Stability of hemoglobin mass over 100 days in active men

Annette Eastwood,1,2 Will G. Hopkins,3 Pitre C. Bourdon,1 Robert T. Withers,2 and Christopher J. Gore4,2

1Sport Science Unit, South Australian Sports Institute, Adelaide; 2Exercise Physiology Laboratory, Flinders University, Adelaide, Australia; 3Institute for Sport and Recreation Research, AUT University, Auckland, New Zealand; and 4Department of Physiology, Australian Institute of Sport, Canberra, Australia

Submitted 5 July 2007; accepted in final form 16 January 2008

The purpose of this study was to investigate the presence of biological variation in hemoglobin mass (Hbmass) in men over a period of 100 days. Hemoglobin mass of six active men was monitored, using the carbon monoxide (CO) rebreathing method every 1–6 days for 100–114 days. Measurement error comprises both biological variation and technical error (11). The technical error for determining Hbmass in our hands is 1.9% for the CO rebreathing procedure of Burge and Skinner (3), and the sources of technical error should be independent of time (7). It follows that the increase in error over 1 mo reported in the meta-analysis may be due to biological variation. The purpose of this study was, therefore, to determine and possibly quantify the presence of biological variation in Hbmass in men over a period of ~100 days.

METHODS

Subjects. Six healthy recreationally active men (means ± SD: age, 33 ± 7 yr; height, 182 ± 5 cm; body mass, 82 ± 10 kg) volunteered to participate in the study after being informed of potential risks. During the 3 mo before the study commenced, all subjects were exposed to altitude in the year before or during the study. All experimental procedures were approved by the ethics committees of the Australian Institute of Sport and Flinders University, and written, informed consent was obtained from the subjects before any testing was conducted.

Protocols. Hbmass was measured every 1–6 days for 100–114 days, (42 ± 3 measurements, mean ± SD) using a modified version of the CO rebreathing technique first reported by Schmidt and Prommer (18). This modified method has been described in detail by Gore and colleagues (6). Briefly, this procedure comprised inhalation of a bolus of 99.5% chemically pure CO (BOC gasses, Sydney, Australia) in a dose of 1 ml/kg of body mass. The CO bolus was administered via a 100-ml plastic syringe (Omnifix, B Braun, Melsungen, Germany) connected to a glass spirometer with a 3.5-liter anaesthetic bag attached and was rebreathed for 2 min. Arterialized blood samples (200 μl) were taken from a prewarmed finger tip and analyzed in quintuplicate for percent carboxyhemoglobin (%HbCO) using a diode array spectrophotometer (OSM3 Hemoximeter Radiometer) before as

Hbmass (3, 18) and has an error of measurement similar to that of red-cell volume via 51Cr (7).

In a recent meta-analysis, the error of measurement for Hbmass using the CO rebreathing method was apparently considerably higher over 30 days (~3.5%) than over 1 or 2 days (~2.2%) (7). Measurement error comprises both biological variation and technical error (11). The technical error for determining Hbmass in our hands is 1.9% for the CO rebreathing procedure of Burge and Skinner (3), and the sources of technical error should be independent of time (7). It follows that the increase in error over 1 mo reported in the meta-analysis may be due to biological variation. The purpose of this study was, therefore, to determine and possibly quantify the presence of biological variation in Hbmass in men over a period of ~100 days.

THE VOLUME OF THE BLOOD and red blood cells have been studied extensively by clinical and exercise physiologists for over 100 years, beginning with Haldane and Smith (9). Hemoglobin, which is essential for oxygen transport, comprises approximately one-third of each red blood cell’s volume. Being able to accurately measure hemoglobin mass (Hbmass) is therefore important when acute or chronic changes in oxygen transport capacity are monitored, particularly changes that may arise from illegal blood-doping practices in athletes.

