JAMA Cardiology | Original Investigation

Thermodilution vs Estimated Fick Cardiac Output Measurement in Clinical Practice

An Analysis of Mortality From the Veterans Affairs Clinical Assessment, Reporting, and Tracking (VA CART) Program and Vanderbilt University

Alexander R. Opotowsky, MD, MPH; Edward Hess, MS; Bradley A. Maron, MD; Evan L. Brittain, MD, MSc; Anna E. Barón, PhD; Thomas M. Maddox, MD, MSc; Laith I. Alshawabkeh, MD, MSc; Bradley M. Wertheim, MD; Meng Xu, MS; Tufik R. Assad, MD; Jonathan D. Rich, MD; Gaurav Choudhary, MD; Ryan J. Tedford, MD

IMPORTANCE Thermodilution (Td) and estimated oxygen uptake Fick (eFick) methods are widely used to measure cardiac output (CO). They are often used interchangeably to make critical clinical decisions, yet few studies have compared these approaches as applied in medical practice.

OBJECTIVES To assess agreement between Td and eFick CO and to compare how well these methods predict mortality.

DESIGN, SETTING, AND PARTICIPANTS This investigation was a retrospective cohort study with up to 1 year of follow-up. The study used data from the Veterans Affairs Clinical Assessment, Reporting, and Tracking (VA CART) program. The findings were corroborated in a cohort of patients cared for at Vanderbilt University, an academic referral center. Participants were more than 15 000 adults who underwent right heart catheterization, including 12 232 in the Veterans Affairs cohort between October 1, 2007, and September 30, 2013, and 3391 in the Vanderbilt cohort between January 1, 1998, and December 31, 2014.

EXPOSURES A single cardiac catheterization was performed on each patient with CO estimated by both Td and eFick methods. Cardiac output was indexed to body surface area (cardiac index [CI]) for all analyses.

MAIN OUTCOMES AND MEASURES All-cause mortality over 90 days and 1 year after catheterization.

RESULTS Among 12 232 VA patients (mean [SD] age, 66.4 [9.9] years; 3.3% female) who underwent right heart catheterization in this cohort study, Td and eFick CI estimates correlated modestly (r = 0.65). There was minimal mean difference (eFick minus Td = −0.02 L/min/m^2, or −0.4%) but wide 95% limits of agreement between methods (−1.3 to 1.3 L/min/m^2, or −50.1% to 49.4%). Estimates differed by greater than 20% for 38.1% of patients. Low Td CI (<2.2 L/min/m^2 compared with normal CI of 2.2-4.0 L/min/m^2) more strongly predicted mortality than low eFick CI at 90 days (Td hazard ratio [HR], 1.71; 95% CI, 1.47-1.99; χ^2 = 49.5 vs eFick HR, 1.42; 95% CI, 1.22-1.64; χ^2 = 20.7) and 1 year (Td HR, 1.53; 95% CI, 1.39-1.69; χ^2 = 71.5 vs eFick HR, 1.35; 1.22-1.49; χ^2 = 35.2). Patients with a normal CI by both methods had 12.3% 1-year mortality. There was no significant additional risk for patients with a normal Td CI but a low eFick CI (12.9%, P = .51), whereas a low Td CI but normal eFick CI was associated with higher mortality (15.4%, P = .001). The results from the Vanderbilt cohort were similar in the context of a more balanced sex distribution (46.6% female).

CONCLUSIONS AND RELEVANCE There is only modest agreement between Td and eFick CI estimates. Thermodilution CI better predicts mortality and should be favored over eFick in clinical practice.

JAMA Cardiol. 2017;2(10):1090-1099. doi:10.1001/jamacardio.2017.2945
Published online September 6, 2017.

Supplemental content

Author Affiliations: Author affiliations are listed at the end of this article.

Corresponding Author: Alexander R. Opotowsky, MD, MPH, Department of Cardiology, Boston Children's Hospital, 300 Longwood Ave, Boston, MA 02115 (alexander.opotowsky@cardio.chboston.org).

Accurate measurement of cardiac output (CO) is central to diagnosis, management, and definition of prognosis across a diverse range of medical contexts. Cardiac output measurement is most commonly performed using either the thermodilution (Td) technique or the estimated oxygen uptake Fick (eFick) method.[1-5] The accuracy of eFick hinges on correctly estimating oxygen consumption ($\dot{V}O_2$). The $\dot{V}O_2$ estimates, and consequently eFick CO, are particularly prone to error in the setting of pulmonary hypertension, heart failure, or abnormal body habitus.[6-9] The accuracy of Td CO, where a thermal registering device measures changes in temperature distal to proximal injection of saline with known temperature and volume, can be confounded in the presence of an intracardiac shunt or tricuspid regurgitation.[10-14] Discrepancies in CO estimates between methods are common and may be associated with diagnostic misclassification.[15] Imprecise CO measurements have ramifications for a wide range of clinical decision making.[8,16-19] However, there are few data, to our knowledge, comparing the performance of Td and eFick against clinical end points, such as mortality.

Research comparing CO measurement techniques has focused on experimental validity in highly selected study samples.[5,12,14,15,20] Neither the agreement between Td and eFick CO nor the comparative prognostic value of these methods as applied in clinical practice is known to date. To bridge this knowledge gap, we systematically analyzed correlation and agreement between Td and eFick, as well as the association of each method with survival in 2 independent, large patient cohorts. This study aimed to address the pressing need for an evidence-based approach to clinical measurement of CO, with important implications for patients with cardiopulmonary disease, sepsis, and other disorders.

Key Points

Question Which commonly applied method to measure cardiac output, thermodilution or Fick using estimated oxygen consumption, performs better in routine clinical practice?

Findings Among more than 15 000 adults in this cohort study who underwent right heart catheterization, thermodilution and estimated Fick cardiac output measurements agreed poorly, with estimates differing by greater than 20% in well over one-third of patients. Thermodilution estimates of cardiac output were more strongly associated with mortality than estimated Fick cardiac output estimates.

Meaning Thermodilution and estimated Fick cardiac output estimates should not be considered interchangeable; thermodilution is preferable for most situations in clinical practice.

Methods

Data Sources

The Veterans Affairs Clinical Assessment, Reporting, and Tracking (VA CART) program uses a software application embedded in the VA electronic health record (EHR) for documentation of catheter laboratory procedures. Key patient and procedural data on cardiac catheterizations conducted in the 76 VA catheter laboratories nationwide are linked to the VA EHR.[21] Quality checks are performed regularly to ensure completeness and accuracy. Additional details have been described previously.[21,22] Particulars of the hemodynamic cohort at Vanderbilt University, an academic referral center, are included in the eMethods in the Supplement and elsewhere.[23,24] The Colorado Multiple Institutional Review Board approved this study, and a waiver of the requirement for written informed consent was granted.

