainly the case for the studies that provided estimates for both of these measures (55, 119, 120).

Our estimate of total error for Hct agrees well with Thirup's (140) value of 4.2% for day-to-day variation in centrifuged micro-Hct, which he estimated as consisting of 3% biological variation and 3% analytic variation. Analytic error has been reported for automated Hct to be 2.3% and 0.8% for [Hb] (38). Using the lowest values from above, the likely minimal measurement error for $V_{\text{RBC,CO}}$ with 1 day between measures = $\sqrt{(2.2^2 + 2.3^2 + 0.8^2)} = 3.3\%$. This indicates that, in careful hands, the total error for $V_{\text{RBC,CO}}$ should approach that of $M_{\text{Hb,CO}}$. However, the additional propagated error in $V_{\text{RBC,CO}}$ from the measurement of Hct and [Hb] does little to support the former method if one needs to monitor small changes in the red cell compartment.

Implications for Monitoring Individuals

What are the implications of measurement error for a clinician's uncertainty in the assessment of an individual? Applying statistical first principles, the observed value of a measurement plus or minus the error of measurement is 68% confidence limits for the true value of a normally distributed single measurement, or 52% confidence limits for a change between two such measurements; the observed value plus or minus twice the measurement error is, respectively, 95 and 84% confidence limits (66). For the clinician to be reasonably confident that an observed small but substantial positive value is not, in reality, substantially negative, it follows that the error of measurement has to be less than the least clinically important difference (66, 69, 70). In other words, to measure a signal confidently, the noise must be less than the signal (67). A measure with an error of 7% is, therefore, useful only for characterizing differences $>7\%$. Differences in V_{Blood} of this magnitude are probably the least clinically important difference in most clinical situations, when one considers that a healthy individual can donate ~ 500 ml of blood from a total volume of ~ 5 liters with little risk. Evans blue used carefully would, therefore, be suitable in many clinical settings, as an alternative to ^{131}I or ^{125}I (32). Clinicians should, nevertheless, be aware that the error of measurement with Evans blue will sometimes produce unrealistically large differences in V_{Blood} . Discounting such differences to an extent guided by clinical experience is sensible and an appropriate application of Bayesian reasoning (22, 70).

Implications for Controlled Trials

Adequacy of sample size is one of the most important issues in quantitative studies. With the traditional requirement of 80% chance of statistical significance at the 5% level for the smallest clinically important difference, statistical first principles show that the sample size of each group in a randomized controlled trial needs to be $32 d^2/e^2$, where d is the smallest difference and e is the standard error of measurement. In studies of athletes, a 2% increase in M_{Hb} or V_{RBC} would be important, because it would likely produce a useful change in endurance performance. A researcher interested in detecting such an increase in M_{Hb} using the CO method with an error of measurement of $\sim 2\%$ would, therefore, need 32 subjects in each of the control and experimental groups. Samples in the studies that we reviewed were typically less than one-half this size. Detecting the same change in V_{RBC} using an Evans blue method with an error of $\sim 7\%$ would require an unprecedented nine times as many subjects in each group. Less conservative approaches to sample size estimation based on adequate precision of estimation would make the usual sample sizes adequate for characterizing the smallest changes in $M_{\text{Hb,CO}}$, but these sample sizes would be adequate for changes in $V_{\text{RBC,Evans}}$ only when the changes are greater than the error of measurement for $V_{\text{RBC,Evans}}$: $\sim 7\%$ for the typical laboratory.

In our view, mean changes in V_{RBC} in excess of 7% are likely only following radical interventions, such as exposure to high altitude, administration of large doses of erythropoietin, and direct substantial manipulation of red cell volume by venesection or blood transfusion. There are, nevertheless, claims in the literature for even larger changes in $V_{\text{RBC,Evans}}$ following less severe interventions, for example 4 wk of

training at sea level (88) and 4 wk living and training at moderate altitude (16, 88). The fact that the findings were not consistent in similar studies of $V_{RBC,CO}$ at higher altitude (55, 113, 151) points to type I error as an alternative explanation to such large physiological effects. Publication and other biases, such as deleting apparent outliers before analysis (e.g., a physiologically improbable large negative change in V_{RBC}), would make the rate of reporting of such large effects substantially greater than the usual type I error rate of 5%. It might help to limit such biases if researchers routinely reported the magnitude of the error of measurement in their studies, not only from any reliability study but also from the data in the control and experimental groups.

Methods with the Smallest Error

For researchers and clinicians seeking to improve their measurement error, we have identified the studies that demonstrated the smallest measurement errors for each of the four methods illustrated in Fig. 1.

$M_{Hb,CO}$. The lowest measurement error of 0.9% [90% confidence limits (CL), 0.5–1.5%] was obtained by Burge and Skinner (15).

$V_{RBC,51Cr}$. The smallest measurement error of 1.4% (90% CL = 0.6–2.9%) was obtained by Johnson et al. (79).

$V_{RBC,Evans}$. Although the lowest estimates were $\sim 2\%$, these were derived from studies of only three to four subjects [Faura and Reynefarje (33), and a subset of the 1992–2002 data from Levine and coworkers (16, 42, 86–88, 150)]. The 90% uncertainties in these estimates are, therefore, a factor of approximately $\times/÷ 2.5$. Allowing also for regression to the mean with extreme values when there is such large sampling error, it is likely that the true error in these studies was substantially greater.

$V_{RBC,CO}$. This method would also have acceptable error in a laboratory, with a good technique for $M_{Hb,CO}$ and good reliability for Hct and [Hb]. The lowest errors ($\sim 2\text{--}4\%$) were obtained by Myhre et al. (99), although the small sample size (4–5 subjects) implies considerable uncertainty in these estimates.

Finally, it is critical that, for all measures reliant on Hct or [Hb] and to derive V_{RBC} , V_{Plasma} , or V_{Blood} , subjects should adopt a consistent posture for at least 20 min before blood sampling (15). Venous stasis of any duration should also be avoided because it introduces substantial and unquantifiable error due to regionally increased hydrostatic pressure and localized hemoconcentration.

In conclusion, $M_{Hb,CO}$ has error of measurement similar to that of $V_{RBC,51Cr}$, which is considered the gold standard (32, 75, 106). Both $V_{RBC,51Cr}$ and $M_{Hb,CO}$ have only about one-third of the error of Evans blue dye or $V_{RBC,CO}$ and thus are more suitable to monitor small changes in red cell volume and M_{Hb} , respectively. Given the relative ease of handling CO compared with ^{51}Cr , arising from the fact that $M_{Hb,CO}$ is independent from biological variation in Hct, and the shorter biological half-life of CO (46), this review supports the routine use of CO rebreathing in clinical as well as research situations to monitor changes. Our results also reinforce the importance of researchers estimating and reporting the error of measurement of the method in their hands to improve the analysis and interpretation of their data.

