

Improved Survival for Rural Trauma Patients Transported by Helicopter to a Verified Trauma Center: A Propensity Score Analysis

Thein Hlaing Zhu, MBBS, FACE, FRCP, Lisa Hollister, RN, MSN, Dazar Opoku, MPH, and Samuel M. Galvagno, Jr., DO, PhD, MS, FCCM

ABSTRACT

Objectives: Recent studies using advanced statistical methods to control for confounders have demonstrated an association between helicopter transport (HT) versus ground ambulance transport (GT) in terms of improved survival for adult trauma patients. The aim of this study was to apply a methodologically vigorous approach to determine if HT is associated with a survival benefit for when trauma patients are transported to a verified trauma center in a rural setting.

Methods: The ascertainment of trauma patients age ≥ 15 years ($n = 469$ cases) by HT and ($n = 580$ cases) by GT between 1999 and 2012 was restricted to the scene of injury in a rural area of 10 to 35 miles from the trauma center. The propensity score (PS) was determined using data including demographics, prehospital physiology, intubation, total prehospital time, and injury severity. The PS matching was performed with different calipers to select a higher percentage of matches of HT compared to GT patients. The outcome of interest was survival to discharge from hospital. Identical logistic regression analysis was done taking into account for each matched design to select an appropriate effect estimate and confidence interval (CI) controlling for initial vital signs in the emergency department, the need for urgent surgery, intensive care unit admission, and mechanical ventilation.

Results: Unadjusted mortalities for HT compared to GT were 7.7 and 5.3%, respectively ($p > 0.05$). The adjusted rates were 4.0% for HT and 7.6% for GT ($p < 0.05$). In a PS well-matched data set, HT was associated with a 2.69-fold increase in odds of survival compared to GT patients (adjusted odds ratio = 2.69; 95% CI = 1.21–5.97).

Conclusions: In a rural setting, we demonstrated improved survival associated with HT compared to GT for scene transportation of adult trauma patients to a verified Level II trauma center using an advanced methodologic approach, which included adjustment for transport distance. The implication of survival benefit to rural population is discussed. We recommend larger studies with multiple trauma systems need to be repeated using similar study methodology to substantiate our findings.

In the United States, trauma is the number one cause of death under age 45;¹ approximately 500 verified or designated Level I and Level II trauma centers have been created for definitive trauma care,² and 1,045 rotary-wing helicopters are currently stationed at 879 bases in hospitals or airports to

From the Department of Trauma Services and Acute Care Surgery, Parkview Adult and Pediatric Level II Trauma Center, Parkview Regional Medical Center (THZ, LH, DO), Fort Wayne, IN; and the Department of Anesthesiology and the Program in Trauma, R. Adams Cowley Shock Trauma Center, University of Maryland School of Medicine (SMG), Baltimore, MD.

Received June 23, 2017; revision received August 29, 2017; accepted August 31, 2017.

The authors have no relevant financial information or potential conflicts to disclose.

Supervising Editor: Manish N. Shah, MD, MPH.

Address for correspondence and reprints: Thein Zhu; e-mail: thein.zhu@parkview.com or Lisa Hollister; email: lisa.hollister@parkview.com

ACADEMIC EMERGENCY MEDICINE 2018;25:44–53.

This is an open access article under the terms of the Creative Commons Attribution-NonCommercial-NoDerivs License, which permits use and distribution in any medium, provided the original work is properly cited, the use is non-commercial and no modifications or adaptations are made.

transport scene patients to trauma centers.³ While capable of rapidly transporting trauma patients to a trauma center, helicopter transport (HT) is more expensive with considerable safety issues compared to ground transport (GT).⁴

Many investigators have conducted studies on the association of HT versus GT and trauma mortality. These studies have demonstrated both positive and negative findings.^{5,6} A recent formal systematic review, which included a risk of bias assessment for study methodology, showed that the overall quality of evidence in these studies was low.⁶ Due to the complex nature of HT, only studies that have employed advanced statistical techniques to control for confounders have demonstrated a significant reduction in trauma mortality. Propensity scoring methods have been recommended as a preferred advanced technique for balancing baseline covariates for effect estimates in future helicopter emergency medical services studies since the assignment of HT versus GT cannot be randomized.⁶

Previous HT studies employed advanced statistical methods for control of confounders and used large databases for case ascertainment by identifying trauma patients from the national trauma data bank⁷⁻¹⁰ and state trauma registry data.¹¹ The mortality of trauma patients in such large data sets could be affected by inclusion of rural and urban areas¹² or different volumes of hospital patient admission in multiple trauma centers.¹³ Moreover, clustering of patients within many trauma centers, if not controlled in the analysis as in one study,⁹ could affect the validity of results.¹⁴ In addition, these data sets could include patients who were injured close to trauma centers and transported by ground ambulance but were ineligible for HT.¹⁵

Some investigators have suggested that the distance traveled by the helicopter and ground ambulance to the trauma center may be a potential confounder.^{8,9} However, capturing accurate distance data can be challenging, as measured road distance from the centroids of 5-digit zip code at the scene of injury to the trauma center can be misleading because some zip codes could be large or missing.¹¹

To minimize variations in case ascertainment, distance traveled, and statistical methodology for improvement of study accuracy, we performed a retrospective observational study at a verified trauma center in a rural setting with the hypothesis that HT is associated with improved survival if admission criteria were restricted and derived from a rural area with a similar transportation radius for both HT and GT.

