Inaccuracy of Estimated Resting Oxygen Uptake in the Clinical Setting

Nikhil Narang, MD; Jennifer T. Thibodeau, MD, MSc; Benjamin D. Levine, MD; M. Odette Gore, MD; Colby R. Ayers, MS; Richard A. Lange, MD, MBA; Joaquin E. Cigarroa, MD; Aslan T. Turer, MD, MHS; James A. de Lemos, MD; Darren K. McGuire, MD, MHS

Background—The Fick principle (cardiac output = oxygen uptake (V\textsubscript{O}\textsubscript{2})/systemic arterio-venous oxygen difference) is used to determine cardiac output in numerous clinical situations. However, estimated rather than measured V\textsubscript{O}\textsubscript{2} is commonly used because of complexities of the measurement, though the accuracy of estimation remains uncertain in contemporary clinical practice.

Methods and Results—From 1996 to 2005, resting V\textsubscript{O}\textsubscript{2} was measured via the Douglas bag technique in adult patients undergoing right heart catheterization. Resting V\textsubscript{O}\textsubscript{2} was estimated by each of 3 published formulae. Agreement between measured and estimated V\textsubscript{O}\textsubscript{2} was assessed overall, and across strata of body mass index, sex, and age. The study included 535 patients, with mean age 55 yrs, mean body mass index 28.4 kg/m\textsuperscript{2}; 53% women; 64% non-white. Mean (±standard deviation) measured V\textsubscript{O}\textsubscript{2} was 241 ± 57 ml/min. Measured V\textsubscript{O}\textsubscript{2} differed significantly from values derived from all 3 formulae, with median (interquartile range) absolute differences of 28.4 (13.1, 50.2) ml/min, 37.7 (19.4, 63.3) ml/min, and 31.7 (14.4, 54.5) ml/min, for the formulae of Dehmer, LaFarge, and Bergstra, respectively (P<0.0001 for each). The measured and estimated values differed by >25% in 17% to 25% of patients depending on the formula used. Median absolute differences were greater in severely obese patients (body mass index > 40 kg/m\textsuperscript{2}), but were not affected by sex or age.

Conclusions—Estimates of resting V\textsubscript{O}\textsubscript{2} derived from conventional formulae are inaccurate, especially in severely obese individuals. When accurate hemodynamic assessment is important for clinical decision-making, V\textsubscript{O}\textsubscript{2} should be directly measured. (Circulation. 2014;129:203-210.)

Key Words: cardiac output ■ catheterization ■ hemodynamics

A ccurate determination of cardiac output (Q\textsubscript{c}) is important in the hemodynamic evaluation of valve area, pulmonary and systemic vascular resistance, and severity of heart failure. The Fick method [Q\textsubscript{c} = oxygen uptake (V\textsubscript{O}\textsubscript{2})/systemic arterio-venous oxygen difference]\textsuperscript{1,2} is the time-honored gold standard for determining Q\textsubscript{c}, and has been used to validate other techniques such as indicator dilution and foreign gas rebreathing.\textsuperscript{3} Application of the Fick method requires measurement of V\textsubscript{O}\textsubscript{2}. However, direct measurement of V\textsubscript{O}\textsubscript{2}, through (1) mass spectrometry analysis of timed Douglas bag collections of exhaled air, or (2) breath-by-breath analysis of exhaled air using indirect calorimetry or metabolic cart analysis\textsuperscript{4,5} is time consuming and involves specific equipment that requires frequent calibration and is expensive to maintain. As a result, resting V\textsubscript{O}\textsubscript{2} is commonly estimated rather than measured using derived formulae available in the peer-reviewed literature.\textsuperscript{6-9} However, the accuracy of the formulae and nomograms most commonly used to estimate resting V\textsubscript{O}\textsubscript{2} is questionable, with most estimating methods derived from limited samples of highly selected, ethnically homogenous populations consisting of similarly aged, lean adults,\textsuperscript{10-12} populations that differ substantially from contemporary adult cardiology practice. Other formulae were derived from clinical populations composed exclusively or primarily of infants and children.\textsuperscript{5,8} Hence, we assessed the accuracy of estimated resting V\textsubscript{O}\textsubscript{2} compared with measured V\textsubscript{O}\textsubscript{2} obtained by the gold-standard analysis of timed collections of exhaled air by the method of Douglas in a large population of consecutive adult patients who underwent right-heart cardiac catheterization for clinical indications at our hospital.

Clinical Perspective on p 210

Received April 25, 2013; accepted September 23, 2013.