APPENDIX: DETAILS OF THE META-ANALYTIC APPROACH AND ITS INTERPRETATION

Meta-Analysis

The main outcome from a meta-analysis is a weighted mean of values of the statistic of interest (measurement error in this instance) from the various studies, where the weighting factor is the inverse of the square of the sampling standard error of the statistic. We performed mixed-linear model meta-analyses, where effects of blood measures, treatment, and subject characteristics on the estimates were estimated as fixed effects, and the remaining unexplained true variation (heterogeneity) within and between studies was estimated as one or more random effects. Mixed-model meta-analysis is more realistic than traditional meta-analysis (in which “outlier” estimates are progressively eliminated until the test of heterogeneity is no longer statistically significant).

Data transformation. Meta-analysis of untransformed errors of measurement was not an option, because it would result potentially in negative values for means and confidence limits, which is not possible in reality. Analysis of the errors as variances was attractive for several reasons but suffers from the same problem. We, therefore, opted for log transformation, which was used successfully in an earlier meta-analysis of error of measurement (assessing reliability of power in physical performance tests) (72). Meta-analysis with mixed linear modeling provides confidence limits based on the assumption that random effects (including the residuals) are uniform and that any nonnormality in the individual observations is normalized in the outcome statistics by the central limit theorem. Gross departures from normality in the distribution of the individual observations are, therefore, to be avoided. We, therefore, used simulation to check that the sampling distribution of a log-transformed standard deviation derived from small sample sizes (>3) had an acceptably near-normal distribution.

Weighting factor. The sampling standard error for the log-transformed error of measurement was derived semi-empirically by initially assuming it was defined approximately by the 68.4% confidence limits of the standard error of measurement. These limits were derived from the usual formula involving the χ^2 distribution and df, and then converted to a single times/divide factor ($\times/÷$ factor) by taking the square root of the upper limit divided by the lower limit. The squared log of this $\times/÷$ factor should approximate the inverse of the weighting factor used in the meta-analysis. The accuracy of the inverse weighting factor was checked by comparing the average of its value computed for a standard deviation from each of 10,000 samples of a given size with the mean value of the variance of the log of the standard deviations of the same samples. The weighting factor was found to be biased low for small sample sizes, and a correction factor of $1 + 1/[2(df + 1)]$ was found by trial and error. This factor corrected the bias to within 1% for $df \geq 2$.

Fixed effects. The fixed-effects model varied with the blood measures under analysis as follows.

M_{Hb} , $V_{RBC,Evans}$, $V_{RBC,CO}$, AND $V_{RBC,51Cr}$. For this main comparison of error of measures, a paucity of data for $V_{RBC,51Cr}$ dictated a simple model, consisting of the interaction of measure (representing these four blood parameters) with treatment (two levels: none and other) and with \log_{10} time (the base-10 logarithm of time between pairs of measurements). Sex and fitness level were excluded from the analysis because the subtle effects of these predictors were masked by the considerably larger errors of measurement for $V_{RBC,Evans}$ and $V_{RBC,CO}$.

M_{Hb} . A more complex fixed-effects model was used in the analysis of M_{Hb} , because of its low error of measurement (see RESULTS) and a relatively large amount of available data. The predictors were sex (proportion of men in the sample), fitness (three levels: athlete, active nonathlete, and inactive), treatment (three levels: none, altitude, and other), method (Burge and Skinner and other), and \log_{10} time. The

interaction of treatment with \log_{10} time was included in a preliminary analysis.

EVANS BLUE VOLUMES. The predictors were measure (three levels: $V_{\text{Blood,Evans}}$, $V_{\text{Plasma,Evans}}$, and $V_{\text{RBC,Evans}}$), treatment (three levels: none, altitude, and other), and \log_{10} time.

EVANS BLUE METHOD VARIATIONS. The error of measurement of V_{Plasma} was the dependent variable in a model where three Evans-blue method variations (three levels: extracted/not back-extrapolated, not extracted/not back-extrapolated, and not extracted/back-extrapolated) were coded as levels of measure. The other predictors were treatment (two levels: none and other) and \log_{10} time.

[Hb] AND HCT. In separate analyses for these parameters, the models were the same as for M_{Hb} , with the inclusion of method (two levels: automated and cyanmethemoglobin) for [Hb].

All meta-analyzed estimates of measurement error for a given blood parameter shown in the RESULTS are values predicted for 1 day between measurements and, where relevant, for equal contributions from each level of predictors in the model and for equal proportions of men and women. Comparisons of effects of different levels of individual predictors (e.g., no treatment vs. all other treatments) on measurement error are shown as \times/\div factors, derived by back-transformation of the log-transformed measurement error. For example, the difference between the error of measurement (ratio of CV between measures) for $M_{\text{Hb,CO}}$ and that for $V_{\text{RBC,Evans}}$ is a ratio of 3.0 (Table 2). Ratios can be reinterpreted as percent differences, for example, a ratio of 1.25 represents 25% more error, but we have opted to present all differences as ratios to minimize potential confusion with the percent units of the error of measurement. Because the time between measurements was \log_{10} transformed, its effect on measurement error is shown as a factor per 10-fold multiple of time.

Random effects. For the main comparison of blood parameters and for the analyses of $M_{\text{Hb,CO}}$, [Hb], and Hct, the random effects coded for each measure were an identifier for the estimate (to provide the pure within-laboratory, between-estimate variation), an identifier for the laboratory technique (to provide pure between-laboratory variation), and the residual. Owing to a paucity of data for some Evans blue volumes and Evans blue methods, the between-laboratory random effect in each analysis was modeled to have the same magnitude for the three volumes and for the three method variations. The between-laboratory variance for [Hb] was negative and, therefore, set to zero; the confidence limits for the between-estimate variation were estimable only by allowing the lower confidence limit of the variance to be negative. In initial analyses, an identifier for each series of estimates of error coming from the same subjects was included as a random effect, but the paucity of data too often resulted in negative variance for one or more of the random effects. This identifier was, therefore, excluded from all analyses.

