Intensive Care Medicine

When More is not Less: A Robust Framework to Evaluate the Value of a Diagnostic Test in Critical Care

--Manuscript Draft--

Manuscript Number:	
Full Title:	When More is not Less: A Robust Framework to Evaluate the Value of a Diagnostic Test in Critical Care
Article Type:	Original
Funding Information:	National Institute of Biomedical Imaging and Bioengineering (R01 EB017205-01A1) — Dr. Leo Anthony Celi
Abstract:	Purpose: While the use of echocardiography in the ICU is rapidly expanding, the impact of transthoracic echocardiography (TTE) on patient outcomes among patients with sepsis has not been examined. The study was designed to examine the value of TTE among critically ill patients with sepsis. Methods and Results: The MIMIC-III database was employed to identify pa-tients with sepsis who had and had not received transthoracic echocardiography. The statistical approaches utilized included multi-variate regression, propensity score analysis, doubly robust estimation, the gradient boosted model and an in-verse probability-weighting model to ensure the robustness of our findings. Signif-icant benefit in terms of 28-day mortality was observed among the TTE patients compared to the control group (Odds Ratio = 0.78, 95% CI = 0.67~0.89 and p-value <0.001). The amount of fluid administered (2.5 liters vs. 1.9 liters on day 1, p<0.001), use of dobutamine (4% vs. 1%, p<0.001) and the maximum dose of norepinephrine (1.76 vs. 0.81 mg/min, p<0.001) were significantly higher for the TTE patients. Significantly greater reductions in serum lactate (1.35 vs. 0.84, p<0.001) and serum creatinine (0.79 vs.0.37, p<0.001) were also observed in the TTE group. Importantly, the TTE patients were weaned off vasopressors more quickly than those in the 'no TTE' group (vasopressor-free days on day 28 of 15.4 vs. 10.9, p<0.001). Conclusion: In a general population of critically ill patients with sepsis, use of TTE is associated with an improvement in 28-day mortality.
Corresponding Author:	Jakob McSparron, MD University of Michigan Ann Arbor, MI UNITED STATES
Corresponding Author Secondary Information:	
Corresponding Author's Institution:	University of Michigan
Corresponding Author's Secondary Institution:	
First Author:	Mengling Feng, PhD
First Author Secondary Information:	
Order of Authors:	Mengling Feng, PhD Jakob McSparron, MD Dang Trung Kien David Stone, MD David Roberts, MD Richard Schwartzstein, MD Antoine Vieillard-Baron, MD, PhD

	Leo Anthony Celi, MD
Order of Authors Secondary Information:	
Author Comments:	Dear Professor Vieillard-Baron, Thank you kindly for your consideration of " When More is not Less: A Robust Framework to Evaluate the Value of a Diagnostic Test in Critical Care" as an original investigation in response to your call for papers for the thematic cardiovascular issues edition of Intensive Care Medicine. In particular, our work responds to the editorial points that "future clinical trials need to address how monitoring of heart function and haemodynamics by echocardiography may impact on delivered therapies and ultimately patient-centred outcomes" as well as now being possible with the "evolution of electronic medical records" and "the potential to analyse big data". Understanding the clinical value of diagnostic tests and interventions performed in the intensive care unit (ICU) has major implications for improved quality of care and cost. The use of real-world data that is captured in electronic health records for this purpose, although clearly complex given the confounding by indication, is nonetheless necessary given that for most of ICU practice, randomized controlled trial (RCT)-based data are lacking and no RCT will be likely be performed to provide evidence. In view of the marked increase in the use of transthoracic echocardiography (TTE) in the critical care setting, we set out to examine the clinical utility of TTE among patients with sepsis using a novel framework to interrogate big data. Our findings suggest that there is a significant mortality benefit when TTE is performed among patients admitted to the ICU with sepsis. We are eager to share this approach to assess value using big data with the medical community, and we look forward to your review. Jakob I. McSparron MD
Suggested Reviewers:	

When More is not Less:
A Robust Framework to Evaluate the Value of a Diagnostic Test in Critical Care

Mengling Feng[1*] PhD, Jakob I. McSparron[2*] MD, Dang Trung Kien[1,] David J. Stone MD[3], David H. Roberts[4] MD, Richard M. Schwartzstein[4] MD, Antoine Vieillard-Baron[5] MD PhD, Leo Anthony Celi [4,6] MD

*co-first authors

1. Saw Swee Hock School of Public Health, National University of Singapore
2. Division of Pulmonary and Critical Care Medicine, University of Michigan
3. Departments of Anesthesiology and Neurosurgery, University of Virginia School of Medicine
4. Division of Pulmonary, Critical Care and Sleep Medicine, Beth Israel Deaconess Medical Center
5. Hospital Ambroise Paré, Assistance Publique-Hôpitaux de Paris, Boulogne-Billancourt, France
6. Institute for Medical Engineering and Science, Massachusetts Institute of Technology

Corresponding author:
Jakob I. McSparron
Division of Pulmonary and Critical Care Medicine, University of Michigan
Email: jmcsparr@med.umich.edu

Word Count: 2929

Conflict of Interest: None

Abstract:

Purpose: While the use of echocardiography in the ICU is rapidly expanding, the impact of transthoracic echocardiography (TTE) on patient outcomes among patients with sepsis has not been examined. The study was designed to examine the value of TTE among critically ill patients with sepsis.