METHODS

Study Design

This study employed a retrospective observational design that used restrictive admission criteria¹⁶ for all trauma scene patients having similar transport distances from a localized selected rural area and transportation to a single verified trauma center. The primary objective was to evaluate association between transport modes (HT vs. GT) and survival to hospital discharge by using advanced statistical techniques for control of confounders including propensity score (PS) matching for balancing baseline covariates between HT and GT patients. The HT and GT patients were assigned as treatment and control groups, respectively. This project was approved as expedited review by the Parkview Health Institutional Review Board.

Study Setting

The verified trauma center is located in Fort Wayne, Allen County, Indiana, and was verified by the American College of Surgeons Committee on Trauma in 2000. It houses the Samaritan Flight Program that was established in 1989. The program has two rotary-wing helicopters. One is located at the trauma center and the second is located 70 miles west of the trauma center. Either one is used for transport when the other is not available. Both can land in either of two landing pads on the roof of the trauma center. Each helicopter is staffed 24 hours a day, 7 days a week with a paramedic and a registered nurse, who provide basic life support (BLS) and advanced life support (ALS) monitoring and interventions, including prehospital intubation and red blood cell transfusions. Annually, there are approximately 400 air transports to the trauma center. Of those, 60% are transported from the scene of injury. Trauma center registry staff began collecting inpatient data, including air transport information in 1991.

The selected area of injury origin is a strip of land consisting of three rural neighboring counties (DeKalb, Noble, and Whitley) from where approximately equal proportion of cases were transported by helicopter and ground ambulance to the trauma center in Allen County in Northeast Indiana. The shortest and farthest distances of each county are 10 and 35 miles from the trauma center. This is to limit the speed benefit and exclude ineligible cases of HT close to the center. The counties are situated in flat land areas with a network of asphalt and gravel county

roads and asphalt highways for transportation. Emergency medical service (EMS) agencies in these counties are staffed 24 hours a day, 7 days a week with a paramedic and emergency medical technician (EMT) or advanced EMT. The staffing model provides BLS and minor ALS monitoring and interventions.

Data were obtained from our trauma registry that was established in 1991. The data entry errors for outliers and coding on abstracted data were regularly validated started in 1999 according to an established data validation scheme.¹⁷ The trauma center was moved to a new location about 5 nautical miles to the north in March 17, 2012. We therefore employed the registry data from January 1, 1999, to March 16, 2012, as a convenience study sample in that period for this study.

Missing Information in Original Data Set

The aim of multiple imputation for missing data is to preserve given sample size, study power, and precision of subsequent point estimates.¹⁸ In addition, PS matching in SPSS software requires a complete data set with no missing observed covariates.¹⁹ First, the rates of missing information of covariates were estimated. Missing covariates for prehospital times, four prehospital, and four initial vital signs recorded in the emergency department (ED), such as, Glasgow Coma Scale (GCS), systolic blood pressure (sBP), respiratory rate (RR), and pulse rate (four physiologic variables) were replaced by statistical technique for multiple imputation together with auxiliary variables that were associated with the variables of interest and/or their pattern of missingness.¹⁸ The auxiliary variables were prehospital intubation, field time, dispatch-to-scene arrival time, scene time, transport mode, sex, age, trauma type, admission to intensive care unit (ICU), the need for urgent surgery, mechanical ventilation, and outcome. The continuous variables with significant skewed distributions, such as field time, (total) EMS time (elapsed time from 9-1-1 dispatch call to hospital arrival), and Injury Severity Score (ISS) were converted into natural log prior to imputation. The covariates that would be used in later analysis were included in the imputation model.¹⁸ A Markov Chain Monte Carlo algorithm known as fully conditional specification (FCS) was used. Continuous variables were modeled with a linear regression and categorical variables with a logistic regression to replace missing data with substituted values. Five multiply imputed data sets were generated together with the original data set in a stacked file format. Diagnostic tests were done

to assess FCS convergence for continuous variables in trace plots by iteration and imputation and relative efficiency of multiple imputation by imputation numbers and fraction of missing information for the quantity being estimated.²⁰

PS-matched Data Sets

A PS is a measure of the likelihood that an individual assigned to a treatment conditional on baseline covariates.²¹ First, an overall PS mean was determined on the five imputed data sets via binary logistic regression with transport mode as outcome, the CDC field triage criteria,^{22,23} sex, prehospital intubation, prehospital heart rate, ISS, and EMS time (transformed logarithmically) were considered as observed baseline covariates. ISS reflects injury severity of the anatomic lesions at the time of injury and was determined at hospital admission. All baseline covariates occurred in the prehospital setting.