When the residual variance is scaled to unity (153), the standard deviation of the estimate identifier gives the typical variation of mean measurement error (CV) obtained for a method performed routinely in a specified laboratory (within-laboratory, between-estimate variation). The standard deviation of combined variances for the estimate and laboratory identifiers represents typical variation in mean method measurement error obtained between laboratories. The df for the combined variances was estimated by using the Satterthwaite (121) approximation and applied to the χ^2 distribution to estimate confidence limits for the combined variance (in the same manner that confidence limits for the individual variances are estimated in the statistics program). The square root of the variances and their confidence limits were back-transformed to times/divide factors. As an example, a between-laboratory typical variation of $\times/\div 1.80$ obtained for a blood measure with mean measurement error (CV) of 6.7% is interpreted as follows: researchers applying the method will obtain, on average, a true CV of 6.7%, but the true value for any given researcher could be typically between 12.1% ($6.7 \times 1.80\%$) and 3.7% ($6.7 \div 1.80\%$). The uncertainty is actually wider, because uncertainty repre-

sented by the confidence limits for the 6.7% have not been taken into account in this example. However, since the focus of the present study is primarily examination of the fixed effects, we have shown the random effects only to illustrate the generally wide variation in the errors of measurement between assays (within laboratories) and between laboratories. These estimates of error of measurement apply only to true values or to values from large sample sizes, because sampling variation will further inflate the variation between observed estimates of error of measurement when sample sizes are small (for example, by a factor of $\times/\div 1.52$ for samples of 10 subjects).

SAS mixed-modeling procedure. The meta-analyses were performed with the mixed-modeling procedure (Proc Mixed) in the Statistical Analysis System (version 8.2, SAS Institute, Cary, NC). Key elements of the code of the Proc Mixed step, adapted from Yang (153), are shown below for the analysis of the four related blood measures: $M_{\text{Hb,CO}}$, $V_{\text{RBC,Evans}}$, $V_{\text{RBC,CO}}$, and $V_{\text{RBC,51Cr}}$. The statements for estimating the fixed effects and for back-transforming the fixed and random effects have been omitted:

```
proc mixed covtest cl alpha = 0.1;
class ErrorID LabID Measure Treatment;
weight inverr;
model LnError = Measure*Treatment Measure*Log10Time/cl
      ddfm = sat outp = pred;
random ErrorID LabID /s group = Measure;
parms(10)(10)(10)(10)(10)(10)(10)(10)(1)/hold = 9;
```

Statistical inferences. We made inferences about population (true) values of statistics via precision of estimation (using confidence limits) rather than via hypothesis testing (using *P* values and statistical significance). We also used 90% rather than 95% confidence limits to discourage attempts to reinterpret the limits in terms of statistical significance at the 5% level (135). Furthermore, 90% confidence limits represent adequate precision for making probabilistic inferences, because the probability that the true value is less than the lower confidence interval and more than the upper confidence interval is both only 0.05, which we interpreted as very unlikely. Inferences about the substantiveness of true differences between two errors of measurement were made by interpreting the confidence limits of the ratio in relation to the thresholds for substantial ratios (0.9 and 1.1). Ratios >1.1 or <0.9 were considered substantial on the basis of the impact of error on sample size: in controlled trials, sample size is inversely proportional to the square of the error of measurement (71), so an increase in error by a factor of 1.1 represents an increase in sample size of 1.21, or 21%. The difference between two errors was inferred to be clear, decisive, or conclusive in relation to one or other threshold, if the confidence interval for the ratio of the errors did not overlap the threshold. Thus one error would be clearly greater than another, if the entire confidence interval of their ratio was >1.1 ; one error would be possibly greater than another, if the confidence interval overlapped 1.1 but not 0.9; the difference in the errors would be clearly or decisively trivial, if the confidence interval was within or abutted 0.9 to 1.1; and the difference would be unclear or indecisive, if the confidence interval overlapped 0.9 and 1.1. Overlap of a threshold that was slight relative to the width of the confidence interval was characterized by terms such as "marginal" and "unlikely," whereas greater overlap was characterized by "possible" and "likely."

ACKNOWLEDGMENTS

We thank all those who contributed de-identified raw data for us to meta-analyze. These include Clarence Alfrey, Michael Ashenden, David Branch, Kevin Davy, Birgit Friedmann, Christopher Gordon, Robert Grover, Katja Heinicke, Benjamin Levine, Jack Loepkky, Jack Reeves, Paul Robach, Philo Saunders, Walter Schmidt, Gary Slater, Ian Stewart, Jim Stray-Gundersen, and Darren Warburton.

A detailed summary table of all data sources used in the meta-analysis is available from the corresponding author on request.

