Equivalence of Bioimpedance and Thermodilution in Measuring Cardiac Index After Cardiac Surgery

W. Scott Sageman, MD, Robert H. Riffenburgh, PhD, and Bruce D. Spiess, MD

Objective: **To compare thoracic electrical bioimpedance (TEB) cardiography versus pulmonary artery thermodilution (TD) derived cardiac index in patients after cardiopulmonary bypass.**

Design: **Prospective, blinded electronic data collection.**

Setting: **Intensive care unit of a military hospital.**

Participants: **Post–cardiopulmonary bypass patients for primary comparison between technologies (n = 20) and patients for comparison of variability within each technology (n = 20).**

Interventions: **None.**

Measurements and Main Results: **Cardiac index values by TEB or TD were collected simultaneously. Linear regression, Lin's concordance correlation coefficient, bias, and precision measures within the large data set group and within each patient over time were calculated. Linearity in regression and Lin's concordance correlation coefficient of 0.99 were shown. A bias of 0.07 L/min/m^2 and precision of 0.40 L/min/m^2 were within acceptable clinical limits, as were equivalence test results.**

Conclusions: **TEB is equivalent to TD-derived cardiac index in postoperative cardiac surgery patients.**

KEY WORDS: monitoring, hemodynamics, cardiac output, cardiac index, pulmonary artery catheter, bioimpedance, noninvasive

DESPITE THE CONTROVERSY regarding the role of pulmonary artery catheterization (PAC) in critically ill patients, the technique is widely practiced and is the standard of care in many intensive care units (ICUs). There are generally accepted indications for PAC monitoring despite the absence of prospective studies showing improvement in patient outcomes with PAC.[1-3] Some studies have suggested an association between PAC usage and increased morbidity and mortality.[4-6] The debate continues today, and resolution requires a definitive, prospective, randomized study. If a less invasive technology capable of providing the same data were available, however, the argument would be moot. Although there have been several promising advances in the area of *substitute technology,* none has been extensively tested in the arenas that PACs are presently used.

The field of thoracic electrical bioimpedance (TEB) monitoring of cardiac output has yielded unclear results. More recent studies using upgraded computer technology and refined algorithms suggest that many of the problems with first-generation technology have been overcome.[7-14]

In the ICU, PACs are used for 2 main purposes: (1) to assess cardiac filling pressures, cardiac output and cardiac index, and vascular resistance at a single time point, and (2) to track hemodynamic performance over time. The second-generation TEB studies and some earlier studies have shown reasonably good agreement between TEB-derived and PAC-derived (thermodilution [TD]) cardiac index.[7,8,10-14] Virtually all previous studies compare pooled groups of patients' measurements, however, and do not distinguish interpatient from intrapatient data. It has not been investigated in individual patients whether the agreement between TEB and TD is maintained throughout a monitoring period. The authors were curious about the ability of a new generation of bioimpedance monitor to measure accurately and follow the hemodynamic state of individual patients after surgery with cardiopulmonary bypass (CPB) (either coronary artery bypass graft [CABG] surgery or valve replacement). This group of patients was chosen for several reasons, as follows: (1) It is well known that patients returning from CPB may show significant hemodynamic instability in the first 12 to 18 hours in the ICU, (2) prior work from one of the authors[14] had shown a first-generation TEB monitor to have limitations in accuracy between TEB and TD in this setting, and (3) CPB surgeries represent one of the most frequent uses of PAC.[15] If equivalency could be shown between TEB and TD, the noninvasive system (TEB) may reduce morbidity, mortality, and overall costs associated with the use of PACs.[16-21] It was hypothesized that a new-generation TEB hemodynamic monitor would show clinically acceptable agreement with PAC-derived cardiac index in post-CPB patients during the initial ICU period.

MATERIALS AND METHODS

After receiving approval of the protocol, including the consent form, by the institutional review board, patients undergoing CPB were considered for enrollment into the study between December 1998 and April 1999. Patients enrolling gave consent. PAC is considered the standard of care in CPB patients in this institution. Patients underwent either CABG or valve replacement surgery. Patients having valve replacements were enrolled if their dysfunctional valves were replaced at surgery (ie, no significant postoperative valvular pathology existed) and their pulse was <150 beats/min when initial ICU assessment was completed.

On entry to the ICU, patients were promptly connected to the TEB system in accordance with manufacturer's instructions (BioZ System 1.52, Cardiodynamics International Corporation, San Diego, CA). The time clock on the TEB monitor and the patient's clinical monitor were synchronized. The BioZ System provides a near-continuous measure of cardiac output and cardiac index by measuring stroke volume of each

From Private Practice, Pulmonary Medicine, Monterey, CA; Department of Biostatistics, Naval Medical Center, San Diego, CA; and Department of Anesthesiology, Virginia Commonwealth University/ Medical College of Virginia, Richmond, VA.