Second, PS matching was done using transport mode as an outcome and the baseline variables as covariates used for PS determination. The ratio of GT patients to HT patients in the original data set ("Effect Measures" under Results) was smaller than a typical ratio of between 2 and 20.²⁴ We employed caliper widths of varying values from 0.00 to 0.35 times the SD of logit of PS (including calipers close to the recommended value of 0.2²⁵ and without the caliper) to generate 10 caliper-specific PS-matched data sets with 1-to-1 nearest neighbor or pair matching and matching without replacement of the GT cases. This was performed to discern the highest possible matched pairs for all HT cases in a relatively small reservoir of GT cases with absolute standardized differences < 0.1, which was taken as the cutoff point for the best balance of measured covariates between the treated and control groups.²⁶ The balance of baseline covariates between the two groups was evaluated by measurement of absolute standardized differences of mean and proportion for each covariate.²⁷

Data Analysis

We employed identical conditional logistic regression model to account for the matched pair design to determine association of outcome with HT compared to GT as done in recent studies.^{28,29} The covariates were selected a priori for known prognostic significance in the outcome after injury, which were not taken into consideration for the PS-matching procedure. The covariates included four initial ED vital signs, ICU

admission, the need for urgent surgery, and mechanical ventilation. A standard (enter) method for entering of the covariates in the regression model was chosen to determine the effect measures for odds ratio (OR) and 95% confidence interval (CI). For presentation of the effect measures and other parameters, a single PS-matched data set was selected based on standardized difference of <0.1 for each baseline covariate along with the highest attainable percentage of HT total cases being matched and the CI width for precision as the ratio of upper to lower 95% confidence limits.³⁰

From this PS-matched data set, we employed a simple method in a Microsoft Excel spreadsheet to estimate sensitivity parameter, Γ (Gamma) value for hidden bias for unknown confounders based on the McNemar test with discordant pairs for exposures, and outcomes using a 2-by-2 table.³¹ The Γ -value is the measure of the degree of departure from a study that is free of hidden bias, which could alter the inference about the treatment effects on survival. A value of 1.0 is assumed to be free of hidden bias for the treatment assignment due to unknown confounding. The number needed to treat (NNT) to save one additional life by HT was estimated from the ORs.³²

The improved survival by HT was expressed using adjusted ORs (AORs) with 95% CIs. Other studies employed the terms, mortality reduction, or mortality benefits for the same interpretation. In our study, we used improved survival or survival benefit throughout the paper except for mortality reduction calculation for comparison with other studies. The Hosmer-Lemeshow test was used to show appropriate goodness of fit for logistic regression models ($p > 0.05$). Multicollinearity was assessed using variance inflation factors (VIF) with a value > 5 considered problematic. An interaction effect³³ was assessed between transport mode and EMS time and prehospital intubation by including an effect modification term in the model. A two-tailed p -value of <0.05 was used to evaluate statistical significance. All statistical analyses were performed using SPSS software version 23.0 (IBM Corp.).

RESULTS

Background Information

Figure 1 illustrates the ascertainment of 469 HT and 580 GT patients in adults (age ≥ 15 years) for study²² after exclusion of cases. The ratio of GT to HT patients was 1.2:1. Overall unadjusted mortality was 7.7% (36/469) in HT and 5.3% (31/580) in GT

patients. The unadjusted rate in HT was 1.5 times higher than GT but the difference was not statistically significant ($p > 0.05$). Differences in transport times, hospital day utilization demographics, injury, and other variables between HT versus GT are shown in Table 1.

The missing data for prehospital vital signs were 23.6% in SBP, 17.9% in GCS score, 17.1% in RR, and 17.0% in pulse rate, and the corresponding missing data for initial ED vital signs were 0.5, 1.1, 1.0, and 0.3%, respectively. Missing data for EMS time and ISS were 3.1 and 0.3%, respectively. By controlling for both CDC triage category and other covariates, the magnitude of association in HT compared with GT patients before and after multiple imputation is shown in Data Supplement S1 (available as supporting information in the online version of this paper, which is available at <http://onlinelibrary.wiley.com/doi/10.1111/acem.13307/full>).

PS-matched Data Sets

Effect Measures. Compared with GT, HT cases had a twofold or more increase in odds of survival in each PS-matched data set including the one without the caliper. Using caliper values from 0.05 to 0.18, 70% or more of HT cases were found in matched datasets with an absolute standardized mean difference of all baseline covariates < 0.1 including PS (Table 2). The full effect measures for each covariate and caliper are shown in Data Supplement S2 (available as supporting information in the online version of this paper, which is available at <http://onlinelibrary.wiley.com/doi/10.1111/acem.13307/full>). A caliper width of 0.18 was found to provide the more conservative estimate of treatment effects with a 2.69-fold increase in odds of survival (AOR = 2.69, 95% CI = 1.21–5.97) by HT and 75.3% of the total HT cases being matched. Inclusion of only helicopters ($n = 330$) from the trauma center (without helicopters from other flight programs) at a 0.18 caliper-derived PS-matched data set provided similar results, HT patients had a 2.79-fold increase in odds of survival (AOR = 2.79, 95% CI = 1.21–6.42; data not presented). Neither the EMS time nor prehospital intubation demonstrated statistically significant evidence of effect modification with the transport mode. The absolute standardized difference for each baseline covariate between HT and GT patients was ≥ 0.182 except the sex variable prior to matching and < 0.1 after matching (Data Supplement S3, available as supporting information in the