Methods and Results: The MIMIC-III database was employed to identify patients with sepsis who had and had not received transthoracic echocardiography. The statistical approaches utilized included multi-variate regression, propensity score analysis, doubly robust estimation, the gradient boosted model and an inverse probability-weighting model to ensure the robustness of our findings. Significant benefit in terms of 28-day mortality was observed among the TTE patients compared to the control group (Odds Ratio = 0.78, 95% CI = 0.67~0.89 and p-value <0.001). The amount of fluid administered (2.5 liters vs. 1.9 liters on day 1, p<0.001), use of dobutamine (4% vs. 1%, p<0.001) and the maximum dose of norepinephrine (1.76 vs. 0.81 mg/min, p<0.001) were significantly higher for the TTE patients. Significantly greater reductions in serum lactate (1.35 vs. 0.84, p<0.001) and serum creatinine (0.79 vs.0.37, p<0.001) were also observed in the TTE group. Importantly, the TTE patients were weaned off vasopressors more quickly than those in the 'no TTE' group (vasopressor-free days on day 28 of 15.4 vs. 10.9, p<0.001).

Conclusion: In a general population of critically ill patients with sepsis, use of TTE is associated with an improvement in 28-day mortality.

Keywords: echocardiography, sepsis, value, critical care

List of Acronyms

CI = Confidence Interval

GBM = Gradient Boosting Model

IPW = Inverse Probabilities Weighting

IQR = Interquartile Range

LOS = Length-of-stay

MICU = Medical Intensive Care Unit

MIMIC-II = Medical Information Mart for Intensive Care-II

PAC= Pulmonary artery catheter

SAPS = Simplified Acute Physiology Score

SOFA = Sequential Organ Failure Assessment score

TTE = Transthoracic Echocardiography

Introduction:

The clinical value of many tests and interventions used in the care of critically ill patients is unproven. While this circumstance is frequent throughout the healthcare system, it is particularly so in the ICU where randomized controlled trial data is sparse [1,2]. This lack of supportive evidence is well recognized, and persists for a number of reasons including difficulty in obtaining informed consent, pathophysiologic variability in patients with superficially similar clinical presentations, and the pitfalls of interpreting treatment effects and outcomes in a very complex setting. Understanding the clinical value of interventions performed for critically ill patients is enormously important: beyond the epidemiologic significance of the ICU, (a care setting in which six million Americans are treated per year, including one in five Americans at the end of life), identifying interventions that have clinical value--and distinguishing them from those that do not--lays a foundation for effective health policy decision-making [3]. It also promises to improve quality of care, increase cost-effectiveness, and enhance the experience of patients and their families in the ICU. Such knowledge may also reduce clinician burnout by reassuring providers that their interventions have clear-cut benefits [4].

Unsuspected cardiac abnormalities are frequently detected by echocardiography in critically ill patients [5]. While there is evidence that bedside transthoracic echocardiography (TTE) leads to management changes in up to 54% of critically ill patients, the importance and impact of these changes on patient outcomes have not been examined [6-8]. Recent evidence not limited to the

critical care setting demonstrates that less than one third of TTEs lead to an active change in care, with inpatient TTE studies even less likely to result in a change in management [9]. In contrast, a recent study using the National Inpatient Sample suggested that for specific diagnostic purposes, TTE is associated with lower odds of inpatient mortality [10]. Studies thus far have primarily focused on management changes due to TTEs, but the outcome impact of these changes is not clear. While the widespread availability and noninvasive nature of TTE make it an appealing diagnostic tool, the marked increase in the use of TTE in the past ten years has significant financial implications. Use of TTE increased by 90% from 1999 to 2008, accounting for over $1.1 billion of Medicare spending in 2010 [9,11]. Given the increasing attention being placed on value-added care and excessive costs in the ICU, the impact of this expanding technology on patient care warrants further investigation.

Although professional societies have published guidelines for appropriate use of TTE based on expert consensus, many clinicians are not familiar with these guidelines [12]. Notably, approximately 15% of studies are inappropriate according to these guidelines [13]. It has been argued that TTE use in the surgical intensive care unit (SICU) is not cost effective due to a high failure rate, and addition of TTE variables to the APACHE II score does not improve prediction of mortality [14,15]. The current study was designed to investigate the impact of TTE performance on the outcomes of critically ill adult patients with sepsis.

Methods:

Study Cohort

This study is reported in accordance with the STrengthening the Reporting of OBservational studies in Epidemiology (STROBE) statement [16]. We conducted a longitudinal, single center, retrospective study of adult patients from the medical (MICU) and surgical (SICU) intensive care units with a diagnosis of sepsis based on the method established by Angus and colleagues to retrospectively identify patients using billing codes [17].

The study aims to investigate whether TTE independently contributes to improvements in mortality and clinically important changes in the management of septic patients in the ICU. The project was approved by the institutional review boards of the Massachusetts Institute of Technology and Beth Israel Deaconess Medical Center (BIDMC) and was granted a waiver of informed consent.