Study Sample

We evaluated all US veterans undergoing right heart catheterization (RHC) in the VA system between October 1, 2007, and September 30, 2013. Patients were included in the analyses if the following were available from a single RHC: height, weight, Td CO estimate, eFick CO estimate, and mean pulmonary artery pressure. For each individual included in the analysis, all hemodynamic data were extracted from the same catheterization. For patients who underwent multiple RHCs during the study period, only hemodynamic and clinical data from the earliest qualifying study were analyzed. For CO, mean pulmonary artery pressure, and pulmonary artery wedge pressure, a small proportion of physiologically implausible values were observed and were either corrected based on ancillary data or were excluded; the internal consistency of hemodynamic data was confirmed as described in a previous study.[22]

Analysis of correlation and agreement between Td and eFick estimates included all patients who fulfilled the above criteria (n = 12 232). Analysis comparing outcomes between those with a low cardiac index (CI) and those with a normal CI excluded patients with elevated Td or eFick CI (>4.0 L/min/m^2 based on prior data on the range of normal CI)[25] because an elevated CI may also reflect abnormal pathophysiology associated with increased risk, such as sepsis, severe anemia, cirrhosis, or other high CO disease states. There was no indication of the timing of the CO measurements during the catheterization relative to interventions performed, so we excluded patients who underwent procedures that could be associated with substantial changes in CO and also with mortality, including intra-aortic balloon pump placement, temporary pacemaker insertion, aortic or mitral valvuloplasty, pericardiocentesis, or alcohol septal ablation. Overall, 806 of the 12 232 patients (6.6%) were excluded from the survival analysis for one of these reasons. Therefore, survival analysis was performed on a sample of 11 426 patients from the VA CART cohort. The approach for the Vanderbilt cohort was similar, with details provided in the eMethods in the Supplement. The overall Vanderbilt cohort comprised 3391 patients of whom 3197 were included in the survival analysis.

Calculation of CI

A single CO estimate for each of the methods, Td and eFick, was recorded directly in the database (ie, the analysis used the recorded values rather than values calculated from component variables). Cardiac index was calculated by dividing either the Td or eFick CO estimate by body surface area calculated using the Du Bois formula.

Statistical Analysis

Pearson product moment correlation was used to analyze the association between Td and eFick CI estimates. The mean difference and 95% limits of agreement between methods were further assessed using Bland-Altman analysis, repeated for absolute (in liters per minute per meters squared) and percentage differences in CI; the results were essentially equivalent for CO, and only data for CI are presented.

Based on a clinically accepted cutoff for differentiating a low Td CI (<2.2 L/min/m^2) and normal CI (2.2-4.0 L/min/m^2), patients were categorized for each measurement method depending on the value of their respective continuous Td and eFick measurements.[26-28] In addition, patients were categorized based on agreement between the Td and eFick CI assessments, resulting in the following 4 groups: concordant low (ie, both Td and eFick <2.2 L/min/m^2), concordant normal, and 2 discordant groups where either Td or eFick was low, while the other estimate was normal. Because there was only a trivial systematic difference between Td and eFick estimates in the VA CART cohort, the same proportion of patients had low CI by Td and eFick. This was not the case for the Vanderbilt cohort (3391 adults between January 1, 1998, and December 31, 2014). As a consequence, for this cohort, we compared patients in the lowest tertile of CI for each method with those in the upper 2 tertiles; this cut point was chosen to provide a similar proportion of low CI to that seen in the VA CART cohort for each group. For each cohort, differences in baseline clinical and hemodynamic characteristics between the 4 resulting CI categories were analyzed using analysis of variance for continuous variables or using χ^2 test for categorical variables. Cumulative incidence (ie, 1 minus Kaplan-Meier survival estimate) curves were created with mortality as an outcome and stratified by CI status.

Time to event data for the outcome measures of all-cause mortality were analyzed using Cox proportional hazards regression models accounting for clustering by RHC site using a random effect (ie, a frailty model). Details are given in the eMethods in the Supplement.

To assess the relationship between mortality and continuous CI (in contrast to categorical CI), a penalized spline function was used, with separate models run for continuous CI calculated by Td or eFick. These models were replicated with both 90-day and 1-year censoring.

Data preparation and analyses were conducted using SAS (version 9.4; SAS Institute Inc) and R (version 3.2.5; R Foundation for Statistical Computing). A 2-sided $P < .05$ was considered statistically significant.

Results

Study Sample Characteristics

A total of 12 232 VA patients met study entry criteria. Demographic, anthropometric, and clinical characteristics of the cohort are listed in **Table 1** (column on the left). The study participants were predominantly male (11 834 patients [96.7%]), with a mean (SD) age of 66.2 (10.0) years. There was a high prevalence of systemic hypertension (87.2%), congestive heart failure (56.4%), diabetes (45.8%), chronic obstructive pulmonary disease (33.2%), chronic kidney disease (28.5%), and liver cirrhosis (7.2%). Left heart catheterization was performed at the time of RHC in 8006 patients (65.5%).

Correlation and Agreement Between Td and eFick CO Estimates

The mean (SD) CO by Td and eFick was 5.3 (1.7) and 5.3 (1.6) L/min, respectively. The mean (SD) CI was 2.5 (0.7) L/min/m^2 for both methods. Correlation and agreement between methods were similar for CO and CI; therefore, only data for CI are presented. Correlation between methods was modest (r = 0.65). There was no important systematic difference for eFick relative to the mean value of Td and eFick estimates (−0.02 L/min/m^2, or −0.4%). However, agreement between the 2 methods was less favorable, with 95% limits of agreement of −1.31 to 1.27 L/min/m^2, or −50.1% to 49.4% (**Figure 1**). Thermodilution and eFick estimates differed by greater than 20% for 38.1% of the catheterizations. The eFick estimates were slightly higher than the Td estimates at the lower end of the CO range (CI<2 L/min/m^2) and at the higher end of the CO range (CI>4.0 L/min/m^2) (eFigure 1 in the Supplement).

Subgroup analysis identified specific patient subsets in which there appeared to be systematic difference between the 2 methods. We observed that eFick CI estimates were on average greater than Td CI estimates for patients older than 75 years, while the converse was true for those 65 years or younger (eFigure 2 in the Supplement). Other variables exhibiting significant difference between CI methods included body mass index category, body surface area, hemoglobin concentration, and oxygen saturation. There was also a modest difference of 0.11 L/min/m^2 (95% CI, 0.02-0.20 L/min/m^2; P = .003) between Td and eFick in patients with tricuspid valve disease.

Low CI (<2.2 L/min/m^2) when measured by Td and eFick was present in 4347 patients (38.0%) and 4366 patients (38.2%), respectively. Overall, 24.2% of patients had a low CI by both methods, and 47.9% of patients had normal CI by both methods. We observed that 14.0% had low eFick but a normal Td CI, while the remaining 13.9% had low Td but normal eFick CI (Table 1). Certain characteristics were associated with increased likelihood of normal eFick with a low Td CI, including older age, male sex, normal or overweight BMI, atrial fibrillation or flutter, chronic kidney disease, and lower hemoglobin concentration. However, there was no difference in the prevalence of tricuspid valve disease between the discordant groups. A clinically meaningful difference was not observed between the mean eFick CI estimates in low eFick groups (1.8 L/min/m^2 for both Td and eFick low vs 1.9 L/min/m^2 for eFick low and Td normal), and similar findings were observed for the 2 groups with low Td estimates (Table 1). Therefore, discordant groups do not appear to simply reflect patients with borderline CI values.