If small changes in Hbmass or the volume of red blood cells are to be quantified, the error of measurement must be at most only a few percent (7). Radioactive methods of determining the volume of red blood cells using isotopic tracers such as radioactive chromium (51Cr) are considered the gold standard (12); but they are costly to administer, time consuming, and pose a radiation risk to subjects, particularly if used repeatedly (7). On the other hand, carbon monoxide (CO) rebreathing has been shown to be a valid and reliable method for estimating
HEMOGLOBIN MASS STABILITY

983

as well as 8 and 10 min after rebreathing was commenced (6). Peak values of %HbCO were similar to those reported by Schmidt and Prommer (18) and Gore et al. (7). After the 2-min rebreathing period, the volume of CO not taken up by the body was calculated as the product of the rebreathing bag volume and the concentration of CO in it, which was measured using a hand-held analyzer with parts-per-million sensitivity (model 220, Fluke Electronics, Mississauga, Canada). Nine minutes after rebreathing was commenced, end-tidal CO was measured with the same CO analyzer and multiplied by the estimated alveolar ventilation (5.25 l/min) (19) to account for the CO exhaled after disconnecting from the spirometer until midway between the final two blood samples.

The total Hbmass was calculated for minute 9 as follows:

\[ \text{Hb}_\text{mass}(g) = \left( K \times \text{MCO} \times 100 / (\Delta \% \text{HbCO} \times 1.39) \right) \]

where \( K \) is (ambient barometric pressure mmHg \( \times 273 \/ 760 \) mmHg \( \times \) ambient temperature \( ^oK \)). MCO is CO volume administered into the system minus CO volume not bound to hemoglobin (calculated as the sum of CO volume remaining in the spirometer and the lung, as well as CO volume exhaled during the 7 min between disconnecting the subject from the spirometer and the final blood sample), which was then multiplied by 0.99 to correct for 1% loss of the CO dose to myoglobin (2, 16). \( \Delta \% \text{HbCO} \) is the difference between %HbCO at baseline and %HbCO in the blood samples 9 min after CO administration; and 1.39 is Hübner’s value for the CO binding capacity of hemoglobin (1.39 ml/g) (8). The 9-min concentration was calculated as the mean of the 8- and 10-min blood samples.

Statistics. Our inferences are based on precision of estimation via confidence limits rather than null-hypothesis testing via \( P \) values. All analyses of Hbmass were performed via log transformation to allow accurate estimation of effects, variabilities, and uncertainties in percent units. The standard error of measurement of Hbmass for each individual’s series was derived from consecutive pairwise change scores. From statistical first principles, the variance of the change scores is unaffected by any time-dependent linear trend in the raw scores and is equal to twice the error variance plus variance arising from any time-dependent nonlinear trend. The contribution of any such nonlinear trend to the variance is almost certainly negligible: even if it were as unrealistically large as approximately \( \pm 1\% \) between consecutive measurements, it would inflate a 2% error of measurement by a factor of only \( \sim 1.06 \). The error of measurement for each individual was therefore estimated by dividing the standard deviation of the consecutive pairwise changes in log-transformed Hbmass by \( \sqrt{2} \), then back transforming. Degrees of freedom for each of these estimates were the number of change scores minus 1. These six estimates were then meta-analyzed using the approach of Gore et al. (7) to derive estimates of the mean error and the between-subject typical variation (standard deviation) free of sampling variation.

The standard deviation of an individual’s set of Hbmass measurements represents what is sometimes defined as the total error of measurement. This measure would be inflated substantially by substantial linear or nonlinear time-dependent trends in Hbmass over the period of monitoring. The ratio of the total error to pairwise error was therefore derived for each individual. The mean of the log-transformed ratio was then compared with those of simulated series of measurements with error of 2% and sinusoidal periodicities of amplitude of 2% and periods of 0, 20, 25, 50, 100, 200, and 400 days (85 series for each period, the maximum that could be generated in the columns of an Excel spreadsheet). The simulated series of measurements were based on a total Hbmass of 1,000 g and an error of 20 g, which equates to a 2% error, similar to the 1.9% error of the CO method in our laboratory (7). Choice of amplitude equal to the error was based on consideration of making inferences about detecting signals (periodicities) equal to noise (error or measurement). Analyses were also conducted with simulated series of measurements with amplitudes of 1% and 0.5%, but other amplitudes were not investigated. The simulated measurements were generated in an Excel spreadsheet with random initial phase, random time between measurements of 1 and 4 days, and total duration of 101–104 days. The \( t \) statistic was used to estimate uncertainty in the ratio of the errors as 90% confidence limits and to calculate the probability that the ratio was less than that of the mean of each of the simulated series.