We utilized the Medical Information Mart for Intensive Care (MIMIC) database, which was developed and is maintained by the Laboratory for Computational Physiology at the Massachusetts Institute of Technology [18]. MIMIC-III contains data from 38,605 ICU patients and includes physiologic information from bedside monitors in the adult ICUs of BIDMC, a tertiary care university hospital, located in Boston, Massachusetts, USA. Hourly physiologic readings from bedside monitors, validated by ICU nurses, were recorded. The database also contains records of demographics, labs, nursing progress notes, intravenous (IV) medications, fluid balance, and other clinical variables. Specialists evaluated radiologic films at the time of patient care, and written

evaluations were stored in the database along with the corresponding time stamps. International Classification of Diseases, Ninth Revision (ICD-9) codes were also documented for specific diseases by hospital staff on patient discharge.

Primary outcome and secondary outcomes

The primary outcome of the study was 28-day mortality from the date of ICU admission. Patient mortality information for discharged patients was gathered from the US Social Security Death Index. Secondary outcomes included number of mechanical ventilation and vasopressor free days within 28 days after ICU admission; use of dobutamine; maximum dose of norepinephrine; IV fluid totals given to patients during their first, second, and third day in the ICU; reduction in serum lactate and serum creatinine between the value recorded nearest to the time stamp of the TTE and 48 hours later for the TTE group, and between the value recorded on days 1 and 3 for the 'no TTE' group.

Statistical Methods

The doubly robust estimation method was applied to infer the independent associations between TTE and patients' primary and secondary outcomes. "Doubly robust estimation combines a multivariate regression model with a propensity score model to estimate the association and causal effect of an exposure on an outcome" [20.21]. Conventionally, when one applies the regression model or the propensity score model individually to estimate a causal effect, both outcome regression and propensity score methods are unbiased only if both of the statistical models are correctly specified. The doubly robust estimator

combines the two approaches such that only one of the two models needs to be correctly specified to obtain an unbiased effect estimator.

The Gradient Boosted Model (GBM) was employed for the estimation of patients' propensity scores for TTE, so that covariate imbalance between the TTE and no TTE groups was minimized. GBM is a machine learning algorithm that consecutively constructs new models and forms an ensemble of models to provide a more accurate estimate of the response variable. The principal idea is to construct the new base-learners to be maximally correlated with the negative gradient of the pre-defined loss function. In our study, regression tree was used as the base learner of the GBM, and a total of thirty-nine covariates were used in the model.

Using the estimated propensity scores as weights, a weighted cohort was generated based on an inverse probabilities weighting (IPW) model [22]. A logistic regression was then performed on the weighted cohort, adjusting for the variables that remained unbalanced between the groups with and without a TTE in the propensity score model, thus the term doubly robust analysis.

To measure the imbalance of covariates for the original and the weighted cohorts, the Wilcoxon signed rank test, a non-parametric test, was used to statistically test the differences among the continuous covariates. A Chi-square test was used to test the differences among the categorical covariates.

The statistical methods in this study were implemented using software R and STATA.

Sensitivity Analysis

We conducted a series of sensitivity analysis to evaluate the robustness of the findings of the study and how our conclusion can be affected by applying various association inference models. We have described our primary model - the doubly robust model - adjusting for unbalanced covariates - in the previous section. In the sensitivity analysis, we applied four model association inferences models: a doubly robust model adjusting for all covariates, a propensity score based IPW model, a propensity score based patient matching model, and a logistic regression based multivariate analysis model. The calculated effect sizes and p-values from all these models were reported and compared.

Covariates

Demographic and admission information: age, gender, weight, day of the week of admission, severity at admission measured by SAPS score, SOFA score and the Elixhauser co-morbidity score [19-24].

Co-morbidities: Congestive heart failure (CHF), atrial fibrillation (AFIB), chronic renal disease, liver disease, chronic obstructive pulmonary disease (COPD), coronary artery disease (CAD), stroke, and malignant tumor. All the co-morbidities were identified based on the recorded ICD9 codes. (A detailed table of ICD9 codes used for each co-morbidity is included in the Appendix.)

Vital Signs: Mean Arterial Pressure (MAP), Heart Rate, Temperature (F) and Central Venous Pressure (CVP) readings at ICU admission.

Interventions: Use of mechanical ventilation, inotropic and vasopressor agents, and sedative drugs during the first 24 hours of ICU admission.

Laboratory results: White blood cell (WBC) count, hemoglobin, platelet, sodium, potassium, bicarbonate, chloride, blood urea nitrogen (BUN), lactate, creatinine, pH, partial pressure of oxygen (PaO2), partial pressure of carbon dioxide (PaCO2), B-type natriuretic peptide (BNP), troponin and creatinine kinase.

We observed that CVP values were not collected for more than half of the patients in our cohort. If we directly used the CVP readings as the co-variate, we would have had a large number of missing values. Instead, we utilized the presence or absence of CVP values as a covariate. Thus, a flag indicating whether CVP was recorded was included as the co-variate in our models. Similarly, laboratory tests for BNP, troponin and creatinine kinase were not ordered in more than half of the cohort. Therefore, flags indicating whether these tests were obtained were used as covariates. (Details around missing values can be found in the Appendix).