From the Department of Medicine, University of Chicago, Chicago, IL (N.N.); Department of Internal Medicine (J.T.T., B.D.L., M.O.G., A.T.T., J.A.d.L., D.K.M.) and Cardiovascular Division (J.T.T., B.D.L., M.O.G., A.T.T., J.A.d.L., D.K.M.), the University of Texas Southwestern Medical Center, Dallas, TX; Institute of Exercise and Environmental Medicine at the Texas Health Presbyterian Hospital, Dallas, TX (B.D.L.); the Department of Clinical Sciences at the University of Texas Southwestern Medical Center, Dallas, TX (C.R.A., D.K.M.); Cardiology Division, University of Texas Health Science Center at San Antonio, San Antonio, TX (R.A.L.); and the Cardiovascular Division, Oregon Health and Science University, Portland, OR (J.E.C.).

Correspondence to Darren K. McGuire, MD, MHS, Division of Cardiology, University of Texas Southwestern Medical Center, 5323 Harry Hines Blvd, Room E5.726, Dallas, TX, 75235–8830. E-mail darren.mcguire@utsouthwestern.edu

© 2013 American Heart Association, Inc.

Circulation is available at http://circ.ahajournals.org

DOI: 10.1161/CIRCULATIONAHA.113.003334
Methods
We conducted a retrospective study of consecutive patients who underwent right heart catheterization with direct measurement of resting VO$_2$ at Parkland Memorial Hospital between 1996 and 2005. Charts were reviewed for demographic, anthropometric, and baseline clinical characteristics. This study was approved by the University of Texas Southwestern Medical Center Institutional Review Board.

Calculating Estimated VO$_2$
Estimated resting VO$_2$ was calculated by the formula of Dehmer, et al.$^2$: VO$_2$ (ml/min) = 125 (ml/min/m$^2$) × body surface area (BSA, m$^2$), with BSA calculated according to the formula of Dubois$^{13}$: BSA (m$^2$) = 0.007184 × Weight (kg)$^{0.425}$ × Height (cm)$^{0.725}$. For sensitivity analysis, estimation of VO$_2$ was also calculated using the formula of LaFarge$^6$: VO$_2$ (ml/min) = 138.1 – (X × log age) + (0.378 × Heart Rate) × BSA (Men: X = 11.49; Women: X = 17.04); and the formula of Bergstra$^9$: VO$_2$ (ml/min) = 157.3 × BSA + X – (10.5 × log age) + 4.8 (Men: X = 10; Women = 0).

Direct Measurement of VO$_2$
Resting VO$_2$ was measured in all patients using the gold-standard technique of Douglas,$^1$ with analysis of a 3-minute collection of exhaled air collected through a properly fitted mouth piece with a 3-way valve. Exhaled volume was measured with a Tissot spirometer and concentrations of oxygen, carbon dioxide, and nitrogen were determined by mass spectrometry (Marquette MGA 1100), calibrated before every measurement, and all testing was completed while patients remained in the supine position.

Statistical Analysis
The magnitude of agreement between directly measured and estimated resting VO$_2$ was assessed by median absolute difference, ordinary least products regression, and typical error analysis.$^{15}$ Median absolute difference (ml/min) for the overall cohort was calculated as a median of the absolute value of the differences between measured and estimated resting VO$_2$ determined for each patient. The degree of disagreement between measured and estimated resting VO$_2$ was calculated as a percent error, dividing the absolute difference by the corresponding measured oxygen uptake, and multiplying by 100. Ordinary least products regression was used to assess both fixed and proportional error for the overall cohort.$^{14}$ Variance of comparative plot data points for estimated versus measured VO$_2$ in the overall cohort was assessed by intraclass correlation coefficient. Typical error estimation, expressed as a coefficient of variation derived from the standard deviation of the mean absolute difference divided by the square root of 2,$^{14}$ is reported for body mass index (BMI) strata. Median absolute difference between measured and estimated resting VO$_2$ was assessed in the overall cohort and in patients stratified by (1) BMI using clinical categories (<25, 25–29.9; 30–34.9; 35–39.9; ≥40 kg/m$^2$); (2) sex; and (3) age stratified by median split. The 1-sample Wilcoxon signed-ranks test was used to determine statistical significance for the median of the raw difference of values amongst BMI categories. The clinical relevance of errors was assessed by receiver operator curve analysis, estimation of VO$_2$ calculation, comparing results using VO$_2$ estimated by each of the 3 formulae versus measured VO$_2$. Diagnostic performance was assessed by receiver operator curve analysis comparing area under the curves, and by calculation of optimal cut points via Youden’s index$^{18}$ for each estimating formula to identify AVO$<_1.0$ cm$^2$ derived from direct VO$_2$ measurement. Sensitivity, specificity, and positive and negative predictive values were conventionally defined and compared to assess diagnostic accuracy. Attempts to develop a more accurate estimating equation using the present data included piecewise linear models, restricted cubic splines, additive models, variable transformations, including interaction terms, and higher-order variables. Cross-validated samples were obtained in 100 iterations, with 75% used as the training set and 25% used as the validation set. All testing was 2-tailed at a significance level of 0.05, with analyses performed using SAS Version 9.1.3 (Cary, NC) and no corrections made for multiple comparisons.