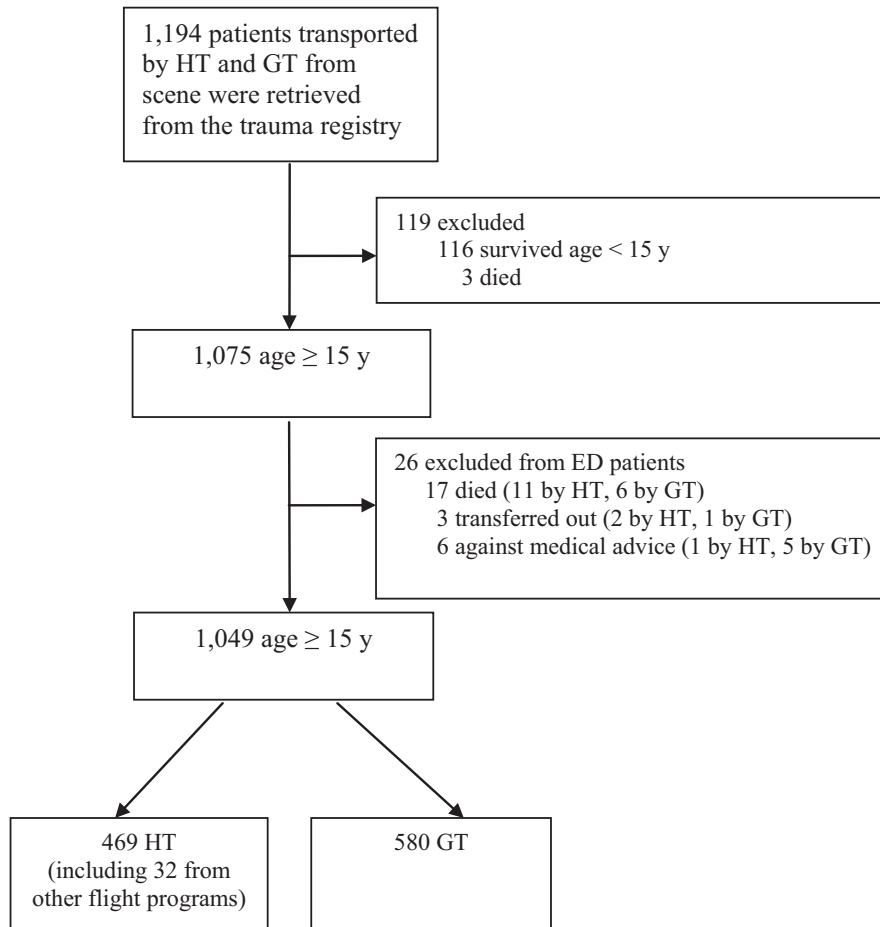


Figure 1. Study participant selection of HT and GT patients age ≤ 15 years from the trauma registry from January 1, 1999, to March 16, 2012. GT = ground ambulance transport; HT = helicopter transport.

online version of this paper, which is available at <http://onlinelibrary.wiley.com/doi/10.1111/acem.13307/full>).

Based on the 0.18 caliper PS-matched data set, the mortality rates in HT and GT patients were 4.0 and 7.6, respectively, and the difference was statistically significant ($p < 0.05$). The GT mortality was 1.93 times higher than that of HT. The Γ -value was 1.196 in the data set with a statistically significant association between transport mode and survival status in McNemar matched pairs test ($p < 0.05$). The NNT to save one additional life by HT was 22 patients (95% CI = 16 to 83 patients) in the same 0.18 caliper PS-matched data set.

DISCUSSION

Starting in 2010, investigators in the United States used large data sets and advanced statistical methodologies to better control for confounders. These studies showed a significant association when HT was compared to GT in terms of adjusted survival for

trauma patients. For instance, some studies employed the National Trauma Data Bank (NTDB) database in all ages (AOR = 1.22, 95% CI = 1.18–1.27)⁷ as well as in adults in one (AOR = 1.64, 95% CI = 1.45–1.87)⁸ and adults in another study (AOR = 1.78, 95% CI = 1.65–1.92).³⁴ One study used a state trauma registry in all ages (AOR = 1.49, 95% CI = 1.19–1.89).¹¹ Another study used NTDB in adults (AOR = 1.16, 95% CI = 1.14–1.17) at Level I trauma centers and (AOR = 1.15, 95% CI = 1.13–1.17) at Level II trauma centers.⁹ Our study showed similar improved survival (AOR = 2.69; 95% CI = 1.21–5.97) as in the aforementioned studies. However, one main difference was their study areas included multiple trauma centers with commingled urban and rural settings whereas our study area was confined to a single verified trauma center in a rural setting.

To our knowledge, this is the first study that addressed the assessment of improved survival in trauma patients transported by helicopter in a rural setting by including admission criteria in the study

Table 1
Background Information of Scene Adult Trauma Patients by Transportation Status