Results

After reviewing 38,605 MIMIC-II adult admissions, sepsis was identified in 17,420 admissions based on the Angus methodology [17]. After eliminating patients with multiple ICU admissions, and excluding admissions to the CCU and the cardiac surgical unit, 6,162 patients were included in our study cohort (Fig 1). TTE was ordered for 49.7% of patients within 48 hours of ICU admission. The characteristics of the cohort are summarized in Tables 1. The TTE patients had significantly higher severity scores on admission: SAPS-I score 20.78 (+/- 5.45) vs. 14.63 (+/-5.28), and SOFA score 6.3 (+/-3.8) vs. 5.3 (+/- 3.62). A larger

percentage of the TTE patients received mechanical ventilation (59% vs. 47%) and vasopressor treatments (38% vs 27%) during the first 24 hours of their ICU stay.

Doubly Robust Analysis

A propensity score model was first constructed employing the thirty-nine covariates with the Gradient Boosting Model (GBM). The contributions of individual covariates to the final propensity score are illustrated in Fig 2. The top three covariates were presence of CHF, heart rate, and SOFA score: unsurprisingly, these covariates represent common factors influencing physicians' decisions on ordering TTE.

Based on the estimated propensity scores, inverse probability weighting (IPW) was applied to standardize the differences between the TTE and no TTE cohorts. As shown in Tables 1, most of the covariates of the weighted cohorts were similar or 'balanced' between the groups with and without echocardiograms. The exceptions were SOFA score; mechanical ventilation; use of inotropic, vasopressor and/or sedative medications; the availability of creatinine kinase values; and two co-morbid conditions (CHF and atrial fibrillation). Under the doubly robust estimation framework, a regression model was developed to adjust for these unbalanced covariates on the weighted cohort.

Primary Outcome and Sensitivity Studies

The doubly robust analysis demonstrated a significant beneficial effect of TTE in terms of the 28-day mortality. The propensity score matched mortalities rates for TTE and non-TTE were 24.9% vs 29.5%. The adjusted odds ratio was

0.78 (95% confidence interval=0.67 to 0.89, p<0.001). For the sensitivity analysis, as summarized in Table 2, all five estimation models led to the same conclusion: patients who had TTE had lower 28-day mortalities.

Secondary Outcomes Studies

We evaluated a number of secondary outcomes to investigate potential factors that might account for the beneficial effects of TTE. Several key differences in secondary outcomes were observed. First, the amount of fluid administered to the TTE group was significantly higher on day 1 (2.5 liters vs. 1.9 liters, p<0.001), day 2 (1.3 liters vs. 0.8 liter, p<0.001) and day 3 (0.7 liter vs. 0.3 liter, p<0.001). Second, the use of dobutamine (4% vs. 1%, p<0.001) and, when administered, the maximum dose of norepinephrine (1.76 vs. 0.81 mg/min, p<0.001) were significantly higher for the TTE patients. Third, the TTE group had a significantly shorter duration of vasopressor use (vasopressor-free days on day 28 of 15.4 vs. 10.9, p<0.001). The duration of mechanical ventilation did not significantly differ between the 2 groups. But what is most interesting is the finding of significantly greater reductions in serum lactate (1.35 vs. 0.84, p<0.001) and serum creatinine (0.79 vs.0.37, p<0.001) for the TTE patients. These comparisons are for those values recorded nearest the time stamp of the TTE with those from 48 hours later for the TTE group, and the values recorded on days 1 and 3 for the no TTE group. The detailed results are summarized in Tables 3.

Discussion

Identifying clinical value is challenging when innovations in healthcare are studied [26-27]. This challenge only increases in complex, dynamic

environments like the ICU. At times, new technologies diffuse rapidly based on theoretical benefits from our understanding of disease pathophysiology, but before rigorous evaluations of benefits and harms are performed. Similarly, innovations which have been found to be beneficial in specific patient populations may be applied to other populations in which they have not been adequately studied, potentially exposing patients to harm without commensurate benefit [4]. Notable examples of this phenomenon include the initial enthusiasm and subsequent decline in pulmonary artery catheter utilization, routine use of invasive cardiac catheterization in the initial evaluation of patients with stable coronary disease, and utilization of cardiac computed tomography angiography [28-36]. Examples of other technologies that are commonly used in the ICU but have received little formal utility assessment include electrolyte repletion, insertion of central venous catheters, and the use of renal replacement therapy [37].

The advent of electronic medical records provides a powerful tool for investigating the clinical effectiveness of technologies using real-world data [38]. In light of the uncertainty surrounding the value of most diagnostic tests and interventions used in the ICU, as well as the implications that this evidence gap has for practice and policy, we describe a novel framework that exemplifies how big data can be employed for measuring impact on clinical and/or patient-centered outcomes.

While the use of TTE has steadily increased over the past decade, the implications for patient outcomes remain unknown [10]. There is limited data available in the literature regarding the utility of TTE in critically ill, septic

patients: A recent study by Papolos et al. found that use of TTE was associated with lower odds of in hospital mortality among patients hospitalized for five specific diagnoses, including sepsis [10].