Results
Patient characteristics for the overall cohort and selected strata are shown in Table 1. The overall cohort comprised 535 patients, with mean age 55 years, 53% were women, and the mean BMI was 28.4 kg/m$^2$. The mean measured VO$_2$ for the overall cohort was 241±56.6 ml/min (mean ± standard deviation; Table 2), with a range of 108 to 457 ml/min. Using the Dehmer formula, the mean estimated VO$_2$ was 235.4 ± 32 ml/min, with a range of 162 to 356 ml/min; it differed significantly from the direct measurement with a median absolute difference of 28.4 (13.1, 50.2) ml/min [median (25th, 75th percentile)] (Table 2 and Figure 1A; $P<0.0001$).

Analysis of agreement in the overall cohort between directly measured VO$_2$ and VO$_2$ estimated by the Dehmer formula using ordinary least products regression demonstrated significant fixed error [reflected by y intercept >0 (95% confidence interval, 92.4–106.1)] and proportional error [reflected by slope <1.0 (95% CI, 0.53–0.61; Figure 1A)]. Poor agreement between directly measured and estimated VO$_2$ was also observed when the LaFarge and Bergstra formulae were used (Table 2; Figure 1B and 1C). Intraclass correlation coefficients were similar and slightly higher in the LaFarge and Bergstra formulae, when compared with the Dehmer formula, demonstrating slightly more consistency overall in intra-test measurement (Figure 1).

The magnitude of error between measured and estimated VO$_2$ expressed as a percent error for the overall cohort is shown in Figure 2. For all 3 formulae used to estimate VO$_2$, the degree of error ranged from 10 to 25% in ≈40% of the overall cohort, and the error was >25% in 17–25% of the cohort, depending on the estimating formula used.

The median absolute difference between measured and estimated VO$_2$ using the 3 formulae was stratified by BMI, sex, and age, as shown in Table 2. When estimating VO$_2$ using the Dehmer formula, the difference in measured and estimated VO$_2$ widened as BMI increased, with significant disagreement between measured and estimated VO$_2$ observed in all BMI strata. Patients with BMI <40 kg/m$^2$ had a median absolute difference ranging from 24.6 (11.6, 43.7) ml/min to 29.3 (14.8, 57.8) ml/min, whereas median absolute difference was significantly higher at 47.0 (29.5, 83.4) ml/min in the ≥40 kg/m$^2$ group when compared with the other strata ($P_{interaction}<0.0001$). Similar results were observed in both the LaFarge and Bergstra formulae ($P_{interaction}=0.001$ and $P_{interaction}=0.005$, respectively). When analyzed using typical error analysis, the error associated with estimating resting VO$_2$ was similarly magnified in the ≥ 40 kg/m$^2$ (Table 2).

Both measured and estimated VO$_2$ were higher in men compared with women (Table 2). In sex-stratified analyses, median absolute differences between measured and estimated resting VO$_2$ were large in both men and women for all
Table 1. Baseline Descriptive Characteristics