	HT (n = 469)*	GT (n = 580)	p-value
Transport times (min)			
Field time	73.45 ± 41.40	94.04 ± 144.12	<0.01
EMS time	59.49 ± 22.89	68.30 ± 18.67	<0.001
Dispatch-to-scene arrival time	12.96 ± 4.36	8.23 ± 4.75	<0.001
Scene time	17.48 ± 11.07	23.11 ± 10.86	<0.001
Hospital stay (days)			
Vent days	6.01 ± 6.89	5.26 ± 5.29	>0.05
ICU LOS	6.53 ± 9.07	4.39 ± 6.76	<0.01
Hospital LOS	8.21 ± 10.08	5.05 ± 6.00	<0.001
Demographics, injury, and outcome			
Age (y)	35.32 ± 16.19	41.98 ± 19.80	<0.001
Age ≥ 55 y	11.3	21.9	<0.001
Male sex	71.2	70.5	>0.05
Blunt injury	98.1	96.7	>0.05
EMS intubation	14.5	2.1	<0.001
Prehospital sBP < 90 mm Hg	12.9	5.4	<0.001
Prehospital GCS < 14	37.1	18.4	<0.001
Prehospital RR, abnormal	13.1	7.4	<0.01
Prehospital pulse rate	95.08 ± 20.62	91.25 ± 21.55	<0.05
Anatomic triage variables, yes†	25.8	18.8	<0.01
Preadmission sBP < 90 mm Hg	3.2	2.8	> 0.05
Preadmission GCS < 14	30.3	11.5	<0.001
Preadmission RR, abnormal‡	25.4	9.2	<0.001
Preadmission pulse rate	97.48 ± 22.72	91.83 ± 20.46	<0.001
ISS	17.88 ± 12.44	12.15 ± 9.31	<0.001
ISS > 15	49.5	25.6	<0.001
ISS < 9	21.6	33.9	<0.001
ICU admission	37.5	24.8	<0.001
Urgent surgery	23.5	20.5	>0.05
Mechanical ventilation	14.3	6.0	<0.001
Discharge to home	62.0	71.7	<0.001

Data are reported as mean ± SD or percent. Field time is the elapsed time from injury to ED arrival. EMS time is the elapsed time from 9-1-1 call to ED arrival.

GCS = Glasgow Coma Scale; GT = ground ambulance transport; HT = helicopter transport; ICU = intensive care unit; ISS = Injury Severity Score; LOS = length of stay; RR = respiratory rate; sBP = systolic blood pressure.

*Sample sizes are shown in parentheses.

†Yes means presence of one or more of eight anatomic variables in a patient. These anatomic criteria included flail chest, long-bone fractures, crushed extremity, amputation proximal to wrist and ankle, pelvis fractures, open or depressed skull fractures, and paralysis.

‡Abnormal RR = RR < 10 or > 29 breaths/min.

design for study accuracy,¹⁶ to similar travel distance by transport vehicles 10–35 miles to the trauma center,³⁵ using data from a rural area and only selecting admissions to a single verified trauma center. These factors prevented or reduced confounding effects. In addition, we used a PS analysis to balance the baseline covariates for eliminating or minimizing selection bias in treating HT or GT trauma patients and the use of the analysis method is a substantial strength of this work.⁶

Our study could have an implication on survival benefit in rural populations in the United States.

According to a study, there were 190 Level I and 255 Level II trauma centers located throughout the United States in 2005 and about 27% of the total U.S. residents (urban and rural) could access a Level I or II trauma center from the scene of injury within 45 and 60 minutes, respectively, by helicopter.³⁶ In 2005, the average U.S. population was 295 million and about 20% of them resided in rural areas.³⁷ Twenty-seven percent of the 2005 U.S. population was approximately equal to 80 million (295 million × 0.27) people. Twenty percent of the 80 million was 16 million (80 million × 0.20) population in rural areas;

Table 2

Conditional Logistic Regression Analysis on PS-matched Samples by Caliper for Survival to Hospital Discharge, Adjusted for Initial ED Vital Signs and Hospital Resource Utilization Variables

Caliper Width Value	Matched Pairs	Percentage Matched		AOR	95% CI	CI (UL/LL)
		GT, <i>n</i> = 580	HT, <i>n</i> = 469			
0.05	329	56.7	70.1	2.54	1.06–6.06	5.70
0.10	338	58.3	72.1	3.45	1.43–8.24	5.78
0.15	349	60.2	74.4	2.98	1.35–6.61	4.91
0.18	353	60.9	75.3	2.69	1.21–5.97	4.93

AOR = adjusted odds ratio; GT = ground ambulance transport; HT = helicopter transport; PS = propensity score; UL/LL = upper-to-lower confidence limit ratio.

therefore, about 16 million people (a sizable portion) may get the survival benefit if they were severely injured at the scene and transported to a verified trauma center (Level I or II) for treatment by helicopter.

Before covariate adjustment, the insignificant differences in mortality rates of HT versus GT (7.7% vs. 5.3%) in the original data set could be due to the baseline covariates, which may have confounded the true association between HT and mortality.³⁸ Such observations have been made in previous work and are likely due to Simpson's paradox: A phenomenon where a trend using unadjusted outcomes may be reversed when groups are combined with appropriate statistical adjustment.⁹ After the baseline covariates were balanced for the PS-matched data set, the adjusted mortality rates changed direction and the rate in HT was significantly lower than GT (4.0% vs. 7.6%). Multicollinearity among covariates in logistic regression analysis in original data set was not evident as each VIF value was <2. Although the missing data for prehospital sBP was relatively high (23.6%), the FCS convergence plots for continuous variables (e.g., EMS time, prehospital pulse rate) showed no discernable patterns indicating the validity of the results from multiple imputation and the imputed data being considered random samples from the posterior distribution of the missing data.³⁹ The relative efficiency value with five imputations was 0.9345 or over 93%.

