Is Doing More, Doing Better? Basic Versus Advanced Life Support Ambulances for Medical Emergencies

Thesis advisor: Professor Alan M. Zaslavsky Prachi Sanghavi

Is Doing More, Doing Better? Basic versus Advanced Life Support Ambulances for Medical Emergencies

Abstract

Deficiencies in the quality of pre-hospital care constitute a serious public health problem that has largely been neglected by the scientific community. Trauma and complications of acute disease produce medical emergencies outside of the hospital setting. Treating patients with these conditions involves an inherent trade-off between providing treatment on-site and reducing time to hospital care. My dissertation compares two models of providing pre-hospital care, and highlights a data-driven approach to identifying potentially fraudulent ambulance claims.

Chapters 1 and 2 compare effects of Advanced Life Support (ALS) and Basic Life Support (BLS) on outcomes after out-of-hospital medical emergencies. Most Medicare patients seeking emergency medical transport are treated by ambulance providers trained in ALS. Evidence supporting the superiority of ALS over BLS is limited. I analyzed claims from a 20% sample of Medicare beneficiaries from non-rural counties between 2006-2011 with cardiac arrest, major trauma, stroke, acute myocardial infarction (AMI), or respiratory failure. To address unmeasured confounding, I exploited variation in geographic penetration in ALS rates across counties, using instrumental variables analysis. In particular, I predicted the probability of ALS use for each patient as a function of ALS rates in each county for patients with other diagnoses, using a multilevel, multivariate model. Survival to 90 days for trauma, stroke, cardiac arrest, and AMI patients was higher with BLS than ALS; respiratory failure patients did not exhibit differences in survival. I conducted a secondary analysis based on propensity score-based balancing weights, and this produced generally similar results. I concluded ALS is associated with substantially higher mortality for several acute medical emergencies compared to BLS, and may harm patients through delayed hospital care and iatrogenic injury.

In Chapter 3, I link patient demographic information and ambulance, outpatient, and inpatient claims to look for the inconsistency of having a claim for an ambulance transport with seemingly no real patient - a 'ghost'. I find 1.9% of emergency transports have this inconsistency. I estimate the distribution of ghost ride rates by suppliers and separately, by counties, using an expectation-maximization algorithm. I find the ghost rides are not evenly distributed across counties or suppliers. Although it is not possible to conclusively distinguish billing anomalies due to fraud from data entry errors and similar explanations, this type of analysis may provide useful starting points for further investigation of Medicare fraud.

Contents

For Mummy, Pappa, and Lalu

Acknowledgments

From Shakespeare in Love (1998)

Phillip Henslowe: Mr. Fennyman, allow me to explain about the theatre business. The natural condition is one of insurmountable obstacles on the road to imminent disaster.
Hugh Fennyman: So what do we do?
Phillip Henslowe: Nothing. Strangely enough, it all turns out well.
Hugh Fennyman: How?
Phillip Henslowe: I don't know. It's a mystery.

THIS DISSERTATION represents the most thrilling and fulfilling years of my life. I am grateful to the people who helped me write it.

Alan Zaslavsky, my advisor, shared the dialogue above with students to describe the 'dissertation business'. In fact, it captures Alan so well - his ability to encourage while acknowledging difficulties, all through a beautiful sense of humor. It also shows Alan's willingness to cross formal boundaries and get to know people intimately. Alan knows my family and upbringing, my worries and fears, and my unrelenting outrage at everything from skin lightening creams to predatory marketing schemes. He not only listens, he actively engages with all of these aspects that make me, me. I have learned more from Alan than any other teacher in my life. My parents taught me teachers are God. If Alan is not God, he certainly speaks to Him (in an atheist sort of way). Alan has spent countless hours over four years explaining complex statistical ideas and scientific writing to me, but we have never discussed his research agenda. Instead, he has worked hard to help me with my goals. I am grateful to Alan for reminding me that as a scientist, I have no side.

Joe Newhouse has changed my life. He supported my dissertation idea at an early stage and then never left my side. His enthusiasm for the work has been a great source of confidence for me, in an otherwise scary process. His impeccable knowledge of the American health care system has been crucial to understanding the broader policy context of the pre-hospital care system. Most importantly, Joe has taught me what it means to be an honest and principled academic. Knowing Joe is part of this academic system has increased my faith in it. For the rest of my life, I will think of Joe each time I have to make a difficult decision so that I can make it with integrity. Joe has been incredibly patient and caring, on a daily basis, but especially during the publishing and job market phases. His support kept the sadness away many times.

My productivity increased tenfold after Bapu Jena joined my committee because he could confidently and quickly address clinical questions. He was creative with research ideas and brilliant as an editor. Most of all, Bapu was my friend on the committee, always available by phone and text to answer questions that I was too shy to ask the senior faculty, and always the most encouraging voice after every rejection.

Gary King has been an inspiration as a methodologist, a scientist, and a teacher. For several years, he has put a roof over my head, quite literally because I have spent more time in my office at the Center for Government and International Studies than I have in my apartment. Gary has played a central role in creating a workspace, including common spaces and computing resources, that has

allowed me to be productive during the most crucial years of graduate school.

The Health Policy Ph.D. Program has made these years as a graduate student the happiest in my life. Several faculty beyond my committee members, including Jim Hammitt, Kathy Swartz, Dan Levy, and Dan Wikler, have been mentors and friends. Debbie Whitney, Ayres Heller, and Jessica Livingston made it possible for me to concentrate on my work by caring for personal, financial, and academic concerns. Their concern for my wellbeing and support for my research was always clear, and I will be forever grateful to the Ph.D. Program administration.

My colleagues and neighbors at the Institute for Quantitative Social Science created a vibrant and engaging environment and were always available for brainstorming and technical help. I especially thank Simo Goshev, Steve Worthington, and Ista Zahn.

I am thankful to my friends in the Health Policy Program for being my partners through this process, including (in no particular order) Natalie Carvalho, Sheila Reiss, Aaka Pande, Portia Cornell, Rebecca Haffajee, Aaron Schwartz, Stephanie Morain, Cleo Samuel, Alecia McGregor, Slawa Rokicki, Matt Frank, Daria Pelech, Dorothy Romanus, and Paula Chu. I am also grateful to my many activist friends at Harvard and in the Cambridge community for never letting me forget the good fight for social justice. My dear friend June Beack played a crucial role at a crucial moment, and I thank her forever for it.

I would not have been able to freely study if my heroic parents, Manoj and Urvashi Sanghavi, had not taken the difficult decision to move from India to the United States. They have worked tirelessly and made tremendous sacrifices. My parents embraced the American spirit and created an environment where my voice mattered - I was always allowed to argue back, challenge ideas and people much bigger than me, and encouraged to take charge of my life. Their love was unconditional during my many unconventional moves.

No one can make me laugh and lift my spirits like my brother, Mihir. I watch him as my teacher to understand how to take life. We are polar opposites, yet no one understands me better.

Lastly, life would be meaningless without my best friend and partner, Kavi. In the words of Rumi, he is my lamp, my lifeboat, and my ladder. Without his support, I might still be a miserable consultant. Without his encouragement to study ambulances, my dissertation and therefore graduate school, may not have been so much fun.

1

Cardiac Arrest

1.1 Abstract

Most out-of-hospital cardiac arrests receiving emergency medical services in the United States are treated by ambulance service providers trained in advanced life support (ALS), but supporting evidence for the use of ALS over basic life support (BLS) is limited. Our objective was to compare the effects of BLS and ALS on outcomes after out-of-hospital cardiac arrest. We conducted an observational cohort study of a nationally representative sample of traditional Medicare beneficiaries from nonrural counties who experienced out-of-hospital cardiac arrest between January 1, 2009, and October 2, 2011, and for whom ALS or BLS ambulance services were billed to Medicare (31,292 ALS cases and 1,643 BLS cases). Propensity score methods were used to compare the effects of ALS and BLS on patient survival, neurological performance, and medical spending after cardiac arrest. Outcomes measures included survival to hospital discharge,to 30days, and to 90days; neurological performance; and incremental medical spending per additional survivor to 1 year. Survival to hospital discharge was greater among patients receiving BLS (13.1% vs 9.2% for ALS; 4.0 [95% CI, 2.3-5.7] percentage point difference), as was survival to 90 days (8.0% vs 5.4% for ALS; 2.6 [95% CI, 1.2-4.0] percentage point difference). Basic life support was associated with better neurological functioning among hospitalized patients (21.8% vs 44.8% with poor neurological functioning for ALS; 23.0 [95% CI, 18.6-27.4] percentage point difference). Incremental medical spending per additional survivor to 1 year for BLS relative to ALS was $154,333. In summary, patients with out-of-hospital cardiac arrest who received BLS had higher survival at hospital discharge and at 90 days compared with those who received ALS and were less likely to experience poor neurological functioning.

1.2 Introduction

American emergency medical services (EMS) respond to an estimated 380,000 out-of-hospital cardiac arrests of primary cardiac etiology annually[1]. Although 90% of these patients do not survive to hospital discharge, community training, rapid and appropriate delivery of prehospital care, and high-quality hospital cardiac care may substantially improve survival rates[2–7]. In the United States and in other developed countries, an important strategy for responding to out- of-hospital cardiac arrest has been the delivery of advanced life support (ALS) by ambulance service providers[8].

Advanced life support providers, or paramedics, are trained to use sophisticated, invasive interventions to treat cardiac arrest, including endotracheal intubation, intravenous fluid and drug delivery, and semiautomatic defibrillation[9]. In contrast, basic life support (BLS) providers, or emergency medical technicians, use simple devices such as bag valve masks and automated external defibrillators. As a result, ALS providers tend to spend substantially more time at the location of the cardiac arrest than BLS providers[10]. Reflecting ALS's additional training and equipment, insurance reimbursement for it is higher[11].

However, ALS has no established benefit over BLS for patients with cardiac arrest[10,12]. Of the few high-quality comparisons that exist, the most robust is a before-after study[10] from Ontario, Canada, which found that ALS did not improve survival to hospital discharge compared with a BLS system that optimized the time to defibrillation. Research from the United States is scant, but observational studies[13,14] from urban areas of other high-income countries have also failed to find a benefit of prehospital ALS. Similarly, studies[15,16] on the effectiveness of airway management favor BLS, and evidence of the benefits of intravenous drug delivery in the prehospital setting is limited[17–21]. Understanding the comparative effects of ALS and BLS on health outcomes and medical spending after out-of-hospital cardiac arrest is important not only for countries such as the United States with developed ALS-based emergency response systems but also for developing countries in the process of designing cost-effective prehospital emergency response systems.

1.3 Methods

1.3.1 Study Population and Data Linkage

This research was approved by institutional review boards at Harvard University and the National Bureau of Economic Research. Informed consent was not required because the analysis is based on deidentified Medicare claims. We analyzed a 20% simple random sample of fee-for-service Medicare beneficiaries from nonrural counties who experienced out-of-hospital cardiac arrest between January 1, 2009, and October 2, 2011. We identified ground emergency ambulance rides by Health Care Financing Administration Common Procedural Coding System codes A0429 (BLS emergency), A0427 (ALS level 1 emergency), and A0433 (ALS level 2)[11] with origin and destination codes RH (residence to hospital), SH (scene of accident or acute event to hospital), NH (skilled nursing facility [SNF] to hospital), or EH (residential, domiciliary, or custodial facility or nursing home other than SNF to hospital). We linked 95.7% of these rides to inpatient and outpatient claims by matching on beneficiary identification numbers and dates of service.

For 43,760 ambulance rides, an International Classification of Diseases, Ninth Revision, Clinical Modification (ICD-9-CM) diagnosis code of 427.5 for cardiac arrest was present on an outpatient claim or an inpatient claim marked as 'present on admission'. To focus on cardiac arrests arising from a non-traumatic etiology and to allow comparison with other studies[10], we removed observations with an injury ICD-9-CM diagnosis code (800-999 or E800-E900). We also removed cases (3.1%) from Connecticut, Delaware, Hawaii, and the District of Columbia, where billing practices make it difficult to determine whether ALS provided the service. For example, in Delaware, ALS is supported by local government funds and does not generally bill Medicare. We excluded observations (approximately 10% of the sample) from rural counties as defined by the US Bureau of the Census because they exhibited large differences on baseline characteristics. Finally, we removed cases from North Dakota, Vermont, and Wyoming because they had no BLS cases in nonrural areas. Our final sample size was 32,935 ambulance rides (Figure 1.1). We linked each observation to beneficiary data on demographics, death, and chronic conditions. Using claims for services during the one year before cardiac arrest, we constructed combined Charlson and Elixhauser comorbidity scores[22]. We ascertained total Medicare spending from claims. We obtained demographic data from the 2009 Population Estimates for Zip Code Tabulation Areas[23], county-level demographic and health information for the most recent year available before 2011 for each variable from the Area Health Resources Files[24], and hospital process measures and mortality rates for 2009 to 2011 from the Hospital

Compare data sets[25].

1.3.2 Comparison Groups

We compared BLS and ALS transports defined by the service level billed on the Medicare ambulance claim, as indicated by the Health Care Financing Administration Common Procedural Coding System code. This code reflects the level of service that was deemed medically necessary. Crucially for our purposes, Medicare allows billing at the ALS level if assessment by ALS-trained providers was considered necessary at dispatch, even if ALS providers delivered only BLS interventions. Medicare pays a single amount for the service level that is inclusive of all items, and there is no itemized list of interventions in the claims. Therefore, although we cannot observe the specific combination of provider training, local protocols, or clinical interventions that a patient experienced, the ambulance crew level is an indicator for the set of interventions and scene and transport times that are characteristic of that level.

Guidelines and training for ALS providers direct them to provide ALS care for cardiac arrest or its antecedent conditions[8,20]. Still, a potential concern may be that, after evaluating a patient, ALS-trained providers will deliver BLS interventions to patients who appear healthier and therefore bill at the BLS level. However, as noted above, ALS providers can still bill at the ALS level in these cases, and it is unlikely that they would not do so given the reimbursement differences. Therefore, it is unlikely that BLS cases in our sample were treated by providers trained in ALS.

A second potential concern with comparing outcomes for patients receiving ALS vs BLS is that, if more severe cases were to be triaged by dispatchers toward ALS, our analyses may be confounded by making ALS outcomes appear worse than they would be if patients were randomized to ALS. However, based on telephone interviews with EMS officials in 45 states, we established that existing dispatch protocols generally lead to BLS dispatch for cardiac arrest or any of its prodromal symptoms (e.g., chest pain, breathing difficulty, or fainting) only if ALS is unavailable within a reasonable

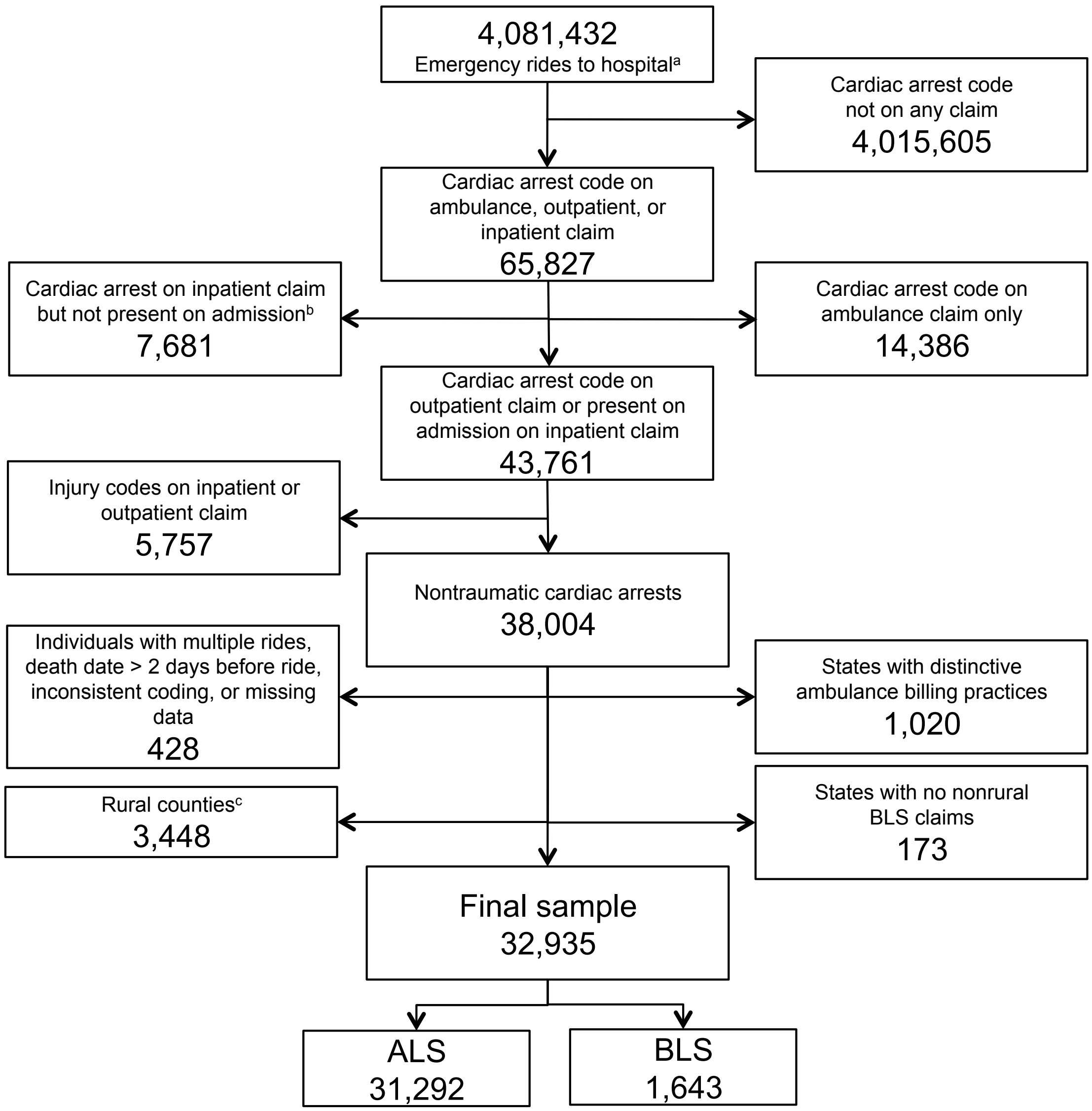


Figure 1.1: Codes refer to International Classification of Diseases, Ninth Revision, Clinical Modification diagnosis codes. ALS indicates advanced life support; BLS, basic life support. [a]Pickup locations included residence, scene of accident or acute event, skilled nursing facility, and non-skilled nursing facility residential, domiciliary, custodial, or nursing home facility. [b]Present on admission status for cardiac arrest is either no or unknown. [c]Rural areas are defined as counties that do not meet the metropolitan or micropolitan criteria as defined by the US Bureau of the Census. Metropolitan counties have at least 1 urbanized area of 50,000 or more population, and micropolitan counties have at least 1 urban cluster of at least 10,000 but less than 50,000 population. Both types have adjacent territory that has a high degree of social and economic integration with the core as measured by commuting ties.

amount of time, either due to travel distance, attendance at another call, or a staffing shortage.

1.3.3 Outcome Measures

Our primary outcome measures were patient survival to hospital discharge, to 30 days, and to 90 days. Our secondary outcomes included neurological performance and medical spending. We inferred Cerebral Performance Categories Scale [26] item 4 (coma or vegetative state) and item 5 (brain death) by the presence of ICD-9-CM diagnosis codes for anoxic brain injury (348.1), coma (700.01), persistent vegetative state (780.03), or brain dead (348.82). We combined these items to create an indicator for poor neurological functioning. For cardiac arrests that occurred in 2009 and 2010, we computed total medical spending up to 1 year after the cardiac arrest or until death.

1.3.4 Statistical Analysis

We first modeled the probability (P) that a beneficiary received ALS using logistic regression. The predicted propensity scores P were used to derive balancing weights [27]. Because ALS cases outnumbered BLS cases, we chose weights to adjust the ALS distribution to the observed BLS distribution over the set of covariates. Therefore, each BLS observation received a weight of 1, and each ALS observation received a weight of $(1 - P)/P$. We chose this approach over propensity score-based matching or stratifying because it provided exact balance most efficiently. Furthermore, unlike using the propensity score as a covariate in a multivariable model, it allowed balance checking.

We tested the following individual-level variables in the propensity score regression: ambulance mileage, history of 27 chronic conditions, and a 6-category zip code-level indicator combining high (>$40,000) or low median household income and racial/ethnic composition (>80% black, >80% white, or integrated) [28]. To account for differences in the quality of hospital care that may be correlated with both outcomes and the propensity of a beneficiary to receive prehospital ALS, we also created zip code-level hospital quality measures, as described in Appendix A.1.

Our final propensity score model adjusted for age (linear spline), sex, race/ethnicity, pickup location, and 3 chronic conditions at the individual level (the model coefficients are summarized in Appendix A.10). At the zip code level, we adjusted for race/ethnicity, the median household income, and hospital quality (Appendix A.1). We also adjusted for urbanicity, percentage older than 25 years with four or more years of college, percentage of primary care practitioners, and the presence of any medical school-affiliated hospital at the county level. We included binary variables for all states with 15 or more BLS observations (i.e., state fixed effects) and created groups by region defined by the US Bureau of the Census for the remaining states. The Hosmer-Lemeshow test was not statistically significant for this model, suggesting that the link function was appropriate.

We used statistical software to construct (SAS version 9.3) and analyze (R version 3.1.0) the sample. All statistical tests were 2-sided at the 5% level. All differences were evaluated using t tests. Kaplan-Meier survival curves were prepared from the weighted observations, with end points defined by death or survival beyond the end of our data on December 31, 2011. Medical spending included Medicare and any non-Medicare primary insurer payments, as well as beneficiary payments, geographically adjusted using the Medicare Hospital Wage Index for an estimated 70% labor share of inputs. For medical spending and survival to 1 year, we used balancing weights estimated for observations in 2009 and 2010, and for survival to 2 years, we used only 2009 data.

1.3.5 Sensitivity Analyses

We conducted several sensitivity analyses, described in Appendices A.2-A.9. First, to assess the extent to which unmeasured disease severity could confound our results, we estimated potential unmeasured confounding by introducing incremental changes to comorbidity scores (Appendix A.2). Second, we assessed the sensitivity of our results to alternative analytic methods by regressing survival on a binary indicator for ambulance type and other variables from our main analysis (Appendix

A.3). Third, we assessed sensitivity to the inclusion of beneficiaries who appeared to have died en route to the hospital (Appendix A.4). We excluded this group in the main analysis because diagnosis is only available from ambulance claims and coding may be inaccurate. Fourth, we used other data sets to check the sensitivity of our results to the exclusion of individuals who may have died at the scene and therefore were not transported (Appendix A.5). Fifth, we estimated the effect of ALS, excluding patients from nursing homes who may have received different on-site care compared with other patients (Appendix A.6). Sixth, we assessed the sensitivity of our results to situations in which BLS called for ALS backup by calculating the number of BLS cases that would have to have been incorrectly attributed to ALS to reverse the direction of our findings (Appendix A.7). Seventh, we estimated the effect of ALS compared with BLS for patients with a primary cardiac etiology by excluding patients with acute respiratory failure codes (Appendix A.8). Eighth, we assessed the robustness of our results to a less sensitive but more specific definition of poor neurological functioning that included only patients with persistent vegetative state or brain death (Appendix A.9).

