



Published in final edited form as:

JAMA. 2012 April 18; 307(15): 1602–1610. doi:10.1001/jama.2012.467.

## Association Between Helicopter vs Ground Emergency Medical Services and Survival for Adults With Major Trauma

Dr. Samuel M. Galvagno Jr, DO, PhD, Dr. Elliott R. Haut, MD, FACS, Dr. S. Nabeel Zafar, MBBS, MPH, Dr. Michael G. Millin, MD, MPH, Dr. David T. Efron, MD, FACS, Dr. George J. Koenig Jr, DO, MS, Ms. Susan P. Baker, MPH, Dr. Stephen M. Bowman, PhD, Dr. Peter J. Pronovost, MD, PhD, FCCM, and Dr. Adil H. Haider, MD, MPH, FACS

Division of Trauma Anesthesiology, Shock Trauma Center, Program in Trauma, Department of Anesthesiology, University of Maryland School of Medicine, University of Maryland Medical Center, Baltimore (Dr Galvagno); Departments of Surgery (Drs Haut, Efron, and Haider), Anesthesiology and Critical Care Medicine (Drs Haut and Pronovost), Emergency Medicine (Drs Haut and Millin), Nursing (Dr Pronovost), Division of Acute Care Surgery, Trauma, Emergency Surgery, and Critical Care (Drs Haut, Millin, and Haider), Center for Innovation in Quality Patient Care, Safety Research Group (Dr Pronovost), and Center for Surgery Trials and Outcomes Research (Dr Haider), Johns Hopkins University School of Medicine, and Graduate Training Program in Clinical Investigation (Dr Haut), Johns Hopkins Bloomberg School of Public Health (Ms Baker and Drs Bowman, Pronovost, and Haider), Baltimore, Maryland; Department of Surgery, Aga Khan University, Karachi, Pakistan (Dr Zafar); and Department of Surgery, Division of Acute Care Surgery, Thomas Jefferson University School of Medicine, Philadelphia, Pennsylvania (Dr Koenig)

### Abstract

**Context**—Helicopter emergency medical services and their possible effect on outcomes for traumatically injured patients remain a subject of debate. Because helicopter services are a limited and expensive resource, a methodologically rigorous investigation of its effectiveness compared with ground emergency medical services is warranted.

**Objective**—To assess the association between the use of helicopter vs ground services and survival among adults with serious traumatic injuries.

---

© 2012 American Medical Association. All rights reserved.

Corresponding Author: Samuel M. Galvagno Jr, DO, PhD, Division of Trauma Anesthesiology, Shock Trauma Center, Program in Trauma, Department of Anesthesiology, University of Maryland School of Medicine, University of Maryland Medical Center, S8B05, 22 South Greene St, Baltimore, MD 21201 (sgalvagno@anes.umm.edu).

**Conflict of Interest Disclosures:** The authors have completed and submitted the ICMJE Form for Disclosure of Potential Conflicts of Interest. Dr Haut reported that he receives royalties from Lippincott Williams & Wilkins for *Avoiding Common ICU Errors*; and has provided expert testimony in various medical malpractice cases. No other disclosures were reported.

**Online-Only Material:** The eFigure, eTables 1–3, and the Author Video Interview are available at <http://www.jama.com>.

**Author Contributions:** Dr Galvagno had full access to all of the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis.

*Study concept and design:* Galvagno, Haut, Millin, Efron, Koenig, Pronovost, Haider.

*Acquisition of data:* Galvagno, Haut, Zafar, Haider.

*Analysis and interpretation of data:* Galvagno, Zafar, Baker, Bowman, Haider.

*Drafting of the manuscript:* Galvagno, Haut, Millin, Efron, Baker, Bowman, Pronovost, Haider.

*Critical revision of the manuscript for important intellectual content:* Galvagno, Koenig, Baker, Bowman, Pronovost, Haider.

*Statistical expertise:* Galvagno, Zafar, Bowman, Haider.

Administrative, technical, or material support: Galvagno, Haut, Efron, Pronovost, Haider.

*Study supervision:* Galvagno, Haut, Haider.

**Design, Setting, and Participants**—Retrospective cohort study involving 223 475 patients older than 15 years, having an injury severity score higher than 15, and sustaining blunt or penetrating trauma that required transport to US level I or II trauma centers and whose data were recorded in the 2007–2009 versions of the American College of Surgeons National Trauma Data Bank.

**Interventions**—Transport by helicopter or ground emergency services to level I or level II trauma centers.

**Main Outcome Measures**—Survival to hospital discharge and discharge disposition.

**Results**—A total of 61 909 patients were transported by helicopter and 161 566 patients were transported by ground. Overall, 7813 patients (12.6%) transported by helicopter died compared with 17 775 patients (11%) transported by ground services. Before propensity score matching, patients transported by helicopter to level I and level II trauma centers had higher Injury Severity Scores. In the propensity score–matched multivariable regression model, for patients transported to level I trauma centers, helicopter transport was associated with an improved odds of survival compared with ground transport (odds ratio [OR], 1.16; 95% CI, 1.14–1.17;  $P < .001$ ; absolute risk reduction [ARR], 1.5%). For patients transported to level II trauma centers, helicopter transport was associated with an improved odds of survival (OR, 1.15; 95% CI, 1.13–1.17;  $P < .001$ ; ARR, 1.4%). A greater proportion (18.2%) of those transported to level I trauma centers by helicopter were discharged to rehabilitation compared with 12.7% transported by ground services ( $P < .001$ ), and 9.3% transported by helicopter were discharged to intermediate facilities compared with 6.5% by ground services ( $P < .001$ ). Fewer patients transported by helicopter left level II trauma centers against medical advice (0.5% vs 1.0%,  $P < .001$ ).

**Conclusion**—Among patients with major trauma admitted to level I or level II trauma centers, transport by helicopter compared with ground services was associated with improved survival to hospital discharge after controlling for multiple known confounders.