REFERENCES

1. **Alfrey CP, Udden MM, Leach-Huntoon C, Driscoll T, and Pickett MH.** Control of red blood cell mass in spaceflight. *J Appl Physiol* 81: 98–104, 1996.
2. **Arbab-Zadeh A, Zuckerman JH, Zhang R, Neimi H, and Levine BD.** Endurance training increases left ventricular distensibility in humans (Abstract). *Med Sci Sports Exerc* 34: S7, 2002.
3. **Asano M, Kaneoka K, Nomura T, Asano K, Sone H, Tsurumaru K, Yamashita K, Matsuo K, Suzuki H, and Okuda Y.** Increase in serum vascular endothelial growth factor levels during altitude training. *Acta Physiol Scand* 162: 455–459, 1998.
4. **Ashenden MJ, Gore CJ, Martin CM, Clough ML, Bourdon PC, Dobson GP, and Hahn AG.** Skin-prick blood samples are reliable for estimating Hb mass with the CO-dilution technique. *Eur J Appl Physiol* 79: 535–537, 1999.
5. **Ashenden MJ, Gore CJ, Dobson GP, and Hahn AG.** “Live high, train low” does not change the total haemoglobin mass of male endurance athletes sleeping at a simulated altitude of 3000 m for 23 nights. *Eur J Appl Physiol* 80: 479–484, 1999.
6. **Ashenden MJ, Gore CJ, Martin DT, Dobson GP, and Hahn AG.** Effects of a 12-day “live high, train low” camp on reticulocyte production and haemoglobin mass in elite female road cyclists. *Eur J Appl Physiol* 80: 472–478, 1999.
7. **Bates DO and Curry FE.** Vascular endothelial growth factor increases microvascular permeability via a Ca^{2+} -dependent pathway. *Am J Physiol Heart Circ Physiol* 273: H687–H694, 1997.
8. **Bent-Hansen L.** Initial plasma disappearance and distribution volume of [^{131}I] albumin and [^{125}I] fibrinogen in man. *Acta Physiol Scand* 136: 455–461, 1989.
9. **Blackmore DJ.** Distribution of HbCO in human erythrocytes following inhalation of CO. *Nature* 227: 386, 1970.
10. **Borisch S, Bärtsch P, and Friedmann B.** Effects of strength endurance training in hypoxia on endurance capacity, blood volume on erythropoietin (Abstract). *Int J Sports Med* 23: S80, 2002.
11. **Branch JD III, Pate RR, Bourque SP, Convertino VA, Durstine JL, and Ward DS.** Effects of exercise mode on hematologic adaptations to endurance training in adult females. *Aviat Space Environ Med* 68: 788–794, 1997.
12. **Branch JD III, Pate RR, Bourque SP, Convertino VA, Durstine JL, and Ward DS.** Exercise training and intensity does not alter vascular volume responses in women. *Aviat Space Environ Med* 70: 1070–1076, 1999.
13. **Brotherhood J, Brozovic B, and Pugh LGC.** Haematological status of middle- and long-distance runners. *Clin Sci Mol Med* 48: 139–145, 1975.
14. **Bruce EN and Bruce MC.** A multicompartment model of carboxyhaemoglobin and carboxymyoglobin responses to inhalation of carbon monoxide. *J Appl Physiol* 95: 1235–1247, 2003.
15. **Burge CM and Skinner SL.** Determination of hemoglobin mass and blood volume with CO: evaluation and application of a method. *J Appl Physiol* 79: 623–631, 1995.
16. **Chapman RF, Stray-Gundersen J, and Levine BD.** Individual variation in response to altitude training. *J Appl Physiol* 85: 1448–1456, 1998.
17. **Chien S, Usami S, Simmons RL, McAllister FF, and Gregersen MI.** Blood volume and age: repeated measurements on normal men after 17 years. *J Appl Physiol* 21: 583–588, 1966.
18. **Christensen P, Eriksen B, and Hennebero SW.** Precision of a new bedside method for estimation of the circulating blood volume. *Acta Anaesthesiol Scand* 37: 622–627, 1993.
19. **Clark BJ and Coburn RF.** Mean myoglobin oxygen tension during exercise at maximal oxygen uptake. *J Appl Physiol* 39: 135–144, 1975.
20. **Connolly DT, Heuvelman DM, Nelson R, Olander JV, Eppley BL, Delfino JJ, Siegel NR, Leimgruber RM, and Feder J.** Tumor vascular permeability factor stimulates endothelial cell growth and angiogenesis. *J Clin Invest* 84: 1470–1478, 1989.
21. **Crandall CG, Wilson TE, Shibusaki M, Cui J, and Levine BD.** Prolonged head-down tilt exposure reduces maximal cutaneous vasodilator and sweating capacity in humans. *J Appl Physiol* 94: 2330–2336, 2003.
22. **D’Agostini G.** *Bayesian Reasoning in Data Analysis, A Critical Introduction.* River Edge, NJ: World Scientific Publishing, 2003.
23. **Dahms TE and Horvath SM.** Rapid, accurate technique for determination of carbon monoxide in blood. *Clin Chem* 20: 533–537, 1974.
24. **Davy KP and Seals DR.** Total blood volume in healthy young and older men. *J Appl Physiol* 76: 2059–2062, 1994.
25. **Dill DB, Horvath SM, Dahms TE, Parker RE, and Lynch JR.** Hemococoncentration at altitude. *J Appl Physiol* 27: 514–518, 1969.
26. **Dor Y, Porat R, and Keshet E.** Vascular endothelial growth factor and vascular adjustments to perturbations in oxygen homeostasis. *Am J Physiol Cell Physiol* 280: C1367–C1374, 2001.
27. **Eklblom B.** Blood doping and erythropoietin. The effects of variation in hemoglobin concentration and other related factors on physical performance. *Am J Sports Med* 24: S40–S42, 1996.
28. **Eklblom B, Goldbarg AN, and Gullbring B.** Response to exercise after blood loss and reinfusion. *J Appl Physiol* 33: 175–180, 1972.
29. **El Sayed H, Goodall SR, and Hainsworth R.** Re-evaluation of Evans blue dye dilution method of plasma volume measurement. *Clin Lab Haematol* 17: 189–194, 1995.
30. **Ernest D, Hartman NG, Deane CP, Belzberg AS, and Dodek PM.** Reproducibility of plasma and extracellular fluid volume measurements in critically ill patients. *J Nucl Med* 33: 1468–1471, 1992.
31. **Fairbanks VF.** Myeloproliferative disease: polycythemia vera: the packed cell volume and the curious logic of the red cell mass. *Hematology* 4: 381–395, 2000.
32. **Fairbanks VF, Klee GG, Wiseman GA, Hoyer JD, Tefferi A, Pettitt RM, and Silverstein MN.** Measurement of blood volume and red cell mass: re-examination of ^{51}Cr and ^{125}I methods. *Blood Cells Mol Dis* 22: 169–186, 1996.
33. **Faura J and Reynafarje C.** Uso de la metandrostenolona para acelerar la eritropoyesis en la exposicion a la altura. *Arch Inst Biol Andina* 3: 94–98, 1970.
34. **Ferrebee JW, Leigh OC, and Berliner RW.** Passage of the blue dye T-1824 from the blood stream into the lymph. *Proc Soc Exp Biol Med* 46: 549–553, 1941.
35. **Fischer CL, Johnson PC, and Berry CA.** Red blood cell mass and plasma volume changes in manned space flight. *JAMA* 200: 579–583, 1967.
36. **Franco F, Morrow MJ, Zuckerman JH, Zhang R, Peshock RM, and Levine BD.** Right ventricular enlargement precedes left ventricular enlargement during long-term endurance training (Abstract). *Circulation* 98: I-85, 1998.
37. **Franco F, Peshock RM, Morrow MJ, Zuckerman JH, Zhang R, and Levine BD.** Left ventricular hypertrophic response to long term endurance exercise (Abstract). *Circulation* 98: I-85, 1998.
38. **Fraser CG, Wilkinson SP, Neville RG, Knox JD, King JF, and MacWalter RS.** Biologic variation of common hematologic laboratory quantities in the elderly. *Am J Clin Pathol* 92: 465–470, 1989.
39. **Frese F, Menold E, Jost J, Jedamsky A, Bartsch P, and Friedmann B.** Interindividuelle Variabilität im Erythropoietinanstieg und in der Zunahme des Gesamtkörperhämglobins bei Höhenttraining (Abstract). *Dtsch Z Sportmed* 54: S78, 2003.
40. **Friedmann B, Jost J, Rating T, Weller E, Werle E, Eckardt KU, Bartsch P, and Mairbaurl H.** Effects of iron supplementation on total body hemoglobin during endurance training at moderate altitude. *Int J Sports Med* 20: 78–85, 1999.
41. **Friedmann B, Weller E, Mairbaurl H, and Bartsch P.** Effects of iron repletion on blood volume and performance capacity in young athletes. *Med Sci Sports Exerc* 33: 741–746, 2001.
42. **Ge RL, Witkowski S, Zhang Y, Alfrey C, Sivieri M, Karlsen T, Resaland GK, Harber M, Stray-Gundersen J, and Levine BD.** Determinants of erythropoietin release in response to short-term hypobaric hypoxia. *J Appl Physiol* 92: 2361–2367, 2002.
43. **Gibson JG and Evans WAJ.** Clinical studies of the blood volume. I. Clinical application of a method employing the azo dye “Evans blue” and the spectrophotometer. *J Clin Invest* 16: 301–316, 1937.
44. **Gillen CM, Lee R, Mack GW, Tomaselli CM, Nishiyasu T, and Nadel ER.** Plasma volume expansion in humans after a single intense exercise protocol. *J Appl Physiol* 71: 1914–1920, 1991.
45. **Gledhill N, Warburton D, and Jamnik V.** Haemoglobin, blood volume, cardiac function, and aerobic power. *Can J Appl Physiol* 24: 54–65, 1999.
46. **Godin G and Shephard RJ.** On the course of carbon monoxide uptake and release. *Respiration* 29: 317–329, 1972.
47. **Gordon CJ, Fogarty AL, Greenleaf JE, Taylor NA, and Stocks JM.** Direct and indirect methods for determining plasma volume during