1.4 Results

Out-of-hospital cardiac arrest mortality rates were high (Table 1.1) and comparable to those of other studies[10,29,30] that used primary data. Beneficiaries who received ALS were slightly younger, were more likely to be male, and were less likely to have most chronic conditions (Table 1.2). They were more often picked up at a residence, whereas patients receiving BLS were more often picked up at a skilled nursing facility. The distributions of household income and race/ethnicity, urbanicity, and the presence of medical school-affiliated hospitals differed (Table 1.3). Beneficiaries receiving ALS services were taken to hospitals that had somewhat better performance on process measures but had slightly worse 30-day mortality from acute myocardial infarction, heart failure, or pneumonia. After applying the propensity score-derived balancing weights to the ALS observations, there were no meaningful differences on any observed measure between the BLS and ALS groups.

Table 1.1: Comparison of Medicare claims-based sample and primary data-based samples on mortality at discharge for individuals brought to a hospital. [a]Discharge status for Medicare outpatient claims was approximated using 2-day mortality because discharge status was poorly coded. CARES, Cardiac Arrest Registry to Enhance Survival [29]; ROC, Resuscitation Outcomes Consortium [30]; OPALS, Ontario Prehospital Advanced Life Support [10].

	Medicare[a]	CARES	ROC	OPALS Study
Number arrived at hospital via EMS	32,935	24,843	7,486	4,247
Inpatients who died before discharge (%)	66	63	NA	NA
Inpatients and outpatients who died before discharge (%)	90	88	87	95

1.4.1 Differences in Patient Survival

Unadjusted survival to hospital discharge was 3.5 (95% CI, 1.9-5.2) percentage points higher among patients receiving BLS (13.1% vs 9.6% for ALS) (Table 1.4). Unadjusted survival after BLS was also greater at 30 days (9.6% vs 6.5% for ALS; 3.1 [95% CI, 1.6-4.5] percentage point difference) and at 90 days (8.0% vs 5.8% for ALS; 2.2 [95% CI, 0.9-3.6] percentage point difference).

After propensity score adjustment, survival to hospital discharge was 4.0 (95% CI, 2.3-5.7) percentage points, or 43%, higher among patients receiving BLS (13.1% vs 9.2% for ALS). Survival after BLS was also greater at 30 days (9.6% vs 6.2% for ALS; 3.4 [95% CI, 1.9-4.8] percentage point difference) and at 90 days (8.0% vs 5.4% for ALS; 2.6 [95% CI, 1.2-4.0] percentage point difference). Kaplan-Meier estimates show that much of the difference in survival between ALS and BLS is explained by higher mortality in the first few days after cardiac arrest for patients receiving ALS (Figure 1.2). After this period, the near constancy in the survival ratios to different time points suggests that patients receiving BLS survive at least as well as those receiving ALS. These findings were unaffected by various sensitivity analyses (Appendices A.2-A.9).

Table 1.2: Differences in individual-level characteristics by ambulance service level. $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$. Differences between BLS and unweighted ALS observations were tested for statistical significance using Student's t-test or chi-square test, as appropriate. Due to missing data, some measures are based on less data than the full sample. a Chi-squared test of independence was used for this categorical variable. [b]Includes non-SNF residential, domiciliary, custodial, or nursing home facilities. [c]Alzheimer's disease/dementia includes Alzheimer's, related diseases, and senile dementia. [d]COPD, chronic obstructive pulmonary disease.

	BLS	Unweighted ALS	Weighted ALS
Mean age	77	75***	77
Female (%)	51	46***	52
Race (%)		a***	
White	72	77	72
Black	21	17	21
Hispanic	3	2	3
Asian	2	2	2
Other	2	2	2
Comorbidity score (mean)	5.5	4.8***	5.5
Chronic conditions (%)			
Acute myocardial infarction	13	14	14
Alzheimer's disease	20	15***	20
Alzheimer's disease/dementia[c]	42	31***	42
Atrial fibrillation	30	29	31
Cataract	66	62*	65
Chronic kidney disease	53	48**	52
COPD[d]	49	49	49
Heart failure	66	62*	67
Diabetes	58	53**	58
Glaucoma	27	22***	25
Hip/pelvic fracture	9	8	9
Ischemic heart disease	75	72**	76
Depression	43	40	43
Osteoperosis	24	20**	23
Rheumatoid arthritis/osteoarthritis	59	55**	58
Stroke/transient ischemic attack	32	27***	31
Breast cancer	5	4	5
Colorectal cancer	6	4	5
Prostate cancer	7	7	7
Lung cancer	5	4	4
Endometrial cancer	1	1	1
Anemia	80	72***	79
Asthma	19	20	19
Hyperlipidemia	76	75	77
Benign prostatic hyperplasia	23	22	21
Hypertension	91	90	92
Acquired hypothyroidism	25	22**	24

Table 1.3: Differences in transport, geographic, and hospital characteristics by ambulance service level. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. Differences between BLS and unweighted ALS observations were tested for statistical significance using Student's t-test or chi-square test, as appropriate. Due to missing data, some measures are based on less data than the full sample. *a* Chi-squared test of independence was used for this categorical variable. [b]High if median income is greater than 40k, low otherwise, and predominantly black if more than 80% black, predominantly white if more than 80% white, and otherwise integrated. [c] Metropolitan areas have at least one urbanized area of 50,000 or more population, and micropolitan areas have at least one urban cluster of at least 10,000 but less than 50,000 population. Both types of area have adjacent territory that has a high degree of social and economic integration with the core as measured by commuting ties. [d] The denominator for these measures is heart attack patients. [e] The denominator for these measures is heart failure patients. LVSD, left ventricular systolic dysfunction; ACE, angiotensin-converting-enzyme; ARB, angiotensin receptor blocker. [f] The denominator for these measures is pneumonia patients.

	BLS	Unweighted ALS	Weighted ALS
Transport level			
Mean mileage (km)	8.7	9.5*	8.5
Pick-up location (%)		*a****	
Residence	55	65	55
Skilled nursing facility	27	14	27
Scene	14	17	14
Non-SNF nursing home[b]	5	4	5
ZIP code level			
Income/race group[b] (%)		*a****	
High/white	37	43	38
Low/white	7	8	7
High/black	2	1	2
Low/black	3	2	3
High/integrated	35	30	34
Low/integrated	16	16	16
Female (%)	51	51***	51
Age 65+ years (%)	14	14	14
County level			
Metropolitan[c] (%)	87	85*	87
Persons with 4+ years of college (%)	24	23***	24
General practice doctors (%)	14	16***	14
Any hospital with med schl affiliation (%)	70	63***	69
Hospital level			
Given aspirin at arrival[d] (%)	98	98	98
Given aspirin at discharge[d] (%)	98	98	98
Given beta blocker at discharge[d] (%)	97	98**	98
Given evaluation for LVSD[e] (%)	97	98***	98
Given ACE inhibitor or ARB for LVSD[e] (%)	94	95	95
Initial blood culture performed prior to first dose of antibiotics[f] (%)	95	96**	96
Given the most appropriate initial antibiotic[f] (%)	93	93*	93
Heart failure (30 day mortality rate)	11	11***	11
Heart attack (30 day mortality rate)	15	16***	15
Pneumonia (30 day mortality rate)	11	12***	11

Table 1.4: Health and payment outcomes by ambulance service level [95% CI]. Unless noted otherwise, estimates are adjusted by propensity-score based balancing weights. Estimates for survival to 1 year used only data from 2009 and 2010, and estimates for survival to 2 years used only data from 2009. Spending includes total payments to the provider by Medicare, the beneficiary, and a non-Medicare primary payer if one exists. Payments are geographically adjusted using the Hospital Wage Index for an estimated 70% labor share of inputs.[a]Discrepancies in differences are due to rounding.

	BLS	ALS	Difference[a]	Ratio
Unadjusted outcomes				
Survival to discharge (%)	13.1 (11.5, 14.8)	9.6 (9.3, 9.9)	3.5 (1.9, 5.2)	1.4 (1.2, 1.5)
Survival to 30 days (%)	9.6 (8.1, 11.0)	6.5 (6.2, 6.8)	3.1 (1.6, 4.5)	1.5 (1.2, 1.7)
Survival to 90 days (%)	8.0 (6.7, 9.3)	5.8 (5.5, 6.1)	2.2 (0.9, 3.6)	1.4 (1.2, 1.6)
Adjusted outcomes				
Survival (%)				
Survival to discharge	13.1 (11.5, 14.8)	9.2 (8.7, 9.7)	4.0 (2.3, 5.7)	1.4 (1.2, 1.6)
Survival to 30 days	9.6 (8.1, 11.0)	6.2 (5.8, 6.6)	3.4 (1.9, 4.8)	1.5 (1.3, 1.8)
Survival to 90 days	8.0 (6.7, 9.3)	5.4 (5.0, 5.8)	2.6 (1.2, 4.0)	1.5 (1.2, 1.8)
Survival to 1 year	6.2 (4.9, 7.6)	4.4 (4.0, 4.8)	1.8 (0.4, 3.3)	1.4 (1.1, 1.8)
Survival to 2 years	6.8 (4.8, 8.9)	3.9 (3.3, 4.5)	2.9 (0.8, 5.0)	1.7 (1.2, 2.4)
Other health measures (%)				
Poor neurological performance	6.1 (5.0, 7.3)	9.7 (9.1, 10.2)	3.5 (2.2, 4.8)	0.6 (0.5, 0.8)
Admission to hospital	25.4 (23.3, 27.5)	20.5 (19.8, 21.2)	4.9 (2.7, 7.1)	1.2 (1.1, 1.4)
Payments (U.S. $)				
Average 1 year spending for all beneficiaries	11,875 (9,754, 13,995)	9,097 (8,527, 9,666)	2,778(582, 4,973)	1.3 (1.1, 1.6)
Average 1 year spending per survivor to 1 year	190,153 (150,041, 230,265)	206,775 (189,909, 223,641)	-	-

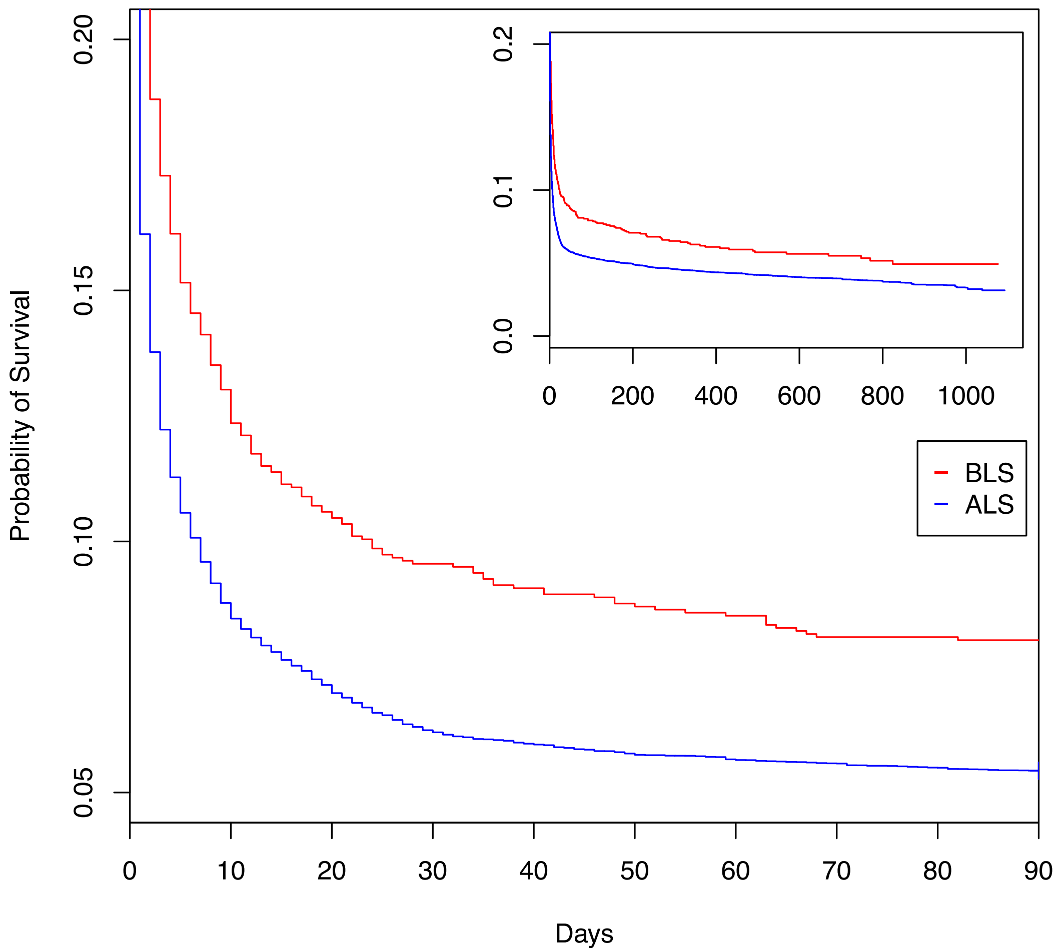


Figure 1.2: Kaplan-Meier analysis of survival after cardiac arrest by ambulance service level. The main plot shows survival probability during the first 90 days, and the inset shows survival probability over the full observational period. Survival analysis was based on cardiac arrests that occurred between January 1, 2009, and October 2, 2011. Mortality was observed until December 31, 2011, when the data were censored; thus, there was follow-up to at least 90 days for each beneficiary. ALS indicates advanced life support; BLS, basic life support.

1.4.2 Differences in Neurological Performance

Among all individuals experiencing an out-of-hospital cardiac arrest, the percentage with poor neurological functioning after cardiac arrest was lower among those who received BLS vs ALS (6.1% vs

9.7%; 3.5 [95% CI, 2.2-4.8] percentage point difference). Among individuals who were admitted to the hospital, rates of poor neurological functioning were markedly lower for BLS compared with ALS (21.8% vs 44.8%; 23.0 [95% CI, 18.6-27.4] percentage point difference).

1.4.3 Differences in Medical Spending

The mean medical spending was higher among beneficiaries receiving BLS ($11,875 vs $9,097 for ALS; $2,778 [95% CI, $582-$4,973] difference), in part because individuals who received BLS survived longer and had more opportunity to receive medical care. Incremental medical spending per additional survivor to 1 year for BLS relative to ALS was $154,333 ([$11,875 - $9,097]/[6.2% - 4.4%]), less than the mean medical spending per survivor to 1 year for ALS ($206,775).

1.4.4 Sensitivity Analyses

With one exception, our results were robust to all the sensitivity analyses described above and in Appendix A.2-A.9. The exception is that, after restricting the definition of poor neurological functioning to only persistent vegetative state or brain death, there was no observed difference in neurological functioning between patients receiving ALS vs BLS.

1.5 Discussion

Using a nationally representative sample of traditional Medicare beneficiaries from nonrural counties who experienced out-of-hospital cardiac arrest between 2009 and 2011 and for whom EMS were billed to Medicare, we compared the effects of out-of-hospital BLS and ALS on survival, neurological performance, and medical spending. Ninety-day survival and neurological performance were substantially better among beneficiaries who received out-of-hospital BLS rather than ALS. Our estimates suggest that each year 1,479 (95% CI, 683-2,276) additional Medicare beneficiaries who experience out-of-hospital cardiac arrest would survive to 90 days if provided BLS instead of ALS. Furthermore, incremental medical spending per additional survivor to 1 year for BLS relative to ALS was $154,333, substantially less than the mean medical spending per survivor to 1 year for ALS ($206,775).

Prehospital care is complex, expensive, and critical to survival after out-of-hospital cardiac arrest, making it crucial to understand the combined effect on morbidity and mortality of the medical interventions, transport time, and training that characterize the two dominant models of prehospital care. Results of our study, to our knowledge the first large-scale systematic comparison of BLS and ALS in the United States, are consistent with those of international studies[10,13,14], which found that ALS does not improve survival to hospital discharge after cardiac arrest. In contrast, our results suggest that the use of ALS is associated with higher mortality than the use of BLS in patients with cardiac arrest. However, most out-of-hospital cardiac arrests treated by EMS in the United States are provided with ALS care.

Although ALS is often assumed to improve clinical outcomes by providing advanced airway management and intravenous drug therapy, other studies have described mechanisms by which ALS may lead to the worse outcomes that we found. First, prehospital endotracheal intubation entails risks, including unrecognized esophageal intubation, aspiration of gastric contents, aggravation of existing injuries such as cervical spine damage, and interference with chest compressions[31]. Furthermore, successful intubation requires high levels of competency and regular practice, but in a Pennsylvania study[32] paramedics performed a median of only one intubation per year. Therefore, bag valve mask ventilation may improve outcomes over endotracheal intubation in out-of-hospital cardiac arrest[15,16]. Consistent with these risks of pre-hospital intubation, a large study[15] of cardiac arrests in Japan found greater neurologically favorable survival with the use of bag valve masks compared with advanced airways. Similarly, an analysis of out-of-hospital cardiac arrests in Los Angeles, California, found that advanced airway methods were associated with decreased survival to hospital discharge compared with bag valve mask ventilation[16]. Second, evidence on the benefits of intravenous drug delivery in out-of-hospital cardiac arrest is limited[17–21]. Third, and perhaps most important, ALS may entail delays in hospital care[10] that would otherwise offer definitive clinical management of the underlying disease (e.g., percutaneous coronary intervention for acute myocardial infarction).

Because a randomized controlled trial of ALS vs BLS is unlikely to occur, we performed an observational analysis. Although our analysis is the largest to date in the United States to our knowledge, it has several limitations. Patients receiving ALS may be at higher risk of mortality irrespective of the intervention, which would confound our estimates. This would be most likely to occur if ALS was dispatched to patients with higher preexisting mortality risk based either on symptoms or preexisting conditions. However, telephone interviews with 45 state EMS agencies demonstrated that if ALS was available it would always be provided in cases of known cardiac arrest or for any typical prodromal symptoms (e.g., chest pain, syncope, etc) that would be known to the dispatcher at the time of dispatch. In other words, BLS would only be dispatched when ALS is unavailable, leaving no clear remaining mechanisms to explain why less severely ill patients would be preferentially dispatched BLS. Moreover, beneficiaries who received BLS had on average more preexisting comorbidities than those who received ALS, suggesting that outcomes among patients receiving BLS would (if anything) be worse and not better. Finally, in analyses of sensitivity to unmeasured confounding (Appendix A.2), our findings that outcomes under BLS were better than under ALS would continue to hold unless an implausibly high difference in unobserved severity was postulated.

An additional source of confounding may be that individuals who can be more easily resuscitated at the scene (e.g., those with ventricular fibrillation) might be overrepresented among BLS cases, while individuals who cannot be resuscitated by BLS wait to be treated by ALS rather than undergoing direct transport to the hospital. Advanced life support would then be spuriously associated with worse outcomes that should have been attributed to BLS. However, our sensitivity analysis of situations in which BLS waits for ALS backup found that this would have to occur in an implausibly high proportion of BLS cases to change the direction of our effect (Appendix A.7).

Additional factors that influence outcomes after cardiac arrest may potentially confound our analysis. For example, shorter ambulance response times to the scene[33] and the presence of a shockable rhythm[29] are associated with improved outcomes. However, no evidence exists that these factors differ between patients receiving ALS vs BLS. However, ALS providers on average spend significantly more time at the scene[10], which suggests how BLS may improve outcomes over ALS via rapid transport to the hospital. Other factors such as the quality of cardiopulmonary resuscitation (CPR) and the use of endotracheal intubation or intravenous drugs are similarly potential mediators of ALS and BLS treatment effects and, like scene and travel time, should not be viewed as confounders. Finally, although bystander-initiated CPR has been associated with improved outcomes[28], we could not directly control for bystander-initiated CPR and defibrillation. However, we adjusted for area-level race/ethnicity and household income, which have been shown to be important determinants of bystander-initiated treatment[28].

An additional limitation is that we used administrative claims, which may be inaccurate and subject to coding errors in diagnoses and procedures. For example, our identification of ALS and BLS exposures may not accurately reflect the service level of the ambulance. However, Medicare policy allows billing at the ALS level if assessment by an ALS-trained crew was considered necessary at dispatch. Based on telephone interviews with state EMS officials, we found some instances of joint BLS and ALS response in which Medicare is billed for only BLS. However, states with distinctive billing practices such as this comprise about 3% of the sample, and our findings were unaffected by their exclusion. Nonetheless, services provided by EMS may differ across areas, which may not be reflected in the level of billing to Medicare. Because we could not identify specific interventions provided to each patient, our conclusions are limited to differences in outcomes associated with the overall practices of BLS and ALS providers.

Our study calls into question the widespread assumption that advanced prehospital care improves outcomes of out-of-hospital cardiac arrest relative to care following the principles of BLS, including rapid transport and basic interventions such as effective chest compressions, bag valve mask ventilation, and automated external defibrillation. It is crucial to evaluate BLS and ALS use in other diagnosis groups and settings and to investigate the clinical mechanisms behind our results to identify the most effective prehospital care strategies for saving lives and improving quality of life conditional on survival.

2

Stroke, Major Trauma, Acute Myocardial Infarction, and Respiratory Failure

2.1 Abstract

Most Medicare patients seeking emergency medical transport are treated by ambulance providers trained in Advanced Life Support (ALS). Evidence supporting the superiority of ALS over Basic

Life Support (BLS), however, is limited, and some studies suggest ALS may harm patients. Our objective was to compare effects of ALS and BLS on health outcomes after out-of-hospital medical emergencies. We analyzed claims from a 20% sample of Medicare beneficiaries from non-rural counties between 2006-2011 with major trauma, stroke, acute myocardial infarction (AMI), or respiratory failure. We compared survival and neurological functioning among patients receiving ALS versus BLS. To address unmeasured confounding, we exploited variation in geographic penetration in ALS rates across counties, using an instrumental variables approach. We conducted a second analysis that balanced characteristics using propensity scores. The measurements included survival to 30 days, 90 days, 1 year, and 2 years, and neurological performance. In the instrumental variables analyses, survival to 90 days among trauma, stroke, and AMI patients was higher with BLS than ALS (4.1 [1.3, 6.9] percentage points for trauma; 4.3 [1.3, 7.3] percentage points for stroke; and 5.9 [2.2, 9.6] percentage points for AMI). For stroke and AMI, these differences persisted for one and two years, respectively. Respiratory failure patients did not exhibit differences in survival between BLS and ALS. Neurological functioning was not significantly different between BLS and ALS in any diagnosis group. Results from the propensity score analyses were broadly consistent. We concluded ALS is associated with substantially higher mortality for several acute medical emergencies compared to BLS, and may harm patients through delayed hospital care and/or iatrogenic injury.