<table>
<thead>
<tr>
<th></th>
<th>Overall (n=535)</th>
<th>Men (n=252)</th>
<th>Women (n=283)</th>
<th>&lt;25 (n=187)</th>
<th>25–29.9 (n=174)</th>
<th>30–34.9 (n=94)</th>
<th>35–39.9 (n=41)</th>
<th>≥40 (n=39)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age, mean (SD), y</td>
<td>55 (13.5)</td>
<td>55 (13.6)</td>
<td>55 (13.5)</td>
<td>54 (15.4)</td>
<td>57 (12)</td>
<td>54 (13.1)</td>
<td>54 (12.7)</td>
<td>55 (11.2)</td>
</tr>
<tr>
<td>Women, n (%)</td>
<td>283 (53)</td>
<td>—</td>
<td>—</td>
<td>93 (49.7)</td>
<td>88 (47.1)</td>
<td>51 (54.3)</td>
<td>25 (61)</td>
<td>32 (82.1)</td>
</tr>
<tr>
<td>Race/Ethnicity, n (%)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>White</td>
<td>192 (36)</td>
<td>101 (40)</td>
<td>91 (32)</td>
<td>65 (34.8)</td>
<td>65 (37.4)</td>
<td>38 (40.4)</td>
<td>11 (26.8)</td>
<td>13 (33.3)</td>
</tr>
<tr>
<td>Black</td>
<td>168 (31)</td>
<td>72 (29)</td>
<td>96 (34)</td>
<td>46 (24.6)</td>
<td>53 (30.4)</td>
<td>32 (34)</td>
<td>19 (46.3)</td>
<td>18 (46.2)</td>
</tr>
<tr>
<td>Hispanic</td>
<td>143 (27)</td>
<td>64 (25)</td>
<td>79 (28)</td>
<td>58 (31)</td>
<td>47 (27)</td>
<td>20 (21.3)</td>
<td>10 (24.4)</td>
<td>8 (20.5)</td>
</tr>
<tr>
<td>Other</td>
<td>32 (6)</td>
<td>15 (6)</td>
<td>17 (6)</td>
<td>18 (9.6)</td>
<td>9 (5.2)</td>
<td>4 (4.3)</td>
<td>1 (2.4)</td>
<td>—</td>
</tr>
<tr>
<td>Height, mean (SD), in</td>
<td>66.0 (4.2)</td>
<td>68.9 (3.4)</td>
<td>63.4 (2.9)</td>
<td>66.2 (4.1)</td>
<td>66.2 (4.1)</td>
<td>66.1 (4.5)</td>
<td>65.9 (4)</td>
<td>64.5 (3.9)</td>
</tr>
<tr>
<td>Weight, mean (SD), kg</td>
<td>79.9 (21.3)</td>
<td>83.7 (19.7)</td>
<td>76.5 (22.1)</td>
<td>61.9 (10.1)</td>
<td>78 (10)</td>
<td>90.8 (12.7)</td>
<td>104 (13)</td>
<td>123.2 (19.7)</td>
</tr>
<tr>
<td>BMI, mean (SD), kg/m²</td>
<td>28.4 (7.1)</td>
<td>27.2 (5.5)</td>
<td>29.4 (8.2)</td>
<td>21.8 (2.1)</td>
<td>27.5 (1.4)</td>
<td>32.1 (1.4)</td>
<td>37.1 (1.3)</td>
<td>45.8 (5.4)</td>
</tr>
<tr>
<td>Valvular heart disease, n (%)</td>
<td>383 (72)</td>
<td>177 (70.2)</td>
<td>213 (65.3)</td>
<td>136 (72.7)</td>
<td>135 (77.6)</td>
<td>59 (62.8)</td>
<td>26 (63.4)</td>
<td>27 (69.2)</td>
</tr>
<tr>
<td>Pulmonary hypertension, n (%)</td>
<td>166 (31)</td>
<td>68 (27)</td>
<td>100 (35.3)</td>
<td>65 (34.8)</td>
<td>49 (28.2)</td>
<td>26 (27.7)</td>
<td>10 (24.4)</td>
<td>16 (41)</td>
</tr>
<tr>
<td>CVD, n (%)</td>
<td>143 (27)</td>
<td>90 (35.7)</td>
<td>57 (20.1)</td>
<td>50 (26.7)</td>
<td>52 (29.9)</td>
<td>26 (27.7)</td>
<td>10 (24.4)</td>
<td>9 (22)</td>
</tr>
<tr>
<td>Heart failure, n (%)</td>
<td>102 (19)</td>
<td>64 (25.4)</td>
<td>40 (14.1)</td>
<td>33 (17.6)</td>
<td>37 (21.3)</td>
<td>21 (22.3)</td>
<td>5 (12.2)</td>
<td>6 (15.4)</td>
</tr>
<tr>
<td>Hypertension, n (%)</td>
<td>114 (23.1)</td>
<td>55 (21.8)</td>
<td>59 (20.8)</td>
<td>23 (12.3)</td>
<td>37 (21.3)</td>
<td>27 (28.7)</td>
<td>13 (31.7)</td>
<td>14 (35.9)</td>
</tr>
<tr>
<td>Diabetes mellitus, n (%)</td>
<td>63 (11.8)</td>
<td>32 (12.7)</td>
<td>31 (11)</td>
<td>13 (7)</td>
<td>16 (9.2)</td>
<td>17 (18.1)</td>
<td>6 (14.6)</td>
<td>10 (25.6)</td>
</tr>
<tr>
<td>Primary indication for RHC (%)</td>
<td>292 (54.6)</td>
<td>131 (52)</td>
<td>161 (56.9)</td>
<td>106 (56.7)</td>
<td>100 (57.5)</td>
<td>43 (45.8)</td>
<td>20 (48.8)</td>
<td>23 (59)</td>
</tr>
<tr>
<td>Valvular heart disease</td>
<td>177 (33.1)</td>
<td>97 (38.5)</td>
<td>80 (28.3)</td>
<td>54 (28.9)</td>
<td>59 (33.9)</td>
<td>32 (34)</td>
<td>20 (48.8)</td>
<td>12 (30.8)</td>
</tr>
<tr>
<td>Heart failure</td>
<td>46 (8.6)</td>
<td>21 (8.3)</td>
<td>25 (8.8)</td>
<td>22 (11.8)</td>
<td>8 (4.6)</td>
<td>13 (13.8)</td>
<td>—</td>
<td>3 (7.7)</td>
</tr>
<tr>
<td>Congential heart disease</td>
<td>20 (3.7)</td>
<td>3 (1.2)</td>
<td>17 (6)</td>
<td>5 (2.7)</td>
<td>7 (4)</td>
<td>6 (6.4)</td>
<td>1 (2.4)</td>
<td>1 (2.6)</td>
</tr>
</tbody>
</table>

BMI indicates body mass index; CVD, cardiovascular disease; RHC, right heart catheterization; and SD, standard deviation.