- thermoneutral and cold-water immersion. *Eur J Appl Physiol* 89: 471–474, 2003.
48. **Gore C, Craig N, Hahn A, Rice A, Bourdon P, Lawrence S, Walsh C, Stanef T, Barnes P, Parisotto R, Martin D, and Pyne D.** Altitude training at 2690m does not increase total haemoglobin mass or sea level $\dot{V}O_{2\max}$ in world champion track cyclists. *J Sci Med Sport* 1: 156–170, 1998.
 49. **Gore CJ, Hahn AG, Burge CM, and Telford RD.** $\dot{V}O_{2\max}$ and haemoglobin mass of trained athletes during high intensity training. *Int J Sports Med* 18: 477–482, 1997.
 50. **Gore CJ, Stray-Gundersen J, Rodríguez FA, Truijens MJ, Townsend NE, and Levine BD.** Comparison of blood volume via CO re-breathing and Evans blue dye (Abstract). *Med Sci Sports Exerc* 36: S336, 2004.
 51. **Gray SJ and Sterling K.** Determination of circulating red cell volume by radioactive chromium. *Science* 112: 179–180, 1950.
 52. **Greenleaf JE, Convertino VA, and Mangseth GR.** Plasma volume during stress in man: osmolality and red cell volume. *J Appl Physiol* 47: 1031–1038, 1979.
 53. **Greenleaf JE, Van Beaumont W, Brock PJ, Morse JT, and Mangseth GR.** Plasma volume and electrolyte shifts with heavy exercise in sitting and supine positions. *Am J Physiol Regul Integr Comp Physiol* 236: R206–R214, 1979.
 54. **Gregersen MI and Rawson RA.** Blood volume. *Physiol Rev* 39: 511–518, 1959.
 55. **Grover RF, Selland MA, McCullough RG, Dahms TE, Wolfel EE, Butterfield GE, Reeves JT, and Greenleaf JE.** β -Adrenergic blockade does not prevent polycythemia or decrease in plasma volume in men at 4300 m altitude. *Eur J Appl Physiol* 77: 264–270, 1998.
 56. **Gurland J and Tripathi RC.** A simple approximation for unbiased estimation of the standard deviation. *American Statistician* 25: 30–32, 1971.
 57. **Haldane J and Smith JL.** The absorption of oxygen by the lungs. *J Physiol* 22: 231–258, 1897.
 58. **Haldane J and Smith JL.** The mass and oxygen capacity of the blood in man. *J Physiol* 25: 331–343, 1899.
 59. **Hannon JP, Chinn KSK, and Shields JL.** Effects of acute high-altitude exposure on body fluids. *Fed Proc* 28: 1178–1184, 1969.
 60. **Hansen JM, Kanstrup IL, Richalet JP, and Olsen NV.** High altitude-induced albuminuria in normal man is enhanced by infusion of low-dose dopamine. *Scand J Clin Lab Invest* 56: 367–372, 1996.
 61. **Hansen JM, Olsen NV, Feldt-Rasmussen B, Kanstrup IL, Déchaux M, Dubray C, and Richalet JP.** Albuminuria and overall capillary permeability of albumin in acute altitude hypoxia. *J Appl Physiol* 76: 1922–1927, 1994.
 62. **Haurie C, Dale DC, and Mackey MC.** Cyclical neutropenia and other periodic hematological disorders: a review of mechanisms and mathematical models. *Blood* 92: 2629–2640, 1998.
 63. **Heinicke K, Heinicke I, Schmidt W, and Wolfarth B.** A three-week traditional altitude training increases hemoglobin mass and red cell volume in elite biathlon athletes. *Int J Sports Med* 26: 350–355, 2005.
 64. **Heinicke K, Prommer N, Cajigal J, Viola T, Behn C, and Schmidt W.** Long-term exposure to intermittent hypoxia results in increased hemoglobin mass, reduced plasma volume, and elevated erythropoietin plasma levels in man. *Eur J Appl Physiol* 88: 535–543, 2003.
 65. **Heinicke K, Wolfarth B, Winchenbach P, Biermann B, Schmid A, Huber G, Friedmann B, and Schmidt W.** Blood volume and hemoglobin mass in elite athletes of different disciplines. *Int J Sports Med* 22: 504–512, 2001.
 66. **Hopkins WG.** Applications of reliability [Online]. Sport Science. <http://sportssci.org/resource/stats/relyappl.html> [2000].
 67. **Hopkins WG.** Measures of reliability in sports medicine and science. *Sports Med* 30: 1–15, 2000.
 68. **Hopkins WG.** A spreadsheet for analysis of straightforward controlled trials [Online]. Sport Science. <http://www.sportssci.org/jour/03/wgtrials.htm> [2003].
 69. **Hopkins WG.** Clinical significance and decisiveness [Online]. Sport Science. <http://www.sportssci.org/jour/04/inbrief.htm#clinical> [2004].
 70. **Hopkins WG.** How to interpret changes in an athletic performance test [Online]. Sport Science. <http://www.sportssci.org/jour/04/wgtests.htm> [2004].
 71. **Hopkins WG and Hewson DJ.** Variability of competitive performance of distance runners. *Med Sci Sports Exerc* 33: 1588–1592, 2001.
 72. **Hopkins WG, Schabert EJ, and Hawley JA.** Reliability of power in physical performance tests. *Sports Med* 31: 211–234, 2001.
 73. **Hopper J Jr, Tabor H, and Winkler AW.** Simultaneous measurements of the blood volume in man and dog by means of Evans Blue dye, T1824, and by means of carbon monoxide I normal subjects. *J Clin Invest* 23: 628–635, 1944.
 74. **International Committee for Standardization in Haematology.** Standard techniques for the measurement of red-cell and plasma volume. A report by the International Committee for Standardization in Hematology (ICSH): Panel on Diagnostic Applications of Radioisotopes in Haematology. *Br J Haematol* 25: 801–814, 1973.
 75. **International Committee for Standardization in Haematology.** Recommended methods for measurement of red-cell and plasma volume: International Committee for Standardization in Haematology. *J Nucl Med* 21: 793–800, 1980.
 76. **International Committee for Standardization in Haematology.** Recommendations for reference method for haemoglobinometry in human blood (ICSH standard 1986) and specifications for international haemoglobin cyanide reference preparation (3rd edition) International Committee for Standardization in Haematology; Expert Panel on Haemoglobinometry. *Clin Lab Haematol* 9: 73–79, 1987.
 77. **Iwasaki K, Zhang R, Perhonen MA, Zuckerman JH, and Levine BD.** Reduced baroreflex control of heart period after bed rest is normalized by acute plasma volume restoration. *Am J Physiol Regul Integr Comp Physiol* 287: R1256–R1262, 2004.
 78. **Iwasaki K, Zhang R, Zuckerman JH, and Levine BD.** Dose-response relationship of the cardiovascular adaptation to endurance training in healthy adults: how much training for what benefit? *J Appl Physiol* 95: 1575–1583, 2003.
 79. **Johnson PC, Driscoll TB, and Fischer CL.** Blood volume changes in divers of Tektite I. *Aerospace Med* 42: 423–426, 1971.
 80. **Jung RC, Dill DB, Horton R, and Horvath SM.** Effects of age on plasma aldosterone levels and hemoconcentration at altitude. *J Appl Physiol* 31: 593–597, 1971.
 81. **Kaner RJ and Crystal RG.** Pathogenesis of high altitude pulmonary edema: does alveolar epithelial lining fluid vascular endothelial growth factor exacerbate capillary leak? *High Alt Med Biol* 5: 399–409, 2004.
 82. **Kleger GR, Bärtsch P, Vock P, Heilig B, Roberts LJ II, and Ballmer PE.** Evidence against an increase in capillary permeability in subjects exposed to high altitude. *J Appl Physiol* 81: 1917–1923, 1996.
 83. **Lange RD, Jones JB, and Johnson PC Jr.** Comparative aspects of hematological responses in animal and human models in simulations of weightlessness and space flight. *Physiologist* 30: S113–S116, 1987.
 84. **Lawson HC.** The volume of blood—a critical examination of methods for its measurement. In: *Handbook of Physiology. Circulation*. Washington, DC: Am. Physiol. Soc., 1962, sect. 2, vol. I, chapt. 3, p. 23–49.
 85. **LeVeen HH and Fishman WH.** Combination of Evans Blue with plasma protein: its significance in capillary permeability studies, blood dye disappearance curves, and its use as a protein tag. *Am J Physiol* 151: 26–33, 1947.
 86. **Levine BD.** Intermittent hypoxic training: fact and fancy. *High Alt Med Biol* 3: 177–193, 2002.
 87. **Levine BD and Stray-Gundersen J.** A practical approach to altitude training: where to live and train for optimal performance enhancement. *Int J Sports Med* 13: S209–S212, 1992.
 88. **Levine BD and Stray-Gundersen J.** “Living high-training low”: effect of moderate-altitude acclimatization with low altitude training on performance. *J Appl Physiol* 83: 102–112, 1997.
 89. **Levine BD, Zuckerman JH, Morrow MJ, Franco F, and Zhang R.** Gender differences in the cardiac adaptation to endurance training (Abstract). *Circulation* 98: I-86, 1998.
 90. **Levine BD, Zuckerman JH, and Pawelczyk JA.** Cardiac atrophy after bed-rest deconditioning: a nonneural mechanism for orthostatic intolerance. *Circulation* 96: 517–525, 1997.
 91. **Levine BD, Zuckerman JH, Zhang R, and Morrow M.** Prolonged intensive endurance training improves ventricular performance via the Frank-Starling mechanism (Abstract). *Med Sci Sports Exerc* 30: S27, 1998.
 92. **Lewis DM, Bradwell AR, Shore AC, Beaman M, and Tooke JE.** Capillary filtration coefficient and urinary albumin leak at altitude. *Eur J Clin Invest* 27: 64–68, 1997.
 93. **Loeppky JA, Luther DK, Maes D, Riboni K, Hinghofer-Szalkay H, Charlton GA, and Icenogle MV.** Plasma volume by Evans blue: effects