Work performed at Naval Medical Center, San Diego, CA.

W.S.S. is a member of the medical advisory board of Cardiodynamics International Corporation.

B.D.S. is a member of the medical advisory board of Cardiodynamics International Corporation and has a financial relationship with Cardiodynamics International Corporation.

Address reprint requests to Bruce D. Spiess, MD, Cardiothoracic Anesthesiology, VCU/MCV, PO Box 980695, 1200 East Broad Street, Richmond, VA 23298-0695.

1053-0770/02/1601-0003$35.00/0
doi:10.1053/jcan.2002.29635

Table 1. Sample Sizes (Number of Readings per Sample), Means, and SDs for the Pooled Set of 20 Patients and for the Individual Patients

Patient No.	Sample Size	Thermodilution		Bioimpedance		Difference	
		Mean	SD	Mean	SD	Mean	SD
Pooled	216	2.641	0.480	2.575	0.505	0.066	0.190
1	5	2.514	0.147	2.726	0.115	−0.212	0.150
2	6	2.100	0.178	2.090	0.289	0.010	0.211
3	4	3.175	0.263	3.200	0.115	−0.025	0.050
4	5	2.776	0.170	2.840	0.182	−0.064	0.059
5	10	2.750	0.321	2.410	0.331	0.340	0.227
6	7	2.986	0.418	2.800	0.493	0.186	0.107
7	6	3.217	0.271	3.100	0.141	0.117	0.279
8	4	2.612	0.248	2.600	0.258	0.013	0.063
9	13	2.346	0.315	2.415	0.365	−0.069	0.125
10	10	3.720	0.308	3.790	0.238	−0.070	0.149
11	6	2.933	0.266	2.967	0.197	−0.033	0.082
12	9	2.167	0.112	2.144	0.142	0.022	0.067
13	11	2.718	0.328	2.755	0.466	−0.036	0.169
14	25	2.534	0.531	2.348	0.560	0.186	0.168
15	15	2.707	0.237	2.640	0.196	0.067	0.145
16	18	2.367	0.243	2.394	0.221	−0.028	0.145
17	25	2.534	0.531	2.348	0.260	0.186	0.168
18	15	2.707	0.237	2.640	0.196	0.067	0.145
19	11	2.791	0.234	2.718	0.232	0.073	0.168
20	11	2.718	0.266	2.191	0.187	0.055	0.151

heartbeat and averaging that over a number (chosen by the user) of cardiac cycles (usually 16 to 30 beats). For this study, the recording interval was set at every minute, and the TEB device was left connected but unavailable for display of data to the ICU team. During the next 12 to 18 hours, TD-derived hemodynamic measurements were made and recorded whenever the clinical situation warranted. The next day, the TEB data were downloaded into a laptop computer. The bedside hemodynamic information was also stored in a computer 24-hour summary format with exact times noted when hemodynamic data were collected.

Required sample size could not be estimated at the outset because there were no established data on the TEB readings for inputs to a power analysis. After some data had been collected, the authors used the pilot data ($\alpha = 0.05$, power $= 0.80$) and a clinically significant mean difference of 0.1 to show a means test requirement of 16 at minimum; the authors took 20.

After electronic collection, the data on cardiac index from TEB and TD were converted to an Excel Spreadsheet (Microsoft Corp, Redmond, WA) and analyzed partly with the aid of a statistical software package (Stata, Stata Corp, College Station TX) and partly by hand calculation. Descriptive statistics regarding the patient population were obtained to characterize the group studied. Because equivalence of technologies remains difficult to prove by any single statistical method, the authors considered it important to examine many questions. These were broken down into 2 major questions with subquestions for each. Question 1 was: *On average for the entire group of patients,* does TEB provide the same information as TD? Question 2 was: *For each individual patient,* does TEB provide the same information as TD? Within each of these broad questions, 3 to 5 subordinate questions were asked. The first question was wholly statistical, dealing with measurement equivalence as a whole between groups of patients and multiple measurements within each patient. To answer this first question, the pooled data from all patients were used to examine TEB and TD and the difference (DIF = TD − TEB) between technologies. The second question included a clinical judgment component, requiring that clinicians establish a range of measurement variability that they are willing to tolerate. The subordinate questions are listed as 1a, 1b, and so forth. Details of the statistical reasoning used for each subordinate question can be found in the Appendix.

RESULTS

Twenty patients were enrolled in this study, generating the 216 matched data points used in the study. All data from the 20 patients were used. Descriptive statistics are shown in Table 1.