Trauma remains the leading cause of death and disability among young people around the world. In the United States, more than 50 million people are injured per year, resulting in approximately 169 000 annual deaths and a lifetime cost of \$406 billion.<sup>1,2</sup>

Over the past several years, significant improvements in survival after trauma have been achieved. One reason for this has been improvements in emergency medical services (EMS) and life-saving transport of trauma patients to a center capable of providing definitive care. The utility of helicopter EMS and its possible effect on outcomes for traumatically injured patients remains the subject of debate.<sup>3–8</sup> Because helicopter transport is a limited and expensive resource, a methodologically rigorous investigation of its effectiveness compared with ground EMS is warranted.

Several studies have used the National Trauma Data Bank (NTDB) to assess outcomes for traumatically injured adults transported by either option.<sup>9–11</sup> Although each of these studies concluded that helicopter transport was associated with improved odds of survival, they also reported limitations to their conclusions because of the lack of testing for the assumptions of a regression model, and none was able to account for the differences in how patients were assigned treatments (helicopter or ground transport). Additionally, the high proportion of missing data in the NTDB may have introduced bias.<sup>12</sup>

The purpose of this study was to compare the association between the 2 transport modes and survival among adults with traumatic injuries by performing a robust analysis and controlling for known confounders.

## METHODS

The NTDB is the largest repository of trauma data in the world, with data collected from more than 900 centers in the United States.<sup>13</sup> Since 2007, the quality of data in the NTDB has markedly improved following adoption of the National Trauma Data Standard. After receiving approval from the Johns Hopkins School of Medicine's institutional review board, a merged data set, using data from the 2007–2009 NTDB, was created with all available variables. The data sets were merged using source codes provided by the NTDB.<sup>12</sup>

All hospitalized patients with an *International Classification of Diseases, Ninth Revision, Clinical Modification (ICD-9-CM)* code of 800–959 were eligible for inclusion. Adults older than 15 years and admitted to a level I or II trauma center were included. The analysis was restricted to records with complete information regarding transport and disposition information. Other forms of transportation to trauma centers, such as private conveyance, police, and walk-ins, were excluded. The Injury Severity Score (ISS) was used to quantify the severity of trauma. An ISS higher than 15 was used for inclusion because this has been shown to be associated with a greater need for specialized traumacare.<sup>14, 15</sup> Injury Severity Score and sex were also found to be statistically significant covariates in a previous study comparing urban and rural helicopter transport of patients with blunt trauma injuries.<sup>16</sup>

The primary intervention was transport by either helicopter or ground EMS. The outcome of interest was survival to discharge from the hospital. This outcome was evaluated with 3 analytical models: a multivariate logistic regression, a multivariate logistic regression model with generalized estimating equations and robust variance calculations to control for clustering by trauma center, and a logistic regression model incorporating the results of propensity score matching.

Covariates included demographic, physiologic, and hospital data. Demographics included information on age, sex, and race. Covariates were carefully selected based on the assumption that none was affected directly by the intervention. Other a priori selected variables were planned for inclusion in the final statistical models. These variables included the type of trauma (blunt vs penetrating), initial recorded vital signs (systolic blood pressure, respiratory rate, heart rate), Glasgow Coma Scale (GCS; motor component) score, and the locally calculated ISS or ISS calculated from *ICD-9-CM* admission codes. Illicit drug use, alcohol use, and comorbidities were also tabulated but were found to have a prevalence of missing data greater than 40% (eFigure available at <http://www.jama.com>) so were not included. Patients who were dead on arrival to the emergency department (ED) were excluded.

Variables were considered for inclusion in the final models after calculating correlation coefficients, examining scatterplot matrices, and ensuring that the proportion of missing data was below 20%. Final models included the following independent variables: systolic blood pressure, respiratory rate, heart rate, motor component of GCS, mechanism of injury derived from *ICD-9-CM* e-codes for primary and secondary diagnoses, age, sex, type of trauma (blunt vs penetrating), and transport mode. Total GCS scores were excluded because of significant colinearity with the motor component and because more than 30% of data for the verbal component was missing. Moreover, a previous study has shown that the motor component is equally predictive of the main outcome measure of survival in the NTDB.<sup>17</sup> All vital signs consisted of those first recorded in the ED because of a high proportion of missing data for prehospital vital signs. A sensitivity analysis was conducted to ensure that initial ED vital signs had a high correlation with pre-hospital vital signs (eTable 1). Total elapsed EMS times from dispatch to ED arrival were excluded as a variable because of a

57.8% prevalence of missing data (eFigure). A sensitivity analysis was performed for all complete cases to examine the role of total EMS times as an independent variable.

In the absence of a posited assignment mechanism for how patients are assigned to helicopter or ground transport, formal causal inference is not credible.<sup>18</sup> The goal of causal inference is to assess the average effect of a treatment on a subsequently measured outcome. The goal of an observational study should be to create an analysis that resembles conditions that would otherwise be achieved under a randomized design.<sup>19, 20</sup> Systematic differences between treatment groups should be balanced at the beginning of observational studies to control for bias.

Propensity score methods, as first described by Rubin,<sup>21</sup> are 1 way of creating subgroups of treated units (helicopter) and control units (ground) that are similar with respect to distributions of observed background characteristics and potential confounders.<sup>22</sup> Propensity scores reflect the likelihood of a study participant being assigned to a particular treatment group, conditional on multiple variables thought to influence such an assignment.<sup>21</sup> Once a propensity score is calculated, the score can be used in a multivariable model as an independent variable or the score can be used to match participants with different treatment assignments, thereby creating a matched patient cohort, reducing the risk of confounding by indication.