2.2 Introduction

The predominant response to out-of-hospital medical emergencies by ambulance providers in the United States is Advanced Life Support (ALS) rather than Basic Life Support (BLS). ALS accounts for 65% of emergency medical care among Medicare beneficiaries[34], and even more among patients with high-acuity conditions such as stroke. ALS provides sophisticated care on site ('stay and play'), whereas BLS emphasizes rapid transport to the hospital, providing only minimal treatment at the scene ('scoop and run')[35–37]. Whereas ALS providers use invasive interventions, such as endotracheal intubation for airway management and intravenous catheters for drug and fluid delivery, BLS-trained providers use non-invasive interventions, such as bag-valve masks for respiratory support. ALS providers spend more time at the scene on average[10,36,38,39] and receive higher reimbursement[11].

Despite the predominance of ALS, previous studies, mostly from outside the United States, show some evidence of similar or longer survival with BLS[10,12–15,35–38,40–42]. As a result, the World Health Organization has advised countries without ALS not to implement it for trauma until there is greater evidence of its benefits[43,44]. Recent evidence among Medicare beneficiaries experiencing out-of-hospital cardiac arrest suggests ALS is associated with lower 30-day survival and poorer neurologic recovery[45].

Randomized trials of ALS versus BLS are probably infeasible, but American counties differ markedly in their rates of ALS use. This between-county variation in the use of ALS, which is plausibly exogenous to an individual's health and healthcare, provides an opportunity to assess the comparative effectiveness of ALS versus BLS on outcomes after major trauma, stroke, acute myocardial infarction (AMI), and respiratory failure.

2.3 Methods

2.3.1 Data

We analyzed claims between January 1, 2006 and October 2, 2011 from a 20% random sample of Medicare fee-for-service beneficiaries from non-rural counties who were transported to a hospital for out-of-hospital trauma, stroke, AMI, or respiratory failure (Appendices B.1-B.2). We identified ground emergency ambulance rides by Healthcare Common Procedure Coding System (HCPCS) codes (Appendix B.3). We linked 96% of ambulance rides to inpatient and outpatient claims by matching on beneficiary identification number and service date.

We linked each observation to validated death dates and demographic data in the Medicare Denominator/Beneficiary Summary File and to chronic medical conditions in the Chronic Conditions Warehouse File. We used demographic data for ZIP Code Tabulation Areas in 2009[23] and county-level demographic and health information from the Area Health Resources Files[24].

Using claims during the year prior to the emergency event, we calculated Charlson/Elixhauser comorbidity scores[22]. For trauma cases we computed New Injury Severity Scores (NISS) from hospital claim diagnosis codes[46] (Appendix B.4). We generated risk-adjusted hospital quality scores based on non-emergent surgical survival (Appendix B.5).

2.3.2 Sample Construction

We based patient diagnoses on hospital-assigned ICD-9CM diagnosis codes rather than ambulance-assigned codes, which are less likely to be accurate. The Appendix provides flowcharts for the sample construction of each diagnosis group (Figures B.1-B.4) and the diagnosis codes used to define the sample (Appendix B.2).

Among trauma patients, we focused on major trauma, defined as a NISS score above 15 (7% of scored cases)[47,48]. This left 79,687 cases (30,919 BLS, 48,768 ALS). The sample sizes for the other diagnoses were 119,989 for stroke (19,985 BLS, 100,004 ALS), 114,469 for AMI (14,434 BLS, 100,035 ALS), and 82,530 for respiratory failure (9,502 BLS, 73,028 ALS).

2.3.3 Identification of ALS and BLS Services

We identified whether a patient received BLS or ALS using the HCPCS code on the claim. Although provider training, local protocols, and clinical interventions are not recorded, the ambulance provider level indicates the set of interventions and transport times characteristic of that level.

Importantly, even if ALS providers delivered only BLS interventions, Medicare allows billing at the ALS level if assessment by ALS-trained providers was considered necessary at dispatch. Based on telephone interviews with Emergency Medical Services officials in 45 states, we established that the symptoms that generally precede the conditions we studied, such as chest pain or difficulty breathing, would only result in BLS dispatch if ALS were unavailable within a reasonable amount of time, either due to travel distance, attendance at another call, or a staffing shortage. Given the high severity of the medical conditions under study, the policy of allowing ALS billing if ALS was considered necessary at dispatch, and the reimbursement differences between BLS and ALS, it is unlikely that the BLS cases in our sample were actually treated by providers trained in ALS.

2.3.4 Outcome Measures

Our primary outcome measures were survival at 30, 90, 365, and 730 days after ambulance transport. We also created an indicator for poor neurological functioning based on the presence of ICD-9CM diagnosis codes for anoxic brain injury (348.1), coma (700.01), persistent vegetative state (780.03), or brain death (348.82), and so inferred Cerebral Performance Categories 4 and 5[26].

2.3.5 Statistical Analysis

We used two methodological approaches. Our main analysis was a quasi-experimental design that relied on variation in ALS penetration at the county level (Figure 2.1). In particular, we predicted the probability of ALS use for each patient as a function of ALS rates in the county for patients with other diagnoses. Our approach thus estimated survival effects using variation in county-level rates of ALS use that are presumptively driven by local ambulance supply and dispatch protocols common to all included diagnoses, rather than by unobserved characteristics of individuals in the focal diagnosis, such as acuity. In other words, this analysis should not be confounded by individual-level associations between acuity and ambulance type used, since the predicted probability of ALS

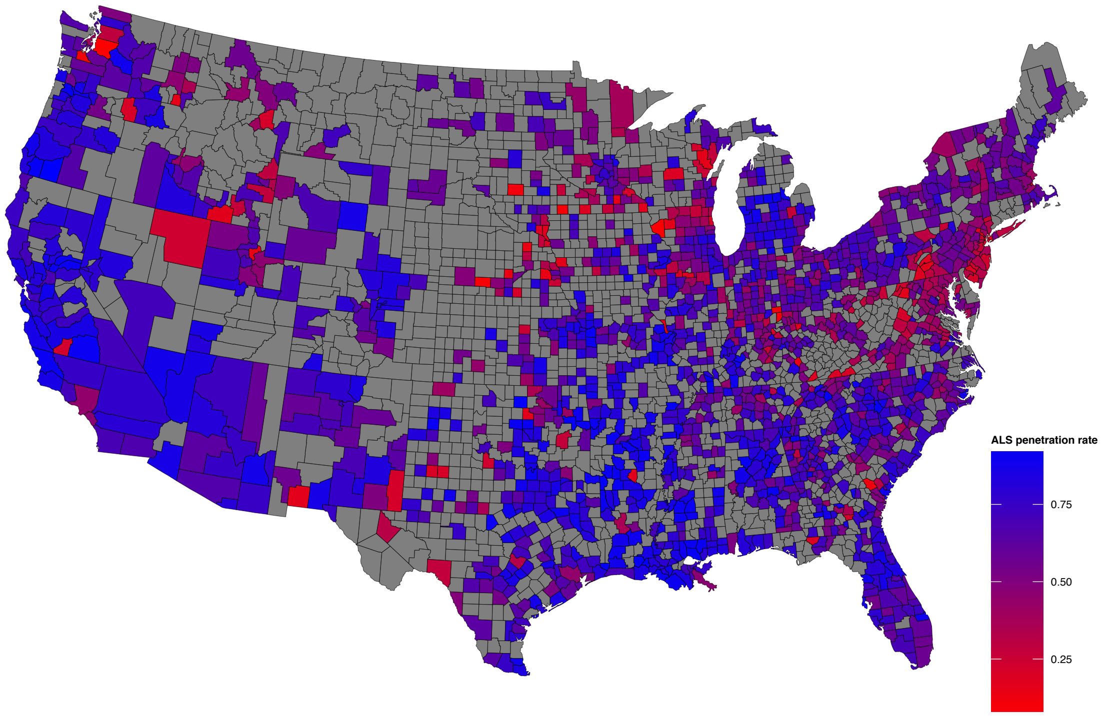


Figure 2.1: County-level ALS penetration rates for major trauma. Rates are for a standardized population but are not derived from characteristics of trauma patients. Rather, they are predicted from ALS use rates in *other* diagnosis groups in each county.

for an individual does not depend on the ambulance type serving that individual.

This approach formally constitutes an instrumental-variables analysis[49–51]. ALS probabilities for an individual case predicted from rates in other diagnosis groups meet the conditions for an instrumental variable, and we use those predicted rates to predict outcomes. We describe our estimation methods in detail in Appendix B.6. We adjusted estimates in the first stage (probability of ALS transport) and second stage (probability of outcome) of the two-stage least-squares procedure for pick-up location type (e.g., residence, scene), mileage from pick-up location to hospital, and a range of individual, county, and ZIP code characteristics and hospital quality (Appendices B.6-B.7).

The instrumental-variables estimates yield the average effect of receiving ALS rather than BLS for those individuals who received ALS but would instead have received BLS in counties with lower

ALS utilization[49]. Because emergency patient outcomes reflect quality of both pre-hospital and hospital care, we conducted a falsification test (Appendix B.15) to address the possibility that ALS prevalence in a county affects patient outcomes through association with quality of hospital care[52]. To do so, we repeated the instrumental variables analysis with an outcome of risk-adjusted nonemergency inpatient surgical mortality, which should be unaffected by the nature of emergency services so long as ALS penetration in a county is not correlated with unobserved quality of hospital care.

We also conducted a secondary analysis, which compared outcomes between ALS and BLS using propensity-score-based balancing weights to adjust for potential confounders (Appendix B.8). This analysis thus exploited individual-level variation in ALS and BLS assignment. We first modeled the probability that a beneficiary received ALS using logistic regression, adjusting for the variables described above. We used diagnosis-specific propensity scores from these models to derive weights[27] that balanced the BLS and ALS distributions over the observed set of covariates, and compared weighted BLS and ALS outcomes. A key assumption in this analysis is that this balance also applies to potential unobserved confounders. Unlike the instrumental-variable analysis, this analysis estimates the magnitude of survival effects from variation in the use of ALS both between and within counties. As a result, the size of estimated effects can be expected to differ between the two methods of analysis.

In both the instrumental-variable and propensity-score analyses, we compared BLS and ALS outcomes with 5% level t tests. We generated Kaplan-Meier survival curves, censored at the end of our data. Finally, we conducted pre-specified subgroup analyses of trauma patients who experienced falls (Appendix B.2), the most common external cause of trauma (75% of patients), and patients with relatively low (16 - 24) and high (25 - 75) injury severity scores. The Appendix gives modeling details (B.6-B.8) and multiple sensitivity analyses (B.9-B.15).

We used SAS 9.3, R 3.0.2, and Stata 13.1 to carry out our analysis.

2.3.6 IRB Approvals.

The research protocol was approved by the relevant institutional review boards at Harvard University and the National Bureau of Economic Research.

2.4 Results

2.4.1 Patient Characteristics

On average, patients who received BLS were older, more likely female, had higher comorbidity scores, and, except for trauma, more likely black (Appendix B, Table B.1). They were also more likely to be picked up at a skilled nursing facility and live in metropolitan areas.

2.4.2 Trauma

In instrumental-variable analysis, patients receiving BLS were 4.1 [1.3, 6.9] percentage points more likely to survive to 90 days (Table 2.1). At 1 year and 2 years, survival was higher with BLS, but not significantly so.

Table 2.1: BLS - ALS differences in health outcomes from county-level analysis. Instrumental variables estimates represent the effect on survival (in percentage points [95% confidence interval]) of receiving BLS rather than ALS for a 'switcher', who would receive BLS in an area with a higher rate of BLS utilization but ALS in an area with lower BLS utilization.

	Trauma	Stroke	AMI	Respiratory failure
30 day survival	3.7 [1.3, 6.0]	5.3 [2.7, 8.0]	4.8 [1.2, 8.4]	4.2 [-0.9, 9.4]
90 day survival	4.1 [1.3, 6.9]	4.3 [1.3, 7.3]	5.9 [2.2, 9.6]	0.2 [-4.7, 5.1]
1 year survival	1.8 [-1.4, 5.0]	3.6 [0.4, 6.8]	7.1 [2.6, 11.6]	-2.9 [-7.8, 1.9]
2 year survival	2.4 [-1.3, 6.1]	3.2 [-0.2, 6.7]	8.4 [2.7, 14.2]	-2.4 [-7.2, 2.3]
Poor neurological performance	-0.3 [-0.6, 0.04]	0.2 [-0.1, 0.5]	-0.7 [-1.5, 0.2]	-0.6 [-2.5, 1.2]

In propensity-score analysis, survival after BLS was 6.0 [5.5, 6.5] percentage points higher at 30 days, and remained higher at intervals up to two years (Table 2.2). Much of the difference in survival between ALS and BLS patients is explained by higher mortality among ALS patients in the days immediately following trauma (Figure 2.2A). After this period the near constancy of survival ratios over time suggests that BLS patients survive as well as ALS patients. Patients receiving BLS were 0.22 [0.15, 0.30] percentage points less likely to experience poor neurological functioning by the time of hospital discharge or hospital death, though there was no such statistically significant difference in the county-level analysis.

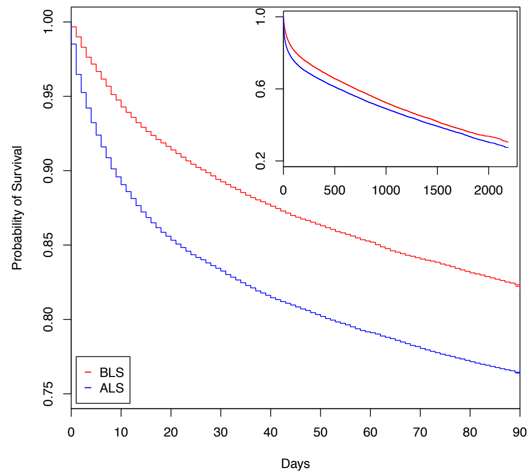
Survival differences between BLS and ALS were more pronounced for patients with more severe trauma (12.5 [4.7, 20.2]) than less severe trauma (2.7 [-0.2, 5.5]) in both instrumental-variable and propensity-score analyses (Table 2.3). For patients experiencing falls, survival to 90 days was higher with BLS in both instrumental-variable (7.1 [2.0, 12.3] percentage points) and propensity-score (4.7 [3.4, 5.9] percentage points) analyses.

2.4.3 Stroke

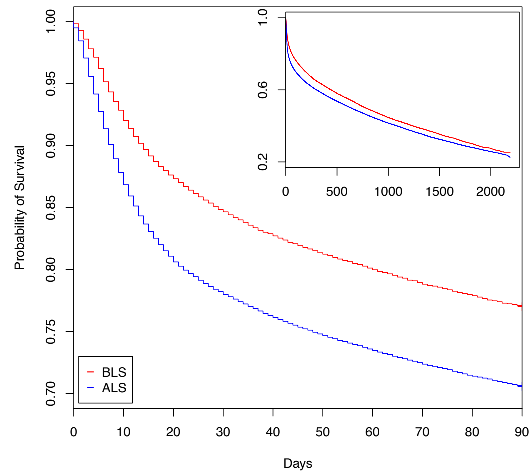
In instrumental-variable analysis, BLS survival was 3.6 [0.4, 6.8] percentage points higher at one year (Table 2.1). Survival was 5.0 [4.2, 5.9] percentage points higher in propensity-score analysis at one year (Table 2.2). As with trauma, this difference was largely explained by higher survival among BLS patients in the initial period following the event (Figure 2.2B). BLS patients were 0.24 [0.16, 0.33] percentage points less likely to experience poor neurological functioning in the propensity-score analysis, but there was no statistically significant difference between BLS and ALS patients in instrumental-variable analysis.

2.4.4 Acute Myocardial Infarction

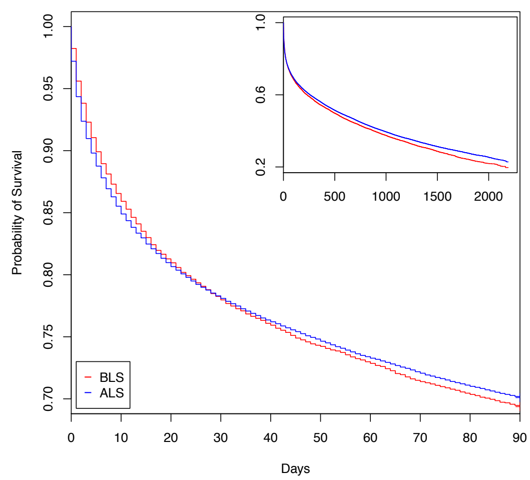
In instrumental-variable analysis the receipt of BLS versus ALS was associated with higher survival at all intervals (Table 2.1). By two years, survival with BLS was 8.4 [2.7, 14.2] percentage points



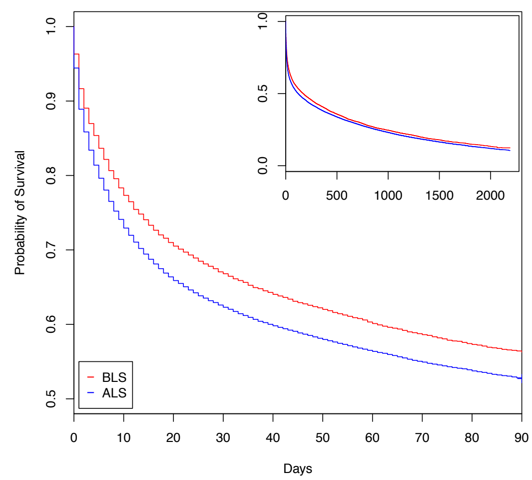
(A) Trauma



(B) Stroke



(C) AMI



(D) Respiratory failure

Figure 2.2: Kaplan-Meier analysis of survival after emergency event by ambulance service level. The inset shows the survival probability over the full observational period, while the main graph shows it for the first 90 days. Data include emergency medical events between January 1, 2006 and October 2, 2011. Mortality was observed until December 31, 2011, when the data were censored, and thus there was follow-up to at least 90 days for each beneficiary. Plots use different y-axis scales.

Table 2.2: Major trauma and stroke health outcomes by ambulance service level, from individual-level analysis. Unless noted otherwise, estimates are adjusted by propensity-score based balancing weights. Estimates for survival to 1 year used data from 2006-2010, and estimates for survival to 2 years used data from 2006-2009.

	BLS [95% CI]	ALS [95% CI]	Difference [95% CI]
Major trauma			
N	30,919	48,768	
Unadjusted outcomes (%)			
Survival to discharge	95.5 [95.3, 95.7]	90.1 [89.8, 90.3]	5.4 [5.1, 5.8]
Survival to 30 days	89.2 [88.9, 89.5]	83.1 [82.8, 83.5]	6.1 [5.6, 6.5]
Survival to 90 days	81.7 [81.3, 82.2]	76.8 [76.5, 77.2]	4.9 [4.3, 5.5]
Poor neurological performance	0.17 [0.12, 0.21]	0.43 [0.37, 0.49]	-0.26 [-0.34, -0.19]
Adjusted outcomes (%)			
Survival to discharge	95.4 [95.2, 95.7]	90.7 [90.4, 91.0]	4.7 [4.3, 5.1]
Survival to 30 days	89.3 [88.9, 89.6]	83.2 [82.9, 83.6]	6.0 [5.5, 6.5]
Survival to 90 days	82.2 [81.8, 82.7]	76.4 [76.0, 76.8]	5.8 [5.2, 6.4]
Survival to 1 year	69.9 [69.3, 70.5]	65.0 [64.5, 65.5]	4.9 [4.1, 5.7]
Survival to 2 years	59.3 [58.6, 60.0]	54.9 [54.3, 55.5]	4.4 [3.4, 5.3]
Poor neurological performance	0.17 [0.12, 0.22]	0.39 [0.33, 0.45]	-0.22 [-0.30, -0.15]
Stroke			
N	19,985	100,004	
Unadjusted outcomes (%)			
Survival to discharge	94.6 [94.3, 95.0]	91.9 [91.8, 92.1]	2.7 [2.4, 3.1]
Survival to 30 days	84.4 [83.9, 84.9]	79.3 [79.1, 79.6]	5.1 [4.5, 5.6]
Survival to 90 days	76.6 [76.0, 77.2]	72.2 [72.0, 72.5]	4.4 [3.7, 5.0]
Poor neurological performance	0.23 [0.16, 0.30]	0.46 [0.41, 0.50]	-0.23 [-0.30, -0.15]
Adjusted outcomes (%)			
Survival to discharge	94.8 [94.5, 95.1]	91.4 [91.2, 91.6]	3.4 [3.0, 3.8]
Survival to 30 days	84.7 [84.2, 85.2]	78.0 [77.7, 78.3]	6.7 [6.1, 7.3]
Survival to 90 days	77.0 [76.4, 77.6]	70.6 [70.2, 70.9]	6.4 [5.7, 7.1]
Survival to 1 year	62.7 [61.9, 63.4]	57.7 [57.2, 58.1]	5.0 [4.2, 5.9]
Survival to 2 years	51.6 [50.7, 52.4]	47.6 [47.1, 48.0]	4.0 [3.0, 5.0]
Poor neurological performance	0.22 [0.15, 0.28]	0.46 [0.41, 0.51]	-0.24 [-0.33, -0.16]

higher. By contrast, survival to 30 or 90 days did not significantly differ between ALS and BLS in propensity-score analysis (Table 2.4). At 1 year, however, survival was 1.7 [0.7, 2.6] percentage points higher with ALS in the propensity-score analysis, and the Kaplan-Meier plot (Figure 2.2C) shows the BLS and ALS curves remain separate after this period. Neurological performance did not statistically differ between BLS and ALS patients in the IV analysis, but in propensity-score analysis

Table 2.3: BLS - ALS differences in 90-day survival by trauma subgroups. Individual-level estimates are the difference in 90-day survival between individuals who received BLS and ALS, adjusted by propensity-score based balancing weights. The county-level estimates represent the effect on survival of receiving BLS rather than ALS for a 'switcher', who would receive BLS in an area with a higher rate of BLS utilization but ALS in an area with lower BLS utilization. Differences are in percentage points [95% confidence intervals]. Accidental falls were analyzed only for 2010 and 2011, in which separate external cause code fields exist and are complete for 92% of observations.

Subgroup	N BLS	N ALS	Survival difference from county-level analysis	Survival difference from individual-level analysis
New Injury Severity Scores 16 - 24	27,297	39,341	2.7 [-0.2, 5.5]	4.4 [3.7, 5.0]
New Injury Severity Scores 25 - 75	3,622	9,427	12.5 [4.7, 20.2]	14.7 [12.9, 16.5]
Accidental falls	7,568	11,947	7.1 [2.0, 12.3]	4.7 [3.4, 5.9]

patients receiving BLS were 0.9 [0.7, 1.1] percentage points less likely to experience poor neurological functioning.

2.4.5 Respiratory Failure

In instrumental-variable analysis there were no statistically significant differences in survival between ALS and BLS (Table 2.1), but in propensity-score analysis survival with BLS was higher at all time intervals (Table 2.4). Early survival gains among patients receiving BLS narrowed with time (Figure 2.2D). In propensity-score analysis, patients receiving BLS were 2.9 [2.6, 3.3] percentage points less likely to experience poor neurological functioning.