To demonstrate the clinical importance of error in resting VO₂ estimation, we calculated hypothetical Fick-derived Qc based on resting and estimated VO₂ from all 3 formulae for each patient in the study. Fick-derived Qc values were used to derive hypothetical AVA by the Hakki equation,¹⁸ as a clinical example to determine the effect of errors in resting VO₂ estimation. For this analysis, each patient was assigned an hypothetical mean aortic valve gradient of 40 mm Hg and an arteriovenous oxygen difference of 4.5. An AVA of <1.0 cm² is classified as severe stenosis and warrants consideration for surgical correction. Using AVA <1.0 cm² derived from directly measured VO₂ as the outcome, the Dehmer formula had a sensitivity of 93% (95% CI, 90–95%) with a specificity of 33% (95% CI, 25–40%) and an area under the curve (AUC) of 0.79 (95% CI 0.75–0.84; Table 3). The LaFarge formula had similar sensitivity, specificity, and AUC when compared with the Dehmer formula; the Bergstra formula had the lowest sensitivity of 75% (95% CI, 71–79%) of the formulae tested, but was significantly more specific (77%; 95% CI, 68–86%) with a slightly greater AUC of 0.82 (95% CI, 0.77–0.87) when compared with the other formulae (Figure 3). Similarly, the optimal ROC cut point for the Bergstra formula (0.94 cm²) was the closest of the 3 formulae to the clinically important diagnostic cutoff of 1.0 cm² (Table 3). Expanded clinical application of

3 formulae tested. Although statistically significant, the difference in error between sexes was small (<10 ml/min) and varied in direction depending on the estimating formula used.

In analyses stratified by median age, median absolute difference between estimated and measured VO₂ comparisons for age groups ≤55 years and >55 years were both large [31.7 (13.5, 55.2) ml/min and 26.5 (12.7, 46.6) ml/min, respectively, P<0.0001 for each] when calculated by the Dehmer formula, though the differences between the 2 groups did not reach statistical significance (P=0.13). In assessment of both LaFarge and Bergstra formulae, the median absolute difference for both age strata was >30 ml/min, and the differences between age groups similarly did not reach statistical significance (Table 2; P=0.92 and P=0.20, respectively).

We were unable to resolve or improve upon the discordance between measured and estimated resting VO₂ using data-derived estimating equations, which we explored using a variety of methods of multivariable linear regression. In all cases of the exploratory models, the mean bias was unacceptably large, and the median percent difference was no less than 15%. Additionally, the predicted R² based on the PRESS statistic¹⁷ yielded values no larger than 0.27 (data not shown), demonstrating unacceptably poor model performance.
the errors encountered in resting VO₂ estimation are shown in Figures 4 and 5 represented by theoretical Fick cardiac output and systemic vascular resistance (hypothetical mean arterial pressure of 75 mm Hg, central venous pressure of 8 mm Hg, arteriovenous oxygen difference of 4.5), where a large majority of data points deviate notably from the line of equality.

Table 2. Comparison of Median Absolute Difference and Typical Error (BMI Only) by Estimating Formulae for Overall Cohort, and Subgroups Stratified by Sex, Age, and Body Mass Index

<table>
<thead>
<tr>
<th>Group</th>
<th>Dehmer Formula†</th>
<th>LaFarge Formula‡</th>
<th>Bergstra Formula§</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean Measured VO₂ (ml/min)</td>
<td>Mean Calc. VO₂ (SD), ml/min</td>
<td>Median* Absolute Difference [IQR], ml/min</td>
</tr>
<tr>
<td>Overall (n=535)</td>
<td>241 (56.6)</td>
<td>235.4 (32)</td>
<td>28.4</td>
</tr>
<tr>
<td>Men (n=252)</td>
<td>260.6 (58.3)</td>
<td>248.1 (29.7)</td>
<td>31.3</td>
</tr>
<tr>
<td>Women (n=283)</td>
<td>224.2 (49.1)</td>
<td>224.1 (29.6)</td>
<td>24</td>
</tr>
<tr>
<td>Age ≤ 55 years (n=255)</td>
<td>250.1 (61.6)</td>
<td>239.9 (34.3)</td>
<td>31.7</td>
</tr>
<tr>
<td>Age &gt; 55 years (n=280)</td>
<td>233.4 (50.3)</td>
<td>234 (29.6)</td>
<td>26.5</td>
</tr>
<tr>
<td>BMI, kg/m²</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt;25</td>
<td>218.6 (48)</td>
<td>212.6 (23.3)</td>
<td>24.6</td>
</tr>
<tr>
<td>25.29.9</td>
<td>242.4 (51.3)</td>
<td>235 (22.8)</td>
<td>26.8</td>
</tr>
<tr>
<td>30–34.9</td>
<td>258.1 (55.5)</td>
<td>250.4 (26.9)</td>
<td>29.3</td>
</tr>
<tr>
<td>35–39.9</td>
<td>265.3 (63.3)</td>
<td>264.6 (25.7)</td>
<td>27</td>
</tr>
<tr>
<td>≥40</td>
<td>280.1 (67.8)</td>
<td>279.5 (28.8)</td>
<td>47</td>
</tr>
</tbody>
</table>