Propensity score–based analyses were conducted with subclassification matching. Subclassification matching was selected after assessing standardized mean differences compared with nearest neighbor and full matching protocols. Multiple imputation, using techniques previously established for use with the NTDB,<sup>23</sup> was performed for each variable associated with missing values. All variables, including the missing data indicators, were included in a propensity score model. Multiple imputation was performed using the multiple imputation (*mi*) suite of commands available in Stata version 11 (Stata Corp). Imputation was used for the following missing data: systolic blood pressure, heart rate, motor component of GCS score, and ISS. First, the proportion of missing data for variables of interest was calculated. The *mi* set of commands was used to generate a regression model to impute missing data based on other available variables. This process was repeated 5 times, creating 5 separate imputed data sets. These 5 data sets were combined using commands from the *mi* suite to create a full data set with no missing values. Multiple imputation was used only for variables with less than 20% missing data. The proportion of missing data for variables used in all regression analyses is described in eTable 2.

The following independent variables were used to calculate the propensity score: age, sex, ISS, systolic blood pressure, respiratory rate, heart rate, type of trauma, GCS motor component score, mechanism of injury (*ICD-9-CM* e-codes), and trauma facility. Propensity score matching was performed after the propensity scores were estimated. For the subclassification method, 5 subclasses were used because this has been shown to remove at least 90% of bias in the estimated treatment effect due to the covariates used to estimate a propensity score.<sup>24</sup> Balance among the covariates after propensity score matching was assessed with numerical diagnostics, jitter plots, histograms, quantile-quantile plots, and standardized bias plots. Standardized mean differences were reduced by 91.3% for the level I and 91.8% for the level II groups, following matching by propensity scores. The jitter plot for the samples demonstrated few outliers, and all available data were used for both groups. Overall, the matching quality achieved was excellent for level I and level II trauma center data sets, presumably due to the large number of patients in the control group. A logistic regression analysis was performed on the matched samples, and effects were estimated within subclasses and then combined for a final effect estimate with an associated standard error.

The data mergers and construction of the imputed data sets were conducted using Stata version 11. Stata was also used to calculate odds ratios (ORs) with 95% confidence intervals for the logistic regression with generalized estimating equations analyses. Propensity scores and propensity score matching, as well as all propensity score–based logistic regression analyses, were performed using the *Match It* and *opt-match* packages available in the 64-bit version of R 2.12.1 (R Foundation for Statistical Computing). With 90% power and an  $\alpha$  level of .05, a total of 8802 patients were required to detect a 2% mortality difference between groups. *P* values of less than .05 were considered statistically significant, and all tests were 2-sided. Absolute risk reduction (ARR) calculations were made after calculating the number needed to treat, based on the adjusted ORs as previously described by Lindenauer et al.<sup>25</sup>

Regression diagnostics were performed for all multivariate logistic regression models. Leverage was assessed with hat matrices, and influence assessed with changes in Pearson residuals when covariates were fitted to the model. The Hosmer-Lemeshow and Pearson goodness-of-fit tests were used to confirm that the models adequately fit the data ( $P > .10$ ). Colinearity among the independent variables was assessed by calculating variance inflation factors.

## RESULTS

The Figure depicts the study profile. A total of 1 816 982 records were initially available in the 2007, 2008, and 2009 NTDB data sets. After stratification and confirmation of data availability for disposition, transport mode, type of trauma, and injury severity and after excluding patients who died before reaching the ED (1897 by ground; 324 by helicopter), there were 159 511 patients transported to level I and 63 964 patients transported to level II trauma centers available for analysis. Of the NTDB covariates initially considered for inclusion in the models, 38% had more than 40% missing values (eFigure). The final study population included 61 909 patients transported by helicopter and 161 566 by ground.

Patient demographics and characteristics are summarized in Table 1 and Table 2. The mean age was similar between level I and level II trauma centers. There were more men in the level I group (70.1%) than in the level II group (56.3%). Unadjusted mortality was significantly higher for those transported by helicopter ( $n=7813$ ; 12.6%) than those by ground ( $n=17\ 775$ ; 11%); however, a higher proportion of both level I and level II patients transported by helicopter had an ISS higher than 24. Patients transported by helicopter had statistically significantly higher heart rates and lower GCS motor scores, lower respiratory rates, and lower systolic blood pressure compared with patients transported by ground EMS ( $P < .001$  for all 4 variables).

The results of the logistic regression models are listed in Table 3. Unadjusted mortality was higher for patients transported by helicopter to level I (ARR, 1.6%; 95% CI, 1.4%–1.7%) and level II trauma centers (ARR, 1.7%; 95% CI, 1.4%–1.8%). However, in all regression models, helicopter transportation was associated with a statistically significantly greater odds of survival. For level I patients, standard logistic regression revealed a greater odds of survival (OR, 1.31; 95% CI, 1.27–1.38;  $P < .001$ ; ARR, 2.9%). This association remained stable when generalized estimating equations with robust variance calculations were applied (OR, 1.32; 95% CI, 1.20–1.45; ARR, 3.0%). In the propensity score–matching analysis, patients in the helicopter group had an attenuated increased odds of survival (OR, 1.16;  $P < .001$ ) and narrower confidence intervals were observed (95% CI, 1.14–1.17) with a smaller ARR (1.5%, 95% CI, 1.4%–1.6%). For level II patients, similar results were observed in each model, with a greater odds of survival in the standard regression model (OR, 1.37; 95% CI, 1.28–1.48;  $P < .001$ ; ARR, 4.3%) and the model accounting for clustering by trauma



center (OR, 1.37; 95% CI, 1.23–1.53;  $P < .001$ ; ARR, 3.4%). The survival benefit associated with helicopter transport remained statistically significant but was attenuated in the propensity score analysis (OR, 1.15; 95% CI, 1.13–1.17;  $P < .001$ ; ARR 1.4%; Table 3 and eTable 3).

When considering patient disposition, the results in Table 4 suggest a higher injury severity in the helicopter group than in the ground transport group. Fewer patients in the helicopter groups were discharged home without services (47.6%) than in the ground transport group (57.3%;  $P < .001$ ) at level I centers. A higher proportion of those transported by helicopter to level I trauma centers were discharged to rehabilitation (18.2% vs 12.7% in ground transport group) and to intermediate facilities (9.3% vs 6.5%, respectively). Fewer patients in the helicopter group left level II trauma centers against medical advice. Patients transported by ground services were more likely to be discharged from level I centers to a nursing home.