Relationship Between CO and Mortality

One-year mortality was lower among patients with a concordant normal CI by both methods compared with those who had a low CI by both methods (12.3% vs 19.9%, P < .001). Those classified as having discordant estimates but with a normal CI by

Table 1. Demographic, Anthropometric, and Clinical Characteristics of Patients, Overall and by CO Estimate Classification Methods in the VA CART Cohort[a]

Variable	All (N = 12 232)	All, Excluding High CI (N = 11 426)[b]	Both Low (n = 2764)	eFick Low and Td Normal (n = 1602)	eFick Normal and Td Low (n = 1583)	Both Normal (n = 5477)	*P* Value for Overall	*P* Value for Discordant Groups
Demographic and Anthropometric Data								
Age, mean (SD), y	66.2 (10.0)	66.4 (9.9)	66.8 (10.2)	64.7 (9.6)	68.6 (10.1)	66.1 (9.7)	<.001	<.001
Female sex, No. (%)	398 (3.3)	360 (3.2)	81 (2.9)	74 (4.6)	41 (2.6)	164 (3.0)	.003	.003
Race, No. (%)								
White	9464 (77.4)	8857 (77.5)	2038 (73.7)	1246 (77.8)	1186 (74.9)	4387 (80.1)		
Black	1852 (15.1)	1711 (15.0)	533 (19.3)	232 (14.5)	276 (17.4)	670 (12.2)	<.001	.07
Other	916 (7.5)	858 (7.5)	193 (7.0)	124 (7.7)	121 (7.6)	420 (7.7)		
BMI category, No. (%)								
Underweight, <18.5	115 (0.9)	103 (0.9)	24 (0.9)	16 (1.0)	12 (0.8)	51 (0.9)		
Normal, 18.5 to <25	2338 (19.1)	2169 (19.0)	588 (21.3)	234 (14.6)	371 (23.4)	976 (17.8)		
Overweight, 25 to <30	3946 (32.3)	3682 (32.2)	975 (35.3)	451 (28.2)	530 (33.5)	1726 (31.5)	<.001	<.001
Obese, 30 to <35	3070 (25.1)	2893 (25.3)	691 (25.0)	437 (27.3)	393 (24.8)	1372 (25.1)		
Severely obese, ≥35	2763 (22.6)	2579 (22.6)	486 (17.6)	464 (29.0)	277 (17.5)	1352 (24.7)		
Body surface area, mean (SD), m^2	2.1 (0.2)	2.1 (0.2)	2.1 (0.2)	2.1 (0.2)	2.1 (0.2)	2.1 (0.2)	<.001	<.001
Clinical Data								
Tricuspid valve disease, No. (%)	240 (2.0)	223 (2.0)	71 (2.6)	30 (1.9)	34 (2.1)	88 (1.6)	.03	.67
Other valve disease, No. (%)	6186 (50.6)	5833 (51.1)	1360 (49.2)	773 (48.3)	832 (52.6)	2868 (52.4)	.003	.02
Atrial fibrillation or flutter, No. (%)	3235 (26.4)	3095 (27.1)	1066 (38.6)	390 (24.3)	523 (33.0)	1116 (20.4)	<.001	<.001
Systemic hypertension, No. (%)	10672 (87.2)	9984 (87.4)	2426 (87.8)	1395 (87.1)	1365 (86.2)	4798 (87.6)	.45	.51
Mean pulmonary artery pressure >25 mm Hg, No. (%)	6907 (56.5)	6442 (56.4)	1956 (70.8)	935 (58.4)	879 (55.5)	2672 (48.8)	<.001	.11
Pulmonary vascular resistance, mean (SD), WU[c]	2.4 (1.9)	2.4 (1.9)	3.6 (2.7)	2.5 (1.7)	2.5 (1.7)	1.8 (1.1)	<.001	.13
Congestive heart failure, No. (%)	6897 (56.4)	6511 (57.0)	2055 (74.3)	874 (54.6)	986 (62.3)	2596 (47.4)	<.001	<.001
Chronic obstructive pulmonary disease, No. (%)	4063 (33.2)	3809 (33.3)	927 (33.5)	578 (36.1)	483 (30.5)	1821 (33.2)	.01	.001
Chronic kidney disease, No. (%)	3485 (28.5)	3177 (27.8)	871 (31.5)	355 (22.2)	477 (30.1)	1474 (26.9)	<.001	<.001
Diabetes, No. (%)	5597 (45.8)	5227 (45.7)	1241 (44.9)	765 (47.8)	690 (43.6)	2531 (46.2)	.08	.02
Liver cirrhosis, No. (%)	884 (7.2)	749 (6.6)	168 (6.1)	108 (6.7)	94 (5.9)	379 (6.9)	.35	.39
Heart rate, mean (SD), beats/min	71.9 (14.2)	71.5 (14.1)	73.8 (15.7)	71.5 (14.1)	69.9 (14.2)	71.0 (13.2)	<.001	.01
Hemoglobin concentration, mean (SD), g/dL	12.8 (2.0)	12.9 (2.0)	13.3 (1.9)	13.4 (1.9)	12.7 (1.9)	12.5 (2.0)	.26	<.001
Oxygen saturation, mean (SD), %	94.7 (3.9)	94.6 (5.4)	94.5 (5.6)	94.3 (4.8)	94.7 (7.0)	94.6 (4.9)	.14	.85

(continued)

Table 1. Demographic, Anthropometric, and Clinical Characteristics of Patients, Overall and by CO Estimate Classification Methods in the VA CART Cohort[a] (continued)

Variable	All (N = 12 232)	All, Excluding High CI (N = 11 426)[b]	Both Low (n = 2764)	eFick Low and Td Normal (n = 1602)	eFick Normal and Td Low (n = 1583)	Both Normal (n = 5477)	P Value for Overall	P Value for Discordant Groups
CO, mean (SD), L/min								
eFick	5.3 (1.6)	5.1 (1.3)	3.8 (0.7)	4.1 (0.7)	5.4 (0.9)	6.0 (1.1)	<.001	<.001
Td	5.3 (1.7)	5.1 (1.5)	3.7 (0.8)	5.7 (1.1)	4.1 (0.6)	6.0 (1.2)	<.001	<.001
CI, mean (SD), L/min/m^2								
eFick	2.5 (0.7)	2.4 (0.6)	1.8 (0.3)	1.9 (0.3)	2.6 (0.3)	2.8 (0.4)	<.001	<.001
Td	2.5 (0.7)	2.4 (0.6)	1.7 (0.3)	2.6 (0.4)	1.9 (0.2)	2.8 (0.4)	<.001	<.001
Mortality, No. (%)								
90 d	794 (6.5)	716 (6.3)	252 (9.1)	82 (5.1)	110 (6.9)	272 (5.0)	<.001	.04
1 y	1838 (15.0)	1673 (14.6)	550 (19.9)	207 (12.9)	244 (15.4)	672 (12.3)	<.001	.049

Abbreviations: BMI, body mass index (calculated as weight in kilograms divided by height in meters squared); CI, cardiac index; CO, cardiac output; eFick, estimated oxygen uptake Fick; Td, thermodilution; VA CART, Veterans Affairs Clinical Assessment, Reporting, and Tracking; WU, wood units.