2.5 Discussion

We compared health outcomes after pre-hospital ALS versus BLS for Medicare patients with major trauma, stroke, acute myocardial infarction, or respiratory failure. Because these high-acuity conditions necessitate early optimization of care, one would expect any advantage of ALS over BLS to manifest itself in these conditions. It is, however, clinically uncertain whether on-site care improves outcomes due to early treatment or worsens them due to iatrogenic injury or delays in definitive

Table 2.4: Acute myocardial infarction and respiratory failure health outcomes by ambulance service level, from individual-level analysis. Unless noted otherwise, estimates are adjusted by propensity-score based balancing weights. Estimates for survival to 1 year used data from 2006-2010, and estimates for survival to 2 years used data from 2006-2009.

	BLS [95% CI]	ALS [95% CI]	Difference [95% CI]
Acute myocardial infarction (AMI)			
N	14,434	100,035	
Unadjusted outcomes (%)			
Survival to discharge	87.9 [87.4, 88.5]	87.6 [87.4, 87.8]	0.3 [-0.3, 0.9]
Survival to 30 days	77.5 [76.8, 78.2]	80.3 [80.0, 80.5]	-2.8 [-3.5, -2.1]
Survival to 90 days	68.6 [67.8, 69.3]	73.6 [73.3, 73.9]	-5.1 [-5.9, -4.3]
Poor neurological performance	0.71 [0.58, 0.85]	2.07 [1.98, 2.15]	-1.35 [-1.51, -1.19]
Adjusted outcomes (%)			
Survival to discharge	88.2 [87.6, 88.7]	87.2 [86.9, 87.4]	1.0 [0.4, 1.6]
Survival to 30 days	78.0 [77.3, 78.7]	78.1 [77.8, 78.4]	-0.1 [-0.9, 0.6]
Survival to 90 days	69.4 [68.6, 70.1]	70.1 [69.7, 70.5]	-0.7 [-1.6, 0.1]
Survival to 1 year	54.4 [53.5, 55.3]	56.0 [55.6, 56.5]	-1.7 [-2.6, -0.7]
Survival to 2 years	44.0 [43.0, 45.0]	45.5 [45.1, 46.0]	-1.5 [-2.6, -0.4]
Poor neurological performance	0.74 [0.60, 0.89]	1.63 [1.54, 1.72]	-0.88 [-1.05, -0.72]
Respiratory failure			
N	9,502	73,028	
Unadjusted outcomes (%)			
Survival to discharge	77.0 [76.1, 77.8]	75.2 [74.9, 75.5]	1.8 [0.9, 2.7]
Survival to 30 days	66.4 [65.4, 67.3]	64.3 [63.9, 64.6]	2.1 [1.1, 3.1]
Survival to 90 days	55.6 [54.6, 56.6]	55.4 [55.0, 55.7]	0.2 [-0.9, 1.2]
Poor neurological performance	2.39 [2.08, 2.70]	5.86 [5.69, 6.03]	-3.47 [-3.83, -3.12]
Adjusted outcomes (%)			
Survival to discharge	77.4 [76.5, 78.2]	73.7 [73.3, 74.1]	3.7 [2.7, 4.6]
Survival to 30 days	66.8 [65.8, 67.8]	62.3 [61.9, 62.7]	4.5 [3.4, 5.5]
Survival to 90 days	56.3 [55.3, 57.3]	52.7 [52.3, 53.2]	3.6 [2.5, 4.7]
Survival to 1 year	40.4 [38.9, 41.1]	37.4 [37.0, 37.9]	2.5 [1.4, 3.7]
Survival to 2 years	29.0 [27.9, 30.1]	27.6 [27.2, 28.1]	1.4 [0.1, 2.6]
Poor neurological performance	2.42 [2.11, 2.73]	5.36 [5.16, 5.55]	-2.94 [-3.31, -2.57]

hospital management.

We used two methodological approaches to compare outcomes between BLS and ALS. Our primary approach was an instrumental-variable analysis of associations exploiting county-level variation in overall ALS prevalence to predict patients' outcomes. This approach is not susceptible to confounding by associations of ALS use with individual patient characteristics. Although it is potentially susceptible to confounding by unmeasured county-level factors such as hospital quality, in a falsification analysis we found no association between county ALS use and survival of non-ambulance surgical patients, suggesting such confounding is not present (Appendix B.15).

Our second analysis used propensity score methods and so is susceptible to confounding by any unobserved patient characteristics associated with both survival and ALS use. However, it is less subject to county-level confounding than the instrumental-variable analysis since individuals are compared both within and between counties. Furthermore, because ambulance dispatch protocols prioritize ALS for these conditions, such individual-level confounding is likely to be minimal.

The two methodological approaches we employed rely on different comparisons and so will generally estimate different effect sizes. However, using two distinct approaches allows us to test the robustness of our inferences to a particular methodological strategy. We found that these approaches generally delivered similar qualitative results. The findings from both methods for trauma and stroke showed that survival was higher at most intervals following the event for patients receiving BLS. For AMI and respiratory failure patients, however, the results differed. For AMI, survival was higher with BLS for all intervals examined in the instrumental-variable analysis, whereas in the propensity-score analysis, there was no detectable difference at 30 and 90 days and ALS survival was greater at one and two years. For respiratory failure, survival was not different between BLS and ALS in the instrumental-variable analysis, but survival was significantly higher with BLS for all time intervals in the propensity-score analysis. Poor neurological performance was significantly more likely among ALS than BLS patients for all diagnoses in propensity-score analyses, but there were no significant differences in the instrumental-variable analyses.

In sum, with the possible exception of AMI, BLS produced significantly better or similar outcomes than ALS. These findings are consistent with other evidence for cardiac arrest (Appendix B.16) and trauma[10,13–15,37,38,40–45]. Little prior evidence, however, exists for stroke, AMI, and respiratory failure. Based on reimbursement levels of ALS and BLS, our findings suggest that Medicare would have spent $322 million less on ambulance services in 2011 if all rides had been BLS[34], without little or no detriment in patient outcomes.

We conducted several sensitivity analyses (Appendix B.9-B.15), none of which changed the direction or significance of our main findings.

How might ALS result in worse outcomes? First, pre-hospital endotracheal intubation by ALS providers has risks[32]. Successful intubation requires high competency and practice, but in one large state the median paramedic performed only one intubation annually[53]. Bag-valve mask ventilation, commonly performed as part of BLS, may improve outcomes in comparison[13,16,54–57]. Second, administration of pre-hospital intravenous fluids may actually harm trauma patients by disrupting hemostasis, either by directly disrupting an already-formed clot or reflexively reducing peripheral vascular resistance through the expansion of intravascular volume[58,59]. Third, ALS may delay hospital care that would otherwise offer definitive clinical management[10,36,38,39]. Even when clinical guidelines recommend not delaying transport for pre-hospital interventions, delays may still result from the on-site provision of optional interventions that are intended to be performed en route to the hospital[60].

Our study has potential limitations. A key limitation of any observational study is the possibility of selection bias. To guard against such bias, we conducted two types of analyses that are subject to different types of confounding. Our instrumental-variable analysis would be confounded if counties with poorer quality health care or higher severity had higher ALS penetration. We conducted a falsification test (Appendix B.15), however, that showed no association of ALS penetration with non-emergent surgical mortality at the county level. Our propensity-score analysis would be biased if ALS were more likely to be dispatched for more severely ill patients. Interviews we conducted with 45 state EMS representatives, however, confirmed that BLS would only be dispatched for the types of symptoms our non-trauma patients would exhibit if ALS were unavailable. We did not ask the EMS representatives about dispatch decisions for trauma, but we controlled for trauma severity so our analysis of individuals with major trauma is unlikely to be confounded by unobserved severity differences between BLS and ALS.

Our propensity-score analysis would be subject to a selection bias if ALS-trained providers evaluated a patient and then provided care and billed at the BLS level. Given the substantial reimbursement differences between ALS and BLS, however, ALS providers' billing at BLS rates is unlikely since Medicare allows billing at the ALS level if assessment by ALS-trained providers was considered necessary at dispatch. Furthermore, analysis of survival differences in 2005 claims, which distinguish ALS claims billed at the BLS level, showed little sensitivity to inclusion of this small group in ALS or BLS categories (Appendix B.14).

Because we limited our samples to patients with hospital claims, another potential concern may be that more BLS patients died at the scene or en route to the hospital. In sensitivity analyses that considered these cases, however, the direction and significance of our findings was unchanged (Appendix B.9-B.10).

Finally, our results are limited to the Medicare population, and the administrative data we relied on may not always accurately reflect diagnoses, comorbidities, or neurological performance.

In conclusion, our study questions whether advanced life support improves clinical outcomes after out-of-hospital emergency medical events. We studied medical conditions for which ALS would be expected to manifest life-saving benefits over BLS, if these benefits exist. Our findings suggest survival is longer with BLS and that BLS may also offer benefits for non-fatal outcomes.

3

Ghosts on Ambulances, Courtesy of Medicare

Fraud in health care is much discussed, but these discussions necessarily rely on incomplete data[61]. Medicare fraud recovery efforts returned $3.3 billion to the federal treasury in fiscal year 2014, about half a percent of Medicare spending in that year[62]. The extent of fraud, however, is thought to be considerably larger, with one well known estimate that it accounts for at least 3 percent and possibly as much as 10 percent of all US health care spending[63]. Geographic areas that are extreme outliers for specific services have been suggested to indicate fraud. For example, in 2006 Miami Medicare home health spending per beneficiary was more than 6 times the national average and durable medical equipment spending per beneficiary was more than 7 times the national average[64] .

Most 'big-data' driven investigations have relied on single sources of data to infer patterns of anomalous billing, e.g., identifying excessively high rates of durable medical equipment spending in Part B claims submitted to Medicare. However, relatively little use has been made of claim linkages to identify patterns that might signal improper or incomplete billing. In this Perspective, we highlight the potential importance of claim linkages to identify anomalous and plausibly fraudulent billing in Medicare.

To demonstrate this strategy we investigated possible fraud in Medicare emergency ambulance transports, using a 20% simple random sample of Medicare claims between 2006 and 2011. Emergency ambulance transports, which are generated by a 911 dispatch protocol, account for the majority of ground ambulance rides, 55%, and $3.05 billion in spending in 2011. In 2006, the Office of the Inspector General (OIG) in the Department of Health and Human Services estimated that a quarter of overall ambulance transports in 2002 (non-emergency and emergency) did not meet Medicare coverage requirements and had been improperly paid . In particular, the OIG and in a more recent report, the Medicare Payment Advisory Commission, raised fraud and abuse concerns over non-emergency ambulance transports, in part due to uneven growth in this sector across providers and geography[34]. Aberrant billing in emergency ambulance transports has not been investigated to our knowledge.

Linking together patient demographic information with ambulance, outpatient, and inpatient claims, we looked for the inconsistency of having a claim for an ambulance transport with seemingly no real patient - a 'ghost'. We found a non-trivial number of such seemingly fraudulent claims.

We focused on Basic and Advanced Life Support ground emergency ambulance claims (Healthcare Common Procedure Coding System codes A0429, A0427, and A0433) from non-institutional suppliers with a destination of a hospital and a pick-up location of residence, scene of accident or acute event, skilled nursing facility (SNF), or non-SNF nursing home, and removed claims for which payment was denied. Beneficiaries with more than one ride in the same day (1%) were dropped. By matching beneficiary identification numbers, we linked ambulance claims to outpatient and inpatient claims that had a date of service between two days before and up to seven days after the transport. Therefore, we only used claims for ambulance transports that occurred between January 3rd and December 24th of each year. Most beneficiaries (96.5%) had a matching claim within two days of the transport, but the extended date range allowed for date errors in billing. We also dropped beneficiaries who were enrolled in a Medicare Advantage plan at any point in the year, to exclude the possibility the ambulance claim was submitted to Medicare and hospital claims to the plan. Finally, we excluded patients with death dates that were not validated with the Social Security Administration or death dates prior to or within two days of the ambulance transport.

After these exclusions, 1.9% of claims billed to Medicare for emergency ambulance rides to a hospital had no linked hospital claim in the subsequent days. These 'ghost rides' accounted for $309 million of ambulance spending over the 6 years of our sample.

As one might expect of fraudulent behavior, these 'ghost rides' were not evenly distributed across suppliers or across counties. We examined data for the 8,163 largest suppliers and the 2,591 counties that account for 99.5% percent of the rides. We excluded 2.5% of observations for which a provide identifier was unavailable. We estimated the underlying distribution of ghost rides rates for counties, and separately for suppliers, by nonparametric maximum likelihood using an expectation-maximization (EM) algorithm. This approach accounted for sampling variation in our 20% sample, and the mean and standard deviations reported below describe these distributions. We defined outliers as counties and suppliers that were 1.5 standard deviations or more above the mean. To estimate the percentage of counties and suppliers that were above this cutoff, we summed the estimated probability distribution above the cutoff. To identify specific outliers with a high probability (0.95 or higher) of being above the cutoff, we summed the posterior probabilities (that is, the product of the likelihood and the estimated underlying distribution) for each county and supplier above the cutoff.

The mean ghost ride rates for both providers and counties equaled 1.7%, with standard deviations (SD) equaling 1.4% and 1.3% respectively. We estimated that 355 (4.3%) suppliers had rates exceeding 3.7% (1.5 standard deviations above the mean for suppliers) and 128 (4.9%) counties had rates exceeding 3.6% (1.5 standard deviations above the mean for counties). We identified 123 suppliers and 61 counties with at least a 0.95 probability of having a ghost ride rate exceeding 1.5 standard deviations above the mean. Figure 3.1 shows the geographic distribution of raw ghost ride rates in Texas and Florida, both of which had several outlying counties.

With 100% of claims, the Centers for Medicare and Medicaid Services (CMS) would be able to identify many more outlying providers. Since suppliers often operate within a county and county government agencies can be suppliers of emergency ambulance services too, there was overlap in the two lists. For example, among the nine counties with more than 10% estimated ghost rides, five were also represented in the provider list.

A few suppliers and counties caught our attention. County A has about 146,000 emergency transports in our sample, of which 7% are ghost rides. Supplier B is a private company based out of a southern state with approximately 43,000 emergency transports, of which 11%, or 4,900, are ghost rides. Supplier C is a large city with almost 74,000 emergency transports, of which 8%, or 5,900, have no hospital claim.

To summarize, our analysis identified a substantial number of claims for emergency ambulance transports that are strongly suggestive of fraud. Through data linkages, the particular billing inconsistency we detected should be relatively easy for CMS and other payers to pursue. It may be useful

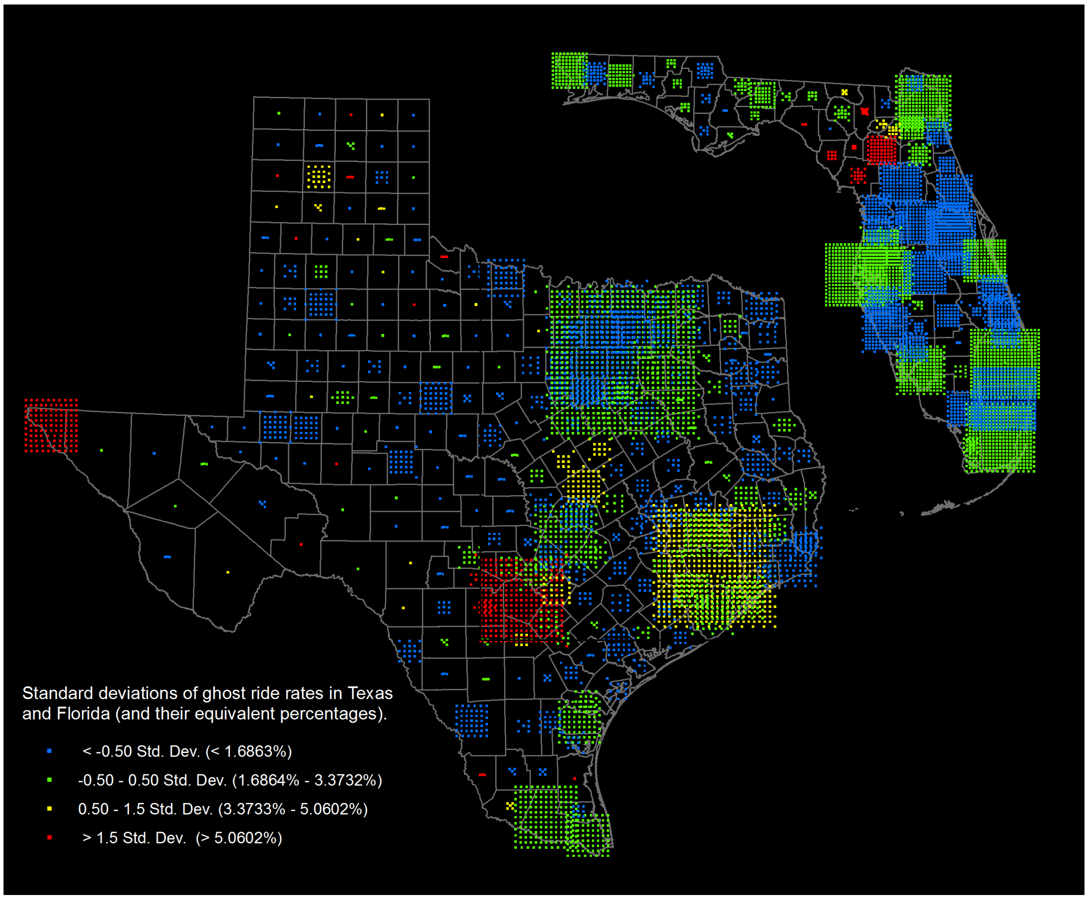


Figure 3.1: Each individual square represents 100 rides (counties with less than 100 rides are represented with one square). Groups of squares representing the total rides in a county are placed at the county's geographic center. The national mean ghost ride rate for counties was 1.7%, with standard deviations (SD) corrected for sampling variation equaling 1.3%. The mean raw ghost ride rate for counties in Texas and Florida is 1.9%, with SD of 1.9%.

for payers to check for other inconsistencies in ambulance billings by linking to hospital claims, e.g., using hospital-assigned diagnosis codes to assess the likelihood that ambulance claims were up-coded from Basic to Advanced Life Support. Although such analyses cannot by themselves conclusively distinguish billing anomalies due to fraud from data entry errors and similar explanations, they may identify anomalous patterns with a high probability of confirmed fraud or inappropriate use upon further investigation. Unlocking this information through data linkages can help Medicare to promote proper and efficient use of health care resources.

A

Supplementary Materials for Chapter 1

A.1 Hospital quality measures

Because the quality of hospitals to which ALS and BLS ambulances transport patients might differ systematically, we allow hospital quality to be part of the ambulance effect, while still controlling for the average quality of available hospital options. We constructed this measure of hospital quality as follows.

First, we selected all patients who had ALS ambulance service to each hospital in 2009 - 2011, with hospital-assigned diagnosis codes for acute myocardial infraction (AMI), congestive heart failure (CHF), or pneumonia (PN) from our Medicare claims, excluding cardiac arrest cases. Second, we obtained AMI, CHF, and PN 30-day mortality measures from Hospital Compare for 2009 - 2011, and averaged over the three years within each measure. Third, we linked each Medicare observation from the first step with the hospital mortality measures from the second step for the hospital to which the patient was taken. Fourth, we averaged the hospital mortality measures in the newly linked dataset by ZIP codes. This created ZIP-code level hospital mortality averages, weighted by the number of people visiting each hospital from the ZIP code.

In our analysis, we linked the ZIP code of each cardiac arrest patient with its corresponding ZIP code level hospital mortality rates. In cases where no ZIP code level hospital mortality rates were found, we linked with rates in ZIP codes that were nearby, based on sharing the same first four digits and numerically nearest fifth digit. We averaged in cases where mortality rates were found for two equidistant ZIP codes. Finally, for each observation, we computed an average of the ZIP code hospital mortality rates for AMI, CHF, and PN, weighted by the overall distribution of these cases in the Medicare sample. This is the final measure we used to control for the quality of available hospitals in a ZIP code.

A.2 Sensitivity analysis: Unmeasured severity in ALS transports

Unmeasured severity differences between patients might have led to differential ambulance dispatch and treatment, and also affected outcomes. Though this is unlikely in cardiac arrest cases, we used comorbidity scores to estimate this potential bias.

We regressed survival to 90 days on a binary indicator for ambulance type. Our logistic regression was specified similarly to our propensity score model in the main analysis. We incremented the average comorbidity score for ALS cases until the coefficient for ambulance type was not statistically significant at the 5% level.

The mean comorbidity score among BLS patients was 5.5 with a standard deviation of about 4, and among ALS patients was 4.8 with a standard deviation of also about 4. The observed difference in survival could be explained by an unobserved factor affecting ALS mortality that has an effect equivalent to an average comorbidity score that is 5 units higher, or 9.8, which is about 1.3 standard deviations above the observed mean ALS comorbidity score. It is unlikely that there was an unobserved difference in severity of this magnitude. Thus, our main findings are not sensitive to unobserved differences in severity. However, a limitation of this analysis was that comorbidity scores may not be good constructs for measuring severity in acute events.

A.3 Sensitivity analysis: Adjustment using logistic regression for outcomes

In our main analysis, we balanced the covariate distributions between Basic Life Support (BLS) and Advanced Life Support (ALS) by generating weights based on propensity scores. We developed our propensity score model systematically and used likelihood ratio tests to compare model specifications. As an additional alternative, we used logistic regression to estimate survival to 30 days and to 90 days. This allowed us to check the modeling dependency of our results.

We regressed the outcomes, survival to 30 days and to 90 days, on a binary indicator for ambulance type. Otherwise, our logistic regression used the same variables as our propensity score model in the main analysis. To estimate the difference in outcomes, we predicted the probabilities of survival for the population that was transported by BLS for both types of ambulances. Thus, we simulate the average effect of ALS for the BLS population as we did in the main analysis.

Survival to 30 days was 3.6 percentage points (95% CI: 1.1, 7.8) higher and to 90 days was 2.8 percentage points higher (95% CI: 0.7, 6.9) with BLS level of service (Table A.1). The direction or significance of our main findings did not change.

Table A.1: Survival outcomes by ambulance service level, adjusted by logistic regression model.

	BLS (95% CI)	ALS (95% CI)	Difference (95% CI)
Survival to 30 days (%)	9.6 (3.2, 21.3)	6.1 (2.0, 13.9)	3.6 (1.1, 7.8)
Survival to 90 days (%)	8.1 (2.2, 19.9)	5.4 (1.4, 13.4)	2.8 (0.7, 6.9)

A.4 Sensitivity analysis: Death en route to hospital

Ambulance diagnosis coding is generally of poor quality, and thus we did not use it in our main analysis. However, this may have excluded some beneficiaries who died prior to arrival at a hospital and thus do not have hospital claims. According to Medicare rules, if a patient dies after dispatch but prior to loading onto the truck, the ambulance service may only bill at the BLS level and indicate this situation with a HCPCS modifier code. Thus, it is not possible to know the service level in these cases. These cases are likely to often involve individuals who would not be considered revivable. If a patient was transported and the ambulance correctly coded cardiac arrest, we would expect the patient to have a death date on the same day as the ride, or at most, on the day after the ride. In this analysis, we check the sensitivity of our main findings to the inclusion of this group.