## COMMENT

The results from this study indicate that helicopter EMS transport is independently associated with improved odds of survival for seriously injured adults. In regression analyses performed after propensity score matching for adult trauma patients, a 1.5% increased absolute rate of improved survival for 159 511 patients transported by helicopter vs ground emergency service to level I trauma centers was observed. For 63 964 patients transported to level II trauma centers, an absolute survival advantage of 1.4% was found for those transported by helicopter compared with ground transport. Thus, for patients transported to level I trauma centers by helicopter, 65 patients would need to be transported to save 1 life; for patients transported to level II trauma centers, the number needed to treat is 69.

These results are congruent, although more conservative, compared with the results of the few multivariate analyses comparing these 2 modes of EMS. Frankema et al<sup>26</sup> found that after adjustment for injury severity, time of day, and other physiological variables, transportation of a highly trained medical crew to the scene was associated with a nonstatistically significant difference in survival (OR, 2.2; 95% CI, 0.92–5.9;  $P = .08$ ). Blunt trauma patients transported by helicopter in the study by Frankema et al had a statistically significant survival improvement (OR, 2.8; 95% CI, 1.07–7.52;  $P = .04$ ). However, Frankema et al only investigated the effect of helicopter delivery by a highly trained medical crew—including a physician—and did not evaluate the specific effect of helicopter transport of trauma patients to trauma centers.

Brown et al<sup>9</sup> used a logistic regression model including the GCS score, demographics, prehospital times, vital signs, and other hospital variables such as intensive care unit admission and length of hospital stay to calculate the association of helicopter transport with a greater odds of survival (OR, 1.22; 95% CI, 1.18–1.27;  $P < .01$ ). The authors acknowledged major limitations with this study because there were no adjustments for missing data or consideration for clustering by trauma center.

In another study using the 2007 NTDB data, Sullivent et al<sup>10</sup> reported a lower odds of death for patients 18 to 54 years old transported by helicopter (OR, 0.51; 95% CI, 0.44–0.60;  $P < .001$ ). However, Sullivent et al did not control for missing data. In a trauma-related ISS study by Mitchell et al,<sup>8</sup> patients with an ISS of 12 or higher had a W score of 6.4, indicating 6.4 more survivors per 100 patients for patients transported by helicopter compared with patients transported by ground services. Although efforts were made to control for selection bias because all trauma patients who received tertiary care were accounted for, 84% of all patients transported by helicopter were transferred to a trauma center and not flown directly from the scene.

Thomas et al<sup>27</sup> performed a retrospective, registry-based cohort study with 16 699 patients to investigate the role of helicopter transport for patients with blunt trauma. It was found to be associated with a statistically significant mortality reduction (OR, 0.96; 95% CI, 0.59–0.98;  $P=.03$ ); however, Thomas et al only examined blunt trauma patients in the state of Massachusetts over a 4-year period.

In another registry-based study conducted by Stewart et al,<sup>11</sup> propensity scores were used to calculate hazard ratios for patients transported by both helicopter and ground EMS in the state of Oklahoma. The propensity score in the study by Stewart et al was used as a composite variable in a multivariable regression analysis and was calculated by using multiple covariates, including prehospital vital sign data. Although overall mortality was 33% lower for patients transported by helicopter (hazard ratio [HR], 0.67; 95% CI, 0.54–0.84), no significant difference was found for patients with an ISS between 16 and 24 (HR, 0.96; 95% CI, 0.62–1.48).

Each of the above mentioned studies has methodological limitations related to the use of a regression model to estimate causal effects. In many cases, it is not clear how the assumptions for each of their models were tested because regression diagnostics were not reported. The potential correlation of patient outcomes across different trauma centers—ie, clustering—was not assessed in any of the studies except for Thomas et al.<sup>27</sup> Thus, although helicopter transportation was shown to be beneficial, it is possible that the analyses did not fully adjust for known confounders. Indeed, when propensity scores were used in this study to adjust for known confounders, the strength of the association between helicopter and an improved odds of survival was diminished.

In observational studies, the use of logistic regression without balancing background covariates between the treatment and control group can produce biased estimates, especially if the imbalance of covariates is extreme or if the treatment effect is not constant across values of the covariates.<sup>20</sup> A benefit in using propensity scores is that covariates that are speculated to be causing an imbalance between treated and control groups can be balanced and used in a postmatching analysis to control for selection bias and known confounders. Propensity scores have the advantage of producing more accurate effect estimates, especially when the outcomes of interest are relatively rare.<sup>28</sup> For observational, nonrandomized studies, propensity scores represent one of the best available methods to adjust for baseline differences and to simulate the results of a randomized trial.<sup>29</sup>

Some limitations to this research are worth noting. The proportion of missing data in the NTDB is high for many variables. At least one study has suggested that some missing data in the NTDB are not missing at random<sup>30</sup>; this has implications for the imputation methods used to provide plausible values for missing data. Multiple imputation may be more advantageous for valid statistical inference and prevention of type I errors, especially when using large data sets,<sup>31</sup> but this technique requires significant computing power and knowledge about how to combine the results of imputed values.<sup>32,33</sup> Although multiple imputation has been used with the NTDB to impute missing physiological data<sup>34</sup> and has been used in other studies to impute missing baseline covariate data,<sup>33</sup> there is no substitute for actual data. In our study, we limited the use of multiple imputation to variables with no more than 20% missing data because this was the threshold established in a previous study that validated the use of multiple imputation for missing variables in the NTDB.<sup>23</sup>

In addition, the NTDB is described as a “convenience sample,” and contains a disproportionate number of larger hospitals with younger and more severely injured patients.<sup>12</sup> Hence, the NTDB is not a population-based sample, so selection bias may be a problem because not all types of hospitals or patients may be represented.