In our study, patients who had TTE had higher severity of illness scores, more co-morbid conditions, and were more likely to receive mechanical ventilation, inotropic, vasopressor and sedative agents. Despite these factors pointing to a sicker group of patients, we found a significantly lower 28-day mortality among patients who had TTE after adjustment for confounding. Considering the factors displayed in figure two, clinicians may particularly want to consider TTE early in the ICU stay for patients with sepsis.

We tested several hypotheses to account for the mortality benefit, and compared several variables between the patients with and without TTE. More fluids were administered to the TTE group on days 1, 2 and 3 in the ICU. Dobutamine was used more often in the group who received TTE, but this might be because a history of CHF was more frequent among this group i.e. it is not certain whether the TTE triggered the use of dobutamine or if it had already been in place. Those who had TTE also had a higher maximum dose of norepinephrine, but surprisingly, were weaned off vasopressors earlier compared to the no TTE group. But perhaps what is most fascinating and physiologically consistent is the finding of greater reduction in serum lactate and serum creatinine among the patients who had TTE. Whether the physiological and mortality improvements are entirely due to the differences in the volume of fluid administered, dobutamine use

and/or maximum dose of norepinephrine is impossible to assess given the sample size.

Our findings raise the possibility that TTE provides information to physicians that may aid in the management of critically ill septic patients. We fully realize that observational, database studies of this kind require careful, multifaceted, and rigorous statistical approaches in order to produce valid, reliable, and actionable results. We believe that we have done so in this regard for the subject at hand, and intend to pursue further such analyses in the future in order to minimize the ambiguity of clinical decision-making in the confounding and complex environment posed by the ICU.

Conclusions

The performance of TTE is associated with a 28-day mortality benefit in a general population of septic, critically ill patients. The mechanism of this benefit remains to be explored but may be related to the increased use of inotropic and vasopressor agents. Given that for most of ICU practice, randomized controlled trial (RCT)-based data are lacking and no RCT will likely be performed to provide evidence, the application of the real-world data that is captured in electronic health records will be necessary to assess the clinical effectiveness of interventions such as TTE. While these investigations must be performed with full awareness of and attention to the complexity, and confounding by indication, of such data applications, they are now possible and we feel, absolutely necessary, in the future development and evolution of optimal clinical care.

References:

1. Angus DC. Caring for the critically ill patient: challenges and opportunities. JAMA 2007;298:456–458.
2. Singer M, Matthay MA. Clinical review: Thinking outside the box--an iconoclastic view of current practice. Crit Care 2011;15:225.
3. Angus DC, Barnato AE, Linde-Zwirble WT, Weissfeld LA, Watson RS, Rickert T, Rubenfeld GD. Use of intensive care at the end of life in the United States: an epidemiologic study. Crit Care Med 2004;32:638–643.
4. Moss M, Good VS, Gozal D, Kleinpell R, Sessler CN. An Official Critical Care Societies Collaborative Statement—Burnout Syndrome in Critical Care Health-care Professionals. Chest 2016;150:17–26.
5. Bossone E, DiGiovine B, Watts S, Marcovitz PA, Carey L, Watts C, Armstrong WF. Range and prevalence of cardiac abnormalities in patients hospitalized in a medical ICU. Chest 2002;122:1370–1376.
6. Stanko LK, Jacobsohn E, Tam JW, Wet CJ De, Avidan M. Transthoracic echocardiography: impact on diagnosis and management in tertiary care intensive care units. Anaesth Intensive Care 2005;33:492–496.
7. Tam JW, Nichol J, MacDiarmid AL, Lazarow N, Wolfe K. What is the real clinical utility of echocardiography? A prospective observational study. J Am Soc Echocardiogr 1999;12:689–697.
8. Orme RML, Oram MP, McKinstry CE. Impact of echocardiography on patient management in the intensive care unit: an audit of district general hospital practice. Br J Anaesth 2009;102:340–344.

9. Matulevicius SA, Rohatgi A, Das SR, Price AL, DeLuna A, Reimold SC. Appropriate use and clinical impact of transthoracic echocardiography. JAMA Intern Med 2013;173:1600–1607.

10. Papolos A, Narula J, Bavishi C, Chaudhry FA, Sengupta PP. U.S. Hospital Use of Echocardiography: Insights From the Nationwide Inpatient Sample. J Am Coll Cardiol 2016;67:502–511.

11. Andrus BW, Welch HG. Medicare services provided by cardiologists in the United States: 1999-2008. Circ Cardiovasc Qual Outcomes 2012;5:31–36.

12. Douglas PS, Garcia MJ, Haines DE, Lai WW, Manning WJ, Patel AR, Picard MH, Polk DM, Ragosta M, Parker Ward R, Weiner RB. ACCF/ASE/AHA/ASNC/HFSA/HRS/SCAI/SCCM/SCCT/SCMR 2011 Appropriate Use Criteria for Echocardiography. A Report of the American College of Cardiology Foundation Appropriate Use Criteria Task Force, American Society of Echocardiography, American Heart Associat. J Am Soc Echocardiogr 2011;24:229–267.