- of eating and comparison with other methods at altitude. *Aviat Space Environ Med* 73: 902–906, 2002.
94. **Loeppky JA, Scotto P, Charlton GC, Gates L, Icenogle M, and Roach RC.** Ventilation is greater in women than men, but the increase during acute altitude hypoxia is the same. *Respir Physiol* 125: 225–237, 2001.
 95. **Luomanmäki K and Coburn RF.** Effects of metabolism and distribution of carbon monoxide on blood and body stores. *Am J Physiol* 217: 354–363, 1969.
 96. **Mather K, Bowler RG, Crooke AC, and Morris CJOR.** The precision of plasma volume determinations by the Evans blue method. *Br J Exp Pathol* 28: 12–24, 1947.
 97. **Maw GJ, Mackenzie IL, and Taylor NA.** Human body-fluid distribution during exercise in hot, temperate and cool environments. *Acta Physiol Scand* 163: 297–304, 1998.
 98. **Myhre LG, Brown DK, Hall FG, and Dill DB.** The use of carbon monoxide and T-1824 for determining blood volume. *Clin Chem* 14: 1197–1205, 1968.
 99. **Myhre LG, Dill DB, Hall FG, and Brown DK.** Blood volume changes during three-week residence at high altitude. *Clin Chem* 16: 7–14, 1970.
 100. **Nomof N, Hopper J, Brown E, Scott K, and Wennesland R.** Simultaneous determinations of the total volume of red blood cells by use of carbon monoxide and chromium⁵¹ in healthy and diseased subjects. *J Clin Invest* 33: 1382–1387, 1954.
 101. **Palmer D, Perhonen M, Zuckerman J, Witkowski S, Zhang R, and Levine BD.** Prolonged exercise training during head-down tilt bedrest preserves upright exercise capacity (Abstract). *Med Sci Sports Exerc* 33: S299, 2001.
 102. **Palmer D, Perhonen M, Zuckerman JH, Witkowski S, Zhang R, and Levine BD.** Prolonged exercise training during head-down tilt bedrest coupled with Dextran infusion prevents orthostatic intolerance (Abstract). *Med Sci Sports Exerc* 34: S112, 2002.
 103. **Parisotto R, Gore CJ, Emslie KR, Ashenden MJ, Brugnara C, Howe C, Martin DT, Trout GJ, and Hahn AG.** A novel method utilising markers of altered erythropoiesis for the detection of recombinant human erythropoietin abuse in athletes. *Haematologica* 85: 564–572, 2000.
 104. **Pawelczyk JA, Zuckerman JH, Blomqvist CG, and Levine BD.** Regulation of muscle sympathetic nerve activity after bed rest deconditioning. *Am J Physiol Heart Circ Physiol* 280: H2230–H2239, 2001.
 105. **Pearson TC.** Evaluation of diagnostic criteria in polycythemia vera. *Semin Hematol* 38: 21–24, 2001.
 106. **Pearson TC, Guthrie DL, Simpson J, Chinn S, Barosi G, Ferrant A, Lewis SM, and Najean Y.** Interpretation of measured red cell mass and plasma volume in adults: Expert Panel on Radionuclides of the International Council for Standardization in Haematology. *Br J Haematol* 89: 748–756, 1995.
 107. **Perhonen MA, Zuckerman JH, Hinojosa JR, Zhang R, Iwasaki K, and Levine BD.** Exercise training reverses cardiac atrophy during head down tilt bed rest (Abstract). *Med Sci Sports Exerc* 93: S53, 2000.
 108. **Perhonen MA, Zuckerman JH, and Levine BD.** Deterioration of left ventricular chamber performance after bed rest: “cardiovascular deconditioning” or hypovolemia? *Circulation* 103: 1851–1857, 2001.
 109. **Petersen PH, Fraser CG, Jorgensen L, Brandslund I, Stahl M, Gowans E, Libeer JC, and Ricos C.** Combination of analytical quality specifications based on biological within- and between-subject variation. *Ann Clin Biochem* 39: 543–550, 2002.
 110. **Petersen PH, Stockl D, Westgard JO, Sandberg S, Linnert K, and Thienpont L.** Models for combining random and systematic errors. Assumptions and consequences for different models. *Clin Chem Lab Med* 39: 589–595, 2001.
 111. **Poulsen TD, Klausen T, Richalet JP, Kanstrup IL, Fogh-Andersen N, and Olsen NV.** Plasma volume in acute hypoxia: comparison of a carbon monoxide rebreathing method and dye dilution with Evans’ blue. *Eur J Appl Physiol* 77: 457–461, 1998.
 112. **Rawson RA.** The binding of T-1824 and structurally related diazo dyes by the plasma proteins. *Am J Physiol* 138: 708–717, 1943.
 113. **Reeves JT, Zamudio S, Dahms TE, Asmus I, Braun B, Butterfield GE, McCullough RG, Muza SR, Rock PB, and Moore LG.** Erythropoiesis in women during 11 days at 4,300 m is not affected by menstrual cycle phase. *J Appl Physiol* 91: 2579–2586, 2001.
 114. **Reilly WA, French RM, Lau FYK, Scott KG, and White WE.** Whole blood volume determined by radiochromium-tagged red cells. Comparative studies on normal and congestive heart failure patients. *Circulation* 9: 571–580, 1954.