SI conversion factor: To convert hemoglobin concentration to grams per liter, multiply by 10.0.

[a] The left most column provides data for the overall cohort while the second column from the left presents data excluding the subset of patients with elevated CI of >4.0 L/min/m^2, and the remaining columns to the right present data for patients classified as low or normal CI. Differences between the 4 columns on the right were assessed using analysis of variance or χ^2 test. Differences between discordant estimates (eFick low and Td normal vs eFick normal and Td low) were analyzed using independent *t* test or χ^2 test. Data were missing for the following variables (number for the overall cohort/number excluding CI>4.0 L/min/m^2): pulmonary vascular resistance (61/56), heart rate (4616/4309), hemoglobin concentration (1930/1808), and oxygen saturation (757/695).

[b] High CI was defined as exceeding 4.0 L/min/m^2.

[c] Calculated using the mean of Td and eFick CO estimates.

Td had essentially the same mortality as a concordant normal CI (12.9%, *P* = .51); those with discordant estimates where eFick suggested normal CI had higher mortality (15.4%, *P* = .001). The pattern was similar for 90-day and 1-year mortality (**Figure 2**A) and in the Vanderbilt cohort (Figure 2B); cumulative incidence plots are shown in eFigure 3 in the Supplement.

For the study sample taken as a whole, there was no improvement in predictive power with the measurement of CI by both methods compared with Td alone. The C statistic for a Cox proportional hazards regression model that included only continuous Td CI as a predictor variable for mortality over 90 days was 0.587; the equivalent value for continuous eFick CI was 0.548. There was no improvement in overall model performance when both Td and eFick were included in the same model (0.584; likelihood ratio χ^2 = 0.03, *P* = .87). The pattern was equivalent for 1-year mortality (C statistic, 0.566 for Td alone, 0.547 for eFick alone, and 0.566 for both; *P* = .12).

Analyzing the data simply stratified by a CI less than 2.2 L/min/m^2, patients with a low CI had a higher risk of mortality compared with patients with a normal CI. This was true for both methods, but the relationship was stronger for Td CI than eFick CI (**Table 2** and **Figure 3**A and B); equivalent data for the Vanderbilt cohort are provided in Table 2 and Figure 3C and D. This finding was consistent among all subgroups analyzed (see eFigure 2 in the Supplement for a list of subgroups). Even among patients with tricuspid valve disease, low Td CI was more strongly associated with mortality than low eFick CI (hazard ratio [HR] for 90-day mortality, 3.4; 95% CI, 1.4-8.0 for Td vs 1.8; 95% CI, 0.8-3.9 for eFick).

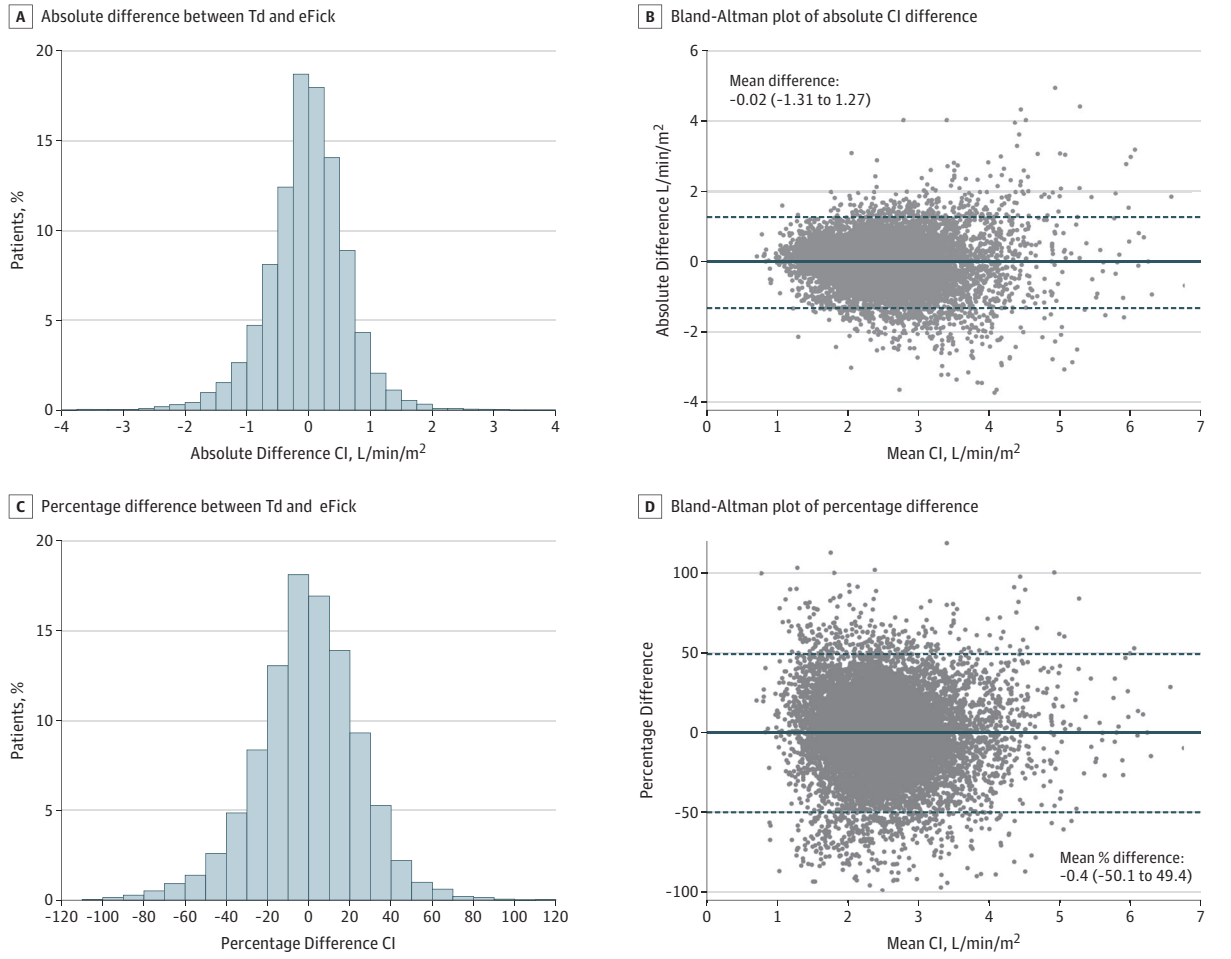
The hazard for mortality appeared to increase for both methods at a CI just under 2.2 L/min/m^2, with substantially higher risk of mortality as CI decreased below 2 L/min/m^2 (eFigure 4 in the Supplement). Very low Td CI estimates (<1.6 L/min/m^2) were associated with a greater incremental increase in risk than similarly low eFick CI estimates. There was also a suggestion of slightly increased risk for eFick CI estimates at the upper part of normal range (CI approximately 3-4 L/min/m^2). This U-shaped relationship was absent for Td estimates.