Although the methods used for propensity score matching resulted in good balance among covariates, as assessed with numerical and graphical diagnostics, it is possible that differences attributed to helicopter vs ground EMS transport might have been due to unobserved confounders. This study represents an attempt to estimate the average treatment effect for patients transported by helicopter vs ground transport using national-level data from a large database and controlling for known confounders. Invariably, any beneficial effect of helicopter over ground transport is the result of some combination of speed, crew expertise, and disposition to a designated trauma center.<sup>4</sup>

Crew configuration and distance data are not available in the NTDB. If distance information were available, alternative methods, such as the use of instrumental variables, might be considered, potentially using distance as the instrument.<sup>35</sup> Crew configuration varies regionally, ranging from the presence of a physician and nurse on the helicopter to a single paramedic. The effect of clinical interventions remains unknown, and in some studies, the presence of a physician failed to confer a survival or quality-of-life benefit.<sup>5, 36</sup> Clustering of patients within each trauma center might be another source of residual bias; previous work with the NTDB has shown that failure to account for clustering may lead to artificially narrow confidence intervals.<sup>30</sup> We attempted to control for this by using deidentified facility identifiers in both our propensity score analysis and generalized estimating equations analysis. We examined the role of total EMS time for helicopter and ground transportation (ie, time from dispatch to arrival in the ED) as an independent variable in a sensitivity analysis that included 42.2% of the data from the level I and level II groups. The results were not qualitatively different from our primary analyses, but any analysis based on EMS times from the 2007–2009 NTDB is highly likely to be biased due to the large amount of missing data, which cannot be assumed to be missing at random.<sup>30</sup>

The survival benefit found in this study may be the result of not only the covariates available in the NTDB but also other unmeasured variables. It is not clear which aspect of helicopter transport is responsible for the mortality benefit in this highly stratified sample. Future studies should investigate specific components of helicopter EMS such as prehospital interventions, total prehospital time, crew configuration, and distance as factors that may in part or whole explain the benefit of helicopter EMS for adults with major trauma because understanding the effectiveness of each may help determine which patients benefit most from this resource. To date, the development and use of effective prehospital triage tools that can identify adults with a high ISS have remained elusive.<sup>37</sup> Future studies should focus on efficient and user-friendly prehospital assessment tools to properly identify injured adults who will be the most likely to benefit from helicopter transport.

Additional outcomes besides mortality, such as health-related quality of life, should also be considered in future helicopter EMS transportation studies.<sup>38</sup> Because it one of the most expensive interventions in contemporary health care, cost must be considered. In a recent systematic review, the annual cost of helicopter transportation ranged from \$114 777 to \$4.5 million per institution.<sup>39</sup> Five studies showed helicopter transportation to be a more expensive transport modality whereas 8 studies indicated that the cost per life-year saved ranged from \$2227 to \$3292 per trauma.<sup>35</sup> In the state of Maryland, the average estimated cost of a helicopter transport is \$5000.<sup>40</sup> Using the number needed to treat estimate of 65 for patients transported to level I trauma centers, approximately \$325 000 would have to be spent to save 1 life. However, this figure does not account for the number needed to treat to prevent disability or other health-related quality-of-life outcomes. Indeed, helicopter transportation represents 1 of the highest cost prehospital modalities used in contemporary trauma care. Hence, policy makers should consider funding a formal cost-effectiveness analysis to help inform policy decisions regarding its use.



Because it is highly unlikely that a randomized clinical trial will be practicable or endorsed by the public to further study the effectiveness of helicopter EMS for adults with major trauma, future studies to estimate its treatment effects for trauma patients will need to rely on designs and techniques that control for selection biases and confounding. Propensity score–based methods appear to be viable in achieving balance among covariates and in making valid statistical inference. Given the attenuation of the association between helicopter EMS and survival observed with propensity score matching, future studies should use this technique and account for missing data or risk overestimating treatment effects.

## CONCLUSION

Among patients with major trauma admitted to level I or level II trauma centers, transport by helicopter compared with ground EMS was associated with improved survival to hospital discharge after controlling for multiple known confounders.

## Acknowledgments

**Funding/Support:** Dr Galvagno was funded in part by an institutional training grant (T-32 Ruth Kischstein grant) from the National Institutes of Health when this study was initiated as part of his PhD program at Johns Hopkins Bloomberg School of Public Health. Dr Haut receives support from a mentored clinical scientist development award (1K08HS017952-01) from the Agency for Healthcare Research. Dr Haider receives support from the American College of Surgeons (C. James Carrico fellowship in trauma and critical care) and from the National Institute of General Medical Sciences (career development award K23GM093112-01).

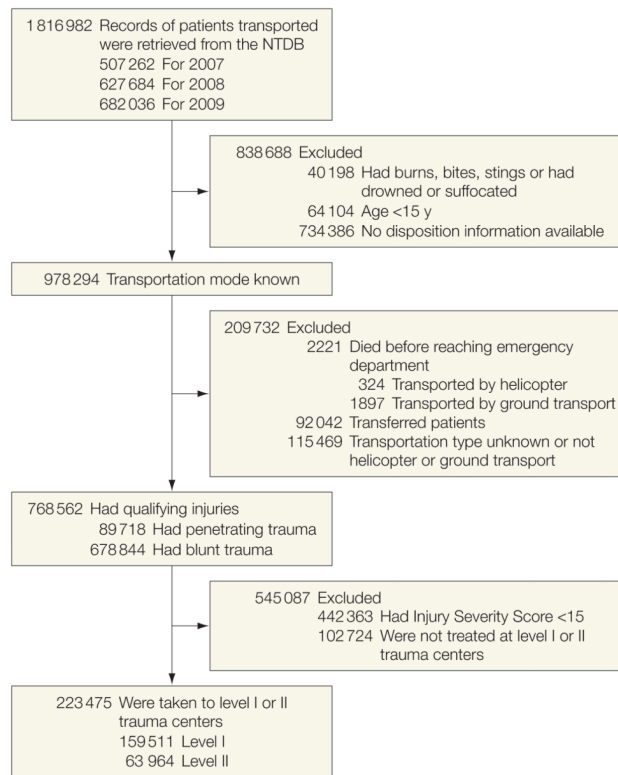
**Role of the Sponsors:** None of the funding sources had a role in the design and conduct of the study; collection, management, analysis, and interpretation of the data; and preparation, review, or approval of the manuscript.