The analysis addressed to question 1 on equivalency of average TEB and average TD was: (a) Can the data be pooled across patients? (b) Is the relationship between TEB and TD linear? (c) Are TEB and TD readings equivalent for the entire patient pool? (d) If so, how well do TEB and TD agree in the patient pool? (e) Are TEB readings as dependable as TD?

Question 1a was: May the data for several readings each on multiple patients be pooled without violating the statistical assumption of independence? An almost universal assumption underlying statistical tests is the independence of residuals (remainders of readings after removing components resulting from known causal factors). Readings between patients possess at least the potential for patient-to-patient influence not present in readings from a single patient. Will the pooling of 5 readings for the first patient, 6 for the second, and so forth, violate this assumption? The deviations from the mean with or without this potential influence were compared by an F test. The null hypothesis was not rejected ($p = 0.12$) (details are in the Appendix).

Question 1b was: Is the relationship between TEB and TD linear? In Fig 1, TEB is plotted against TD with a 45° angle of identity shown. A close agreement is apparent to the eye, but to verify linearity, quadratic regression was performed, testing significance to the second degree. The test yielded $p = 0.835$ for the second-degree term and the coefficient of determination (R^2) remained the same (0.859) on adding the second-degree term.

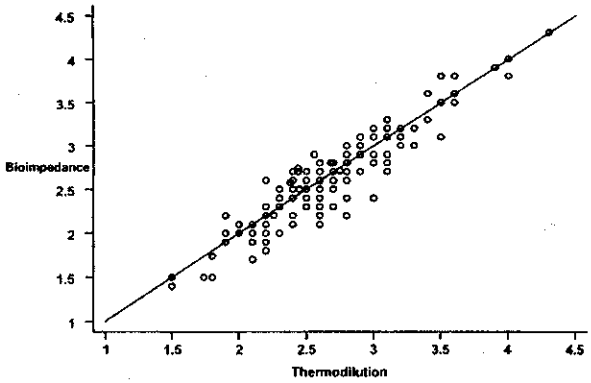

Fig 1. Bioimpedance measures of cardiac index plotted against those from thermodilution for 216 readings.

Question 1c was: Is mean TEB equivalent to mean TD for pooled patients? The authors used the approach of bioequivalence testing, in which the null hypothesis poses that TEB *is different* from TD by a clinically chosen amount. Failure to show this difference leads to the conclusion of the alternate hypothesis—that the 2 methods are not different. (This method is explained further in the Appendix.) TEB and TD were decided to be clinically different if they were ≥20% apart. (Although there is no established difference to be adopted generally, the closest is the 20% Food and Drug Administration ([FDA] standard for pharmaceutical equivalence testing.)[22] An equivalence *t*-test yielded $p < 0.001$. (The same test, but addressing only question 1c, with 76 additional patients [96 total] of similar character at the same hospital has been conducted,[23] leading to the same conclusion for a larger sample.)

For question 1d, to address whether the measures differ significantly, bias and precision analysis was performed on the data set as a whole. The bias (mean difference between measures) was −0.07 L/min/m^2, and the precision (95% confidence interval of the mean difference) was 0.4 L/min/m^2. Fig 2 shows differences, with the mean confidence bounds superposed, as a function of the average of the 2 measures. Lin's concordance correlation coefficient[24] was calculated as a further measure of equivalence and found to be 0.990 (details are in the Appendix).

Question 1e was: If TEB and TD agree well on average, is the variability of one greater than the other, indicating less dependability? The authors followed equivalence testing logic similar to that used in question 1c and chose 20% for an acceptable SD. The null hypothesis became that the TEB SD (the larger one in the sample) was ≥1.20 times the SD of TD. Rejection of this null hypothesis provides statistical evidence for the alternate hypothesis of no difference. The variance ratio was tested by F test yielding a p value = 0.027 (details are in the Appendix).

Additional information in addressing only question 1e is available. A different group of 10 post-CPB patients was randomly selected to establish the contemporary variance in TD determination within the institution; this was done by recording every TD measurement obtained by injecting 10 mL of room-temperature saline as rapidly as possible. If 3 outputs were within 10% of each other, they were averaged, and that number was used as the *actual* output. Failure to obtain 3 readings within 10% required reshooting outputs, and higher values were discarded as outliers. All outputs were shot by the clinical nursing staff in the ICU. Results of this substudy showed a 95% confidence interval for TD cardiac indices of ±17%.