BM indicates body mass index; Calc, calculated; CV, coefficient of variation; IQR, interquartile range; SD, standard deviation; and VO₂, oxygen uptake.

*P<0.005 for median raw difference for overall cohort and subgroups determined by one-sample Wilcoxon signed-ranks test
†Dehmer Formula: VO₂ (ml/min) = 125 × Body Surface Area
‡LaFarge Formula: VO₂ (ml/min) = 138.1 – (11.49 × loge age) + (0.378 × Heart Rate) × Body Surface Area (Men); VO₂ (ml/min) = 138.1 – (17.04 × loge age) + (0.378 × Heart Rate) × Body Surface Area (Women).
§Bergstra Formula: VO₂ (ml/min) = 157.3 × Body Surface Area + 10 – (10.5 × loge age) + 4.8 (Men); VO₂ (ml/min) = 157.3 × Body Surface Area – (10.5 × loge age) + 4.8 (Women).

Figure 1. Comparative plot of measured VO₂ vs estimated VO₂ for overall cohort using (A) Dehmer, (B) LaFarge, and (C) Bergstra formulae. Dashed line indicates line of equality; solid line, ordinary least products regression line. CI indicates confidence interval; ICC, intraclass correlation coefficient; and VO₂, oxygen uptake.
Discussion

This study demonstrates that in a large, consecutive sample of adult patients undergoing right-heart cardiac catheterization, estimation of resting \( \dot{V}O_2 \) by commonly used formulae is inaccurate compared with the gold-standard analysis of timed collections of exhaled air. Estimated \( \dot{V}O_2 \) was most inaccurate in morbidly obese patients with BMI \( \geq 40 \) kg/m\(^2\). To put the degree of inaccuracy of resting \( \dot{V}O_2 \) estimation into clinical context, the observed error and variance in estimated \( \dot{V}O_2 \) applied to Fick-calculated \( \dot{Q}_c \) yields up to 38% error within 1 standard deviation (SD) and up to 64% error within 2 SDs. Because Fick-calculated \( \dot{Q}_c \) and resting \( \dot{V}O_2 \) are directly proportional, error in estimation of resting \( \dot{V}O_2 >25\% \), a magnitude of error observed in 17% of the present study cohort, can dramatically alter clinical decision-making.

Formulae and nomograms for the estimation of resting \( \dot{V}O_2 \), published as early as 1954, were derived from highly selected cohorts of homogeneous ethnicity and age. More recently published estimating formulae presently in broader clinical use, such as that by LaFarge and colleagues,\(^6\) were derived primarily from pediatric patients. Similarly, many of the patients used to derive the Bergstra formula\(^7\) were infants and children with congenital heart disease, possibly confounding the application of these methods in adult patients. Previous studies have demonstrated errors in \( \dot{V}O_2 \) estimation using these and other formulae in pediatric\(^19\text{--}22\) and adult\(^23\text{--}25\) populations. However, these analyses have limitations, including small sample size with homogeneity of patient population, lack of sub-group analysis in determining other variables that may influence errors in estimation, and use fewer analytic methods to explain the errors observed. Our study builds on previous analyses by better addressing the above-mentioned limitations, allowing for a comprehensive assessment of errors in resting \( \dot{V}O_2 \) estimation not seen to date. In the present study, we observed a substantial fixed and proportional error when \( \dot{V}O_2 \) was estimated by each of the 3 formulae compared with analysis of timed collections of exhaled air, with error in excess of 25% in many patients.