## References

1. World Health Organization. The Global Burden of Disease: 2004 Update. Geneva, Switzerland: World Health Organization Press; 2008.
2. Corso P, Finkelstein E, Miller T, Fiebelkorn I, Zaloshnja E. Incidence and lifetime costs of injuries in the United States. *Inj Prev*. 2006; 12(4):212–218. [PubMed: 16887941]
3. Thomas SH, Cheema F, Wedel SK, Thomson D. Trauma helicopter emergency medical services transport: annotated review of selected outcomes-related literature. *Prehosp Emerg Care*. 2002; 6(3):359–371. [PubMed: 12109585]
4. Thomas SH, Biddinger PD. Helicopter trauma transport: an overview of recent outcomes and triage literature. *Curr Opin Anaesthesiol*. 2003; 16(2):153–158. [PubMed: 17021454]
5. Ringburg AN, Thomas SH, Steyerberg EW, van Lieshout EM, Patka P, Schipper IB. Lives saved by helicopter emergency medical services: an overview of literature. *Air Med J*. 2009; 28(6):298–302. [PubMed: 19896582]
6. Bledsoe BE, Wesley AK, Eckstein M, Dunn TM, O’Keefe MF. Helicopter scene transport of trauma patients with nonlife-threatening injuries: a meta-analysis. *J Trauma*. 2006; 60(6):1257–1265. [PubMed: 16766969]
7. Bledsoe BE. EMS myth #6: air medical helicopters save lives and are cost-effective. *Emerg Med Serv*. 2003; 32(8):88–90. [PubMed: 12942916]
8. Mitchell AD, Tallon JM, Sealy B. Air versus ground transport of major trauma patients to a tertiary trauma centre: a province-wide comparison using TRISS analysis. *Can J Surg*. 2007; 50(2):129–133. [PubMed: 17550717]
9. 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(5):1030–1034. [PubMed: 21068607]
10. Sullivent EE, Faul M, Wald MM. Reduced mortality in injured adults transported by helicopter emergency medical services. *Prehosp Emerg Care*. 2011; 15(3):295–302. [PubMed: 21524205]

11. 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(11):1208–1216. [PubMed: 22092906]
12. American College of Surgeons. Annual Report. Chicago, IL: American College of Surgeons; 2007. National Trauma Data Bank: NTDB Research Data Set Admission Year 2007.
13. Haider AH, Saleem T, Leow JJ, et al. Influence of the National Trauma Data Bank on the study of trauma outcomes: is it time to set research best practices to further enhance its impact [published online February 7, 2012]? *J Am Coll Surg*. 10.1016/j.jamcollsurg.2011.12.013
14. Baker SP, O'Neill B. The injury severity score: an update. *J Trauma*. 1976; 16(11):882–885. [PubMed: 994270]
15. Kane G, Engelhardt R, Celentano J, et al. Empirical development and evaluation of prehospital trauma triage instruments. *J Trauma*. 1985; 25(6):482–489. [PubMed: 4009748]
16. McCowan CL, Swanson ER, Thomas F, Handrahan DL. Outcomes of blunt trauma victims transported by HEMS from rural and urban scenes. *Prehosp Emerg Care*. 2007; 11(4):383–388. [PubMed: 17907020]
17. Oyetunji TA, Crompton JG, Efron DT, et al. Simplifying physiologic injury severity measurement for predicting trauma outcomes. *J Surg Res*. 2010; 159(2):627–632. [PubMed: 20036392]
18. Little RJ, Rubin DB. Causal effects in clinical and epidemiological studies via potential outcomes: concepts and analytical approaches. *Annu Rev Public Health*. 2000; 21:121–145. [PubMed: 10884949]
19. Schafer JL, Kang J. Average causal effects from nonrandomized studies: a practical guide and simulated example. *Psychol Methods*. 2008; 13(4):279–313. [PubMed: 19071996]
20. D'Agostino RB Jr, D'Agostino RB Sr. Estimating treatment effects using observational data. *JAMA*. 2007; 297(3):314–316. [PubMed: 17227985]
21. Rubin DB. Estimating causal effects from large data sets using propensity scores. *Ann Intern Med*. 1997; 127(8 pt 2):757–763. [PubMed: 9382394]
22. Rubin DB. The design versus the analysis of observational studies for causal effects: parallels with the design of randomized trials. *Stat Med*. 2007; 26 (1):20–36. [PubMed: 17072897]
23. Oyetunji TA, Crompton JG, Ehanire ID, et al. Multiple imputation in trauma disparity research. *J Surg Res*. 2011; 165(1):e37–e41. [PubMed: 21067775]
24. Stuart EA. Matching methods for causal inference: a review and a look forward. *Stat Sci*. 2010; 25(1):1–21. [PubMed: 20871802]
25. Lindenauer PK, Pekow P, Wang K, Gutierrez B, Benjamin EM. Lipid-lowering therapy and in-hospital mortality following major noncardiac surgery. *JAMA*. 2004; 291(17):2092–2099. [PubMed: 15126437]
26. Frankema SPG, Ringburg AN, Steyerberg EW, Edwards MJR, Schipper IB, van Vugt AB. Beneficial effect of helicopter emergency medical services on survival of severely injured patients. *Br J Surg*. 2004; 91(11):1520–1526. [PubMed: 15455361]
27. Thomas SH, Harrison TH, Buras WR, Ahmed W, Cheema F, Wedel SK. Helicopter transport and blunt trauma mortality: a multicenter trial. *J Trauma*. 2002; 52(1):136–145. [PubMed: 11791064]
28. Braitman LE, Rosenbaum PR. Rare outcomes, common treatments: analytic strategies using propensity scores. *Ann Intern Med*. 2002; 137(8):693–695. [PubMed: 12379071]
29. Hemmila MR, Birkmeyer NJ, Arbabi S, Osborne NH, Wahl WL, Dimick JB. Introduction to propensity scores: a case study on the comparative effectiveness of laparoscopic vs open appendectomy. *Arch Surg*. 2010; 145(10):939–945. [PubMed: 20956761]
30. Roudsari B, Field C, Caetano R. Clustered and missing data in the US National Trauma Data Bank: implications for analysis. *Inj Prev*. 2008; 14(2):96–100. [PubMed: 18388229]
31. Stuart EA, Azur M, Frangakis C, Leaf P. Multiple imputation with large data sets: a case study of the Children's Mental Health Initiative. *Am J Epidemiol*. 2009; 169(9):1133–1139. [PubMed: 19318618]
32. Rubin DB. Multiple imputation after 18+ years. *J Am Stat Assoc*. 1996; 91(434):473–489.
33. Arnold AM, Kronmal RA. Multiple imputation of baseline data in the cardiovascular health study. *Am J Epidemiol*. 2003; 157(1):74–84. [PubMed: 12505893]