Each individual’s set of measurement was analyzed for linear trends and sinusoidal periodicities. Linear trends were derived by simple linear regression of log-transformed Hbmass, using a spreadsheet (available at http://www.sportsci.org/resource/stats/xvalid.xls). The slope and its confidence limits were back-transformed and expressed as percent change in Hbmass over 100 days. Periodicities in each individual’s series were quantified using the Lomb-Scargle algorithm in a Fourier analysis program (Peranso, version 2.1, http://www.peranso.com). The algorithm produces a spectral analysis consisting of a plot of power vs. period of the frequency components of a data stream with irregular time intervals. The magnitudes and frequencies of peaks in the spectrum of each subject’s series of Hbmass measurements were measured and plotted along with those of randomly chosen simulated series of measurements (12 for each sinusoidal period simulated) with periods of 25, 50, 100, and 400 days. Comparison of the real and simulated series in the plots was qualitative.

RESULTS

The number of measurements for each subject, as well as their respective Hbmass values, are shown in Table 1.

<table>
<thead>
<tr>
<th>Subject</th>
<th>( n )</th>
<th>Mean</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>47</td>
<td>1,175</td>
<td>19</td>
</tr>
<tr>
<td>2</td>
<td>41</td>
<td>867</td>
<td>22</td>
</tr>
<tr>
<td>3</td>
<td>40</td>
<td>1,056</td>
<td>16</td>
</tr>
<tr>
<td>4</td>
<td>41</td>
<td>896</td>
<td>22</td>
</tr>
<tr>
<td>5</td>
<td>40</td>
<td>1,003</td>
<td>27</td>
</tr>
<tr>
<td>6</td>
<td>45</td>
<td>806</td>
<td>38</td>
</tr>
</tbody>
</table>

Hbmass, hemoglobin mass; \( n \), number of measurements.

Our inferences are based on precision of estimation via confidence limits rather than null-hypothesis testing via \( P \) values. All analyses of Hbmass were performed via log transformation to allow accurate estimation of effects, variabilities, and uncertainties in percent units. The standard error of measurement of Hbmass for each individual’s series was derived from consecutive pairwise change scores. From statistical first principles, the variance of the change scores is unaffected by any time-dependent linear trend in the raw scores and is equal to twice the error variance plus variance arising from any time-dependent nonlinear trend. The contribution of any such nonlinear trend to the variance is almost certainly negligible: even if it were as unrealistically large as approximately \( \pm 1\% \) between consecutive measurements, it would inflate a 2% error of measurement by a factor of only \( \sim 1.06 \). The error of measurement for each individual was therefore estimated by dividing the standard deviation of the consecutive pairwise changes in log-transformed Hbmass by \( \sqrt{2} \), then back transforming. Degrees of freedom for each of these estimates were the number of change scores minus 1. These six estimates were then meta-analyzed using the approach of Gore et al. (7) to derive estimates of the mean error and the between-subject typical variation (standard deviation) free of sampling variation.

The standard deviation of an individual’s set of Hbmass measurements represents what is sometimes defined as the total error of measurement. This measure would be inflated substantially by substantial linear or nonlinear time-dependent trends in Hbmass over the period of monitoring. The ratio of the total error to pairwise error was therefore derived for each individual. The mean of the log-transformed ratio was then compared with those of simulated series of measurements with error of 2% and sinusoidal periodicities of amplitude of 2% and periods of 0, 20, 25, 50, 100, 200, and 400 days (85 series for each period, the maximum that could be generated in the columns of an Excel spreadsheet). The simulated series of measurements were based on a total Hbmass of 1,000 g and an error of 20 g, which equates to a 2% error, similar to the 1.9% error of the CO method in our laboratory (7). Choice of amplitude equal to the error was based on consideration of making inferences about detecting signals (periodicities) equal to noise (error or measurement). Analyses were also conducted with simulated series of measurements with amplitudes of 1% and 0.5%, but other amplitudes were not investigated. The simulated measurements were generated in an Excel spreadsheet with random initial phase, random time between measurements of 1 and 4 days, and total duration of 101–104 days. The \( t \) statistic was used to estimate uncertainty in the ratio of the errors as 90% confidence limits and to calculate the probability that the ratio was less than that of the mean of each of the simulated series.