We included in our sample those beneficiaries who were transported by ambulance, were identified as being in cardiac arrest by the ambulance crew, do not have a hospital claim, and have a death date on the same day or the day after the ride. We identified 1,538 cases that met this criteria, of which 151 were provided with BLS service and 1,387 were provided with ALS service. It was not possible to exclude injury cases as codes for these diagnoses were generally not reported on ambulance claims. We applied the same propensity score model specification and weighting approach as in our main analysis, and estimated survival to 30 days and 90 days.

Survival to 30 days was 2.9 percentage points (95% CI: 1.5, 4.2) higher and to 90 days was 2.2 percentage points higher (95% CI: 0.9, 3.5) with BLS than with ALS (Table A.2). The direction or significance of our main findings did not change. However, this analysis was limited by the quality of ambulance diagnosis coding.

Table A.2: Survival outcomes by ambulance service level, with beneficiaries who died prior to hospital arrival.

	BLS (95% CI)	ALS (95% CI)	Difference (95% CI)
Adjusted survival to 30 days (%)	8.8 (7.4, 10.1)	5.9 (5.5, 6.3)	2.9 (1.5, 4.2)
Adjusted survival to 90 days (%)	7.4 (6.1, 8.6)	5.2 (4.8, 5.5)	2.2 (0.9, 3.5)
Unadjusted survival to 90 days (%)	7.4 (6.1, 8.6)	5.6 (5.3, 5.8)	1.8 (0.6, 3.0)

A.5 Sensitivity analysis: Death in the field

We excluded patients with only an ambulance claim, and therefore individuals who died at the scene. If patients receiving BLS are more likely to die at the scene, our results may be confounded. However, for two key reasons, it is not possible to use the Medicare claims data to assess the sensitivity of our results to this exclusion. First, in cases where an individual is treated at the scene but pronounced dead before being loaded into the truck, both ALS and BLS providers are paid at the BLS level and therefore bill at this level. Second, these observations have only ambulance diagnosis coding, which is unlikely to be accurate in general, but even more so in cases where there was little time to observe the patient.

Therefore, we have used data sources other than the claims to estimate how deaths in the field may have affected our estimates. While these datasets likely differ in key ways from the Medicare sample, these approximate calculations provide reassurance.

In an analysis by the Resuscitation Outcomes Consortium (ROC)[30], approximately 63% of cardiac arrest cases where resuscitation was attempted by EMS were transported to a hospital. In Table A.3, we apply this figure to our Medicare sample to estimate the BLS/ALS distribution that would be required among cases that died in the field in order to eliminate our observed effect.

Table A.3: BLS/ALS distribution required among additional field deaths to remove observed effect using ROC data. ROC, Resuscitation Outcomes Consortium.

	Medicare sample size	Medicare sample 90-day mortality	Additional estimated deaths in field	Overall mortality rate
BLS	1,643 (5%)	1,511	1,934 (10%)	$\frac{1,511+1,934}{1,643+1,934} = 96\%$
ALS	31,292 (95%)	29,477	17,409 (90%)	$\frac{29,477+17,409}{31,292+17,409} = 96\%$
Total	32,935 (63%)	-	19,343 (37%)	-

To remove our observed effect, 10% of field deaths would have to have been treated by BLS, which is twice the overall percent of BLS in our sample. Further, the BLS mortality rate in the field (56%) would have to be 1.5 times the ALS mortality rate in the field (37%). This does not seem plausible.

We repeated the above analysis using data from the Cardiac Arrest Registry to Enhance Survival (CARES)[29], in which 22% of cases treated by EMS died in the field. In Table A.4, we show that to remove the observed difference between BLS and ALS, about 13% of field deaths would have to be treated by BLS. This is more than two times the overall percent of BLS in our sample. Also, the BLS mortality rate in the field (44%) would have to be twice the ALS mortality rate in the field (22%). Therefore, we do not believe accounting for deaths in the field would change the direction of our observed effect.

Table A.4: BLS/ALS distribution required among additional field deaths to remove observed effect using CARES data. CARES, Cardiac Arrest Registry to Enhance Survival.

	Medicare sample size	Medicare sample 90-day mortality	Additional estimated deaths in field	Overall mortality rate
BLS	1,643 (5%)	1,511	1,208 (13%)	$\frac{1,511+1,208}{1,643+1,208} = 95\%$
ALS	31,292 (95%)	29,477	8,081 (87%)	$\frac{29,477+8,081}{31,292+8,081} = 95\%$
Total	32,935 (78%)	-	9,289 (22%)	-

A.6 Sensitivity analysis: Nursing homes

Although we control for pickup location in the main analysis, there may be concern about residual confounding related with interactions between being in a nursing home and other covariates. For example, nursing home staff may selectively treat some patients with a defibrillator or CPR and therefore be able to request BLS service to the hospital. This would attribute survival to BLS instead of the nursing home staff. To study the sensitivity of our results to this potential source of confounding, we repeat our analysis for 30 day and 90 day survival using only observations that did not originate at a nursing home.

After removing nursing home pickups, our sample includes 1,205 BLS and 26,896 ALS cases. Survival to 30 days was 3.5 percentage points (95% CI: 1.7, 5.3) higher and to 90 days was 3.2 percentage points higher (95% CI: 1.5, 4.9) with BLS level of service (Table A.5). The direction or significance of our main findings did not change.

Table A.5: Survival outcomes by ambulance service level for non-nursing home pickups.

	BLS (95% CI)	ALS (95% CI)	Difference (95% CI)
Survival to 30 days (%)	10.5 (8.7, 12.2)	7.0 (6.5, 7.5)	3.5 (1.7, 5.3)
Survival to 90 days (%)	9.4 (7.7, 11.0)	6.2 (5.7, 6.7)	3.2 (1.5, 4.9)

A.7 Sensitivity analysis: BLS requested ALS backup

In areas with two-tier response, it may be that BLS providers request ALS backup when BLS is unable to resuscitate a patient. In these cases, ALS would be spuriously associated with worse outcomes that otherwise should have been attributed to BLS.

Our sample includes only rides in which a transport occurred and a hospital bill was generated. Thus, in order for these cases to be included in our sample, the patient would have to survive until

ALS arrives, be considered appropriate for transport, and be provided with service in the Emergency Department.

We estimate the number of BLS cases that would have to have been incorrectly attributed to ALS as described above in order to change the direction of our findings. In our sample, 1,511 of 1,643 BLS and 29,477 of 31,292 ALS patients did not survive to 90 days. For the calculation, we simply moved patients who died under ALS to the group of patients who died under BLS until the proportion of survivors was the same in both groups. We found this occurred when about 600 cases were removed from the sample of ALS patients who had died by 90 days after the arrest and added to the sample of BLS cases that had died by 90 days after the arrest. Thus, to change the direction of our findings, 600/(1,643+600) or 27% of BLS cases would have to have been in the situation where BLS could not resuscitate and called ALS for backup, ALS treated the patient and transported the patient to the hospital, and the Emergency Department provided service to the patient. This does not seem plausible.

A.8 Sensitivity analysis: Removal of respiratory failure observations

It is possible that outcomes after ALS and BLS are different for patients with cardiac arrest that originates from a cardiac etiology versus patients with a root respiratory cause. To study the sensitivity of our results to this potential source of confounding, we repeat our analysis for 30 day and 90 day survival for a sample that excludes patients with acute respiratory failure ICD-9CM diagnosis codes (518.4, 518.5x, 518.81, and 518.82).

After removing acute respiratory failure cases, our sample includes 1,373 BLS and 25,999 ALS cases. Survival to 30 days was 3.0 percentage points (95% CI: 1.7, 4.4) higher and to 90 days was 2.4 percentage points higher (95% CI: 1.2, 3.7) with BLS level of service (Table A.6). Though the overall mortality rates are higher in these cases, the direction or significance of our main findings did not change.

Table A.6: Survival outcomes by ambulance service level for observations with likely primary cardiac etiology.

	BLS (95% CI)	ALS (95% CI)	Difference (95% CI)
Survival to 30 days (%)	6.7 (5.4, 8.0)	3.7 (3.3, 4.0)	3.0 (1.7, 4.4)
Survival to 90 days (%)	5.8 (4.5, 7.0)	3.3 (3.0, 3.6)	2.4 (1.2, 3.7)

A.9 Sensitivity analysis: Narrower definition of poor neurological performance

We inferred Cerebral Performance Categories (CPC) scale items 4 and 5 based on the presence of ICD-9CM diagnosis codes for anoxic brain injury (348.1), coma (700.01), persistent vegetative state (780.03), or brain dead (348.82). Since individuals with anoxic brain injury and coma can recover, defining poor neurological performance using only diagnosis codes for persistent vegetative state and brain dead may be more precise. Therefore, we repeat our analysis of poor neurological functioning following ALS and BLS using this narrower specification.

After restricting the definition to only persistent vegetative state and brain death, a higher percentage of ALS than BLS patients experienced poor neurological functioning, both overall and among only admitted patients, but the difference between ALS and BLS was not statistically significant (Table A.7).

Table A.7: Neurological performance outcomes by ambulance service level using narrower definition of poor neurological functioning.

	BLS (95% CI)	ALS (95% CI)	Difference (95% CI)
Poor neurological performance, overall (%)	0.06 (-0.05, 0.2)	0.2 (0.1, 0.2)	0.1 (-0.02, 0.3)
Poor neurological performance, admitted patients (%)	0.2 (-0.2, 0.7)	0.7 (0.4, 1.0)	0.5 (-0.06, 1.0)

A.10 Propensity score regression parameters

Below, Table A.8 shows the regression parameters from the logistic regression model that was used to predict the probability of receiving ALS. This was used to generate propensity scores and hence balancing weights for the analysis.

Table A.8: Coefficients from logistic regression model for predicting the propensity to receive ALS (log-odds ratios are shown). [a]Includes non-SNF residential, domiciliary, custodial, or nursing home facilities. [b]Alzheimer's disease/dementia includes Alzheimer's, related diseases, and senile dementia. [c]High if median household income > $40,000, low otherwise, and predominantly black if more than 80% black, predominantly white if more than 80% white, and otherwise integrated. [d]Metropolitan areas have at least one urbanized area of 50,000 or more population, and micropolitan areas have at least one urban cluster of at least 10,000 but less than 50,000 population. Both types of area have adjacent territory that has a high degree of social and economic integration with the core as measured by commuting ties. [e]Measure described in detail in Section A.1.

Variable	Coefficient	95% CI
Intercept	0.198	-0.921, 1.336
State fixed effects (not shown)	-	-
Female -0.134	-0.240, -0.029	
Linear age spline: 0 - 65 years	0.888	0.253, 1.498
Linear age spline: 65 - 75 years	0.973	0.419, 1.501
Linear age spline: 75 - 80 years	0.642	0.063, 1.195
Linear age spline: 80 - 85 years	0.850	0.278, 1.395
Linear age spline: 85 years and over	0.344	-0.383, 1.061
Race Reference: White	-	-
Race: Asian	-0.306	-0.665, 0.081
Race: Black	-0.167	-0.323, -0.009
Race: Hispanic	-0.224	-0.538, 0.110
Race: Other	-0.303	-0.634, 0.054
Pickup Reference: Residence	-	-
Pickup: Non-SNF Nursing Home[a]	-0.400	-0.648, -0.140
Pickup: SNF	-0.797	-0.928, -0.665
Pickup: Scene	0.203	0.050, 0.359
Chronic condition: Alzheimer's/Dementia[b]	-0.192	-0.312, -0.072
Chronic condition: Diabetes	-0.069	-0.176, 0.039
Chronic condition: Asthma	0.162	0.030, 0.298
Race/Income ZIP Mix[c] Reference: Black, High Income	-	-
Race/Income ZIP Mix: Black, Low Income	-0.051	-0.568, 0.446
Race/Income ZIP Mix: Integrated, High Income	0.198	-0.262, 0.622
Race/Income ZIP Mix: Integrated, Low Income	0.163	-0.303, 0.596
Race/Income ZIP Mix: White, High Income	0.516	0.047, 0.952
Race/Income ZIP Mix; White, Low Income	0.315	-0.188, 0.792
Metropolitan County[d]	0.166	-0.015, 0.345
Percent of Persons with 4 Plus Years of College in County	0.007	-0.001, 0.014
Percent General Practice Doctors in County	-0.011	-0.017, -0.006
Any Medical School-Affiliated Hospital in County	-0.079	-0.225, 0.065
Hospital Quality ZIP Measure[e]	0.218	0.161, 0.275

B

Supplementary Materials for Chapter 2

B.1 Sample Construction

We modeled outcomes of multiple rides for a beneficiary separated by at least the period of interest (90 days, 1 year, or 2 years) in a given analysis. For example, in studying 90-day survival, the index ambulance transports for patients with multiple rides were restricted to those at least 90 days after the last ride for the same diagnosis.

We removed cases (about 3% of the sample) from Connecticut, Delaware, Hawaii, and the District of Columbia where billing practices make it difficult to determine whether ALS provided the service. For example, in Delaware, ALS is supported by local government funds and does not generally bill Medicare. We excluded observations (about 11% of the sample) from rural counties, those not meeting standard metropolitan or micropolitan criteria[24], because they exhibited large differences on baseline characteristics from the metropolitan and micropolitan sample. Sample construction flow charts are show in Figures B.1-B.4. For outcomes beyond 90 days, we lacked data for individuals who entered our sample late in the observation period and so we appropriately trimmed the end of our dataset (i.e., we dropped the last year of data for 1 year outcomes for those for whom a year would have extended past the end of our period of observation and similarly for those for whom the last two years of data for 2 year outcomes would have extended past the period of observation) and refit the models using the reduced datasets.

We linked ambulance claims to the nearest in time hospital claims using the beneficiary's identification number and the date of service. This allowed us to use diagnosis codes that described the medical emergency, rather than any subsequent developments. Our algorithm prioritized linking to the nearest hospital claim (up to two days after the ride) and to inpatient over outpatient claims. In each diagnosis group, at least 91% of ambulance transports were linked to inpatient claims. The vast majority of ambulance transports linked to outpatient claims (at least 94% in each diagnosis group) had either died in the emergency department and therefore were not admitted to the hospital or were transferred to another health facility, according to discharge status codes.

B.2 Diagnosis Codes

Trauma cases were identified by ICD-9CM codes 800 to 959.9, excluding late effects of injury, foreign bodies, complications, and burns. Falls were identified by external cause codes E880-E888 for

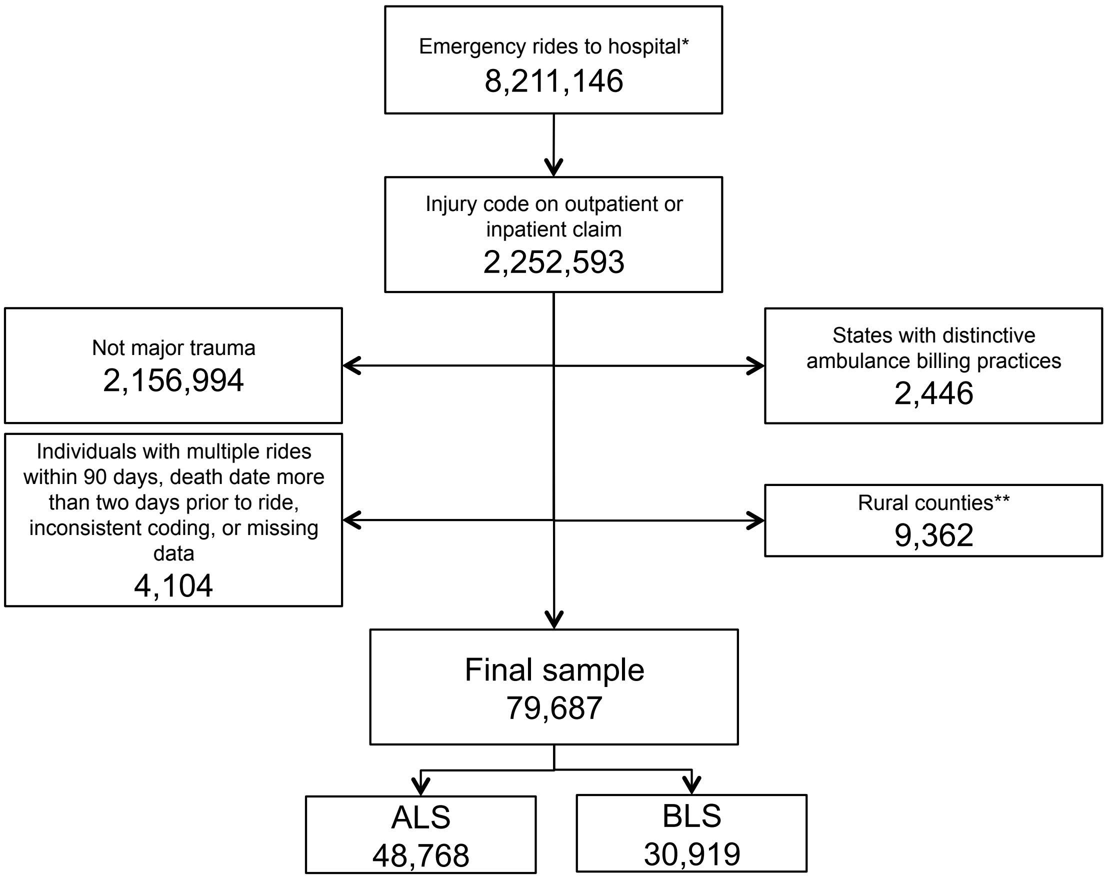


Figure B.1: Flowchart for trauma observations. Codes refer to ICD-9CM diagnosis codes. *Pick-up locations included residence, scene of accident or acute event, skilled nursing facility (SNF), and non-SNF residential, domiciliary, custodial, or nursing home facility. **Rural areas are defined as counties that do not meet the metropolitan or micropolitan criteria as defined by the U.S. Bureau of the Census. Metropolitan counties have at least one urbanized area of 50,000 or more population, and micropolitan counties have at least one urban cluster of at least 10,000 but less than 50,000 population. Both types have adjacent territory that has a high degree of social and economic integration with the core as measured by commuting ties.

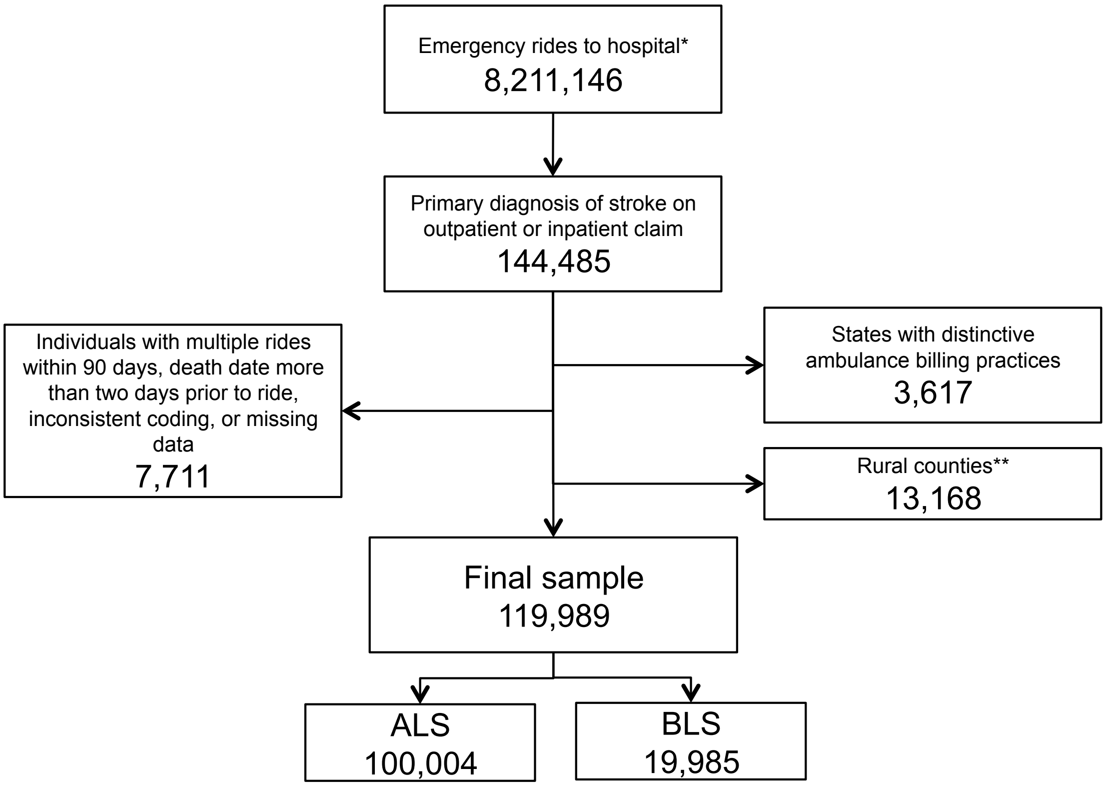


Figure B.2: Flowchart for stroke observations. Codes refer to ICD-9CM diagnosis codes. *Pick-up locations included residence, scene of accident or acute event, skilled nursing facility (SNF), and non-SNF residential, domiciliary, custodial, or nursing home facility. **Rural areas are defined as counties that do not meet the metropolitan or micropolitan criteria as defined by the U.S. Bureau of the Census. Metropolitan counties have at least one urbanized area of 50,000 or more population, and micropolitan counties have at least one urban cluster of at least 10,000 but less than 50,000 population. Both types have adjacent territory that has a high degree of social and economic integration with the core as measured by commuting ties. An additional 74 observations were dropped from the final stroke sample for the county-level analysis due to missing data on two county-level covariates.