34. Moore L, Hanley JA, Lavoie A, Turgeon A. Evaluating the validity of multiple imputation for missing physiological data in the national trauma data bank. *J Emerg Trauma Shock*. 2009; 2(2): 73–79. [PubMed: 19561964]
35. Hernán MA, Robins JM. Instruments for causal inference: an epidemiologist's dream? *Epidemiology*. 2006; 17(4):360–372. [PubMed: 16755261]
36. Iiro TT, Laaksonen MI, Vahlberg TJ, Pälve HK. Effect of physician-staffed helicopter emergency medical service on blunt trauma patient survival and prehospital care. *Eur J Emerg Med*. 2006; 13(6):335–339. [PubMed: 17091054]
37. Stewart KE, Cowan LD, Thompson DM, Sacra JC. Factors at the scene of injury associated with air versus ground transport to definitive care in a state with a large rural population. *Prehosp Emerg Care*. 2011; 15(2):193–202. [PubMed: 21208038]
38. Galvagno SM Jr. Assessing health-related quality of life with the EQ-5D: is this the best instrument to assess trauma outcomes? *Air Med J*. 2011; 30(5):258–263. [PubMed: 21930080]
39. Taylor CB, Stevenson M, Jan S, Middleton PM, Fitzharris M, Myburgh JA. A systematic review of the costs and benefits of helicopter emergency medical services. *Injury*. 2010; 41(1):10–20. [PubMed: 19853251]
40. Salamon, SJ.; Cowdry, RW. Air Ambulance Study: Required Under Senate Bill. Annapolis: Maryland Health Care Commission; 2006. p. 770



**Figure.**  
Study Flow Diagram Detailing the Stratification and Selection of Patients in the 2007–2009 National Trauma Data Bank (NTDB)

**Table 1**

Characteristics of Patients Transported by Emergency Medical Services to Level I Trauma Center

	All Patients, No. (%) (n = 159 511)	Patients by Transportation Method, No. (%)	
		Helicopter (n = 47 637)	Ground (n = 111 874)
Sex			
Men	111 869 (70.1)	20 422 (42.9)	62 749 (56)
Women	47 642 (29.9)	27 215 (57.1)	49 125 (44)
Age, y			
15–55	120 798 (75.7)	37 182 (78.4)	83 616 (75.7)
>55–65	16 101 (10.1)	4998 (10.5)	11 103 (10.3)
>65	15 648 (9.8)	4833 (10.1)	15 648 (14)
Injury type			
Blunt	144 907 (90.8)	45 155 (94.8)	99 752 (89.2)
Penetrating	14 604 (9.2)	2482 (5.2)	12 122 (10.8)
Cause of injury			
Motor vehicle crashes	109 938 (68.9)	38 388 (80.5)	71 550 (64)
Falls	12 225 (7.7)	1731 (3.6)	10 494 (9.4)
Pedestrians struck by motor vehicles	4423 (2.8)	1206 (2.5)	3217 (2.9)
Unarmed assaults	3818 (2.4)	352 (0.7)	3466 (3.1)
Assaults with a weapon, other than stabbing or shooting	12 089 (7.6)	1512 (3.2)	10 577 (9.5)
Stabbings	3584 (2.3)	509 (1.1)	3075 (2.7)
Gunshot wounds	4788 (3)	669 (1.4)	4119 (3.7)
Other	8646 (5.4)	2162 (4.5)	6484 (5.8)
Injury Severity Score			
15–24	113 730 (71.3)	29 250 (61.4)	80 756 (72.2)
25–34	33 036 (20.7)	11 758 (24.6)	21 278 (19)
35–44	10 476 (6.6)	4360 (9.2)	6116 (5.5)
Died on arrival	1333 (0.8)	194 (0.4)	1139 (1.0)
Died	18 290 (11.5)	5997 (12.6)	12 293 (11)
<b>Mean (SD)</b>			
Systolic blood pressure, mm Hg	129.6 (30.9)	128.3 (29.1)	130.1 (31.7)



	All Patients, No. (%) (n = 159 511)	Patients by Transportation Method, No. (%)	
		Helicopter (n = 47 637)	Ground (n = 111 874)
Respiratory rate, breaths/min	18.1 (7.4)	17.1 (9.0)	18.5 (6.5)
Heart rate, beats/min	93.9 (23.4)	96.4 (23.3)	92.9 (23.4)
Glasgow Coma Scale motor score	4.9 (1.8)	4.4 (2.1)	5.1 (1.6)

**Table 2**

Characteristics of Patients Transported by Emergency Medical Services to Level II Trauma Center