Vanderbilt Cohort

Characteristics of the 3391 patients included in the independent Vanderbilt University hemodynamic cohort are listed in the eTable in the Supplement. The mean (SD) CO by Td and eFick, respectively, was 5.0 (1.6) and 5.6 (1.8) L/min. The respective corresponding mean (SD) values for CI were 2.6 (0.8) and 2.8 (0.9) L/min/m^2. Data presented below are on CI; the results for CO were similar. Correlation between the 2 methods was modest (*r* = 0.63). In contrast to the VA CART cohort, there was a systematic difference between methods, with eFick, on average, estimating CI 0.3 L/min/m^2 (9.3%) greater than Td (eFigure 5 in the Supplement). Agreement between the 2 methods was poor, similar to that seen in the VA CART cohort, with 95% limits of agreement of −1.1 to 2.0 L/min/m^2, or −40.8% to 59.9%. Thermodilution and eFick estimates differed by greater than 20% for 36.4% of the catheterizations.

Among the overall sample of 3391 patients, a CI less than 2.2 L/min/m^2 was present in 1171 patients (34.5%) by Td and 785 patients (23.1%) by eFick. This discrepancy reflected the

Figure 1. Distribution of Differences and Bland-Altman Analysis of Thermodilution (Td) vs Estimated Oxygen Uptake Fick (eFick) Cardiac Index (CI) Estimates in the Veterans Affairs Clinical Assessment, Reporting, and Tracking (VA CART) Cohort



Shown are the proportions of patients with a given absolute difference between Td and eFick estimates (A). Shown is a Bland-Altman plot of the mean of the Td and eFick CI estimates on the x-axis against the difference (eFick minus Td) on the y-axis (B). Shown are the percentage differences between Td and eFick estimates (C).There was no substantial systematic difference between Td and eFick, but the 95% limits of agreement were wide (−1.31 to 1.27 L/min/m^2, or −50.1% to 49.4%) (D).

systematic difference between methods in this cohort; as a result, we grouped patients into lowest tertile vs second and third tertiles for each method to yield a proportion of low CI similar to the VA CART analysis.

The 2 methods showed concordance in terms of low vs normal CI in 74.0% of cases (19.8% for both low and 54.2% for both normal) and showed discordance in the remaining 26.0%, evenly split 13.0% each for eFick low and Td normal and eFick normal and Td low (eTable in the Supplement). Patients who had concordant normal CI by both methods were less likely to die over the following year (15.0%) compared with those who had a low CI by both methods (28.5%). Those classified as discordant but with a normal CI by Td had marginally higher mortality compared with a concordant normal CI (16.9%); however, those with discordant estimates where the eFick CI was normal had substantially higher 1-year mortality (20.6%). The pattern was similar for 90-day mortality (Figure 2B and eFigure 3B in the Supplement).

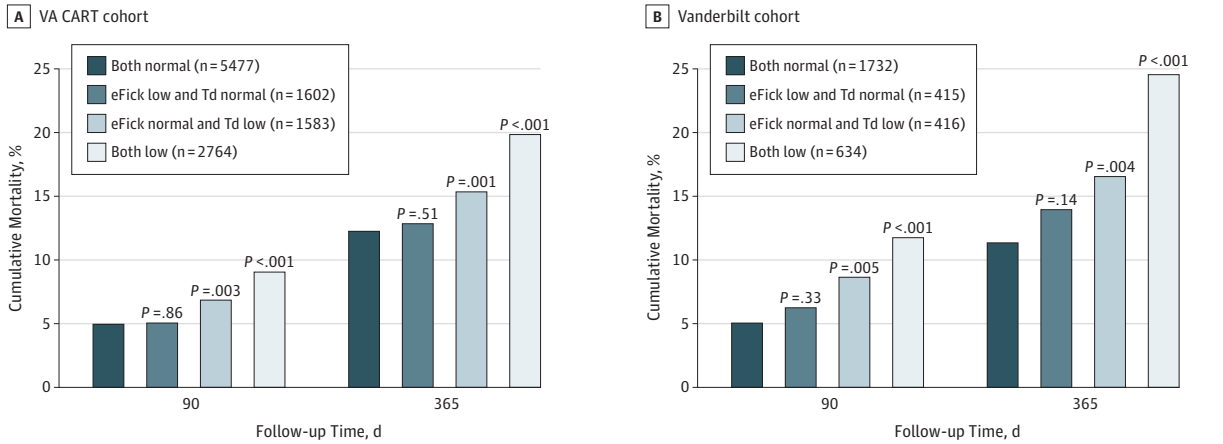
A finding of low a CI by either method was associated with higher mortality, but a low CI by Td was a stronger predictor. These results are summarized in Table 2 and Figure 3C and D.

Discussion

Measurement of CO is critical to appropriate clinical assessment of patients across a broad spectrum of disease commonly encountered in general medical practice. The present analysis of invasive clinical hemodynamic data from a large, national cohort with validation of the findings in a second, independent cohort demonstrates that (1) there is poor agreement between CO measured by Td and eFick methods and (2) Td CO estimates are superior to eFick for predicting all-cause mortality.

Prior reports comparing methods to estimate CO have generally involved applying a rigorous research protocol to a small number of highly selected individuals.[14,20,29-31] Reported pre-

Figure 2. Cumulative Mortality Through 90 Days and 1-Year Follow-up, Classified by Normal and Low Thermodilution (Td) and Estimated Oxygen Uptake Fick (eFick) Cardiac Index Categories



A, Shown are the cumulative proportions of patients who had died by 90 days and 1 year after catheterization, stratified by cardiac index classification in the Veterans Affairs Clinical Assessment, Reporting, and Tracking (VA CART) cohort. All patients who survived had at least 1 year of follow-up. Patients with concordantly low cardiac index estimates by both Td and eFick had the highest mortality risk. Those with either normal cardiac index by both methods or isolated low eFick cardiac index had the lowest risk. The incidence of death was intermediate for patients with low cardiac index by Td but normal eFick cardiac index. B, Shown are the corresponding Vanderbilt cohort results. The findings are qualitatively equivalent (eg, isolated low Td was associated with worse prognosis than isolated low eFick). However, in this cohort, there was marginally increased risk of death for those with isolated low eFick cardiac index relative to normal cardiac index by both methods.

Table 2. Survival Analysis for All-Cause Mortality Predicted by Low vs Normal Cardiac Index, as Measured by eFick and Thermodilution (Td) Methods[a]

	eFick		Thermodilution	
Variable	HR (95% CI)	χ^2 Statistic	HR (95% CI)	χ^2 Statistic
VA CART Cohort				
Univariate				
90 d	1.42 (1.22-1.64)	20.7	1.71 (1.47-1.99)	49.5
1 y	1.35 (1.22-1.49)	35.2	1.53 (1.39-1.69)	71.5
Adjusted model, age/sex/race				
90 d	1.43 (1.23-1.66)	21.4	1.60 (1.38-1.87)	37.3
1 y	1.36 (1.24-1.51)	37.9	1.44 (1.30-1.59)	51.7
Adjusted model, age/sex/race and clinical variables[b]				
90 d	1.36 (1.17-1.58)	15.7	1.45 (1.24-1.69)	22.1
1 y	1.31 (1.19-1.45)	29.0	1.32 (1.19-1.46)	28.5
Vanderbilt Cohort				
Univariate				
90 d	1.63 (1.25-2.13)	13.0	1.98 (1.52-2.58)	25.8
1 y	1.67 (1.39-2.00)	30.5	1.85 (1.54-2.21)	43.7
Adjusted model, age/sex/race				
90 d	1.74 (1.33-2.28)	16.1	1.82 (1.39-2.37)	19.1
1 y	1.75 (1.45-2.11)	34.9	1.72 (1.43-2.06)	33.1
Adjusted model, age/sex/race and clinical variables[b]				
90 d	1.64 (1.14-2.35)	7.2	1.77 (1.23-2.53)	9.5
1 y	1.51 (1.19-1.91)	11.8	1.59 (1.25-2.00)	14.8

Abbreviations: eFick, estimated oxygen uptake Fick; HR, hazard ratio; Td, thermodilution; VA CART, Veterans Affairs Clinical Assessment, Reporting, and Tracking.