A similar attempt at establishing variance with TEB cardiac index was performed. Data were recorded every 5 seconds for 2 minutes in another series of 10 random post-CPB patients. This group of patients was thought to be hemodynamically stable (pulse, 70 to 90 beats/min; systolic blood pressure, 95 to 120 mmHg; diastolic blood pressure, 55 to 80 mmHg; and temperature, 98.0°F to 100.0°F). The 95% confidence interval for cardiac index measurement in TEB was ±5%.

Question 2 on agreement by the methods for individual patients was addressed by asking: (a) Are TEB readings equivalent to TD readings for each patient? (b) If so, how well do TEB and TD agree for each patient? (c) In light of inherent measurement error in both technologies, is the time-dependent pattern of TEB and TD the same for each patient (ie, do they move in the same direction with the same magnitude in each patient at successive readings)?

Question 2a was: Is TEB equivalent to TD for individual patients? Following the same logic as for question 1c, the authors performed a bioequivalence test for each patient. The FDA clinical tolerance of ±20% was selected for each patient, and Fig 3 shows means and confidence intervals superimposed on the 20% tolerance lines for each patient.

For question 2b, Lin's concordance correlation coefficient was calculated as in Appendix question 1d for each patient. Of 20, 15 (75%) exceeded a coefficient of 0.9, and 18 (90%) exceeded a coefficient of 0.7. Fig 4 shows the frequency distribution. Patients exhibiting coefficients of <0.9 were those in whom the matched data set was relatively small (4 to 10).

Question 2c was: Do the patterns of increase and decrease over time agree for each patient? For each patient, the TEB and TD readings were graphed. A plus sign was assigned if the difference from the first to second and second to third (and so forth) measurements was the same (both techniques moved up or down), and a minus sign was assigned for the opposite (ie,

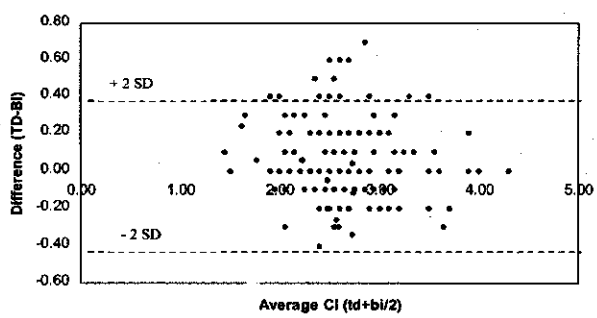

Fig 2. Bias and precision plot for all 216 readings (bias = 0.07 L/min/m², precision = ±0.40 L/min/m²). Abbreviations: TD, thermodilution; BI, bioimpedance; CI, confidence interval.

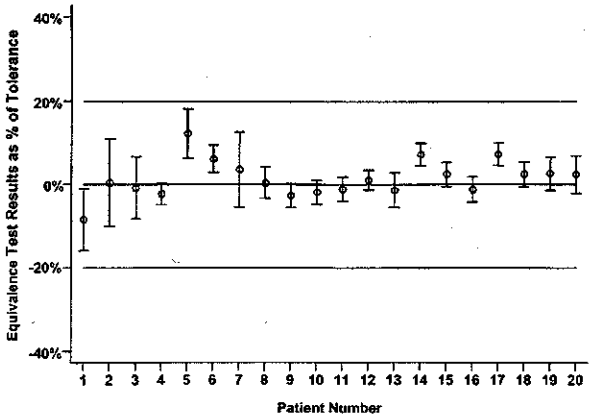

Fig 3. A plot of the equivalence tests for each of the 20 patients, converted to percent of clinical tolerance to be shown on the same axis. Note no whiskers cross the ±20% tolerance lines; cardiac index by bioimpedance is accepted as equivalent to cardiac index by thermodilution for all patients using this test.

one technique moved up, while the other moved down). The percentage of pluses was recorded, and the distribution of these percentages is shown in Fig 5. There was perfect (100%) agreement in 40%, and in 55% the agreement was ≥90%. Agreement was >75% for all patients.

The question of whether the differences between TEB and TD remain clinically insignificant in each patient over time was addressed by plotting successive TEB and TD values calculated as a percentage of the initial reading for each technique. The authors reasoned that clinicians were willing to accept modest differences between TD and TEB as long as both measures showed the same degree of relative movement over time. The results of this comparison are plotted as an identity plot in Fig 6. The coefficient of determination from a regression fit, calculated to be 87%, represents the goodness of fit, whereas the correlation coefficient, calculated to be 0.95, shows how closely matched the data are.