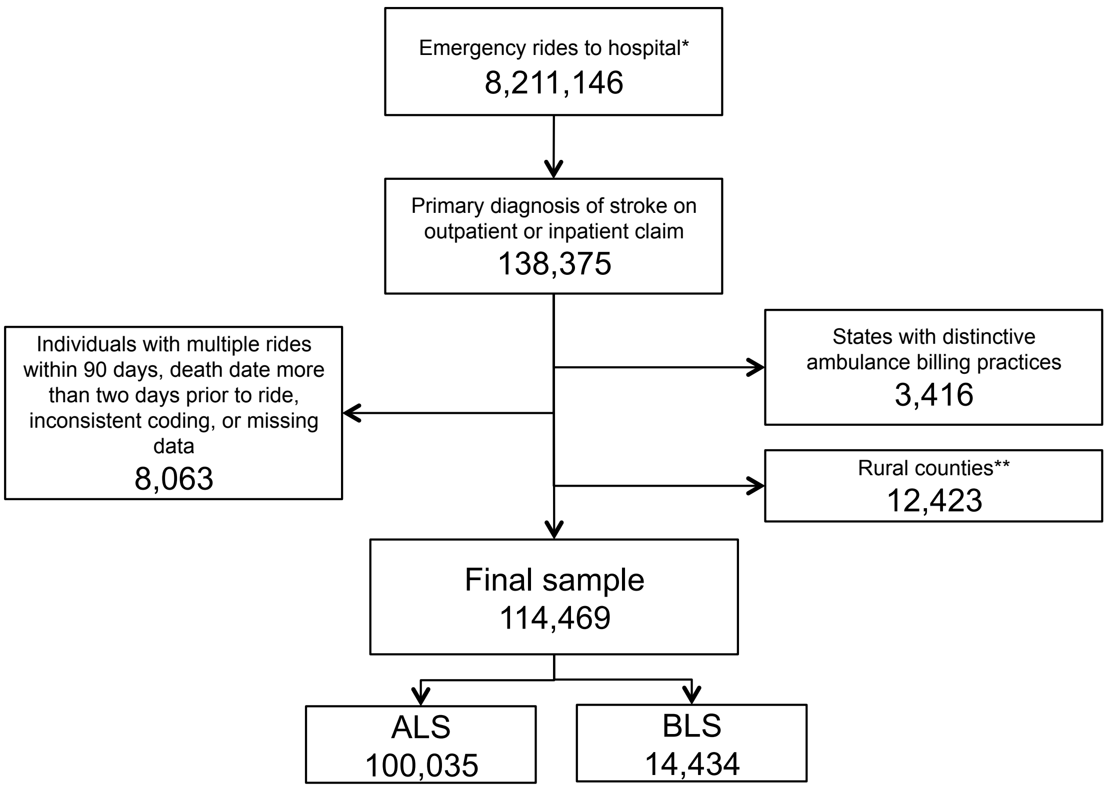


Figure B.3: Flowchart for AMI observations. Codes refer to ICD-9CM diagnosis codes. *Pick-up locations included residence, scene of accident or acute event, skilled nursing facility (SNF), and non-SNF residential, domiciliary, custodial, or nursing home facility. **Rural areas are defined as counties that do not meet the metropolitan or micropolitan criteria as defined by the U.S. Bureau of the Census. Metropolitan counties have at least one urbanized area of 50,000 or more population, and micropolitan counties have at least one urban cluster of at least 10,000 but less than 50,000 population. Both types have adjacent territory that has a high degree of social and economic integration with the core as measured by commuting ties.

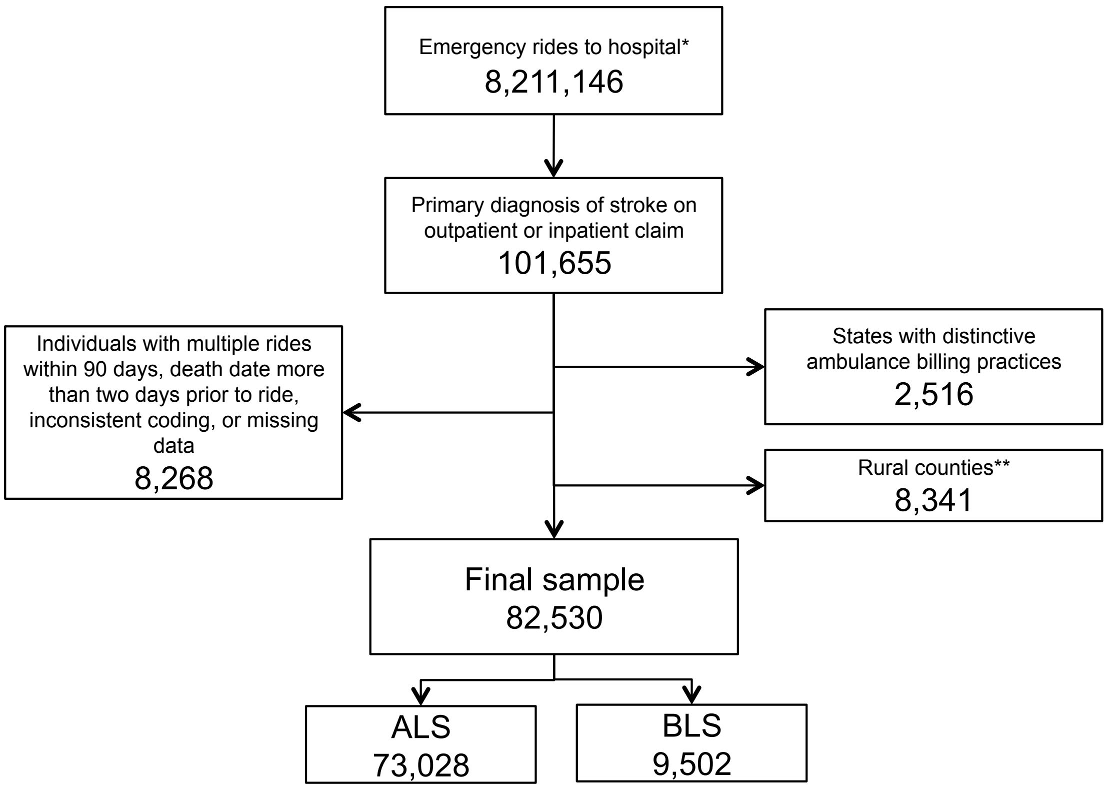


Figure B.4: Flowchart for respiratory failure observations. Codes refer to ICD-9CM diagnosis codes. *Pick-up locations included residence, scene of accident or acute event, skilled nursing facility (SNF), and non-SNF residential, domiciliary, custodial, or nursing home facility. **Rural areas are defined as counties that do not meet the metropolitan or micropolitan criteria as defined by the U.S. Bureau of the Census. Metropolitan counties have at least one urbanized area of 50,000 or more population, and micropolitan counties have at least one urban cluster of at least 10,000 but less than 50,000 population. Both types have adjacent territory that has a high degree of social and economic integration with the core as measured by commuting ties.

accidental falls (excluding E887), and were analyzed only in 2010 and 2011, which, unlike earlier years, include separate external cause code fields. Those fields were completed for 92% of observations.

We used primary diagnosis codes for AMI (only initial episodes, 410.x1), stroke (433, 434, or 436), and respiratory failure (518.4, 518.81, or 518.82). Table B.1 shows differences in characteristics of BLS and ALS patients by diagnosis group.

B.3 Healthcare Common Procedure Coding System (HCPCS) Codes for Ambulance Services

We identified ground emergency ambulance rides by HCPCS codes A0429 (BLS Emergency), A0427 (ALS Level 1 Emergency), and A0433 (ALS Level 2)[11], with origin and destination modifier codes RH, SH, NH, or EH, indicating a ride to a hospital from a residence, scene of accident or acute event, skilled nursing facility (SNF), or non-SNF residential, domiciliary, custodial, or nursing home facility. Providers can bill at the ALS2 level if certain advanced ALS procedures are performed. Medicare pays a single amount for the service level and does not require an itemized list of interventions. Based on conversion factors for 2012, and prior to any adjustments or mileage payments, reimbursement is about $343 for BLS Emergency, $407 for ALS1 Emergency, and $590 for ALS2, which is always considered an emergency.

B.4 Injury Severity Scores

We used the New Injury Severity Score (NISS)[48,65–67] to identify major trauma cases (scores above 15)[47,48] and to adjust for severity differences between ALS and BLS patients. The NISS is the sum of squares of the three highest scores assigned to a patient on the Abbreviated Injury Scale (AIS), which is a widely recognized system for classifying injuries by body region and severity[46]. However,

Table B.1: Differences in characteristics by ambulance service level before adjustments. Differences between BLS and ALS observations were tested for statistical significance using Student's t-test or chi-square test, as appropriate. *a* Chi-squared test of independence was used for this categorical variable. *b* Includes non-SNF residential, domiciliary, custodial, or nursing home facilities. *c* High if median household income > $40,000, low otherwise, and predominantly black if more than 80% black, predominantly white if more than 80% white, and otherwise integrated. *d* Average difference between actual and predicted surgical survival for hospitals, weighted by number of patients transported to hospital within ZIP code. Details are provided in supplementary materials. *e* Metropolitan counties have at least one urbanized area of 50,000 or more population, and micropolitan areas have at least one urban cluster of at least 10,000 but less than 50,000 population. Both types of area have adjacent territory that has a high degree of social and economic integration with the core as measured by commuting ties.

	Trauma		AMI		Stroke		Resp. failure	
	BLS	ALS	BLS	ALS	BLS	ALS	BLS	ALS
N	30,919	48,768	14,434	100,035	19,985	100,004	9,502	73,028
Mean age	82	79***	81	78***	80	79***	75	74***
Female (%)	69	61***	60	53***	63	61***	59	56***
Race (%)		*a*		*a****		*a****		*a****
White	91	91	83	87	79	84	78	81
Black	5	5	12	9	15	12	16	14
Hispanic	2	1	3	2	3	2	3	2
Asian	1	1	2	1	2	1	1	1
Other	1	1	1	1	2	1	2	1
Comorbidity score [mean]	3.3	2.9***	4.2	3.3***	3.6	3.1***	5.9	5.3***
Mean mileage	6.2	7.8***	6.4	7.5***	6.4	7.5***	6.0	6.5***
Pick-up location (%)		*a****		*a****		*a****		*a****
Residence	57	60	61	71	63	71	49	64
Skilled nursing facility	18	10	23	11	22	12	36	19
Scene	18	24	10	14	10	13	9	11
Non-SNF nursing home[b]	6	5	5	4	5	4	6	5
ZIP: Income/race mix[c] (%)		*a****		*a****		*a****		*a****
High/white	51	52	45	49	44	48	39	43
Low/white	7	8	8	9	7	8	9	9
High/black	0	0	1	0	1	0	1	1
Low/black	1	1	2	1	2	2	2	2
High/integrated	30	27	31	27	33	28	33	29
Low/integrated	10	12	13	13	14	14	16	16
ZIP: Hospital quality measure[d]	0.0001	-0.0011***	0.0003	-0.0009***	-0.0001	-0.0011***	-0.0001	-0.0011***
County: Metropolitan[e] (%)	87	84***	87	84***	87	85***	87	85***
County: Any hospital with med schl affiliation (%)	68	61***	69	62***	69	63***	69	63***
County: Any trauma center (%)	72	67***	74	67***	75	68***	74	68***
County: General practice doctors (%)	14	16***	14	16***	14	16***	14	16***
County: Persons with 4+ years of college (%)	25	24***	24	23***	25	23***	24	23***

since direct AIS coding requires access to medical records, we used ICDPIC software to map ICD-9CM discharge diagnosis codes to AIS scores[68,69]. We included up to 30 diagnosis codes for each observation and removed duplicate codes. Scores are assigned to observations with a valid trauma ICD-9CM code (800 to 959.9 excluding 905-909 (late effects of injury), 930-939 (foreign bodies), 958 (complications), and 940-949 (burns)) that has both a known severity and body region in the mapping tables. Since the NISS is neither normally distributed nor continuous, we specified it as a categorical variable using a similar breakdown to other studies (<16, 16 - 24, >25)[47,48,69]. We also included a numerical variable in the regressions to adjust for within-category severity differences. We analyzed the sensitivity of our trauma results to alternative measures of injury severity (see Appendix B.12).

B.5 Hospital Surgical Quality Scores

Quality of hospital care may be correlated with both outcomes and the propensity of a beneficiary to receive pre-hospital ALS, and thus may be a confounder. However, hospital choice may also be part of the ambulance effect. For example, ALS may be able to reach higher quality hospitals that are further away due to field stabilization, or conversely, BLS may be able to reach a higher quality hospital because it prioritized rapid transport. We wished to control for the pre-treatment hospital quality effect (quality of hospitals available to ambulances picking up patients residing in a particular area), but not the post-treatment quality effect (quality of the specific hospital chosen for each patient). To do so, we created measures of the surgical quality for non-ambulance patients of hospitals used by ambulance patients from each ZIP code. We did this by first regressing 30-day survival on age, sex, comorbidity score and fixed effects for surgical Diagnosis-Related Groups (DRGs), using all inpatient claims with a surgical DRG code between 2006 and 2011. We removed all transfers from other hospitals, admissions from the Emergency Department, and beneficiaries overlapping with our ambulance cases from the sample. Therefore, this sample did not overlap with our ambulance sample. Then, we subtracted the modeled survival probability from the binary 30-day survival indicator for each observation. Finally, we averaged these residuals for each hospital, and then averaged the hospital averages over beneficiaries in the same ZIP code, weighted by the number of individuals transported to each hospital from that ZIP code. This weighting was diagnosis specific so that the final measures for a given diagnosis reflected the average hospital quality experienced by ambulance patients of that group. Individuals were matched to the hospital quality measure for their ZIP code. In a small number of cases, an exact match was not found, so measures from the nearest or nearest equidistant ZIP codes (determined by fixing the first three digits and considering ZIP codes within a range of $\pm$ 5 of the last two digits), were used directly or averaged.

B.6 County-level Models

The propensity score weighting approach in the individual-level analysis (Appendix B.8) is useful for balancing observed covariates, but it does not account for unobserved confounders. Therefore, we conducted an instrumental variables (IV) analysis. If the requirements for a valid IV analysis are met, the instruments represent exogenous randomization in the treatment assignment that is uncorrelated (after adjustment) with unobserved covariates. Thus, in theory, an IV analysis can produce causal inferences comparable to randomized controlled trials over the range of support in the data.

Our IV analysis replaced the individual-level ambulance assignment with a county-level probability of ALS use, predicted from measures of ALS use in other, non-overlapping diagnosis groups. The IV analysis was conducted in several steps for each diagnosis group. In the first step, we fitted a multivariate, multilevel model, with county-level random intercepts for each diagnosis group from a multivariate normal distribution:

$$ALS_{ic} \sim Bernoulli(\lambda_{ic})$$

$$\lambda_{ic} = Pr(ALS_{ic}|\alpha, \beta) = logit^{-1} \sum_{d=1}^{4} D_{icd}[\alpha_{cd} + X_{ic}\beta_d]$$

$$\alpha_c = \begin{pmatrix} \alpha_{c1} \\ \alpha_{c2} \\ \alpha_{c3} \\ \alpha_{c4} \end{pmatrix} \sim N(0, \Sigma)$$

$$D_{icd} = \begin{cases} 1 & \text{if case } i \text{ in county } c \text{ has diagnosis } d, \\ 0 & \text{otherwise.} \end{cases}$$

where i, c, and d denote an individual, county, and diagnosis, respectively.

Above, ALS_{ic} is an indicator variable for whether individual i in county c received ALS or BLS and follows a Bernoulli distribution with probability λ_{ic}. To ease computational problems due to the size of the dataset and number of covariates, we first modeled the fixed effects using a logistic regression, $logit\ Pr(ALS_{ic}) = \sum_{d=1}^{4} D_{icd}[X_{ic}^T\beta_d]$. We included the linear predictor, $X_{icd} = X_{ic}^T\beta_d$, from this model as a covariate in the first step. Therefore, X_{ic}, the linear predictor, represents the individual- and county-level covariates that are modeled as fixed effects, including year, age, sex, race, comorbidity score, pickup location, and mileage from scene to hospital at the individual level, and urbanicity (metropolitan/micropolitan), percent over 25 years of age with four or more years of college, percent of primary care practitioners, presence of any medical school-affiliated short-term hospital, and presence of a trauma center at the county level. We also included the six-category measure for racial makeup and household income at the ZIP code level, described in the individual-level models section (Appendix B.8). In short, this model includes county-diagnosis-specific random intercepts and diagnosis-specific fixed effects for each covariate.

We used this fitted nonlinear model to predict the probability of receiving ALS for each individual in the sample from observed ALS use in the county for other diagnoses. We implemented this by duplicating the dataset, trivially modifying the county labels in the copy, removing the target diagnosis ambulance data from the copy, and fitting the model to a dataset that included both the original full data and the duplicated data. This allowed the model to estimate the covariance structure of the random effects, and then this fitted model to make predictions for the diagnosis group of interest (the set with trivially modified county labels whose outcomes were removed from the copy). This procedure guaranteed that the predicted probabilities of ALS-use for the target diagnosis cases were independent of unobserved characteristics of these cases and depended solely on shared resources and policy, and not any unobserved characteristics of the patients.

The next step was to use the predicted probability of ALS from the nonlinear model as the instrument in the two-stage least-squares IV analysis. We implemented the first stage of the two-stage-least-squares analysis by regressing the actual binary indicator of whether an individual received ALS or BLS on the predicted probability of ALS use and other covariates from the model above, using a separate linear probability model for each diagnosis. Using the fitted values of the nonlinear model as the instruments in a linear model addressed potential concerns about inconsistent estimates. Specifically, we compared predicted and empirical ALS rates within deciles and found the logistic model fitted the data significantly better than a linear probability model. Nonetheless, the use of a nonlinear model in the first stage of an IV analysis will not generate consistent estimates in the second stage if the nonlinear model is misspecified[50]. It is possible, however, to generate consistent estimates by feeding predictions from a nonlinear model into a linear first-stage model, which was how we proceeded. The F- statistics for testing the null hypothesis that the instruments are not strong predictors of ALS use were extremely high (F-statistic > 1,000 in every diagnosis group), indicating that the predicted ALS probabilities from the non-linear model were very strong predictors in the linear probability models. Also, the coefficients of the predicted probability of ALS in the linear models were close to 1 (between 0.97 and 1.06), indicating our logistic regression fit the data well.

The two key requirements for a valid IV analysis are that the instruments be good predictors of the treatment, and that there be no causal pathway between the instruments and the final outcome except through the treatment. Tables B.2 and B.3 show both that the random effects were highly correlated across diagnoses (if a county was likely to use ALS for one diagnosis it was likely to use it for others) and that the instruments had good predictive value, given the range of predictions produced. Our sensitivity analysis (Appendix B.15) indicates it is improbable that our findings are biased by county-level confounding.

Table B.2: Correlations of county-level random effects in multilevel model for instruments.

	Trauma	AMI	Stroke
AMI	0.90		
Stroke	0.91	0.97	
Respiratory failure	0.86	0.98	0.96

Table B.3: Distributions of predicted probabilities of ALS from first stage instrumental variables model by diagnosis group.

Diagnosis groups	Min	1st Quartile	Median	Mean	3rd Quartile	Max
Trauma	0.01	0.49	0.64	0.61	0.77	1.01
Stroke	0.00	0.78	0.88	0.83	0.94	1.01
AMI	0.10	0.84	0.92	0.87	0.95	1.00
Respiratory failure	0.11	0.86	0.92	0.88	0.96	1.00

In the next step, the second stage of the two-stage-least-squares procedure, we estimated the effect of the instrumented probability of ALS use, $pALS_i$, on survival by fitting separate linear probability models for each diagnosis group:

$$survival_i \sim \mathcal{N}(\mu_i, \sigma^2)$$

$$\mu_i = Pr(survival_i | \beta, \gamma) = pALS_i\beta + X_i\gamma$$

Above, $survival_i$ is a binary variable for whether individual i survived to a specific number of days. Here, we included all of the variables from the first stage and also added other covariates from the final individual-level models for each diagnosis group of interest (Table B.4). These are denoted by X_i. We interpret β as the average change in survival probability for an individual who would have received ALS in a higher ALS-use county but would have received BLS in a lower-ALS use county. The standard errors in the second stage were adjusted, using the design effect (the ratio of the actual variance under clustering to the variance under simple random sampling, estimated from a single-stage model equivalent to the second stage of 2SLS), to account for the fact that individuals are clustered within counties where they face common ALS penetration rates.

B.7 Mediators

We adjusted for pre-treatment variables that are potential confounders of the treatment-outcome relationship, but not for post-treatment variables, which are potential mediators of treatment effects. For example, we controlled for distance from the scene to the hospital, which is a confounder, but did not study transport time from the scene to the hospital, which is a mediator. Other factors, such as the quality of CPR and the use of endotracheal intubation or specific intravenous drugs, are also mediators of ALS and BLS effects, rather than confounders.

B.8 Individual-level Models

We aimed to estimate the difference between ALS and BLS in an individual's survival probability after a high-acuity medical emergency. However, since individuals were not randomly assigned to ALS and BLS, a direct comparison might produce a biased estimate of this relationship. Therefore, in addition to the instrumental variable analysis described in Appendix B.6, we conducted a propensity score analysis by balancing the distributions of ALS and BLS over observed covariates

Table B.4: Coefficients from linear probability models for 90-day survival from second stage county-level analysis. [a]Includes non-SNF residential, domiciliary, custodial, or nursing home facilities. [b]Alzheimer's disease/dementia includes Alzheimer's, related diseases, and senile dementia. [c]COPD = chronic obstructive pulmonary disease. [d]High if median household income $>$ $40,000, low otherwise, and predominantly black if more than 80% black, predominantly white if more than 80% white, and otherwise integrated. [e] Metropolitan counties have at least one urbanized area of 50,000 or more population, and micropolitan areas have at least one urban cluster of at least 10,000 but less than 50,000 population. Both types of area have adjacent territory that has a high degree of social and economic integration with the core as measured by commuting ties.