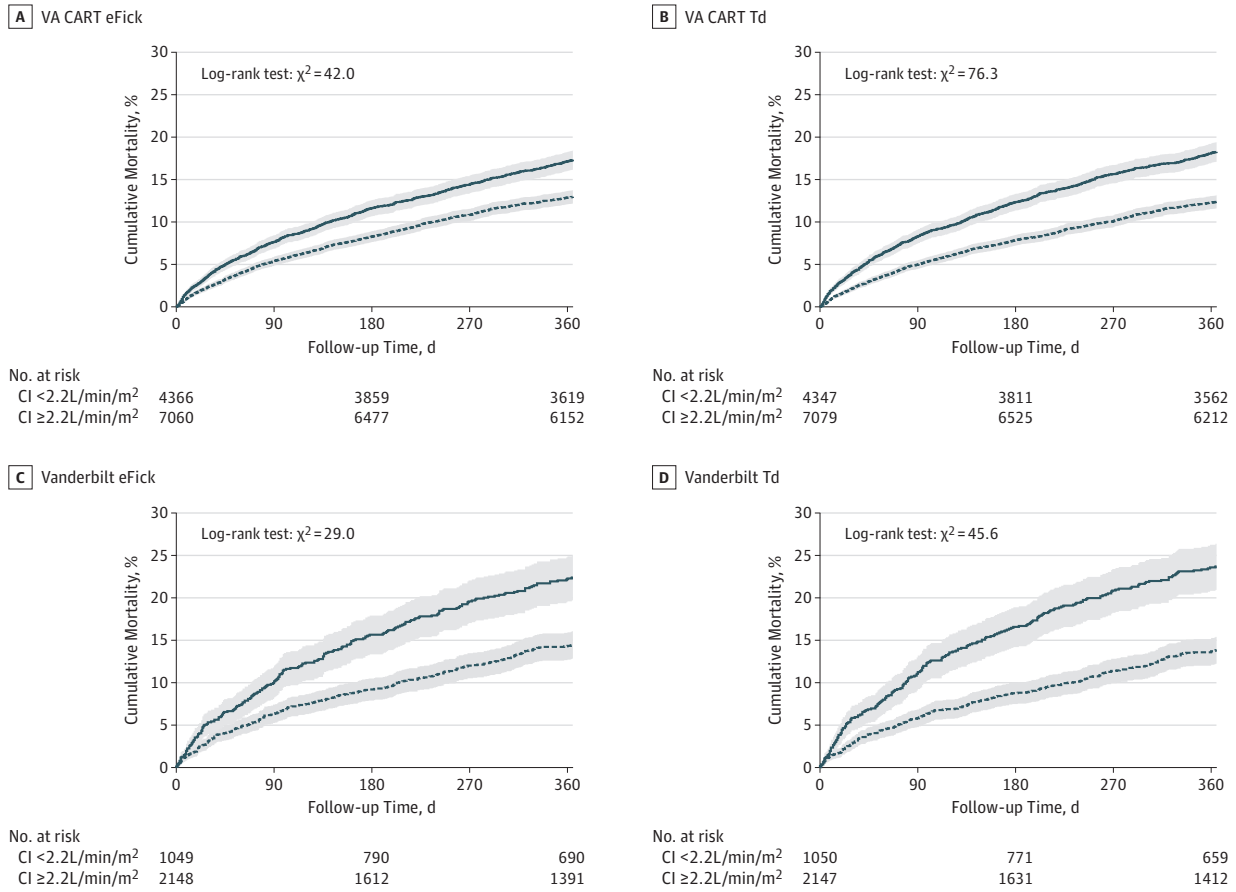
[a] The HRs with 95% CIs and χ^2 statistics are given for the prediction of all-cause mortality by low CI (<2.2 L/min/m^2) vs normal CI (2.2-4.0 L/min/m^2) by eFick and Td methods. In the Vanderbilt cohort, low is defined as the lowest tertile of cardiac index for each method. The univariate results are from unadjusted Cox proportional hazards regression analysis. For both methods, a low cardiac index was more strongly predictive of 90-d mortality than 1-y mortality. Compared with eFick, Td estimates were consistently associated with better model performance in terms of both HRs and χ^2 statistics.

[b] Also adjusting for body mass index category, atrial fibrillation or flutter, systemic hypertension, chronic obstructive pulmonary disease, chronic kidney disease, diabetes, liver cirrhosis, and tobacco use.

cision and accuracy are unlikely to be reproduced in usual clinical practice. Such experimental studies may not be able to identify patient subsets at risk for measurement error, and they do not address the relative value of different methods to predict clinical outcomes. The present study provides novel information on real-world testing in 2 large, independent, clinical cohorts with follow-up that included ascertainment of tangible clinical end points.

There are multiple reasons one might expect that Td estimates would be superior to eFick estimates. First are the known

Figure 3. Cumulative Mortality Through 1-Year Follow-up, Stratified by Estimated Oxygen Uptake Fick (eFick) and Thermodilution (Td) Cardiac Index (CI)



Patients with low CI (<2.2 L/min/m^2) in the VA cohort had a higher incidence of death than those with normal CI, whether estimated by eFick (A) or Td (B). However, the relative hazard for mortality and model fit was better for Td CI. C and D, Equivalent plots are presented for the Vanderbilt cohort, stratifying CI by lowest tertile vs second and third tertiles. The solid and dotted line in each plot represents the cumulative mortality for those with low and normal CI, respectively. Gray area reflects the 95% confidence interval.

inaccuracies in $\dot{V}O_2$ estimation.[6-9] Narang and colleagues[9] compared the accuracy of estimated $\dot{V}O_2$ with directly measured $\dot{V}O_2$ in 535 patients; estimates differed greater than 25% from true $\dot{V}O_2$ in up to one-quarter of patients, with some variability between estimating formulas. This error in $\dot{V}O_2$ translates into a proportionally equivalent error in the CO estimate. Second, although there has been debate regarding the validity of Td in patients with low CO or severe tricuspid regurgitation,[10-13] other studies[14,30] have reported that agreement between Td and direct Fick (ie, using measured $\dot{V}O_2$) estimates remains robust even in these scenarios. The findings from the present data sets agree with that assertion because we found only a small difference between Td and eFick in either of these settings. Notably, eFick was not significantly associated with mortality at 90 days in the presence of tricuspid regurgitation, whereas Td remained predictive. While Td may be affected by other variables, including the timing of injection during the respiratory cycle and the volume and temperature of injectate,[32,33] Td proved more predictive of mortality than eFick. Overall, our study lends further support to recent consensus statements suggesting that Td rather than eFick alone should be the preferred method to measure CO in clinical practice.[34]