 115. **Remington JW and Baker CH.** Evaluation of blood volume measurement techniques. *Circ Res* 9: 60–68, 1961.
 116. **Richardson RS, Noyszewski EA, Saltin B, and Gonzalez-Alonso J.** Effect of mild carboxy-hemoglobin on exercising skeletal muscle: intravascular and intracellular evidence. *Am J Physiol Regul Integr Comp Physiol* 283: R1131–R1139, 2002.
 117. **Robach P, Brugniaux JV, Schmitt L, Nicolet G, Fouillot JP, Lasne F, Moutereau S, Olsen NV, and Richalet JP.** *Living High-Training Low: Effect On Red Cell Mass And Aerobic Performance In Elite Middle-Distance Runners* (Abstract). Congress of the European College of Sports Sciences Proceedings, 2004, p. 11–12.
 118. **Robach P, Dechaux M, Jarrot S, Vaysse J, Schneider JC, Mason NP, Herry JP, Gardette B, and Richalet JP.** Operation Everest III: role of plasma volume expansion on $\dot{V}O_{2\max}$ during prolonged high-altitude exposure. *J Appl Physiol* 89: 29–37, 2000.
 119. **Robach P, Lafforgue E, Olsen NV, Dechaux M, Fouquieray B, Westerterp-Plantenga M, Westerterp K, and Richalet JP.** Recovery of plasma volume after 1 week of exposure at 4,350 m. *Pflügers Arch* 444: 821–828, 2002.
 120. **Robach P, Schmitt L, Roels B, Millet G, Brugniaux JV, Nicolet G, Duvallet A, Fouillot JP, Moutereau S, Lasne F, Hellard P, Olsen NV, and Richalet JP.** *Vivre en Haut–S’entraîner en Bas: Effect Sur le Volume de Globules Rouges et La Performance Aérobique Chez L’athlète de Haut Niveau* (Abstract). XII Congrès de l’ACAPS, 2003, p. 146–147.
 121. **Satterthwaite FW.** An approximate distribution of estimates of variance components. *Biometrics* 2: 110–114, 1946.
 122. **Saunders PU, Telford RD, Pyne DB, Cunningham RB, Gore CJ, Hahn AG, and Hawley JA.** Improved running economy in elite runners after 20 days of moderate simulated altitude exposure. *J Appl Physiol* 96: 931–937, 2004.
 123. **Sawka MN, Convertino VA, Eichner ER, Schnieder SM, and Young AJ.** Blood volume: importance and adaptations to exercise training, environmental stresses, and trauma/sickness. *Med Sci Sports Exerc* 32: 332–348, 2000.
 124. **Sawka MN and Coyle EF.** Influence of body water and blood volume on thermoregulation and exercise performance in heat. *Exerc Sport Sci Rev* 27: 167–218, 1999.
 125. **Sawka MN, Young AJ, Muza SR, Gonzalez RR, and Pandolf KB.** Erythrocyte reinfusion and maximal aerobic power. An examination of modifying factors. *JAMA* 257: 1496–1499, 1987.
 126. **Sawka MN, Young AJ, Pandolf KB, Dennis RC, and Valeri CR.** Erythrocyte, plasma, and blood volume of healthy young men. *Med Sci Sports Exerc* 24: 447–453, 1992.
 127. **Schmidt W, Heinicke K, Rojas J, Manuel GJ, Serrato M, Mora M, Wolfarth B, Schmid A, and Keul J.** Blood volume and hemoglobin mass in endurance athletes from moderate altitude. *Med Sci Sports Exerc* 34: 1934–1940, 2002.
 128. **Sear H, Allen TH, and Gregersen MI.** Simultaneous measurement in dogs of plasma volume with I¹³¹ human albumin and T-1824 with comparisons of their long term disappearance from the plasma. *Am J Physiol* 175: 240–242, 1953.
 129. **Semple RE, Thomsen AET, and Ball AJ.** Description and evaluation of a dextran-dilution technique for the determination of plasma volume in the dog and man. *Clin Sci (Colch)* 17: 511–518, 1958.
 130. **Senger DR, Galli SJ, Dvorak AM, Perruzzi CA, Harvey VS, and Dvorak HF.** Tumor cells secrete a vascular permeability factor that promotes accumulation of ascites fluid. *Science* 219: 983–985, 1983.
 131. **Senger DR, Van de Water L, Brown LF, Nagy JA, Yeo KT, Yeo TK, Berse B, Jackman RW, Dvorak AM, and Dvorak HF.** Vascular permeability factor (VPF, VEGF) in tumor biology. *Cancer Metastasis Rev* 12: 303–324, 1993.
 132. **Shibasaki M, Wilson TE, Cui J, Levine BD, and Crandall CG.** Exercise throughout 6 degrees head-down tilt bed rest preserves thermoregulatory responses. *J Appl Physiol* 95: 1817–1823, 2003.
 133. **Sjöstrand T.** Blood volume. In: *Handbook of Physiology. Circulation*. Washington, DC: Am. Physiol. Soc., 1962, sect. 2, vol. I, chapt. 4, p. 51–62.
 134. **Stachenfeld NS, Mack GW, DiPietro L, Morocco TS, Jozsi AC, and Nadel ER.** Regulation of blood volume during training in post-menopausal women. *Med Sci Sports Exerc* 30: 92–98, 1998.
 135. **Sterne JA and Davey Smith G.** Sifting the evidence—what’s wrong with significance tests? *Br Med J* 322: 226–231, 2001.
 136. **Stevenson ET, Davy KP, and Seals DR.** Maximal aerobic capacity and total blood volume in highly trained middle-aged and older female endurance athletes. *J Appl Physiol* 77: 1691–1696, 1994.