Each individual’s set of measurement was analyzed for linear trends and sinusoidal periodicities. Linear trends were derived by simple linear regression of log-transformed Hbmass, using a spreadsheet (available at http://www.sportsci.org/resource/stats/xvalid.xls). The slope and its confidence limits were back-transformed and expressed as percent change in Hbmass over 100 days. Periodicities in each individual’s series were quantified using the Lomb-Scargle algorithm in a Fourier analysis program (Peranso, version 2.1, http://www.peranso.com). The algorithm produces a spectral analysis consisting of a plot of power vs. period of the frequency components of a data stream with irregular time intervals. The magnitudes and frequencies of peaks in the spectrum of each subject’s series of Hbmass measurements were measured and plotted along with those of randomly chosen simulated series of measurements (12 for each sinusoidal period simulated) with periods of 25, 50, 100, and 400 days. Comparison of the real and simulated series in the plots was qualitative.

RESULTS

The number of measurements for each subject, as well as their respective Hbmass values, are shown in Table 1.

Between-subject differences in measurement error. The errors of measurement derived from consecutive pairwise changes for the six subjects and expressed as coefficients of variation were 1.4, 1.7, 2.2, 2.3, 2.6 (the subject who had given a transfusion), and 2.7%. The 90% confidence limits for each subject’s true (large-sample) error was a \( \times 1/\varepsilon \) factor of \( \sim 1.20 \). (For example, the subject with an observed error of 1.4% could have a true error as high as 1.4 \( \times 1.2 = 1.7\% \) or as low as 1.4 \( \div 1.2 = 1.2\% \).) The mean error of measurement derived by meta-analysis of all six estimates was 2.1% (90% confidence interval = 1.7–2.6%), the between-subject typical variation in the error of measurement free of sampling variation was a \( \times 1/\varepsilon \) factor of 1.24 (90% confidence interval = 1.15–1.84).

Pairwise error vs. total error. Excluding the man who gave blood, five of the men showed little difference between the pairwise error (range 1.4–2.7%) and the total error of measurement (range 1.6–2.7%) for Hbmass. The mean ratio of these two errors was 1.06 (90% confidence interval = 0.96–1.17), which was probably less than the ratios from simulated series with periodicities of 20–100 days and 2% amplitude (ratios 1.13–1.21). However, when the amplitude was \( \leq 1\% \), or when the period was <20 or >100 days with a 2% amplitude, the
Hb\textsubscript{mass} data of the five men were indistinguishable from the simulated data.

**Periodicities and linear trends.** Figure 1 shows the raw values for Hb\textsubscript{mass} and corresponding periodograms for one normal subject, a randomly generated series with known sinusoidal periodicity, and the man who donated blood. Several periods are apparent in Fig. 1A, but their amplitudes are small compared with that of the peak in Fig. 1B, which has a period centered on the value that was simulated (50 days). The series of the blood donor produced a broad peak centered on a period of 200 days.

Plots of the amplitudes and periods of the peaks in the periodograms are shown in Fig. 2 for the real and simulated series. The plots in Fig. 2A show that the simulated series with amplitudes of 20 g (2%), noise of 20 g (2%), and periods of 25, 50, and 100 days produce obvious peaks corresponding to their periods, whereas the peaks that occurred in the real series of the five normal subjects had smaller amplitudes. The series of the blood donor (Fig. 1C) resembled a quarter cycle of a sinusoidal wave with period \sim 400 days and amplitude of \sim 16%. Figure 2B shows that this series and simulated series with the same amplitude, period, and noise produced spectra with peaks at \sim 200 days.

Linear regression analysis of the series of the five normal subjects produced estimates of linear change in Hb\textsubscript{mass} ranging from \sim 0.6 to 2.3% (90% confidence limits for each of the estimates is approximately \pm 1.9%) over the 100-day period. The blood donor showed a linear increase of 12.5% (90% confidence limits \pm 1.8%) over the same period.

**DISCUSSION**

This is the first study to investigate Hb\textsubscript{mass} with serial measurements every few days for several months using a validated technique (3, 18) that has the sensitivity to identify minor perturbations of just a few percent. The main finding was little evidence of biological variation, either cyclical or linear, in Hb\textsubscript{mass} >2% over 100 days. Our results imply the efficacy of Hb\textsubscript{mass} measurements to identify athletes who may seek to gain a performance advantage by infusion of blood cells or erythropoietic drugs.