These data argue that eFick CI measurements provide no substantive additive prognostic value beyond Td CI among the global population of patients undergoing cardiac catheterization. However, one could interpret the results to suggest that, while Td alone may be superior to eFick alone, the methods may provide additive information, at least among the subset of patients with a low Td CI. In that subgroup, those who had normal eFick CI had lower mortality compared with the group having concordantly low CI estimates. Of note, adjustment for demographic, anthropometric, and clinical covariates reduced the difference in prognostic value between Td and eFick. Adjustment for variables like age, sex, and body mass index may serve to improve, indirectly, the accuracy of $\dot{V}O_2$ estimates and act to ameliorate the resulting error as it relates to the relationship between eFick CO and outcomes. Cardiac output is measured in clinical practice primarily to define patho-

physiology rather than predict outcomes; in that context, unadjusted measurements are more relevant than adjusted estimates. However, the multivariable results suggest that measuring $\dot{V}O_2$ or more accurately estimating $\dot{V}O_2$ could address the disparity between Td and eFick.

Limitations

Some limitations of this study merit consideration in interpreting these findings. We do not know how either technique was performed across centers (eg, temperature of injectate used for Td). This could influence the accuracy of CO measurement. However, clinical decisions were made using only the available data; therefore, our analysis is likely to be a valid comparison of these methods in the context of standard clinical applications. The consistent findings in 2 large cohorts from distinct settings strongly support broad generalizability. For example, women are unrepresented in the VA CART cohort but not in the Vanderbilt cohort (3.3% vs 47.5%). Direct Fick, with measurement of $\dot{V}O_2$, would presumably be more accurate than Td and may be superior for risk prediction; unfortunately, the additional effort and equipment required have prevented widespread use of this approach in clinical practice. The use of administrative data to define comorbidities limits the granularity of clinical phenotype; for example, while it seems likely that most patients assigned a code of tricuspid valve disease had clinically significant tricuspid regurgitation, we do not have confirmatory data on the severity or type of this disease. Information on the presence of shunting was not available for the present analysis. Although significant intracardiac shunts are uncommon in the general adult population, it should be reiterated that Td is inaccurate in the setting of shunting.[5] A preference for Td over eFick to estimate CO should also not come at the expense of appropriate comprehensive oximetric shunt evaluations. Finally, we did not compare Td and eFick with a criterion standard CO measurement but rather used the association between CO and mortality both as an independent indicator of clinical value and, implicitly, as a surrogate marker of accurate CO measurement.

Conclusions

There is only modest agreement between Td and eFick CO estimates. In the setting of a low CI, Td predicts incident mortality better than eFick alone and should be favored in clinical practice.

ARTICLE INFORMATION

Accepted for Publication: July 12, 2017.

Published Online: September 6, 2017. doi:10.1001/jamacardio.2017.2945

Author Affiliations: Division of Cardiovascular Medicine, Department of Medicine, Brigham and Women's Hospital and Harvard Medical School, Boston, Massachusetts (Opotowsky, Maron, Alshawabkeh); Department of Cardiology, Boston Children's Hospital, Boston, Massachusetts (Opotowsky, Alshawabkeh); Veterans Affairs Eastern Colorado Health Care System, Denver (Hess, Barón, Maddox); Veterans Affairs Boston Healthcare System, Boston, Massachusetts (Maron); Division of Cardiovascular Medicine, Department of Medicine, Vanderbilt University School of Medicine, Nashville, Tennessee (Brittain); Vanderbilt Translational and Clinical Cardiovascular Research Center, Vanderbilt University Medical Center, Nashville, Tennessee (Brittain); University of Colorado School of Medicine, Denver (Maddox); Division of Pulmonary and Critical Care Medicine, Department of Medicine, Brigham and Women's Hospital and Harvard Medical School, Boston, Massachusetts (Wertheim); Department of Biostatistics, Vanderbilt University, Nashville, Tennessee (Xu); Division of Allergy, Pulmonary, and Critical Care Medicine, Department of Medicine, Vanderbilt University School of Medicine, Nashville, Tennessee (Assad); Division of Cardiology, Department of Medicine, Northwestern University, Chicago, Illinois (Rich); Providence Veterans Affairs Medical Center, Providence, Rhode Island (Choudhary); Alpert Medical School of Brown University, Providence, Rhode Island (Choudhary); Division of Cardiology, Department of Medicine, Medical University of South Carolina, Charleston (Tedford).

Author Contributions: Drs Brittain and Maddox had full access to all of the data (for the Vanderbilt cohort and the Veterans Affairs Clinical Assessment, Reporting, and Tracking [VA CART] cohort, respectively) in the study and take responsibility for the integrity of the data and the accuracy of the data analysis.
Study concept and design: Opotowsky, Maron, Brittain, Maddox, Wertheim, Tedford.
Acquisition, analysis, or interpretation of data: All authors.
Drafting of the manuscript: Opotowsky, Maron, Brittain, Maddox, Wertheim, Tedford.
Critical revision of the manuscript for important intellectual content: All authors.
Statistical analysis: Opotowsky, Hess, Brittain, Barón, Xu, Tedford.
Obtained funding: Maron, Maddox.
Administrative, technical, or material support: Opotowsky, Maron, Brittain, Maddox, Assad, Tedford.
Study supervision: Opotowsky, Maron, Barón, Maddox, Tedford.

Conflict of Interest Disclosures: All authors have completed and submitted the ICMJE Form for Disclosure of Potential Conflicts of Interest. Dr Opotowsky reported investigator-initiated research supported by Actelion Pharmaceuticals and Roche Diagnostics. Dr Maron reported investigator-initiated research supported by Gilead Sciences Inc. Dr Brittain reported investigator-initiated research supported by Gilead Sciences Inc. Dr Maddox reported being national director of the Veterans Affairs Clinical Assessment, Reporting, and Tracking (VA CART) program. Dr Choudhary reported investigator-initiated research supported by Novartis. No other disclosures were reported.

Funding/Support: Dr Opotowsky is supported by the Dunlevie Family Fund and by grant 17IRG33370023 from the American Heart Association. Dr Maron is supported by grant 1K08HL11207-01A1 from the National Institutes of Health, grant 15GRNT25080016 from the American Heart Association, the Pulmonary Hypertension Association, and The Cardiovascular Medical Research and Education Fund. Dr Brittain is supported by grant 13FTF16070002 from the American Heart Association. Dr Choudhary is supported by grant R01HL128661 from the National Institutes of Health.

Role of the Funder/Sponsor: None of the funding organizations were involved in the design and conduct of the study; collection, management, analysis, and interpretation of the data; preparation, review, or approval of the manuscript; or decision to submit the manuscript for publication.

Additional Contributions: At Vanderbilt University, Quinn S. Wells, MD, and Eric H. Farber-Eger, BS, contributed to the development of the hemodynamic database. No compensation was received.

REFERENCES

1. Ganz W, Donoso R, Marcus HS, Forrester JS, Swan HJ. A new technique for measurement of cardiac output by thermodilution in man. *Am J Cardiol*. 1971;27(4):392-396.