The calipers affected matched sample sizes and balance of baseline covariates in PS matching. The OR estimates varied in PS-matched data sets with standardized differences < 0.1 (Table 2). The inclusion of EMS time and prehospital pulse rate as continuous baseline covariates may have improved the PS model as Austin in 2011²⁵ indicated that continuous variables in the model could minimize the mean squared error of the resultant estimated treatment effect, denoting that there was less difference between the true

estimator and what is estimated. It is believed that prospective randomized trials are neither practical nor ethical when it comes to impact of HT on trauma mortality.^{40,41} However, analysis using a PS-matched data set mimics a randomized controlled trial by balancing only known covariates.²⁷

The gamma-value of 1.196 denotes highly sensitive to hidden bias for unknown confounders. A mild change in its value to 1.197 could affect the conclusion for association between HT and mortality benefit by producing significance level > 0.05.⁴² We estimated the gamma value by using the Rosenbaum's approach which is easy to implement and gives a conservative result, but the Rosenbaum's approach has some deficiencies.⁴³ For instance, the magnitude of sensitivity analysis value (gamma-value) is a function of sample size. In our study, we had a low number of discordant pairs (*n* = 33) in a small sample of matched pairs (*n* = 353). The gamma-value converted to over 2.0 if we increased the existing matched pairs to 50 times with similar equal increase in each cell of the 2-by-2 table. This value can be compared to a similar study in children with a large data set (*n* = 25,700 matched pairs) where the Γ -value was found to be 2.6, which suggested a fair robustness to potential hidden bias.²⁸ The Rosenbaum's approach assumes a situation in which the unobserved confounder perfectly predicts the outcome of interest. As a result, this method may overstate our study sensitivity. Most of Excel's statistical functions may be slightly imprecise³¹ and may affect the estimated value. However, in this study, the gamma-value was estimable by getting the number of discordant pairs without missing observations in the 2-by-2 table for each pair of subjects and the Rosenbaum's approach was implementable by getting a significant McNemar statistical test (T. Love, personal communication, May 26, 2016). A study could become insensitive if the extreme values (gamma > 5) are required to alter the inference. That is, gamma

values < 5 could explain the association between the treatment and the outcome.³¹ The gamma thresholds could be taken with gamma values < 6 .⁴²

In spite of HT patients being sicker than GT patients because of longer hospital stay and lesser proportion of discharge to home (Table 1) in the former than in the latter patients, our study showed HT could save more lives (one additional life for every 22 transports) than a similar study (one additional life for every 65 transports) in adults.⁹

The mechanism underlying the mortality improvement is considered multifactorial by many studies.^{10–12} Due to our study design and analysis, the mechanism could be narrowed down to fewer factors. The potential effects of transport speed (transport mode), urban location of injury (geographic situation), and variability and quality of patient care at multiple trauma centers (receiving facilities) on mortality was reduced or eliminated. The underlying mechanism may be due to differences in role and structure of EMS agencies in relation to crew capabilities and career experience between HT and GT in prehospital patient care in severely injured patients at the trauma center. For instance, the higher prehospital endotracheal intubation (ETI) rate in HT than GT (14.5% vs. 2.1%; Table 1) after actual data analysis may indicate that HT crews were more skillful and likely to perform ETI more successfully than their GT counterparts.⁴⁴ The higher mortality rate in GT (55.6%) than HT (25.8%) in intubated trauma patients with prehospital GCS scores of 3 to 8 observed in our data (data not presented) might explain improved survival by HT. Moreover, the significant reduction of prehospital hypotension rate from 12.9% to 3.2% at ED (Table 1) may denote that HT personnel had higher level of ALS by possibly providing fluid resuscitation and/or packed red blood cell transfusion that could reduce mortality in patients at the scene.⁴⁵ The combination of a certified flight nurse and an experienced paramedic in HT as seen in most HT flight programs in the United States,^{5,6} compared to paramedic and EMT in GT in our trauma system could probably have a higher effectiveness in assessing inpatient acuity and evaluating clinical conditions during prehospital care at the scene.

LIMITATIONS

This observational study may have inherent defects, such as the possibility of incorrect data entry, missing

data, and incomplete covariates. We addressed the potential deficiencies by performing validation, multiple imputation, PS matching, and estimation of hidden bias for unknown confounders. However, the findings in this study should be interpreted in light of some limitations. Our trauma registry data may have measurement errors since we limited data validation to internal methods and not external validation with those in the medical records in the study period.¹⁷ One main defect of multiple imputation is that multiple imputation will not produce valid results if the missing data are not missing at random.¹⁸ However, the FCS convergence plots show our imputed values were random samples. Our study could have unobserved confounding which could not be adjusted by PS matching. As discussed in detail above, the estimated gamma value that was highly sensitive to hidden bias for unobserved confounding may be due to a small sample study size, type of sensitivity analysis method used, or other factors. When compared to large data set studies, the 95% CIs in our PS-matched data sets were wide (Table 2) due to small sample sizes. Moreover, our study from a single trauma system in rural northeast Indiana may not be generalizable to other geographic locations. For these reasons, larger studies with multiple trauma systems need to be repeated using similar study methodology including a robust sensitivity analysis method for unobserved confounding⁴³ to substantiate our findings.

Investigators often discuss whether HT in mortality benefit was cost-effective on issues of overtriage of selecting patients with minor injury, higher costs, and/or transport risk in HT versus GT.^{8–10,28} In 2013, Delgado and colleagues⁴ addressed these issues on their cost-effectiveness model based on relative risk reduction in mortality, the overtriage, transport cost and risk, and other factors. They expressed cost-effectiveness as incremental cost-effective ratio or cost per quality-adjusted life-year (QALY) gained. They found that reduction of mortality by a minimum of 15% or at least 30% by HT compared to GT would provide a cost of $< \$100,000$ or $< \$50,000/\text{QALY}$ gained, respectively, which were accepted thresholds for cost-effectiveness. The mortality reduction in our study was 17% (1.00–0.83; AOR = 0.83 as reciprocal of lower 95% CI, AOR = 1.21) or 63% (1.00–0.37; AOR = 0.37 as reciprocal of point AOR = 2.69). These mortality reduction figures met their cost-effectiveness threshold values.