Quantitative analysis of pairwise error vs. the total error did not show an increase in error over 100 days compared with a few days, which does not support the suggestion that error increases over time (7). The increase in error identified in the meta-analysis by Gore et al. (7) was therefore probably due to confounding factors arising from a correlation between measurement error and time between measurements in the meta-analyzed estimates. For example, a drop in Hb\textsubscript{mass} of 12% over \sim 32 days in a group of five control subjects in one of the meta-analyzed studies (17) must have been due entirely to measurement error that was much larger than in the present study.

Although qualitative analysis of the individuals’ data compared with the simulated data revealed no periodicity from 25 to 100 days, the individuals’ data could not be distinguished from simulated data with periodicities outside this range or with an error of 1% or less. Furthermore, the range in linear trends cannot exclude the possibility of \sim 1% cyclical variation in Hb\textsubscript{mass} over a longer period (e.g., annual). Consequently, variation in the magnitude of the linear change in Hb\textsubscript{mass} between the men we measured could reflect sampling at different points within an individual’s longer cycle.

The between-subject variation in the error of measurement at its lower confidence limit (a factor of 1.15) was greater than the
factor of 1.1 suggested as the smallest factor difference in a standard deviation derived from a meta-analysis of the reliability of estimating Hb_mass (7). It is therefore very likely that the difference in the error between the subjects was substantial.

**Implications and applications.** In 2000, Cazzola raised the concept of a “hematological passport” to monitor the blood of athletes for signs of manipulation; the passport comprises red blood cell counts, reticulocyte counts, serum ferritin, and soluble transferrin receptor (4). The advent of a relatively effective urine test for recombinant EPO (13) has led to athletes returning to homologous or autologous blood doping to artificially improve their performance, where transfusion of at least one unit of blood (~470 ml) is required to provide a performance benefit (5). Although a test for homologous transfusion is available (14, 15), a test for autologous doping is not. Hb_mass for improved oxygen transport is the key variable of interest for an athlete seeking an unfair advantage from transfusion, and it appears that the CO-rebreathing method has sufficient sensitivity to readily detect the removal and reinfusion of 1–2 units of homologous or autologous blood, since this volume would cause changes in Hb_mass ~5–10% above or below baseline. This magnitude of decrease or increase in Hb_mass is ~2 to 5 times more than the measurement error in experienced hands (6, 18). Moreover, a measure of Hb_mass via CO rebreathing could be readily incorporated into a hematological passport for elite endurance athletes, since it takes only 15 min to complete, uses small samples of blood (~0.6 ml), and is relatively inexpensive. However, variability in the measurement error between subjects and possible linear trends in Hb_mass must be considered when using this method to detect autologous blood doping in elite athletes. Future research is also required to determine the stability of Hb_mass in elite athletes undertaking intensive training.

**Time course of blood restitution.** Although no cyclical variation in Hb_mass of >2% was shown over the 100-day period in the five active males, the blood donor demonstrated gradual restoration in Hb_mass over this period and produced a spectral analysis consistent with a periodicity of 400 days and an amplitude of 16%. (The spectral analysis clearly underestimated the period of a sinusoidoid periodicity that is four times longer than the sampling period.) The only study that we could locate that directly measured red cell volume regeneration following blood withdrawal used 99mTc (10). Ninety-one subjects donated 440 ml of blood on two occasions >56 days apart. There was a 30-ml difference (~2%) in red cell volume after the second withdrawal, suggesting that red cell volume was not completely restored before the second donation. These results are in accord with those of our blood donor.

In conclusion, our results show that Hb_mass is stable (<2% variation) over ~100 days in subjects undergoing moderate training. If serial measurement of Hb_mass demonstrates the same stability with elite athletes, it could be part of a hematological passport to detect abuse of transfusions. Small differences in the error of measurement between subjects and the possibility of cyclical trends over longer periods than measured in this study need to be taken into account.

**ACKNOWLEDGMENTS**

We thank Julie Hill and Rebecca Toone for excellent technical assistance with data collection. We also thank our subjects who cheerfully cooperated with an onerous series of measurements.

**GRANTS**

This project was jointly funded by the South Australian Sports Institute and the Australian Sports Commission.

**REFERENCES**