2. Dehmer GJ, Firth BG, Hillis LD. Oxygen consumption in adult patients during cardiac catheterization. *Clin Cardiol*. 1982;5(8):436-440.

3. LaFarge CG, Miettinen OS. The estimation of oxygen consumption. *Cardiovasc Res*. 1970;4(1):23-30.

4. Bergstra A, van Dijk RB, Hillege HL, Lie KI, Mook GA. Assumed oxygen consumption based on calculation from dye dilution cardiac output: an improved formula. *Eur Heart J*. 1995;16(5):698-703.

5. Guyton AC, Jones CE, Coleman TG. *Cardiac Output and Its Regulation*. 2nd ed. Philadelphia, PA: Saunders; 1973.

6. Fakler U, Pauli C, Hennig M, Sebening W, Hess J. Assumed oxygen consumption frequently results in large errors in the determination of cardiac output. *J Thorac Cardiovasc Surg*. 2005;130(2):272-276.

7. Kendrick AH, West J, Papouchado M, Rozkovec A. Direct Fick cardiac output: are assumed values of oxygen consumption acceptable? *Eur Heart J*. 1988;9(3):337-342.

8. Narang N, Gore MO, Snell PG, et al. Accuracy of estimating resting oxygen uptake and implications for hemodynamic assessment. *Am J Cardiol*. 2012;109(4):594-598.

9. Narang N, Thibodeau JT, Levine BD, et al. Inaccuracy of estimated resting oxygen uptake in the clinical setting. *Circulation*. 2014;129(2):203-210.

10. van Grondelle A, Ditchey RV, Groves BM, Wagner WW Jr, Reeves JT. Thermodilution method overestimates low cardiac output in humans. *Am J Physiol*. 1983;245(4):H690-H692.

11. Nishikawa T, Dohi S. Errors in the measurement of cardiac output by thermodilution. *Can J Anaesth*. 1993;40(2):142-153.

12. Hillis LD, Firth BG, Winniford MD. Analysis of factors affecting the variability of Fick versus indicator dilution measurements of cardiac output. *Am J Cardiol*. 1985;56(12):764-768.

13. Cigarroa RG, Lange RA, Williams RH, Bedotto JB, Hillis LD. Underestimation of cardiac output by thermodilution in patients with tricuspid regurgitation. *Am J Med*. 1989;86(4):417-420.

14. Hoeper MM, Maier R, Tongers J, et al. Determination of cardiac output by the Fick method, thermodilution, and acetylene rebreathing in pulmonary hypertension. *Am J Respir Crit Care Med*. 1999;160(2):535-541.

15. Fares WH, Blanchard SK, Stouffer GA, et al. Thermodilution and Fick cardiac outputs differ: impact on pulmonary hypertension evaluation. *Can Respir J*. 2012;19(4):261-266.

16. Krowka MJ, Mandell MS, Ramsay MA, et al. Hepatopulmonary syndrome and portopulmonary hypertension: a report of the Multicenter Liver Transplant Database. *Liver Transpl*. 2004;10(2):174-182.

17. Tuchschmidt J, Fried J, Astiz M, Rackow E. Elevation of cardiac output and oxygen delivery improves outcome in septic shock. *Chest*. 1992;102(1):216-220.

18. Anderson CB, Codd JR, Graff RA, Groce MA, Harter HR, Newton WT. Cardiac failure and upper extremity arteriovenous dialysis fistulas: case reports and a review of the literature. *Arch Intern Med*. 1976;136(3):292-297.

19. Fincke R, Hochman JS, Lowe AM, et al; SHOCK Investigators. Cardiac power is the strongest hemodynamic correlate of mortality in cardiogenic shock: a report from the SHOCK trial registry. *J Am Coll Cardiol*. 2004;44(2):340-348.

20. Hamilton WF, Riley RL, et al. Comparison of the Fick and dye injection methods of measuring the cardiac output in man. *Am J Physiol*. 1948;153(2):309-321.

21. Maddox TM, Plomondon ME, Petrich M, et al. A national clinical quality program for Veterans Affairs catheterization laboratories (from the Veterans Affairs Clinical Assessment, Reporting, and Tracking program). *Am J Cardiol*. 2014;114(11):1750-1757.

22. Maron BA, Hess E, Maddox TM, et al. Association of borderline pulmonary hypertension with mortality and hospitalization in a large patient cohort: insights from the Veterans Affairs Clinical Assessment, Reporting, and Tracking program. *Circulation*. 2016;133(13):1240-1248.

23. Assad TR, Brittain EL, Wells QS, et al. Hemodynamic evidence of vascular remodeling in combined post- and precapillary pulmonary hypertension. *Pulm Circ*. 2016;6(3):313-321.

24. Brittain EL, Chan SY. Integration of complex data sources to provide biologic insight into pulmonary vascular disease (2015 Grover Conference Series). *Pulm Circ*. 2016;6(3):251-260.

25. Lentner C. *Geigy Scientific Tables, Volume 5: Heart and Circulation*. West Caldwell, NJ: Ciba-Geigy; 1990.

26. Forrester JS, Diamond G, Chatterjee K, Swan HJ. Medical therapy of acute myocardial infarction by application of hemodynamic subsets (first of two parts). *N Engl J Med*. 1976;295(24):1356-1362.

27. Nohria A, Tsang SW, Fang JC, et al. Clinical assessment identifies hemodynamic profiles that predict outcomes in patients admitted with heart failure. *J Am Coll Cardiol*. 2003;41(10):1797-1804.

28. Nohria A, Mielniczuk LM, Stevenson LW. Evaluation and monitoring of patients with acute heart failure syndromes. *Am J Cardiol*. 2005;96(6A):32G-40G.

29. Eliasch H, Lagerlof H, Bucht H, et al. Comparison of the dye dilution and the direct Fick methods for the measurement of cardiac output in man. *Scand J Clin Lab Invest*. 1955;7(suppl 20):73-78.

30. Yung GL, Fedullo PF, Kinninger K, Johnson W, Channick RN. Comparison of impedance cardiography to direct Fick and thermodilution cardiac output determination in pulmonary arterial hypertension. *Congest Heart Fail*. 2004;10(2)(suppl 2):7-10.

31. Cecconi M, Rhodes A, Poloniecki J, Della Rocca G, Grounds RM. Bench-to-bedside review: the importance of the precision of the reference technique in method comparison studies: with specific reference to the measurement of cardiac output. *Crit Care*. 2009;13(1):201.

32. Stevens JH, Raffin TA, Mihm FG, Rosenthal MH, Stetz CW. Thermodilution cardiac output measurement: effects of the respiratory cycle on its reproducibility. *JAMA*. 1985;253(15):2240-2242.

33. Elkayam U, Berkley R, Azen S, Weber L, Geva B, Henry WL. Cardiac output by thermodilution technique: effect of injectate's volume and temperature on accuracy and reproducibility in the critically ill patient. *Chest*. 1983;84(4):418-422.

34. Hoeper MM, Bogaard HJ, Condliffe R, et al. Definitions and diagnosis of pulmonary hypertension. *J Am Coll Cardiol*. 2013;62(25)(suppl):D42-D50.