13. Ward RP, Krauss D, Mansour IN, Lemieux N, Gera N, Lang RM. Comparison of the clinical application of the American College of Cardiology Foundation/American Society of Echocardiography Appropriateness Criteria for outpatient transthoracic echocardiography in academic and community practice settings. J Am Soc Echocardiogr 2009;22:1375–1381.

14. Cook CH, Praba AC, Beery PR, Martin LC. Transthoracic

echocardiography is not cost-effective in critically ill surgical patients. J Trauma 2002;52:280–284.

15. Sawchuk CWT, Wong DT, Kavanagh BP, Siu SC. Transthoracic echocardiography does not improve prediction of outcome over APACHE II in medical-surgical intensive care. Can J Anaesth 2003;50:305–310.

16. STROBE Statement: Home. http://www.strobe-statement.org/index.php?id=strobe-home (12 December 2016)

17. Angus DC, Linde-Zwirble WT, Lidicker J, Clermont G, Carcillo J, Pinsky MR. Epidemiology of severe sepsis in the United States: analysis of incidence, outcome, and associated costs of care. Crit Care Med 2001;29:1303–1310.

18. Johnson AEW, Pollard TJ, Shen L, Lehman L-WH, Feng M, Ghassemi M, Moody B, Szolovits P, Celi LA, Mark RG. MIMIC-III, a freely accessible critical care database. Sci data 2016;3:160035.

19. Vincent JL, Moreno R, Takala J, Willatts S, Mendonça A De, Bruining H, Reinhart CK, Suter PM, Thijs LG. The SOFA (Sepsis-related Organ Failure Assessment) score to describe organ dysfunction/failure. On behalf of the Working Group on Sepsis-Related Problems of the European Society of Intensive Care Medicine. Intensive Care Med 1996;22:707–710.

20. McCaffrey DF, Griffin BA, Almirall D, Slaughter ME, Ramchand R, Burgette LF. A tutorial on propensity score estimation for multiple treatments using generalized boosted models. Stat Med 2013;32:3388–

3414.

21. Funk MJ, Westreich D, Wiesen C, Sturmer T, Brookhart MA, Davidian M. Doubly Robust Estimation of Causal Effects. Am J Epidemiol 2011;173:761–767.

22. Cole SR, Hernan MA. Constructing Inverse Probability Weights for Marginal Structural Models. Am J Epidemiol 2008;168:656–664.

23. Walsh M, Srinathan SK, McAuley DF, Mrkobrada M, Levine O, Ribic C, Molnar AO, Dattani ND, Burke A, Guyatt G, Thabane L, Walter SD, Pogue J, Devereaux PJ. The statistical significance of randomized controlled trial results is frequently fragile: a case for a Fragility Index. J Clin Epidemiol 2014;67:622–628.

24. Walraven C van, Austin PC, Jennings A, Quan H, Forster AJ. A modification of the Elixhauser comorbidity measures into a point system for hospital death using administrative data. Med Care 2009;47:626–633.

25. Gall JR Le, Loirat P, Alperovitch A, Glaser P, Granthil C, Mathieu D, Mercier P, Thomas R, Villers D. A simplified acute physiology score for ICU patients. Crit Care Med 1984;12:975–977.

26. Coye MJ, Kell J. How hospitals confront new technology. Health Aff (Millwood) 25:163–173.

27. Fisher ES, Welch HG. Avoiding the unintended consequences of growth in medical care: how might more be worse? JAMA 1999;281:446–453.

28. Wheeler AP, Bernard GR, Thompson BT, Schoenfeld D, Wiedemann HP, deBoisblanc B, Connors AF, Hite RD, Harabin AL. Pulmonary-artery

versus central venous catheter to guide treatment of acute lung injury. N Engl J Med 2006;354:2213–2224.

29. Yu DT, Platt R, Lanken PN, Black E, Sands KE, Schwartz JS, Hibberd PL, Graman PS, Kahn KL, Snydman DR, Parsonnet J, Moore R, Bates DW. Relationship of pulmonary artery catheter use to mortality and resource utilization in patients with severe sepsis. Crit Care Med 2003;31:2734–2741.

30. Sandham JD, Hull RD, Brant RF, Knox L, Pineo GF, Doig CJ, Laporta DP, Viner S, Passerini L, Devitt H, Kirby A, Jacka M. A randomized, controlled trial of the use of pulmonary-artery catheters in high-risk surgical patients. N Engl J Med 2003;348:5–14.

31. Connors AF, Speroff T, Dawson N V, Thomas C, Harrell FE, Wagner D, Desbiens N, Goldman L, Wu AW, Califf RM, Fulkerson WJ, Vidaillet H, Broste S, Bellamy P, Lynn J, Knaus WA. The effectiveness of right heart catheterization in the initial care of critically ill patients. SUPPORT Investigators. JAMA 1996;276:889–897.

32. Matthay MA, Chatterjee K. Bedside catheterization of the pulmonary artery: risks compared with benefits. Ann Intern Med 1988;109:826–834.