Covariate	Coefficient [95% CI]			
	Trauma	Stroke	AMI	Respiratory failure
Intercept	0.804 [0.668, 0.940]	0.667 [0.474, 0.860]	0.977 [0.881, 1.073]	0.835 [0.749, 0.922]
Probability of ALS	-0.041 [-0.069, -0.013]	-0.043 [-0.073, -0.013]	-0.059 [-0.096, -0.022]	-0.002 [-0.051, 0.047]
State fixed effects (not shown)	included	included	included	included
Linear age splines (not shown)	included	included	included	included
Female	0.088 [0.081, 0.095]	-0.005 [-0.010, 0.000]	0.026 [0.021, 0.031]	0.054 [0.046, 0.062]
Race				
White (Ref)	-	-	-	-
Black	0.027 [0.015, 0.040]	0.045 [0.036, 0.053]	0.013 [0.002, 0.023]	0.048 [0.037, 0.059]
Hispanic	0.028 [0.008, 0.047]	0.033 [0.008, 0.058]	0.010 [-0.006, 0.027]	0.038 [0.013, 0.062]
Asian	0.008 [-0.018, 0.034]	0.039 [0.021, 0.057]	0.016 [-0.009, 0.040]	0.011 [-0.017, 0.038]
Other	0.016 [-0.004, 0.037]	0.011 [-0.009, 0.032]	-0.005 [-0.027, 0.017]	0.000 [-0.028, 0.027]
Comorbidity score linear splines (not shown)	included	included	-	-
Comorbidity score (no splines)	-	-	-0.020 [-0.021, -0.019]	-0.020 [-0.021, -0.019]
ST segment elevation	-	-	-0.137 [-0.143, -0.130]	-
New Injury Severity Score (continuous)	0.007 [0.006, 0.009]	-	-	-
New Injury Severity Score (categorical)				
16-24 (Ref)	-	-	-	-
25-40	-0.192 [-0.210, -0.175]	-	-	-
41-49	-0.418 [-0.463, -0.373]	-	-	-
50-75	-0.664 [-0.746, -0.582]	-	-	-
Mileage linear splines (not shown)	-	included	-	-
Mileage (no splines)	0.000 [0.000, 0.001]	-	0.001 [0.001, 0.002]	0.001 [0.000, 0.001]
Pick-up location				
Residence (Ref)	-	-	-	-
Skilled nursing facility	-0.040 [-0.051, -0.028]	-0.099 [-0.109, -0.090]	-0.109 [-0.119, -0.099]	-0.095 [-0.107, -0.084]
Scene	0.030 [0.023, 0.037]	0.024 [0.016, 0.031]	0.024 [0.017, 0.031]	-0.003 [-0.014, 0.009]
Non-SNF nursing home[a]	-0.037 [-0.052, -0.022]	-0.058 [-0.072, -0.043]	-0.084 [-0.099, -0.068]	-0.063 [-0.080, -0.046]
Chronic conditions				
Acute myocardial infarction	-0.023 [-0.034, -0.011]	-0.028 [-0.038, -0.019]	0.019 [0.011, 0.026]	0.026 [0.016, 0.036]
Alzheimer's disease/dementia[b]	-0.026 [-0.033, -0.019]	-0.039 [-0.045, -0.032]	-0.052 [-0.059, -0.044]	-0.016 [-0.024, -0.008]
Atrial fibrillation	-0.049 [-0.057, -0.042]	-0.065 [-0.072, -0.059]	-0.007 [-0.013, 0.000]	
Chronic kidney disease		-0.008 [-0.015, -0.002]		

Table B.4. (Continued) Coefficients from linear probability models for 90-day survival from second stage county-level analysis

Covariate	Coefficient [95% CI]			
	Trauma	Stroke	AMI	Respiratory failure
COPD[c]				0.049 [0.041, 0.056]
Heart failure	-0.013 [-0.020, -0.007]	-0.021 [-0.027, -0.014]	-0.031 [-0.037, -0.025]	0.034 [0.025, 0.042]
Diabetes	0.002 [-0.005, 0.008]	0.008 [0.002, 0.014]	-0.005 [-0.011, 0.001]	0.032 [0.025, 0.039]
Glaucoma		0.008 [0.002, 0.014]		0.006 [-0.003, 0.014]
Hip/pelvic fracture		-0.021 [-0.030, -0.011]	-0.009 [-0.019, 0.001]	
Ischemic heart disease	-0.003 [-0.010, 0.003]	0.017 [0.011, 0.023]	0.021 [0.015, 0.028]	0.021 [0.012, 0.029]
Depression		0.015 [0.010, 0.021]	0.017 [0.011, 0.022]	0.015 [0.007, 0.022]
Osteoperosis		0.011 [0.005, 0.018]		-0.019 [-0.027, -0.010]
Rheumatoid arthritis/osteoarthritis	0.036 [0.030, 0.042]	0.039 [0.034, 0.044]		0.028 [0.021, 0.036]
Stroke/transient ischemic attack	0.004 [-0.002, 0.010]	0.021 [0.016, 0.027]	-0.001 [-0.006, 0.005]	
Breast cancer			0.017 [0.005, 0.030]	
Colorectal cancer		0.005 [-0.009, 0.018]		-0.014 [-0.032, 0.005]
Prostate cancer	0.025 [0.011, 0.040]			-0.004 [-0.020, 0.011]
Lung cancer	-0.089 [-0.113, -0.064]			
Anemia	-0.002 [-0.009, 0.005]	-0.005 [-0.011, 0.001]	-0.004 [-0.009, 0.002]	-0.010 [-0.019, -0.002]
Asthma				0.054 [0.046, 0.062]
Hyperlipidemia	0.026 [0.020, 0.033]		0.051 [0.044, 0.057]	0.020 [0.012, 0.028]
Hypertension			-0.032 [-0.041, -0.023]	
Acquired hypothyroidism			0.010 [0.003, 0.016]	
ZIP: Income/race mix[d]				
High/black (Ref)	-	-	-	-
High/white	0.056 [0.012, 0.101]	0.006 [-0.017, 0.029]	0.017 [-0.034, 0.069]	0.004 [-0.036, 0.043]
Low/white	0.052 [0.006, 0.098]	-0.007 [-0.031, 0.018]	0.009 [-0.043, 0.060]	0.006 [-0.035, 0.047]
Low/black	0.024 [-0.021, 0.069]	-0.012 [-0.038, 0.014]	0.033 [-0.019, 0.085]	0.006 [-0.045, 0.056]
High/integrated	0.051 [0.006, 0.096]	0.005 [-0.017, 0.028]	0.009 [-0.041, 0.060]	0.007 [-0.034, 0.048]
Low/integrated	0.050 [0.005, 0.095]	0.000 [-0.023, 0.022]	0.015 [-0.035, 0.065]	0.003 [-0.038, 0.044]
ZIP: % Female	0.000 [-0.001, 0.001]	0.000 [-0.001, 0.002]	0.001 [0.000, 0.002]	
ZIP: % Over 65 years	0.000 [0.000, 0.000]	0.000 [-0.001, 0.000]	0.000 [0.000, 0.000]	
ZIP: Hospital quality score linear splines (not shown)	-	included	-	-
ZIP: Hospital quality score (no splines)	0.248 [-0.046, 0.543]	-	0.371 [0.139, 0.604]	0.489 [0.219, 0.759]
County: Metropolitan[e]	-0.003 [-0.012, 0.006]	0.012 [0.004, 0.021]	0.007 [-0.001, 0.015]	0.014 [0.002, 0.026]
County: Any hospital with med schl affiliation	0.004 [-0.004, 0.012]	0.010 [0.003, 0.018]	0.004 [-0.003, 0.011]	0.002 [-0.009, 0.013]
County: Any trauma center	-0.002 [-0.010, 0.005]	0.001 [-0.006, 0.008]	0.002 [-0.005, 0.010]	0.002 [-0.010, 0.013]
County: General practice doctors	0.000 [0.000, 0.001]	0.000 [-0.001, 0.000]	0.000 [0.000, 0.000]	0.000 [0.000, 0.001]
County: Persons with 4+ years of college	0.001 [0.000, 0.001]	0.000 [0.000, 0.001]	0.001 [0.001, 0.001]	0.000 [0.000, 0.001]
Year				
2006 (Ref)	-	-	-	-
2007	0.002 [-0.008, 0.011]	0.002 [-0.006, 0.011]	-0.001 [-0.009, 0.006]	0.009 [-0.001, 0.019]
2008	0.001 [-0.009, 0.011]	0.008 [0.000, 0.016]	0.004 [-0.004, 0.013]	0.020 [0.010, 0.031]
2009	0.010 [0.000, 0.019]	0.016 [0.008, 0.024]	0.008 [-0.001, 0.017]	0.016 [0.006, 0.027]
2010	0.012 [0.002, 0.022]	0.016 [0.008, 0.023]	0.008 [0.000, 0.016]	0.013 [0.003, 0.023]
2011	0.015 [0.005, 0.026]	0.022 [0.013, 0.031]	0.014 [0.006, 0.023]	0.020 [0.007, 0.032]

by weighting to remove the confounding of ambulance assignment with potential outcomes[27]. More specifically, our approach applied a balancing weight to each observation based on a modeled propensity to receive ALS, π_i. Each ALS observation i was weighted by $1 - \pi_i$ and each BLS observation was weighted by π_i. Thus, conditional on π_i, the assignment to ambulance type is random[27]. A key benefit of this approach is that it separates the study design from data analysis, since the weights are derived without knowing the effects on outcomes. Since the distributions of ALS and BLS over the covariates are similar after weighting, it does not require us to drop observations to achieve balance in the observed variables.

To implement this procedure for each diagnosis group, we estimated the probability that a beneficiary received ALS by fitting a logistic regression model:

$$ALS_i \sim Bernoulli(\pi_i)$$
$$\pi_i = Pr(ALS_i|\beta) = logit^{-1}(X_i\beta).$$

ALS_i is an indicator variable for whether individual i received ALS or BLS. Covariates for individual i are denoted by X_i. We used a structured approach to build the model for each diagnosis group by testing groups of covariates with likelihood ratio tests and including those that were jointly significant at the 5% level. At the individual level, these variables included age, sex, race, pickup location type (e.g., residence, scene), mileage from pickup location to hospital, comorbidity score, and indicator variables for 27 chronic conditions. Additionally, for trauma, we included New Injury Severity Scores, described above, and for AMI, an indicator for ST- segment elevation (ST if 410.71; non-ST if 410.x1 other than 410.71). We also tested state fixed effects. At the county level, we tested for inclusion urbanicity (metropolitan/micropolitan), percent over 25 years of age with four or more years of college, percent of primary care practitioners, and the presence of any medical school-affiliated short-term hospital. The ZIP code level covariates tested for inclusion were the hospital surgical quality scores described above, percent of individuals over 65, and percent of females. We also tested a six-category variable at the ZIP code level combining high (>$40,000) and low median household income and racial composition (>80% black, >80% white, or integrated), which have been shown to be important determinants of bystander-initiated CPR[28]. Some of the continuous variables, such as age, were specified as linear splines to allow more flexibility in model fit. Final model fits are shown in Table B.5.

We tested the differences in weighted mean outcomes between ALS and BLS using t tests, as described in the main manuscript. A possible concern with this approach is that there may be insufficient overlap in the propensity scores between ALS and BLS leading to an overreliance on untestable functional form assumptions. However, the propensity score distributions were similar for BLS and ALS cases (Table B.6).

The key assumption of this analysis is that there are no unobserved covariates that are correlated with both ambulance treatment assignment and survival, so the variation remaining after adjustments is random and exogenous to health and healthcare. Because we believe this assumption is satisfied, we interpret the adjusted difference in survival as the effect of receiving ALS versus BLS.

B.9 Sensitivity Analysis: Death in the Field

We excluded patients with only an ambulance claim, and therefore individuals who died at the scene. If patients receiving BLS are more likely to die at the scene, our sample may be biased. However, for two key reasons, it is not possible to use the Medicare claims data to assess the sensitivity of our results to this exclusion. First, in cases where an individual is treated at the scene but pronounced dead before being loaded into the truck, both ALS and BLS providers are paid at the BLS level and therefore usually bill at this level. Second, these observations have only ambulance diagnosis coding, which is unlikely to be accurate in general, but even more so in cases where there was little time to observe the patient.

Table B.5: Coefficients from individual-level logistic regression models for predicting the propensity to receive ALS (log-odds ratios are shown). [a]Includes non-SNF residential, domiciliary, custodial, or nursing home facilities. [b]Alzheimer's disease/dementia includes Alzheimer's, related diseases, and senile dementia. [c]COPD = chronic obstructive pulmonary disease. [d]High if median household income > $40,000, low otherwise, and predominantly black if more than 80% black, predominantly white if more than 80% white, and otherwise integrated. [e] Metropolitan counties have at least one urbanized area of 50,000 or more population, and micropolitan areas have at least one urban cluster of at least 10,000 but less than 50,000 population. Both types of area have adjacent territory that has a high degree of social and economic integration with the core as measured by commuting ties.

Covariate	Coefficient [95% CI]			
	Trauma	Stroke	AMI	Respiratory failure
Intercept	2.694 [1.541, 3.935]	-0.176 [-1.416, 1.079]	2.644 [1.972, 3.326]	2.442 [1.958, 2.933]
State fixed effects (not shown)	included	included	included	included
Linear age splines (not shown)	included	included	included	included
Female	-0.215 [-0.252, -0.178]	0.005 [-0.032, 0.041]	-0.135 [-0.175, -0.095]	-0.006 [-0.057, 0.045]
Race				
White (Ref)	-	-	-	-
Black	-0.008 [-0.088, 0.072]	-0.113 [-0.169, -0.057]	-0.260 [-0.330, -0.189]	-0.064 [-0.137, 0.010]
Hispanic	-0.159 [-0.285, -0.032]	-0.382 [-0.490, -0.273]	-0.285 [-0.409, -0.158]	-0.288 [-0.426, -0.146]
Asian	-0.120 [-0.265, 0.025]	-0.166 [-0.299, -0.031]	-0.274 [-0.427, -0.117]	-0.067 [-0.254, 0.126]
Other	-0.121 [-0.260, 0.019]	-0.307 [-0.439, -0.173]	-0.137 [-0.290, 0.020]	-0.219 [-0.387, -0.046]
Comorbidity score linear splines (not shown)	included	included	-	-
Comorbidity score (no splines)			-0.023 [-0.030, -0.016]	-0.008 [-0.016, -0.001]
ST segment elevation	-	-	0.262 [0.220, 0.304]	-
New Injury Severity Score (continuous)	-0.015 [-0.022, -0.007]	-	-	-
New Injury Severity Score (categorical)				
16-24 (Ref)	-	-	-	-
25-40	0.651 [0.558, 0.743]	-	-	-
41-49	1.184 [0.956, 1.414]	-	-	-
50-75	1.709 [1.267, 2.157]	-	-	-
Mileage linear splines (not shown)	-	included	-	-
Mileage (no splines)	0.019 [0.017, 0.022]		0.002 [-0.001, 0.005]	-0.006 [-0.009, -0.003]
Pick-up location				
Residence (Ref)	-	-	-	-
Skilled nursing facility	-0.513 [-0.562, -0.464]	-0.660 [-0.706, -0.613]	-0.620 [-0.673, -0.567]	-0.827 [-0.882, -0.771]
Scene	0.105 [0.064, 0.146]	0.241 [0.187, 0.295]	0.289 [0.227, 0.350]	0.168 [0.089, 0.249]
Non-SNF nursing home[a]	-0.283 [-0.352, -0.214]	-0.387 [-0.464, -0.308]	-0.390 [-0.478, -0.301]	-0.496 [-0.597, -0.393]
Chronic conditions				
Acute myocardial infarction	0.059 [-0.002, 0.120]	0.061 [0.004, 0.119]	0.132 [0.084, 0.180]	0.183 [0.113, 0.254]
Alzheimer's disease/dementia[b]	-0.104 [-0.141, -0.068]	-0.085 [-0.124, -0.046]	-0.172 [-0.218, -0.127]	-0.074 [-0.127, -0.020]
Atrial fibrillation	0.101 [0.062, 0.141]	0.253 [0.214, 0.291]	0.074 [0.029, 0.119]	
Chronic kidney disease		0.026 [-0.014, 0.065]		

Table B.5 (Continued) Coefficients from individual-level logistic regression models for predicting the propensity to receive ALS (log-odds ratios are shown)

Covariate	Coefficient [95% CI]			
	Trauma	Stroke	AMI	Respiratory failure
COPD[c]				0.049 [-0.005, 0.103]
Heart failure	-0.039 [-0.077, 0.000]	0.026 [-0.014, 0.067]	0.022 [-0.026, 0.070]	-0.036 [-0.096, 0.025]
Diabetes	-0.047 [-0.081, -0.012]	-0.038 [-0.073, -0.004]	-0.014 [-0.054, 0.027]	-0.011 [-0.061, 0.038]
Glaucoma		-0.076 [-0.113, -0.038]		-0.052 [-0.107, 0.003]
Hip/pelvic fracture		-0.013 [-0.072, 0.047]	-0.118 [-0.181, -0.053]	
Ischemic heart disease	0.022 [-0.016, 0.061]	0.009 [-0.032, 0.050]	0.076 [0.023, 0.129]	0.027 [-0.034, 0.088]
Depression		-0.049 [-0.085, -0.013]	-0.078 [-0.119, -0.037]	-0.080 [-0.129, -0.031]
Osteoperosis		0.020 [-0.021, 0.061]		0.009 [-0.047, 0.065]
Rheumatoid arthritis/osteoarthritis	-0.041 [-0.075, -0.007]	-0.046 [-0.083, -0.010]		-0.067 [-0.117, -0.017]
Stroke/transient ischemic attack	0.024 [-0.012, 0.060]	-0.024 [-0.058, 0.011]	-0.052 [-0.094, -0.010]	
Breast cancer			0.077 [-0.008, 0.163]	
Colorectal cancer		-0.099 [-0.177, -0.020]		-0.040 [-0.150, 0.072]
Prostate cancer	-0.046 [-0.118, 0.026]			0.032 [-0.075, 0.140]
Lung cancer	-0.079 [-0.190, 0.034]			
Anemia	-0.080 [-0.118, -0.042]	-0.094 [-0.135, -0.054]	-0.097 [-0.145, -0.048]	-0.077 [-0.139, -0.015]
Asthma				-0.083 [-0.135, -0.031]
Hyperlipidemia	0.076 [0.039, 0.113]		0.112 [0.063, 0.161]	0.026 [-0.027, 0.079]
Hypertension			-0.069 [-0.147, 0.008]	
Acquired hypothyroidism			0.044 [-0.001, 0.088]	
ZIP: Income/race mix[d]				
High/black (Ref)	-	-	-	-
High/white	0.796 [0.508, 1.086]	0.653 [0.459, 0.844]	0.885 [0.664, 1.102]	0.606 [0.350, 0.854]
Low/white	0.773 [0.480, 1.070]	0.633 [0.429, 0.834]	0.722 [0.491, 0.949]	0.383 [0.116, 0.642]
Low/black	0.318 [-0.010, 0.649]	0.225 [0.006, 0.442]	0.439 [0.189, 0.687]	0.009 [-0.277, 0.289]
High/integrated	0.621 [0.334, 0.911]	0.352 [0.161, 0.541]	0.588 [0.369, 0.803]	0.302 [0.049, 0.547]
Low/integrated	0.632 [0.343, 0.923]	0.308 [0.115, 0.499]	0.585 [0.364, 0.802]	0.227 [-0.028, 0.474]
ZIP: % Female	0.002 [-0.005, 0.010]	-0.010 [-0.019, -0.002]	0.001 [-0.008, 0.010]	
ZIP: % Over 65 years	-0.004 [-0.007, -0.001]	-0.003 [-0.006, 0.000]	-0.004 [-0.007, -0.001]	
ZIP: Hospital quality score linear splines (not shown)	-	included	-	-
ZIP: Hospital quality score (no splines)	-3.293 [-5.038, -1.569]		-2.791 [-4.713, -0.935]	-2.318 [-4.349, -0.385]
County: Metropolitan[e]	0.029 [-0.023, 0.082]	0.024 [-0.031, 0.079]	0.101 [0.038, 0.164]	0.184 [0.107, 0.260]
County: Any hospital with med schl affiliation	-0.035 [-0.079, 0.008]	-0.072 [-0.114, -0.031]	-0.054 [-0.105, -0.004]	-0.070 [-0.131, -0.009]
County: Any trauma center	-0.172 [-0.215, -0.129]			
County: General practice doctors	-0.004 [-0.006, -0.002]		-0.002 [-0.004, 0.000]	0.000 [-0.003, 0.002]
County: Persons with 4+ years of college	0.005 [0.002, 0.007]		0.004 [0.001, 0.006]	0.002 [-0.001, 0.005]

Table B.6: Propensity score distributions from individual-level analysis for ALS and BLS by diagnosis group.

Diagnosis groups	Level	Min	1st Quartile	Median	Mean	3rd Quartile	Max
Trauma	ALS	0.08	0.58	0.69	0.67	0.78	1.00
	BLS	0.05	0.38	0.55	0.52	0.67	0.98
Stroke	ALS	0.17	0.82	0.89	0.85	0.92	0.99
	BLS	0.09	0.63	0.79	0.74	0.88	0.98
AMI	ALS	0.21	0.86	0.91	0.89	0.94	1.00
	BLS	0.17	0.72	0.83	0.79	0.90	0.99
Respiratory failure	ALS	0.36	0.86	0.92	0.89	0.94	0.99
	BLS	0.31	0.76	0.85	0.82	0.91	0.99

Therefore, we used data sources other than the claims to estimate how deaths in the field may have affected our estimates. While these datasets likely differ in key ways from the Medicare sample, these approximate calculations provide reassurance. For cardiac arrest, we used data from the Resuscitation Outcomes Consortium (ROC)[30] and the Cardiac Arrest Registry to Enhance Survival (CARES)[29] to estimate the proportion of cases who died at the scene among those patients for whom resuscitation was attempted by the ambulance crew. This analysis is discussed in Appendix A.5 and the Supplementary Materials of our JAMA Internal Medicine article[45], and we concluded that accounting for deaths at the scene would not change the direction of our observed effect in the individual-level analysis.

For other diagnosis groups, we used tabulations generated by the Data Cube of the National EMS Information System, which collates data from participating local and state agencies[70]. Specifically, we used the fields "EMS Primary Impression", which provides diagnosis information, and "Incident Patient Disposition", which indicates whether a patient was transported or died at the scene. In the case of death at the scene, it is not possible to tell whether resuscitation had been attempted. For each diagnosis group, we estimated the proportion of patients who died at the scene, and applied this to the Medicare sample. Then we estimated the BLS/ALS distributions that would be required among cases that died in the field in order to eliminate our observed effects in the individual-level analyses. Tables B.7a through B.7c walk through this process. We were not able to find a suitable diagnosis group in the NEMSIS data for AMIs, but if an AMI did result in cardiac arrest, it would be captured in our cardiac arrest analysis.

Though the NEMSIS data are not a random sample, the proportion of individuals with cardiac arrest who died at the scene (35%) is consistent with estimates from the ROC (37%) and CARES (22%). Therefore, the estimates of the proportion of patients who died at the scene in other diagnosis groups may also be reasonable. Table B.7c shows that even if all of the estimated deaths in field occurred among BLS patients, the overall raw ALS mortality rate for stroke and trauma patients would still be higher than the BLS raw mortality rate. In the case of respiratory arrest, about 21% of the estimated field deaths would have to have occurred among BLS cases to remove the observed effect. This is almost twice the proportion of BLS respiratory failure cases in the sample. Further, the BLS mortality rate in the field would have to be 16%, twice the ALS mortality rate in the field (8%). Therefore, we do not believe accounting for deaths in the field would change the direction of our estimates.

Table B.7a: Estimated additional field deaths in Medicare sample using NEMSIS data.

NEMSIS: Diagnosis	NEMSIS: Dead at scene	NEMSIS: Treated, transported by EMS	NEMSIS: % dead at scene	Medicare sample size	Additional estimated Medicare deaths in field
Stroke / CVA	232	487, 222	0.05	119,989	57
Traumatic Injury	12,597	3,695,228	0.34	79,687	271
Respiratory arrest	1,823	47,924	3.7	82,530	3,024

B.10 Sensitivity Analysis: Death en route to Hospital

Since ambulance diagnosis coding is generally of poor quality, we did not use it in our main analysis. This may have excluded some beneficiaries who died prior to arrival at a hospital and thus do not

Table B.7b: BLS/ALS distributions in Medicare sample

Diagnosis group	BLS Sample	ALS Sample	BLS: Deaths by 30 days	ALS: Deaths by 30 days
Stroke	19,985 (17%)	100,004 (83%)	3,116	20,666
Trauma	30,919 (39%)	48,768 (61%)	3,339	8,222
Respiratory arrest	9,502 (12%)	73,028 (88%)	3,196	26,086

Table B.7c: BLS/ALS distributions required among additional field deaths to remove observed effects

Diagnosis group	Estimated deaths in field	BLS: Field deaths required	ALS: Field deaths required	BLS: Mortality rate(%)	ALS: Mortality rate (%)
Stroke	57	57 (100%)	0 (0%)	16	21
Trauma	271	271 (100%)	0 (0%)	12	17
Respiratory arrest	3,024	625 (21%)	2,399 (79%)	38	38

have hospital claims. The previous sensitivity analysis addresses situations in which patients died at the scene. Since those cases have special billing rules, we provided approximate calculations based on other data sources. However, if a patient died en route, it is possible to use the Medicare data to conduct a sensitivity analysis.