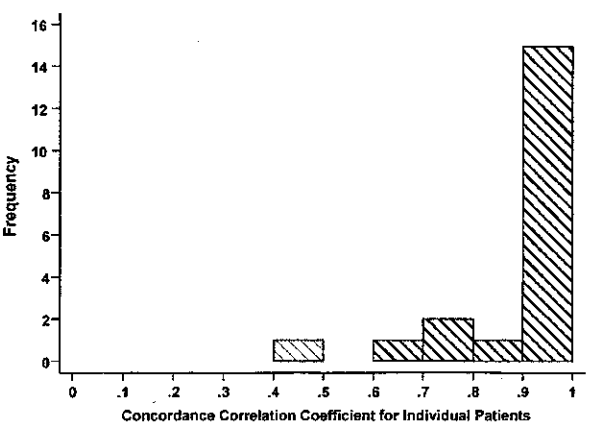

Fig 4. The frequency distribution of Lin's concordance correlation coefficient for 20 patients.

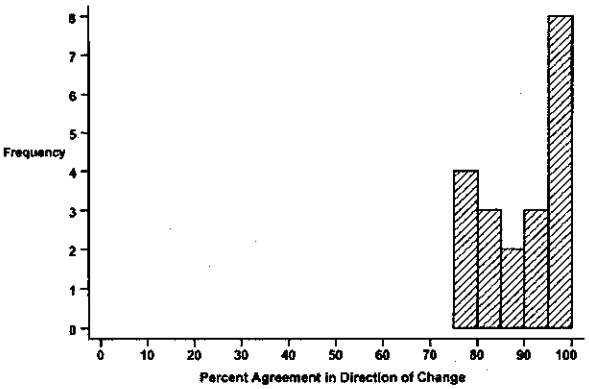

Fig 5. The distribution of percent agreement between bioimpedance and thermodilution in direction of movement from one reading to the next.

DISCUSSION

TEB and TD measurements of cardiac index in post-CPB patients were compared through multiple statistical analyses. From the analysis of question 1a, it was concluded that the readings may be pooled for the purpose of examining overall agreement between TEB and TD. From the analysis of question 1b, it was concluded that the relationship between TEB and TD is linear. From the analysis of question 1c, the null hypothesis was contradicted, and it was concluded that mean TEB for the population was no different from mean TD. For question 1d, it was believed that these TEB measures do not differ clinically from the standard of TD. In support, the remarkable level of concordance of 0.990 was found. In question 1e, the variability of the 2 measures was compared. The authors rejected the hypothesis that population SD of TEB (0.505 in the sample) is greater and concluded that it is equivalent to the population SD of TD (0.480 in the sample). To answer question 2a on TEB equivalence to TD for individual patients, the confidence intervals did not cross clinical tolerance. The null hypothesis was rejected, and it was concluded that TEB and TD are equivalent for all patients. To answer 2b, the authors noted that the concordance coefficients for individual patients are adequately high. To answer 2c, the analyses in 2a and 2b show that

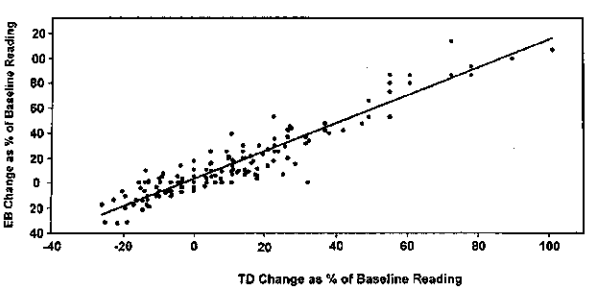

Fig 6. Plot of direction and magnitude of change over time (in percent) with respect to a baseline value (T_0). Percent change TEB = 1.114%; percent change TD = 3.6%; r = 0.95. Abbreviations: TEB, thoracic electrical bioimpedance; TD, thermodilution.

matched data points between TEB and TD within individuals are essentially equivalent.

The analyses of the several ways in which TEB and TD might differ show that these techniques are equivalent. This is the first time in post-CPB patients that bioimpedance cardiography has been shown to be accurate and clinically interchangeable with the existing technique of PAC. The use of the PAC has been criticized, a beneficial impact on patient outcome has never been shown, and questions about its role in increasing morbidity and mortality continue to surround this invasive procedure. Data from the American Heart Association revealed that on average in recent years >600,000 cardiac surgeries were performed in the United States, and PAC has been estimated to add an extra $1,402 in direct costs and prolonged hospital stay (all costs including documented extra ICU time).[19,20] Arguments may occur regarding what percentage of cardiac cases truly require PAC, but the financial impact of their use is large.