We previously demonstrated the inaccuracy of resting \( \dot{V}O_2 \) estimation in a smaller cohort of research participants.\(^25\) The present observations confirm and extend these previous findings of exaggerated error in the most obese patients, now in a larger sample of adult patients in a clinical setting where such methods are commonly applied. Several of the contemporary formulae for estimating \( \dot{V}O_2 \)\(^6\text{,}8\) incorporate weight or body surface area without taking into account the degree of adiposity, which may impact the accuracy of estimation, because fat has little impact on oxygen uptake. In addition, there is a metabolic requirement for accommodating excess weight, including physical support and respiratory efforts, all of which may materially influence resting \( \dot{V}O_2 \). Although predictive formulae for maximal \( \dot{V}O_2 \) have adjusted for the metabolic cost associated with excess adiposity to improve accuracy,\(^26\text{,}27\) whether similar adiposity adjustments for the estimation of \( \dot{V}O_2 \) at rest could improve accuracy remains to be determined.

In contrast to our previous study where estimation with mean absolute difference was exaggerated in men compared with women, we found no significant difference in observed disagreement between measured and estimated

---

Table 3. Diagnostic Performance by Estimating Formula for Calculation of AVA <1.0 cm\(^2\)

<table>
<thead>
<tr>
<th></th>
<th>Dehmer</th>
<th>LaFarge</th>
<th>Bergstra</th>
</tr>
</thead>
<tbody>
<tr>
<td>Optimal ROC cut</td>
<td>0.87</td>
<td>0.78</td>
<td>0.94</td>
</tr>
<tr>
<td>(cm(^2))</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AUC</td>
<td>0.79 (0.75–0.84)</td>
<td>0.79 (0.74–0.84)</td>
<td>0.82 (0.77–0.87)</td>
</tr>
<tr>
<td>Sensitivity</td>
<td>93% (90.1–95.3)</td>
<td>96% (93.6–97.5)</td>
<td>75% (71.2–79.2)</td>
</tr>
<tr>
<td>Specificity</td>
<td>33% (25.2–40.2)</td>
<td>30% (20.3–39.5)</td>
<td>77% (68.2–85.9)</td>
</tr>
<tr>
<td>PPV</td>
<td>78% (74.2–81.7)</td>
<td>88% (84.6–90.5)</td>
<td>94% (92–96.8)</td>
</tr>
<tr>
<td>NPV</td>
<td>64% (52.9–74.4)</td>
<td>57% (42.2–70.9)</td>
<td>38% (30.5–44.8)</td>
</tr>
<tr>
<td>Positive LR</td>
<td>1.4 (1.23–1.55)</td>
<td>1.4 (1.19–1.57)</td>
<td>3.3 (2.22–4.83)</td>
</tr>
<tr>
<td>Negative LR</td>
<td>0.2 (0.15–0.34)</td>
<td>0.15 (0.09–0.26)</td>
<td>0.32 (0.26–0.39)</td>
</tr>
</tbody>
</table>

Parentheses indicate 95% confidence interval. AUC indicates area under the curve; LR, likelihood ratio; NPV, negative predictive value; PPV, positive predictive value; and ROC, receiver operating characteristic.

---

Figure 2. Distribution of error related to use of estimated \( \dot{V}O_2 \) (n=535) across estimating formulae tested.

Figure 3. ROC curves by estimating formula for AVA<1.0 cm\(^2\) (measured \( \dot{V}O_2 \)) as the outcome and AVA<1.0 cm\(^2\) (estimated \( \dot{V}O_2 \)) as the predictor. AVA indicates aortic valve area; cm, centimeters; ROC, receiver operating characteristic; and \( \dot{V}O_2 \), oxygen uptake.
VO₂ at rest when stratified by sex in the present study. We postulate that results from our previous study may be explained by the fact that on average, men have greater absolute and proportional fat-free mass than women, which likely contributed to the sex-based error of the estimating formula. In the present analyses, the error was numerically greater in men compared with women using the Dehmer and Bergstra formulae, whereas with the LaFarge formula the error was numerically greater in women. However, the magnitude of differences was not statistically significant between the groups, and none of the formulae yielded a significant statistical interaction between error and sex, challenging the necessity of inclusion of sex as is done in the LaFarge and Bergstra formulae.

Although early nomograms for estimating VO₂ at rest, some of which were derived from pediatric populations, were stratified by age, we found no age-based error association in this study. No association between the degree of error of the estimating formula and age was present when assessing VO₂ using each of the 3 formulae. Similar results were found in studies of both older adults with mean age 60 years and younger adults with mean age 39 years.

The errors encountered when resting VO₂ is estimated instead of directly measured can potentially impact clinical decision making, including determining the initiation and titration of inotropic support, decision for mechanical ventricular support, determining eligibility and monitoring response of pulmonary vasodilator therapy, and determining candidacy for valve procedures, among others.