137. **Stewart IB, Warburton DE, Hodges AN, Lyster DM, and McKenzie DC.** Cardiovascular and splenic responses to exercise in humans. *J Appl Physiol* 94: 1619–1626, 2003.
138. **Surks MI, Chinn KS, and Matoush LR.** Alterations in body composition in man after acute exposure to high altitude. *J Appl Physiol* 21: 1741–1746, 1966.
139. **Tavassoli M.** Anemia of spaceflight. *Blood* 60: 1059–1067, 1982.
140. **Thirup P.** Haematocrit: within-subject and seasonal variation. *Sports Med* 33: 231–243, 2003.
141. **Udden MM, Driscoll TB, Pickett MH, Leach-Huntoon CS, and Alfrey CP.** Decreased production of red blood cells in human subjects exposed to microgravity. *J Lab Clin Med* 125: 442–449, 1995.
142. **Veillon C, Patterson KY, Nagey DA, and Tehan AM.** Measurement of blood volume with an enriched stable isotope of chromium (^{53}Cr) and isotope dilution by combined gas chromatography-mass spectrometry. *Clin Chem* 40: 71–73, 1994.
143. **Von Porat BTD.** Blood volume determinations with the Evans Blue dye method. *Acta Med Scand* 256: 1–108, 1951.
144. **Walser M and Duffy B.** Body fluids in hypertension and mild heart failure. *JAMA* 160: 858–864, 1956.
145. **Wang W, Dentler WL, and Borchardt RT.** VEGF increases BMEC monolayer permeability by affecting occludin expression and tight junction assembly. *Am J Physiol Heart Circ Physiol* 280: H434–H440, 2001.
146. **Warburton DE, Haykowsky MJ, Quinney HA, Blackmore D, Teo KK, Taylor DA, McGavock J, and Humen DP.** Blood volume expansion and cardiorespiratory function: effects of training modality. *Med Sci Sports Exerc* 36: 991–1000, 2004.
147. **Wenger RH.** Cellular adaptation to hypoxia: O_2 -sensing protein hydroxylases, hypoxia-inducible transcription factors, and O_2 -regulated gene expression. *FASEB J* 16: 1151–1162, 2002.
148. **Wennesland R, Brown E, Hopper J Jr, Scott KG, Hodges JL Jr, and Bradley B.** Experiences with the radiochromium method for determination of red cell volume. *Scand J Clin Lab Invest* 14: 355–367, 1962.
149. **Wilson TE, Shibasaki M, Cui J, Levine BD, and Crandall CG.** Effect of head-down tilt bed rest on human cutaneous vasoconstrictor responses to exogenous norepinephrine (Abstract). *Med Sci Sports Exerc* 34: S218, 2002.
150. **Witkowski S, Karlsen T, Resaland G, Sivieri M, Yates R, Harber M, Ge R, Stray-Gundersen J, and Levine BD.** Optimal altitude for “living high-training low” (Abstract). *Med Sci Sports Exerc* 33: S292, 2001.
151. **Wolfel EE, Groves BM, Brooks GA, Butterfield GE, Mazzeo RS, Moore LG, Sutton JR, Bender PR, Dahms TE, McCullough RE, McCullough RG, Huang S, Sun S, Grover RF, Hultgren HN, and Reeves JT.** Oxygen transport during steady-state submaximal exercise in chronic hypoxia. *J Appl Physiol* 70: 1129–1136, 1991.
152. **Wood JG, Johnson JS, Mattioli LF, and Gonzalez NC.** Systemic hypoxia increases leukocyte emigration and vascular permeability in conscious rats. *J Appl Physiol* 89: 1561–1568, 2000.
153. **Yang M.** A review of random effects modelling in SAS (release 8.2) [Online]. Centre for Multilevel Modelling. Institute of Education, Univ. of London. <http://multilevel.ioe.ac.uk/softrev/reviewsas.pdf> [2003].