33. Boden WE, O'Rourke RA, Teo KK, Hartigan PM, Maron DJ, Kostuk WJ, Knudtson M, Dada M, Casperson P, Harris CL, Chaitman BR, Shaw L, Gosselin G, Nawaz S, Title LM, Gau G, Blaustein AS, Booth DC, Bates ER, Spertus JA, Berman DS, Mancini GBJ, Weintraub WS. Optimal medical therapy with or without PCI for stable coronary disease. N Engl J

Med 2007;356:1503–1516.

34. Mitka M. CT angiography: clearer picture, Fuzzier reception. JAMA 2006;295:1989–1990.

35. Ladapo JA, Horwitz JR, Weinstein MC, Gazelle GS, Cutler DM. Adoption and spread of new imaging technology: a case study. Health Aff (Millwood) 28:w1122-32.

36. Ladapo JA, Jaffer FA, Hoffmann U, Thomson CC, Bamberg F, Dec W, Cutler DM, Weinstein MC, Gazelle GS. Clinical outcomes and cost-effectiveness of coronary computed tomography angiography in the evaluation of patients with chest pain. J Am Coll Cardiol 2009;54:2409–2422.

37. Hsu DJ, Feng M, Kothari R, Zhou H, Chen KP, Celi LA. The Association Between Indwelling Arterial Catheters and Mortality in Hemodynamically Stable Patients With Respiratory Failure: A Propensity Score Analysis. Chest 2015;148:1470–1476.

38. Ghassemi M, Celi L, Stone DJ. State of the art review: the data revolution in critical care. Crit Care 2015;19:118.

Legends

Fig 1: Study cohort. Illustration of exclusion and inclusion criteria as utilized to select the final cohort of 5074 patients.

Fig 2: Relative influence factor of co-variates. The relative influence factor measures how discriminative are the 39 co-variates of the propensity score model when predicting the likelihood of echocardiogram performance.

Table 1. Comparison of the basic demographics, co-morbidity conditions and day of ICU admissions between the original cohort and the adjusted (weighted) cohort.

VARIABLES	Original Cohort			Weighted Cohort		
	TTE	NON-TTE	p	TTE	NON-TTE	p
Age	65.82 (+/-6.62)	66.69 (+/- 17.21)	0.045	65.06 (+/- 16.64)	66.85 (+/- 16.79)	0.07
Gender (Female)	47%	51%	0.014	48%	49%	0.35
Weight (Kg)	82.98 (+/-26.7)	78.56 (+/- 23.58)	<0.001	81.25 (+/- 25.26)	80.24 (+/- 24.31)	0.13
SAPS Score	20.78 (+/-5.45)	14.63 (+/-5.79)	<0.001	20.23 (+/-5.42)	20.01 (+/-5.65)	0.12
SOFA Score	6.32 (+/-3.8)	5.3 (+/-3.62)	<0.001	5.86 (+/-3.70)	5.66 (+/- 3.66)	0.03
Elixhauser Score	10.05 (+/-7.68)	5.41 (+/-6.61)	<0.001	9.44 (+/- 7.64)	9.10 (+/-7.55)	0.08
Service Unit (MICU vs SICU)	74%	68%	<0.001	71%	69%	0.132
Mechanical Ventilation Use (1st 24 Hours)	59%	47%	<0.001	54%	51%	0.007
Vasopressor Use (1st 24 Hours)	38%	27%	<0.001	34%	31%	0.01
Sedative Use (1st 24 Hours)	50%	40%	0.001	46%	43%	0.015
Co-morbid Conditions						
CHF	39%	18%	<0.001	30%	25%	<0.001
AFIB	32%	21%	<0.001	28%	25%	0.007
RENAL	16%	14%	<0.001	15%	15%	0.7
LIVER	11%	10%	0.622	11%	10%	0.4
COPD	17%	15%	0.008	17%	16%	0.3
CAD	16%	12%	<0.001	15%	13%	0.1
STROKE	11%	8%	<0.001	10%	9%	0.09
MALIGANT TUMOR	21%	25%	<0.001	23%	25%	0.25

Day of ICU Admit						
SUNDAY	14%	13%		14%	14%	
MONDAY	15%	14%		14%	14%	
TUEDAY	15%	14%		15%	14%	
WEDESDAY	14%	13%	0.001	14%	14%	0.807
THURSDAY	16%	15%		15%	15%	
FRIDAY	14%	17%		15%	16%	
SATURDAY	12%	14%		12%	14%	

Table 2. Primary outcome analysis with 5 different models: 1) Doubly robust model with unbalanced co-variates 2) Doubly robust model with all co-variates 3) Propensity Score IPW model 4) Propensity Score Matching model 5) Multivariate logistic regression model.