Clinical Implications
The clinical relevance of the observed errors in VO₂ estimation as depicted by hypothetical hemodynamic calculations demonstrates the potential impact on critical clinical decision making, as derivations of AVA by all 3 estimating formulae revealed substantial inaccuracies in diagnostic performance. The Dehmer and Large formulae were adequately sensitive in detection of potential severe aortic stenosis but were poorly specific. Using the Dehmer and LaFarge formulae, a significant number of subjects with AVA >1.0 cm² derived from directly measured VO₂ were found to have estimated AVAs <1.0 cm², reflecting the high false positive rate. The Bergstra formula was significantly more specific than the
other formulae tested, but overall had the largest combined proportion of false negative and false positive results, leading to potential clinical misclassification of AVA when estimated to be both greater and less than 1.0 cm². We similarly observed this critical degree of potential clinical error in our previous study, which was composed of research study participants and not actual patients, contrasted with the present study of patients who underwent cardiac catheterization for clinical indications.

From a drug regulatory standpoint, estimation versus measurement of VO₂ is also an important consideration. In 2010, the FDA Cardiovascular and Renal Drugs Advisory Committee convened to discuss the potential role of using pulmonary vascular resistance index (PVRI) and its change in response to therapy as a surrogate for drug efficacy in pediatric and adult PAH patients, virtually all of the available data from clinical trials had been compiled using estimating formulae for resting VO₂, a key parameter in the calculation of PVRI.²⁹ Given the degree of error of the estimation of resting VO₂ in the present study, and the absence of validation or critical assessment in the pediatric population, the routine use of resting VO₂ estimation in the calculation of hemodynamics and their response to therapy for drug registration studies should be closely scrutinized, if not abandoned.

Limitations
Our study has limitations. Differences in methods between operators with regard to exhaled air collection and analysis could have influenced the accuracy of resting VO₂ measurement using the method of Douglas. However, to address this, a specific protocol was followed by all operators and the equipment was routinely maintained and calibrated. In addition, although patients were stationary and in a steady state condition before timed collections of exhaled air, it is unlikely that the method equates to a true resting VO₂ measurement given the anxiety associated with cardiac catheterization and the variable clinical stability across a population of patients undergoing right heart catheterization for clinical purposes.

Conclusions
Errors in hemodynamic assessment may adversely impact clinical assessment and therapeutic decision-making across a spectrum of serious cardiovascular conditions. Given the imperative in such situations for accuracy of assessment, if the Fick method is to be used for Qc estimation, resting VO₂ should be directly measured and not estimated.

Sources of Funding
This study was supported by the clinical research fellowship from the Doris Duke Charitable Foundation. The funding organization did not participate in study design, data collection, analysis, interpretation, or writing of this report, or the decision to submit this paper for publication. The corresponding author had full access to all of the data and the accuracy of the data analysis, and had final responsibility for the decision to submit for publication. Dr. Narang is supported by a clinical research fellowship from the Doris Duke Charitable Foundation.

Disclosures
None.

References
The Fick method is the gold-standard for determining cardiac output in numerous clinical scenarios. A primary determinant of Fick-derived cardiac output is resting oxygen uptake ($V_{\text{O}_2}$), which if inaccurately estimated will proportionally manifest as commensurate error in estimation of cardiac output, a hemodynamic parameter directly influencing critical clinical decision-making. Our study demonstrates the inaccuracies of estimating $V_{\text{O}_2}$ with commonly used formulae compared with its direct measurement. It is important to consider these limitations when calculating cardiac output based on estimated $V_{\text{O}_2}$, especially in obese and severely obese individuals as commonly encountered in contemporary practice. When accurate hemodynamic assessment is important for clinical decision-making, $V_{\text{O}_2}$ should be directly measured.
Inaccuracy of Estimated Resting Oxygen Uptake in the Clinical Setting
Nikhil Narang, Jennifer T. Thibodeau, Benjamin D. Levine, M. Odette Gore, Colby R. Ayers, Richard A. Lange, Joaquin E. Cigarroa, Aslan T. Turer, James A. de Lemos and Darren K. McGuire

Circulation. 2014;129:203-210; originally published online September 27, 2013; doi: 10.1161/CIRCULATIONAHA.113.003334

Circulation is published by the American Heart Association, 7272 Greenville Avenue, Dallas, TX 75231
Copyright © 2013 American Heart Association, Inc. All rights reserved.
Print ISSN: 0009-7322. Online ISSN: 1524-4539

The online version of this article, along with updated information and services, is located on the World Wide Web at:
http://circ.ahajournals.org/content/129/2/203

Permissions: Requests for permissions to reproduce figures, tables, or portions of articles originally published in Circulation can be obtained via RightsLink, a service of the Copyright Clearance Center, not the Editorial Office. Once the online version of the published article for which permission is being requested is located, click Request Permissions in the middle column of the Web page under Services. Further information about this process is available in the Permissions and Rights Question and Answer document.

Reprints: Information about reprints can be found online at:
http://www.lww.com/reprints

Subscriptions: Information about subscribing to Circulation is online at:
http://circ.ahajournals.org//subscriptions/