	<u>Patients by Transportation Method, No. (%)</u>		
	<u>All Patients, No. (%) (n = 63 964)</u>	<u>Helicopter (n = 14 272)</u>	<u>Ground (n = 49 692)</u>
Sex			
Men	35 985 (56.3)	8147 (57.1)	27 838 (56)
Women	27 979 (43.7)	6125 (42.9)	21 854 (44)
Age, y			
15–55	46 043 (72)	11 263 (78.9)	35 621 (72.0)
>55–65	6659 (10.4)	1572 (11)	5087 (10.0)
>65	10 421 (17.6)	1437 (10.1)	8984 (18.0)
Type			
Blunt	58 297 (91.1)	13 441 (94.2)	44 856 (90.3)
Penetrating	5667 (8.9)	831 (5.8)	4836 (9.7)
Cause of injury			
Motor vehicle crashes	18 235 (28.5)	5450 (38.2)	12 785 (25.7)
Falls	6627 (10.4)	567 (4)	6060 (12.2)
Pedestrians struck by motor vehicles	2247 (3.5)	486 (3.4)	1761 (3.5)
Unarmed assaults	1675 (2.6)	120 (0.8)	1555 (3.1)
Assaults with a weapon, other than stabbing or shooting	4273 (6.7)	495 (3.5)	3778 (7.6)
Stabbings	1390 (2.2)	163 (1.1)	1227 (2.5)
Gunshot wounds	1372 (2.1)	167 (1.2)	1205 (2.4)
Other	28 145 (44)	6824 (47.8)	21 321 (42.9)
Injury Severity Score			
15–24	46 642 (72.9)	9173 (64.3)	37 469 (75.4)
25–34	11 781 (18.4)	3365 (23.6)	8416 (16.9)
35–44	3553 (5.6)	1170 (8.2)	2383 (4.8)
>45	1988 (3.1)	564 (3.9)	1424 (2.9)
Died on arrival	888 (1.4)	130 (0.9)	759 (1.5)
Died	7302 (11.4)	1816 (12.7)	5482 (11)
		<b>Mean (SD)</b>	

	All Patients, No. (%) (n = 63 964)	Patients by Transportation Method, No. (%)	
		Helicopter (n = 14 272)	Ground (n = 49 692)
Systolic blood pressure, mm Hg	130.6 (32.1)	127.6 (29.8)	131.4 (32.7)
Respiratory rate, breaths/min	18.6 (6.8)	17.3 (7.9)	18.9 (6.4)
Heart rate, beats/min	92.4 (23.7)	95.5 (23.9)	91.5 (23.5)
Glasgow Coma Scale motor score	5.0 (1.7)	4.5 (2.1)	5.2 (1.6)

**Table 3**

## Association Between Transportation Method and the Outcome of Survival to Discharge

	OR (95% CI)	P Value	ARR (95% CI), % <sup>a</sup>
Level I centers (n = 159 511)			
Unadjusted mortality	0.88 (0.85–.90)	<.001	1.6 (1.4–1.7) <sup>b</sup>
Logistic regression			
Standard	1.31 (1.27–1.38)	<.001	2.9 (2.6–3.6)
Controlled for clustering <sup>c</sup>	1.32 (1.20–1.45)	<.001	3.0 (1.9–4.2)
After propensity score matching <sup>d</sup>	1.16 (1.14–1.17)	<.001	1.5 (1.4–1.6)
Level II centers (n = 63 964)			
Unadjusted mortality	0.86 (0.83–0.91)	<.001	1.7 (1.4–1.8) <sup>b</sup>
Logistic regression			
Standard	1.37 (1.28–1.48)	<.001	4.3 (2.6–4.5)
Controlled for clustering <sup>c</sup>	1.37 (1.23–1.53)	<.001	3.4 (2.2–5.0)
After propensity score matching <sup>d</sup>	1.15 (1.13–1.17)	<.001	1.4 (1.3–1.6)

Abbreviations: ARR, absolute risk reduction; OR, odds ratio.

<sup>a</sup>Absolute RR calculations were made after calculating the number needed to treat (NNT), based on the adjusted ORs.<sup>25</sup>  $NNT=1-(PEER \times [1-\text{absolute OR}]) / (1-PEER) \times PEER \times (1-OR)$ . PEER is the patient expected event rate (eg, the event rate in the ground transport group [control group]).  $ARR=1/NNT$ .

<sup>b</sup>Indicates greater mortality for transportation by helicopter.

<sup>c</sup>Generalized estimating equations with robust variance calculations were performed. Independent variables included systolic blood pressure, heart rate, respiratory rate, motor component of Glasgow Coma Score, e-code, age, type of trauma, injury severity score, and sex. E-codes refer to types of external injuries according to the *International Classification of Disease, Ninth Edition, Clinical Modification*.

<sup>d</sup>Logistic regression was performed following estimation of a propensity score with the subclassification matching method. The covariates used to estimate the propensity score were systolic blood pressure, heart rate, respiratory rate, motor component of Glasgow Coma Scale score, e-code, age, type of trauma, injury severity score, facility identifier, and sex.

**Table 4**

## Disposition of Patients Transported by Emergency Medical Services

Disposition to	<u>No. (%) of Patients by Transportation Method</u>		P Value
	Helicopter (n = 47 637)	Ground (n = 111 874)	
Level I centers			
Home without services	22 693 (47.6)	64 063 (57.3)	<.001
Rehabilitation center	8663 (18.2)	14 161 (12.7)	<.001
Intermediate facility	4410 (9.3)	7258 (6.5)	<.001
Home with services	1969 (4.1)	4593 (4.1)	.79
Nursing home	2913 (6.1)	6908 (6.2)	.65
Hospice	61 (0.13)	241 (0.22)	<.001
Left against medical advice	121 (0.25)	856 (0.78)	<.001
Level II centers			
	<b>(n = 14 272)</b>	<b>(n = 49 692)</b>	
Home without services	6572 (46.1)	26 458 (53.2)	<.001
Rehabilitation center	2456 (17.2)	6428 (12.9)	<.001
Intermediate facility	1325 (9.3)	3221 (6.5)	<.001
Home with services	648 (4.5)	1982 (4)	.003
Nursing home	763 (5.4)	3593 (7.2)	<.001
Hospice	35 (0.3)	164 (0.3)	.11
Left against medical advice	67 (0.5)	502 (1.0)	<.001