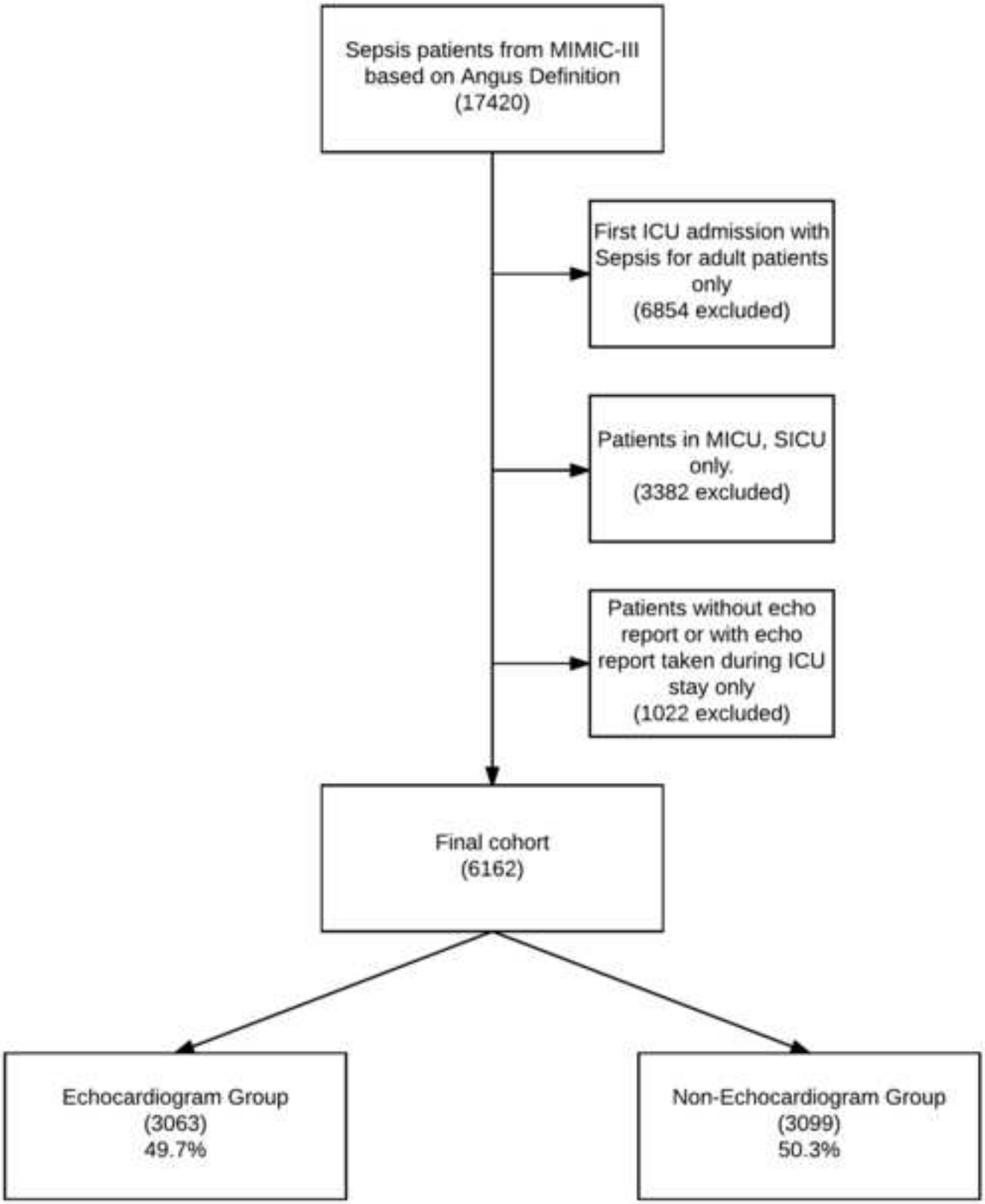
Method	OR	Confidence Interval		P-value
		2.5%	97.5%	
Doubly Robust with Unbalanced Covariates	0.78	0.67	0.89	<0.001
Doubly Robust with All Covariates	0.71	0.53	0.96	0.02
Propensity Score IPW	0.84	0.77	0.91	<0.001
Propensity Score Matching	0.80	0.68	0.94	<0.001
Multivariate	0.74	0.55	0.99	0.04

Table 3 Secondary outcome analysis.

Secondary Outcomes	Non TTE	TTE	Effect Size (95%CI)	p-value
Ventilation free days in 28 days	13.47 (+/- 14.73)	14.72 (+/- 27.21)	1.25 (+/- 1.30)	0.06
Vasopressor free days in 28 days	10.99 (+/- 13.83)	15.41 (+/- 17.26)	4.42 (+/- 1.06)	<0.001
Dobutamine Use	1%	4%	3%	<0.001
Norepinephrine (maximum dosage mg/min)	0.81 (+/- 2.39)	1.76 (+/- 5.77)	0.95 (+/- 0.22)	<0.001
IV Fluid day 1 (mL)	1937 (+/- 3182)	2527 (+/- 3891)	589 (+/-187)	<0.001
IV Fluid day 2 (mL)	835 (+/-2429)	1294 (+/- 2964)	459 (+/- 149)	<0.0001
IV Fluid day 3 (mL)	255 (+/- 2106)	687 (+/-2623)	432 (+/- 146)	<0.0001
Serum Lactate Reduction	0.84 (+/- 1.75)	1.35 (+/- 2.26)	0.51 (+/- 0.12)	<0.0001
Serum Creatinine Reduction	0.37 (+/- 0.75)	0.79 (+/- 2.78)	0.42 (+/- 0.10)	<0.0001

Figure

Click here to download Figure fig1_v2.tiff



Figure

Click here to download Figure fig2.tiff