PAC has been widely used to assess cardiac index; however, it is not a gold standard. Accuracy of PAC depends on certain assumptions. TD has been noted to have 15% to 22% variability in sequential readings, and differences exist between one TD computer module and another within the same patient.[25] Respiratory cycle, temperature of injectate, speed of injection, placement of the catheter within the atrium and pulmonary artery, and intramyocardial blood flow all affect the reproducibility of TD readings.[26,27] At best, TD is an estimate of a physiologic function. Fick measurement of cardiac index has been called the *physiologist's gold standard,* but it requires stable hemodynamics during the collection time and is difficult to apply in the ICU even as a research tool. It was thought that by comparing equivalence of TEB with the admittedly imperfect measure TD, a clinically realistic comparison could be made.

When 2 technologies that are not gold standards are compared, multiple statistical approaches, each with its own limitations, have been used. Regression analysis assumes there is no measurement *error* in the *x axis* and looks at how well a relationship is maintained between the 2 variables throughout a data set. Two methods of measurement may be correlated perfectly but always differ by a constant amount (parallel lines on an identity plot). Bias and precision analysis treat all data pairs as independent points and pool the data. In clinical medicine, this does not reflect how hemodynamic measures are used. Treatment of the data in this way reveals little about the maintenance of between-technique differences over time. Taken alone, each of these tests describes only one part of the complex relationships that require evaluation between 2 estimates of a physiologic function before those measurements can be deemed equivalent. In general, equivalence in clinical comparisons requires the blending of rigorous statistical tests with the uncertainties and generalities of clinical practice and biological variability. In most cases, exact numbers for cardiac index are not necessary for clinicians to feel comfortable; ranges and trends or movements in the data are what drive decision making. There are numerous examples wherein clinician lack of understanding of PAC data significance influences therapeutic decision making.[28-31]

The present study used a systematic method of evaluating 2 *imperfect* measures of cardiac index as they changed over time. The analysis chose the FDA's 20% variability as the tolerance limit for error, then provided a straightforward method of testing equivalence. Concordance testing addresses the level of agreement, bias and precision tests provide perspective on the overall range of differences between measures, and correlation and regression describe the ability of each technique to vary in the same direction with proportional magnitude over time. The analysis carried out in this study satisfactorily considers each element of equivalence testing and shows that TEB and TD are equivalent in the setting of postoperative cardiac surgery.

The interest in and demand for an accurate noninvasive cardiac output monitor remain high despite the problems that bioimpedance monitoring has had in the past. In the 1990s, advances in equipment technology and signal processing proceeded to the point that physiologic measurements (ie, ventricular ejection time and dZ/dT [change of impedance over time]) can proceed in real-time. Artifact rejection has improved as have some of the algorithms regarding patient characteristics. Prior work suggested that TEB could not work in post-CPB patients.[15] Many reasons had been proposed for this with the first-generation TEB systems, including the fact that the chest had been opened and mediastinal hematoma formation had occurred. That clearly is no longer the case.

Despite the encouraging findings from the present study, some limitations must be understood. TEB requires the adherence of chest and neck electrodes, which in post-CPB patients is possible even with the presence of an internal jugular central venous cannulation. The data reported here are from 20 patients, and although problems were not encountered and equivalence was shown, it is possible that in other post-CPB situations, such as aortic or mitral valve regurgitation, equivalence would not be as good. The TEB electrodes must stick well, and if patients have oily skin or are diaphoretic, the electrodes may become dislodged. The ICU is an electronically hostile environment, and patients often have many electronic medical devices attached. Electrical interference may be encountered in other patients or other institutions using equipment different from that used in this study. As with TD, measurement familiarity with the equipment is required for accurate data retrieval. Knowledge of the impedance waveform morphology (just as TD waveforms) is important because waves that appear unreal may yield spurious data. This study showed equivalency with respect to cardiac index (and output) measures, but that does not denote clinical utility. The BioZ monitor gives an estimate of systemic vascular resistance but no preload measurement. It is unclear how important an estimate of preload is for the clinician in the immediate post-CPB patient. Can physicians learn to estimate preload from other clinical signs, or can the cardiac index and systemic vascular resistance data be sufficient that a clinician can test preload adequacy with small fluid challenges? Such questions influence clinical utility of TEB in individual ICUs and patient subgroups.

In summary, this study has shown that 216 paired readings of TEB and TD cardiac index in 20 patients after CPB are linearly related, quantitatively equivalent in mean and variance, and highly concordant across patients and track well over time. TEB and TD follow the same pattern of increase and decrease most of the time and are matched perfectly > 50% of the time. TEB in the ICU after CPB is a promising noninvasive, potentially low-cost alternative to TD and PAC hemodynamic measurement. Much work remains to prove time-tested clinical utility and patient outcome improvement by using TEB. The authors believe that this technology shows a great deal of promise.