CONCLUSIONS

This study utilized a statistically rigorous and methodologic approach including a propensity score matching to demonstrate a survival benefit for adult trauma scene patients by using helicopter transport compared to ground transport at a verified Level II trauma center in a rural setting. A sizable U.S. rural population could get the survival benefit if they were severely injured at the scene and transported to a Level II (or Level I) trauma center for treatment by helicopter. We introduced a new element in the study design components, controlled for distance traveled for both helicopter and ground ambulance by limiting the catchment area for both helicopter transport and ground transport to an appropriate geographic area. Helicopter transport could save one additional life for every 22 trauma scene patients. We suggest a multi-center study using the similar methodologic approach to substantiate these findings.

We acknowledge Felix Thoemmes, PhD, from Cornell University, New York, NY, for assisting with the propensity score analysis. We are grateful to Ohnmar, MBBS, PhD, Lithan University College, Yangon, Myanmar, for her thoughtful suggestions on the initial manuscript.

References

- Haagsma JA, Graetz N, Bolliger I, et al. The global burden of injury: incidence, mortality, disability-adjusted life years and time trends from the Global Burden of Disease study. *Inj Prev* 2013;22:3–18.
- Wang HE, Yealy DM. Distribution of specialized care centers in the United States. *Ann Emerg Med* 2012;60:632–7.
- Association of Air Medical Services (AAMS). Atlas & Database of Air Medical Services (ADAMS). Available at: <http://aams.org/member-services/atlas-database-air-medical-services-adams/>. Accessed Jan 11, 2017.
- Delgado MK, Staudenmayer KL, Wang NE, et al. Cost-effectiveness of helicopter versus ground emergency medical services for trauma scene transport in the United States. *Ann Emerg Med* 2013;62:351–64.
- Thomas SH. Helicopter EMS: Outcomes Research, Cost-effectiveness, and Triage. Available at: <http://www.cctcore.org/wp-content/uploads/2016/07/HEMS-outcomes-12-July-16.pdf>. Accessed Oct 3, 2016.
- Galvagno SM Jr, Sikorski R, Hirshon JM, et al. Helicopter emergency medical services for adults with major trauma. *Cochrane Database Syst Rev* 2015;(12):CD009228.
- Brown JB, Stassen NA, Bankey PE, Sangosanya AT, Cheng JD, Gestring ML. Helicopters and the civilian trauma system: national utilization patterns demonstrate improved outcomes after traumatic injury. *J Trauma* 2010;69:1030–36.
- Sullivent EE, Faul M, Wald MM. Reduced mortality in injured adults transported by helicopters emergency medical services. *Prehosp Emerg Care* 2011;15:295–302.
- Galvagno SM, Haut ER, Zafar SN, et al. Association between helicopter vs. ground emergency medical services and survival for adults with major trauma. *JAMA* 2012;307:1602–10.
- Brown JB, Forsythe RM, Stassen NA, Gestring ML. The national triage protocol: can this tool predict which patients with trauma will benefit from helicopter transport? *J Trauma Acute Care Surg* 2012;73:319–25.
- Stewart KE, Cowan LD, Thompson DM, Sacra JC, Albrecht R. Association of direct helicopter versus ground transport and in-hospital mortality in trauma patients: a propensity score analysis. *Acad Emerg Med* 2011;18:1208–16.
- Brathwaite CE, Rosko M, McDowell R, Gallagher J, Proenca J, Spott MA. A critical analysis of on-scene helicopter transport on survival in a statewide trauma system. *J Trauma* 1998;45:140–6.
- Caputo LM, Salottolo KM, Slone DS, Mains CW, Bar-Or D. The relationship between patient volume and mortality in American trauma centers: a systematic review of the evidence. *Injury* 2014;45:478–86.
- Roudsari B. Clustering and missing data in the US National Trauma Data Bank: implication for analysis. *Injury Prev* 2008;14:96–100.
- Delgado MK, Newgard CD, Hsia RY. Helicopter vs. ground transportation for patients with trauma. *JAMA* 2012;308:563.
- Rothman KJ, Greenland S, Lash TL. *Modern Epidemiology*. 3rd ed. Philadelphia, PA: Lippincott Williams & Wilkins, 2008:168–82.
- Hlaing T, Hollister L, Aaland M. Trauma registry data validation: essential for quality trauma care. *J Trauma* 2006;61:1400–7.
- Newgard CD. The validity of using multiple imputation for missing out-of-hospital data in a state trauma registry. *Acad Emerg Med* 2006;13:314–24.
- Thoemmes F. Propensity score matching in SPSS. arXiv preprint arXiv:1201.6385. 2012. Available at: https://www.human.cornell.edu/hd/qml/upload/Thoemmes_2012.pdf. Accessed Oct 3, 2016.
- Yuan YC. *Multiple Imputation for Missing Data: Concepts and New Development (Version 9.0)*. 2009. Available at: <https://support.sas.com/rnd/app/stat/papers/multipleimputation.pdf>. Accessed Oct 10, 2016.
- D'Agostino RB Jr, D'Agostino RB Sr. Estimating treatment effects using observational data. *JAMA* 2007;297:314–6.

22. American College of Surgeons Committee on Trauma. Resources for Optimal Care of the Injured Patient. Chicago, IL: American College of Surgery, 2014. pp. 23–29, 65–75, 94–99.
23. Sasser SM, Hunt RC, Faul M, et al. Guidelines for field triage of injured patients: recommendations of the National Expert Panel on Field Triage, 2011. *MMWR Recomm Rep* 2012;61:1–20.
24. Rubin DB, Thomas N. Matching using estimated propensity scores: relating theory to practice. *Biometrics* 1996;52:249–64.
25. Austin PC. Optimal caliper widths for propensity-score matching when estimating differences in means and differences in proportions in observational studies. *Pharm Stat* 2011;10:150–61.
26. Harder VS, Stuart EA, Anthony JC. Propensity score techniques and the assessment of measured covariate balance to test causal associations in psychological research. *Psychol Methods* 2010;15:234–49.
27. Austin PC. An introduction to propensity score methods for reducing the effects of confounding in observational studies. *Multivariate Behav Res* 2011;46:399–424.
28. Brown JB, Leeper CM, Sperry JL, et al. Helicopters and injured kids: improved survival with scene air medical transport in the pediatric population. *J Trauma Acute Care Surg* 2016;80:702–10.
29. Abe T, Takahashi O, Saitoh D, Tokuda Y. Association between helicopter with physician versus ground emergency medical services and survival of adults with major trauma in Japan. *Crit Care* 2014;18:R146.
30. Poole C. Low p-value or narrow confidence intervals: which are more durable? *Epidemiology* 2001;12:291–4.
31. Love TE. Spreadsheet-based sensitivity analysis calculations for matched samples. Center for Health Care Research & Policy, Case Western Reserve University, 2008. Available at: <http://www.chrp.org/propensity>. Accessed Jul 19, 2016.
32. Lindenauer PK, Pekow P, Wang K, Gutierrez B, Benjamin EM. Lipid-lowering therapy and in-hospital mortality following major noncardiac surgery. *JAMA* 2004;291:2092–9.
33. Vatcheva KP, Lee M, McCormick JB, Rahbar MH. The effect of ignoring statistical interactions in regression analyses conducted in epidemiologic studies: an example with survival analysis using Cox proportional hazards regression model. *Epidemiology (Sunnyvale)* 2016;6. pii: 216.
34. Ryb GE, Dischinger P, Cooper C, Kufera JA. Does helicopter transport improve outcomes independently of emergency medical system time? *J Trauma Acute Care Surg* 2013;74:149–156.
35. Diaz MA, Hendey GW, Bivins HG. When is the helicopter faster? A comparison of helicopter and ground ambulance times. *J Trauma* 2005;58:148–153.
36. Branas CC, MacKenzie EJ, Williams JC, et al. Access to trauma centers in the United States. *JAMA* 2005;293:2626–33.
37. United States Census Bureau. Geography: 2010 Census Urban and Rural Classification and Urban Area Criteria. Available at: <https://www.census.gov/geo/reference/ua/urban-rural-2010.html>. Accessed Aug 28, 2017.
38. Beaglehole R, Bonita R, Kjellstrom T. *Basic Epidemiology*. Geneva, Switzerland: World Health Organization, 1993:48–51.
39. Dong Y, Peng CY. *Principled missing data methods for researchers*. Springerplus 2013;2:222.
40. Mitchell AD, Tallon JM, Sealy B. Air versus ground transport of major trauma patients to a tertiary trauma center: a province-wide comparison using TRISS analysis. *Can J Surg* 2007;50:129–33.
41. Rhinehart ZJ, Guyette FX, Sperry JL, et al. The association between air ambulance distribution and trauma mortality. *Ann Surg* 2013;257:1147–53.
42. Rosenbaum PR. Sensitivity to Hidden Bias. in: Rosenbaum PR, editor. *Observational Studies*. New York, NY: Springer, 2002:105–70.
43. Liu W, Juramoto SJ, Stuart EA. An introduction to sensitivity analysis for unobserved confounding in non-experimental prevention research. *Prev Sci* 2013;14:570–80.
44. Davis DP, Peay J, Serrano JA, et al. The impact of aeromedical response to patients with moderate to severe traumatic brain injury. *Ann Emerg Med* 2005;46:115–22.
45. Brown JB, Sperry JL, Fombona A, Billiar TR, Peitzman AB, Guyette FX. Pre-trauma center blood cell transfusion, transfusion is associated with improved early outcomes in air medical trauma patients. *J Am Coll Surg* 2015;220:797–808.

Supporting Information

The following supporting information is available in the online version of this paper available at <http://onlinelibrary.wiley.com/doi/10.1111/acem.13307/full>

Data Supplement S1. Logistic regression analysis for survival to hospital discharge by type of dataset, adjusted for CDC triage and other variables.

Data Supplement S2. eTables showing effect measures for each covariate by conditional logistic regression analysis on propensity score matched samples using calipers from 0.05 to 0.18 in trauma scene patients transported by all helicopters and trauma center helicopters.

Data Supplement S3. Comparing baseline covariates in scene adult trauma patients by transportation status and type of dataset.