REFERENCES

1. American Society of Anesthesiologists Task Force on Pulmonary Artery Catheterization: Practice guidelines for pulmonary artery catheterization. Anesthesiology 78: 380-394, 1993
2. Chernow B: Pulmonary artery flotation catheters: A statement by the American College of Chest Physicians and the American Thoracic Society. Chest 111: 261-266, 1997
3. Tuman KS, McCarthy RJ, Spiess BD, et al: Effect of pulmonary artery catheterization on outcome in patients undergoing coronary artery surgery. Anesthesiology 70:199-206, 1989
4. Connors AF, Speroff T, Dawson NV, et al: The effectiveness of right heart catheterization in initial care of the critically ill patients. JAMA 276:889-897, 1996
5. Guyatt G, Ontario Intensive Care Study Group: A randomized controlled trial of right heart catheterization in critically ill patients." J Intensive Care Med 6:91-95, 1991
6. Eisenberg PR, Jaffee AS, Schuster DP: Clinical evaluation compared to pulmonary artery catheterization in the hemodynamic assessment of critically ill patients. Crit Care Med 12:549-553, 1984
7. Wang DH, Tremper KK, Stemmer EA, et al: Noninvasive cardiac output: Simultaneous comparison of two different methods with thermodilution. Anesthesiology 72:784-792, 1990
8. Siegel LC, Shaher SL, Martinez GM, et al: Simultaneous measurement of cardiac output by thermodilution, esophageal Doppler and electrical impedance in anesthetized patients. J Cardiothoracic Vasc Anesth 2:590-595, 1988
9. Shoemaker WC, Wo CC, Bishop MH, et al: Multicenter trial of a new thoracic electrical bioimpedance device for cardiac output estimation. Crit Care Med 22:1907-1912, 1994
10. Ziegler D, Grotti L, Krucke G: Comparison of cardiac output measurements by thoracic electrical bioimpedance versus intermittent bolus thermodilution in mechanically ventilated patient. Chest 116: 281S, 1999 (suppl 4)
11. Raaijmahers E, Faes TJ, Scholten RJPM, et al: A meta analysis of three decades validating thoracic impedance cardiography. Crit Care Med 27:1203-1213, 1999
12. Greenberg BH, Hermann DD, Pranulis MF, et al: Reproducibility of impedance cardiography hemodynamic measures in chemically stable heart failure patients. Congestive Heart Fail 6:19-26, 2000
13. DeMaria AN, Raisinghani A: Comparative overview of cardiac output measurement methods: Has impedance cardiography come of age? Congestive Heart Fail 6:7-18, 2000
14. Sageman WS: Reliability and precision of a new thoracic electrical bioimpedance monitor in a lower body negative pressure model. Crit Care Med 27:1986-1990, 1999
15. Sageman WS, Amundson DE: Thoracic electrical bioimpedance measurement in post aortocoronary bypass patients. Crit Care Med 21:1139-1142, 1993
16. Haley RW, Schaberg DR, Crossley KB, et al: Extra charges and prolongation of stay attributable to nosocomial infections: A prospective inter-hospital comparison. Am J Med 70:51-58, 1981
17. Resweik JM, Engeler CM, Derauf B: Postmortem angiography of catheter-induced pulmonary artery perforation. J Forensic Sci 37: 1346-1351, 1992
18. Purut CM, Scott SM, Parham JV, et al: Intraoperative management of severe endobronchial hemorrhage. Ann Thorac Surg 51:304-306, 1991
19. Ramsey SD, Sanjay S, Sullivan SD, et al: Clinical and economic effects of pulmonary artery catheterization in non-emergent coronary artery bypass graft surgery. J Cardiothoracic Vasc Anesth 14:113-118, 2000
20. American Heart Association: Statistics, 1999. Available at http: www.americanheart.org/statistics/09medicl.html.
21. Spackman TN: A theoretical evaluation of cost effectiveness of pulmonary artery catheters in patients undergoing coronary artery surgery. J Cardiothorac Vasc Anesth 8:570-576, 1994
22. Center for Drug Evaluation and Research (CDER): Guidance for Industry: Statistical Approaches to Establishing Bioequivalence. Rockville, MD, Food and Drug Administration, 2001
23. Riffenburgh RH: Statistics in Medicine. San Diego, CA, Academic Press, 1999, pp 320-321
24. Lin LI-K: A concordance correlation coefficient to evaluate reproducibility. Biometrics 45:255-268, 1989
25. Mathew EB, Vender JS: Comparison of thermodilution cardiac output measured by different computers. Crit Care Med 15:989-998, 1987
26. Stetz CW, Miller RG, Kelly GE, et al: Reliability of thermodilution method in the determination of cardiac output in clinical practice. Am Rev Respir Dis 126:1001-1004, 1982
27. Morris AH, Chapman RH, Garcher RM: Frequency of technical problems encountered in the measurement of pulmonary artery wedge pressure. Crit Care Med 12:164-170, 1984
28. Iberti TJ, Fisher EP, Leibowitz AB, et al: A multicenter study of physician's knowledge of the pulmonary artery catheter. JAMA 263: 2928-2932, 1990
29. Gnaegi A, Feihl F, Perret C: Intensive care physician's insufficient knowledge of the pulmonary artery catheterization at the bedside: Time to act. Crit Care Med 25:213-220, 1997
30. Gattinoui L, Brazzi L, Pelosi P, et al: A trial of goal-oriented hemodynamic therapy in critically ill patients. SrO2 Collaborative Group. N Engl J Med 333:1025-1032, 1995
31. Hayes MA, Timmins AC, Yau EH, et al: Elevation of oxygen delivery in the treatment of critically ill patients. N Engl J Med 330:1917-1922, 1994
32. Dunnett CW, Gent M: Significance testing to establish equivalence between treatments, with special reference to data in the form of 2 × 2 tables. Biometrics 33:593-602, 1977
33. Schuirmann J: A comparison of the two-sided test procedure and the power approach for assessing the equivalence of average bioavailability. J Pharmacokin Biopharm15:657-680, 1987

APPENDIX

Question 1a. The *within-patient* mean square (ie, the pooled variance of deviations from patient means) was calculated. Similarly the *between-patient* mean square (ie, the pooled variance of deviation from the mean of first readings for each patient, second for each, third for each, and so forth) was calculated for the first 15 readings. (Reading sequences 16 to 18 have only 3 readings each, and 19 to 25 have only 2 readings each.) If a patient-to-patient causal factor exists that would violate the pooling, the *between-patient* mean

square would be significantly larger than the *within-patient* mean square. An F test of H_0 was detected: $\sigma_B^2 = \sigma_w^2$ against H_1: $\sigma_B^2 > \sigma_w^2$. F = 1.7782, and $p = 0.121$. H_0 was not rejected.

Question 1c. Bioequivalence methods have become conventional in recent years. The earliest article was by Dunnett and Gent in 1977.[32] Schuirmann's article[33] established the format currently most used. Riffenburgh[23] contains a succinct (2-page). Twenty percent of mean TD is 0.528, and the SEM of the difference TD − TEB = 0.013. The authors tested the difference between mean TEB ($m_B = 2.575$ estimating the population μ_B) and mean TD using a *t*-test with $\alpha = 0.05$. (The assumptions of normality and equal variance are satisfied.) For an equivalence test, H_0: $|\mu_T - \mu_B| = 0.528$ and H_1: $|\mu_T - \mu_B| < 0.528$. H_0 is rejected if $|m_T - m_B| + t_{1-\alpha/2}\,\text{SEM} < 0.528$[2]. The authors substituted to find $|2.641 - 2.575| + 1.971 \times 0.013 = 0.092$, which is far less than the critical value 0.528. $p < 0.001$.

Question 1d. Lin's concordance correlation coefficient[23] is the concordance-measure equivalent of a bioequivalence test. It combines measures of precision and accuracy, and it may be interpreted similar to a correlation coefficient (1 is perfect agreement, 0 is complete independence, and −1 is perfect disagreement). It can be written in the form $C_b = 2s_T s_B/[s_T^2 + s_B^2 + (m_T - m_B)^2]$, where s_T and s_B are the sample SDs 0.480 and 0.505 for TD and TEB. $C_b = 0.990$.

Question 1e. The authors posed null and alternate hypotheses: H_0: $\sigma_B^2/(1.44 \times \sigma_T^2) = 1$ versus H_1: $\sigma_B^2/(1.44 \times \sigma_T^2) < 1$. (Although the question was posed for SDs, the outcome of the test is the same using variances.) The test statistic was $F = s_B^2/(1.44 \times s_T^2)$. The form of the alternate hypothesis implies using the left tail of the F distribution to find the critical value such that the probability of erroneously rejecting the null hypothesis is 5%. For 215, 215 *df*, the left 5% area under the F distribution yields the critical value as 0.799. We calculated $F = 0.505^2/(1.44 \times 0.480^2) = 0.769$. Because this is <0.799, we reject H_0. The *p* value was 0.027.

Question 2a. Logic and formulae similar to those covered in question 1c were followed. The *p* values for patients 1 through 20 were 0.012, 0.005, 0.004, <0.001, 0.017, <0.001, 0.006, and the remainder <0.001.