If a patient was loaded but died en route, we would expect the patient to have a death date on the same day as the ride, or at most, on the day after the ride. In this analysis we checked the sensitivity of our main findings to the inclusion of transported cases without a hospital claim who died within a day after the transport. The key limitation of this analysis is that it depends heavily on ambulance coding. We applied the same propensity score model specification and weighting approach as in our main analysis (we dropped injury severity scores from the trauma model because these were difficult to compute for observations with only ambulance rides), and estimated survival to 90 days for each diagnosis group (Table B.8). AMI is not shown below because no BLS and 13 ALS observations were identified as having died en route, which is too few to affect the main findings. Overall, the direction and significance of our main findings did not change.

Table B.8: Survival outcomes after including cases that died en route to hospital.

	Additional BLS cases died en route	Additional ALS cases died en route	Overall difference (BLS - ALS) in 90-day survival [95% CI]
Trauma	113	368	7.1 [6.4, 7.7]
Stroke	4	48	6.5 [5.8, 7.2]
Respiratory failure	25	102	3.5 [2.4, 4.6]

B.11 Sensitivity Analysis: Out-of-hospital vs. In-hospital Event

We used principal diagnosis codes to identify stroke, AMI, and respiratory failure observations for inclusion in our sample, since these are more likely to be accurate than admitting diagnosis codes (also known as "reason for patient visit" codes). However, this raises the possibility that patients experienced the acute event in-hospital rather than out-of-hospital. Furthermore, if in-hospital events are more likely to have better outcomes, this would introduce selection bias in our individual-level analysis if this occurred differentially among BLS or ALS cases. However, this scenario is highly unlikely given that the prodromal symptoms of the conditions under study would generate an ALS dispatch if it is available. Further, our county-level analysis is not subject to this type of unobserved selection bias.

Nonetheless, as an additional check, we carried out an additional analysis in which we limited our samples to only patients with admitting or "reason for visit" codes for stroke, AMI, and respiratory failure. In 2010 and 2011, reason for visit codes in outpatient claims were newly created fields, which are incomplete for most of our observations, and we therefore limit this analysis to ambulance transports in 2006 - 2009. Table B.9 shows differences in 90-day survival between BLS and ALS based on the same methodology used in the individual-level analyses. The direction and significance of our main findings did not change.

Table B.9: Survival outcomes after limiting samples to patients with admitting or "reason for visit" codes for the relevant diagnoses.

	N BLS	N ALS	BLS 90-day survival	ALS 90-day survival	Difference [95% CI]
Stroke	6,851	35,635	75.2	68.1	7.1 [5.9, 8.3]
AMI	3,426	23,160	70.0	69.4	0.6 [-1.2, 2.3]
Respiratory failure	2,839	24,112	56.5	53.0	3.4 [1.4, 5.4]

B.12 Sensitivity Analysis: Injury Severity Scores

We analyzed the sensitivity of our trauma results to common alternative specifications of injury severity based on different transformations of the ICD-mapped Abbreviated Injury Scale (AIS) scores[46]. These included: the Injury Severity Score (ISS), which is the sum of the squares of the highest AIS scores in the three most severely injured body regions; the maximum overall AIS score; and the modified Anatomic Profile Score (APS), which is the sum of the maximum AIS and for each of three body regions, the square root of the sum of the squares of all AIS scores for all serious injuries (AIS above 3). We created these measures using ICDPIC software[68].

We used the same methodological approaches as in our main analysis, and only replaced the NISS with the alternate specification of injury severity. Specifically, the ISS scores were simultaneously specified categorically and numerically, similar to the NISS in the main analysis, the Max AIS scores were specified categorically and continuously in separate sensitivity analyses, and the APS was specified continuously. For consistency, we used the same sample as in our main analysis, which included observations with NISS scores indicating major trauma. Table B.10 demonstrates our results are not sensitive to alternative specifications of injury severity.

Table B.10: Marginal difference in 90-day survival from receiving BLS instead of ALS by alternative specifications of injury severity.

	Marginal difference in 90-day survival from receiving BLS instead of ALS [95% CI]
Individual-level analysis	
ISS	4.6 [4.0, 5.2]
Max AIS (Numerical)	4.8 [4.2, 5.4]
Max AIS (Categorical)	4.6 [4.0, 5.3]
APS	4.4 [3.8, 5.0]
County-level analysis	
ISS	3.8 [1.2, 6.5]
Max AIS (Numerical)	3.7 [1.0, 6.3]
Max AIS (Categorical)	3.4 [0.8, 6.0]
APS	4.0 [1.4, 6.6]

B.13 Sensitivity Analysis: Identifying Respiratory Failure using Primary and Secondary Diagnoses

The location of acute respiratory failure in the sequence of diagnosis codes may be inconsistent. The underlying cause of illness may be recorded first, with respiratory failure recorded as the secondary diagnosis. However, after the first few diagnosis codes, respiratory failure may be included to record patient history, among other reasons. In this analysis, we repeat the individual and county-level analyses for respiratory failure for a sample that includes primary and secondary diagnosis codes rather than just the primary code.

Our sample included 28,970 BLS and 168,018 ALS observations. The individual-level results were statistically significant and in the same direction as our main results (Table B.11). Our county-level results were also similar to our main analysis. However, with the larger sample size, at 1 year and 2 years confidence intervals are tighter.

Table B.11: Marginal difference in survival from receiving BLS instead of ALS for respiratory failure sample based on primary and secondary diagnosis codes.

	Marginal difference in survival from receiving BLS instead of ALS [95% CI]
Individual-level analysis	
Survival to 30 days	3.0 [2.3, 3.6]
Survival to 90 days	2.0 [1.4, 2.7]
Survival to 1 year	1.7 [1.0, 2.4]
Survival to 2 years	0.8 [0.06, 1.6]
County-level analysis	
Survival to 30 days	3.0 [-0.4, 6.3]
Survival to 90 days	0.5 [-2.9, 4.0]
Survival to 1 year	-0.3 [-3.6, 3.0]
Survival to 2 years	-0.1 [-3.3, 3.0]

B.14 Sensitivity Analysis: ALS billing at the BLS Level

If ALS-trained providers bill at the BLS level for lower acuity patients, BLS would be spuriously associated with better outcomes, and the estimates in the individual-level analysis would be biased. We have addressed this possibility in part with our interviews of EMS officials in 45 states, which confirmed that for our non-trauma observations, BLS would only be dispatched if ALS is unavailable. (As noted in the text, we did not ask about trauma.) We have further argued that since Medicare billing rules allow ALS to bill at the ALS level if assessment by an ALS-trained crew was considered necessary at dispatch, it is unlikely that ALS would bill at the BLS level given the differences in reimbursement rates. Nonetheless, this may still be a concern. Here, we make use of a special HCPCS code in 2005 to estimate the occurrence of ALS billing at the BLS level and its impact on our estimates.

During the implementation of the National Ambulance Fee Schedule in 2002-2005, Medicare blended payments to providers based on reasonable charges and the new Fee Schedule. During this period Medicare allowed ALS-level crews to bill at the ALS level even if the transport did not meet the Fee Schedule rules for the reasonable charge portion of the payment, i.e. that ALS assessment be considered necessary at dispatch or ALS-level treatment be provided. Providers were required to submit a special HCPCS code for these situations, Q3019 ("ALS Vehicle Used, Emergency Transport, No ALS Service Furnished"). Therefore, by identifying observations that billed using code Q3019, we can estimate the occurrence of ALS-trained crews billing at the BLS level.

We identified observations in 2005 that met the key criteria for inclusion in our diagnosis group samples in the main analysis. We included cardiac arrest for comparison purposes. Table B.12 shows that 1-3% of non-trauma cases were cared for by ALS crews that billed at the BLS level in 2005. In trauma cases, this happened in about 13% of cases. However, there is relatively little difference in survival between the ALS groups that billed at the BLS level (ALS-BLS) and the other two groups (ALS-ALS and BLS), and the directions of the survival differences are not consistent across groups. Moreover, when we added the ALS-BLS cases to both the ALS-ALS and BLS groups we found the overall survival for ALS and BLS to change only slightly. The final rows of the table show that estimates of the survival differences between ALS and BLS are similar under both assumptions, i.e., that the ALS-BLS cases were in subsequent years billed at the ALS level or alternatively, at the BLS level.

In sum, the occurrence of ALS billing at the BLS level is low for all diagnosis groups, except trauma. More importantly, the cases where ALS billed for BLS do not exhibit survival patterns that suggest these patients were systematically better off. Finally, there is little difference in the overall estimates of survival differences between ALS and BLS whether the ALS-BLS cases are included in either the ALS or BLS groups. Therefore, we do not believe our individual-level analysis is sensitive to this potential bias. Our county-level analysis is not subject to confounding based on unobserved patient characteristics, and therefore, this potential issue does not apply to that analysis.

Table B.12: Effect of ALS billing at the BLS level on survival in 2005 claims by diagnosis group. ALS-ALS refers to ALS-trained crews billing at the ALS level, ALS-BLS refers to ALS-trained crews billing at the BLS level.

	Stroke	Trauma	Respiratory failure	AMI
Sample sizes				
ALS-ALS	16,182	6,202	12,483	17,354
ALS-BLS	577	913	211	391
BLS	3,171	3,767	1,532	2,288
Survival to 30 days (%)				
ALS-ALS	78.9	82.2	64.2	78.7
ALS-BLS	86.0	89.2	60.2	75.2
BLS	82.9	89.0	63.1	76.4
If ALS-BLS analyzed as ALS	79.1	83.1	64.1	78.6
If ALS-BLS analyzed as BLS	83.4	89.0	62.8	76.3
Estimates of survival differences between ALS and BLS (%)				
If ALS-BLS analyzed as ALS	-3.8	-5.9	1.0	2.1
If ALS-BLS analyzed as BLS	-4.6	-6.9	1.4	2.4
Difference in estimates	0.7	0.9	-0.4	-0.3

B.15 Sensitivity Analysis: Falsification test for county-level instrumental variables analysis

Though the instrumental variables analysis is plausibly immune to confounding from unobserved individual characteristics, it is vulnerable to unobserved variation at the county level that is related to both ALS penetration rates and outcomes. Specifically, the exclusion restriction of instrumental variables analysis requires that after adjusting for observed county-level covariates, there is no pathway between the rates of ALS use in other diagnosis groups and survival in the diagnosis group of interest other than that through the association of both with ALS use in the focal diagnosis. To address this issue we constructed a falsification test to assess the validity of the necessary assumptions. Our test applied the IV analysis approach to an outcome that should *not* be affected by the predicted likelihood of ALS for pre-hospital care. If this analysis were to find a significant effect of ALS on such an outcome, it would imply that the IV analysis is biased due to unmeasured confounding or some other statistical artifact. Conversely, if there were a lack of an effect, our confidence in the validity of the assumptions of the main analysis would increase.

We believe hospital treatment quality is the single most important area-level variable that could be potentially correlated with both overall ALS use and survival, and therefore, the largest source of concern for confounding at the county level. This is because we have no direct measures of the status of patients at the moment when the ambulances dropped them at the hospital; our endpoints all combine the effects of pre-hospital treatment with those of a hospital stay of some duration. Therefore, our falsification test applied the IV analysis approach to outcomes for surgery patients who did *not* arrive at the hospital by ambulance and who therefore should not be affected by the rates of ALS use. More specifically, we used the surgical cases from the hospital quality measure that were inpatient cases and excluded admissions through the Emergency Department and any cases appearing in the ambulance samples.

To conduct this analysis, we generated county-level estimates of ALS use using an approach similar to the step for creating the instrumental variables in the main analysis. We made two key changes to this model. First, since our interest is in county-level variation, we used a standardized sample at the individual level (i.e., we used a fixed set of individual characteristics) to predict ALS use, which allowed us to remove individual-level variation between counties. Second, we specified a single set of county-level fixed and random effects, rather than diagnosis-specific effects, which were not relevant to prediction for non-ambulance diagnoses and in any case are highly correlated.

Next, we created a county-level 30-day survival outcome measure for surgery patients using the same approach as in creating our hospital surgical quality measure (Appendix B.5), but aggregated to the county rather than the ZIP code level. We weighted the averaged hospital scores for a county by the number of surgery patients that attended each hospital from the county, rather than the number of ambulance patients that were transported to each hospital.

We regressed the county-level hospital quality measure based on survival among surgery patients on predicted county-level ALS probabilities in a linear model, adjusted for state fixed effects and the county-level covariates used in our main analysis, but excluding surgical quality. The coefficient of county-level ALS use, which is the main predictor of interest, was exceedingly small (-0.0001) and insignificant ($p = 0.79$). This negative finding gave us additional confidence in the validity of our main analysis.

B.16 Cardiac Arrest

We compared ALS and BLS after out-of-hospital non-traumatic cardiac arrest in an earlier study[45], which used a methodological approach similar to the individual-level analysis. We found survival was higher with BLS than ALS to 30 days by 3.4 [1.9, 4.8] percentage points, to 90 days by 2.6 [1.2, 4.0] percentage points, to 1 year by 1.8 [0.4, 3.3] percentage points, and to 2 years by 2.9 [0.8, 5.0] percentage points. Patients who received ALS were also 3.5 [2.2, 4.8] percentage points more likely to have poor neurological performance. Here, we show the results from applying the instrumental variables approach to cardiac arrest. We only used observations from 2009 to 2011 (because of coding differences before 2009), and removed overlapping cases from the other diagnosis groups. Otherwise, we followed the same procedure as for the other diagnosis groups.

Survival to 30 days was higher with BLS by 4.0 [-1.0, 9.0] percentage points. Survival was also higher with BLS at 90 days by 2.6 [-2.6, 7.9] percentage points, at 1 year by 1.5 [-2.5, 5.5] percentage points, and at 2 years by 3.5 [-2.2, 9.2] percentage points. None of the survival differences are statistically significant, perhaps because of the greater sample size requirements of the county-level instrumental variables analysis compared to the individual-level analysis. However, the directions and magnitudes of estimated survival effects are similar to the individual-level results from our earlier analysis. Patients who received ALS were 8.7 [2.3, 15.2] percentage points more likely to have poor neurological performance.

References

[1] Roger VL, Go AS, Lloyd-Jones DM, et al. Heart disease and stroke statistics - 2012 update: A report from the American Heart Association. *Circulation*, 125(1):e2–e220, 2012.

[2] Hallstrom AP, Ornato JP, Weisfeldt M, et al. Public-access defibrillation and survival after out-of-hospital cardiac arrest. *N Engl J Med*, 351(7):637, 2004.

[3] Stiell IG, Wells GA, Field BJ III, et al. Improved out-of-hospital cardiac arrest survival through the inexpensive optimization of an existing defibrillation program. *JAMA*, 281(13):1175–1181, 1999.

[4] Bernard SA, Gray TW, Buist MD, et al. Treatment of comatose survivors of out-of-hospital cardiac arrest with induced hypothermia. *N Engl J Med*, 346(8):557–563, 2002.

[5] Mild therapeutic hypothermia to improve the neurologic outcome after cardiac arrest. *N Engl J Med*, 346(8):549–556, 2002. PMID: 11856793.

[6] Knafelj R, Radsel P, Ploj T, and Noc M. Primary percutaneous coronary intervention and mild induced hypothermia in comatose survivors of ventricular fibrillation with ST-elevation acute myocardial infarction. *Resuscitation*, 74(2):227–234, 2007.

[7] Sunde K, Pytte M, Jacobsen D, et al. Implementation of a standardised treatment protocol for post resuscitation care after out-of-hospital cardiac arrest. *Resuscitation*, 73(1):29–39, 2007.

[8] Travers AH, Rea TD, Bobrow BJ, et al. Part 4: CPR overview 2010 American Heart Association guidelines for cardiopulmonary resuscitation and emergency cardiovascular care. *Circulation*, 122(18 suppl 3):S676–S684, 2010.

[9] National EMS Scope of Practice Model. National Highway Traffic Safety Administration, 2007. (Accessed January 7, 2014 at http://www.ems.gov/education/EMSScope.pdf).

[10] Stiell IG, Wells GA, Field B, et al. Advanced cardiac life support in out-of-hospital cardiac arrest. *New Engl J Med*, 351(7):647–656, 2004.

[11] MedPAC paymentbasics. Ambulance services payment system, 2012. Accessed April 22, 2015 at http://www.medpac.gov/documents/payment-basics/ambulance-services-payment-system-14.pdf?sfvrsn=0.

[12] Isenberg D and Bissell R. Does advanced life support provide benefits to patients?: A literature review. *Prehospital and Disaster Medicine*, 20(4):265, 2005.

[13] Bur A, Kittler H, Sterz F, et al. Effects of bystander first aid, defibrillation and advanced life support on neurologic outcome and hospital costs in patients after ventricular fibrillation cardiac arrest. *Intensive Care Medicine*, 27(9):1474–1480, 2001.

[14] Ma MHM, Chiang WC, Ko PCI, et al. Outcomes from out-of-hospital cardiac arrest in Metropolitan Taipei: Does an advanced life support service make a difference? *Resuscitation*, 74(3):461–469, 2007.

[15] Hasegawa K, Hiraide A, Chang Y, and Brown DFM. Association of prehospital advanced airway management with neurologic outcome and survival in patients with out-of-hospital cardiac. *JAMA*, 309(3):257–266, 2013.

[16] Hanif MA, Kaji AH, and Niemann JT. Advanced airway management does not improve outcome of out-of-hospital cardiac arrest. *Academic Emergency Medicine*, 17(9):926–931, 2010.

[17] Gueugniaud PY, David JS, Chanzy E, et al. Vasopressin and epinephrine vs. epinephrine alone in cardiopulmonary resuscitation. *New Engl J of Med*, 359(1):21–30, 2008.

[18] Stiell IG, Wells GA, Hebert PC, Laupacis A, and Weitzman BN. Association of drug therapy with survival in cardiac arrest: limited role of advanced cardiac life support drugs. *Academic Emergency Medicine*, 2(4):264–273, 2008.

[19] Nolan JP, de Latorre FJ, Steen PA, Chamberlain DA, and Bossaert LL. Advanced life support drugs: do they really work? *Current Opinion in Critical Care*, 8(3):212, 2002.

[20] Neumar RW, Otto CW, Link MS, et al. Part 8: Adult advanced cardiovascular life support 2010 American Heart Association guidelines for cardiopulmonary resuscitation and emergency cardiovascular care. *Circulation*, 122(18 suppl 3):S729–S767, 2010.

[21] Hagihara A, Hasegawa M, Abe T, Nagata T, Wakata Y, and Miyazaki S. Prehospital epinephrine use and survival among patients with out-of-hospital cardiac arrest. *JAMA*, 307(11):1161–1168, 2012.

[22] Gagne JJ, Glynn RJ, Avorn J, Levin R, and Schneeweiss S. A combined comorbidity score predicted mortality in elderly patients better than existing scores. *Journal of Clinical Epidemiology*, 64(7):749–759, 2011.

[23] The Dartmouth Institute for Health Policy Center for Health Policy Research and Clinical Research. 2009 Population Estimates for Zip Code Tabulation Areas (ZCTAs). http://www.dartmouthatlas.org/tools/downloads.aspx.

[24] U.S Department of Health, Health Resources Human Services, and Services Administration. Area Health Resource File. http://arf.hrsa.gov.

[25] Centers for Medicare and Medicaid Services. Hospital Compare. http://www.medicare.gov/hospitalcompare/Data/About.html.

[26] Safar P. *Brain Failure and Resuscitation*. New York, NY: Churchill Livingstone, 1981.

[27] Hirano K and Imbens GW. Estimation of causal effects using propensity score weighting: An application to data on right heart catheterization. *Health Services and Outcomes Research Methodology*, 2(3-4):259–278, 2001.

[28] Sasson C, Magid DJ, Chan P, et al. Association of neighborhood characteristics with bystander-initiated CPR. *New Engl J of Med*, 367(17):1607–1615, 2012.

[29] McNally B, Robb R, Mehta M, Vellano K, Valderrama AL, Yoon PW, et al. Out-of-hospital cardiac arrest surveillance - cardiac arrest registry to enhance survival (cares), united states, october 1, 2005 - december 31, 2010. *Morbidity and Mortality Weekly Report*, 60(SS08):1–19, 2011.

[30] Nichol G, Thomas E, Callaway CW, et al. Regional variation in out-of-hospital cardiac arrest incidence and outcome. *JAMA*, 300(12):1423–1431, 2008.

[31] Wang HE, Mann NC, Mears G, Jacobson K, and Yealy DM. Out-of-hospital airway management in the united states. *Resuscitation*, 82(4):378–385, 2011.

[32] Wang HE and Yealy DM. Managing the airway during cardiac arrest. *JAMA*, 309(3):285–286, 2013.

[33] Pell JP, Sirel JM, Marsden AK, Ford I, and Cobbe SM. Effect of reducing ambulance response times on deaths from out of hospital cardiac arrest: cohort study. *BMJ*, 322(7299):1385–1388, 2001.

[34] Medicare Payment Advisory Commission. Report to the congress: Medicare and the health care delivery system chapter 7, 2013. Accessed April 22, 2015 at http://www.medpac.gov/documents/reports/chapter-7-mandated-report-medicare-payment-for-ambulance-services-(june-2013-report).pdf?sfvrsn=2.

[35] Haas B and Nathens AB. Pro/con debate: is the scoop and run approach the best approach to trauma services organization. *Crit Care*, 12(5):224, 2008.

[36] Liberman M, Mulder D, and Sampalis J. Advanced or basic life support for trauma: meta-analysis and critical review of the literature. *The Journal of Trauma and Acute Care Surgery*, 49(4):584–599, 2000.

[37] Ryynanen O-P, Iirola T, Reitala J, Palve H, and Malmivaara A. Is advanced life support better than basic life support in prehospital care? a systematic review. *Scand J Trauma Resusc Emerg Med*, 18(11):62, 2010.

[38] Stiell IG, Nesbitt LP, Pickett W, et al. The OPALS Major Trauma Study: impact of advanced life-support on survival and morbidity. *CMAJ*, 178(9):1141–1152, 2008.

[39] Stiell IG, Spaite DW, Field B, et al. Advanced life support for out-of-hospital respiratory distress. *New Engl J Med*, 356(21):2156–2164, 2007.

[40] Jayaraman S and Sethi D. Advanced trauma life support training for ambulance crews. *Cochrane Database Syst Rev*, 1, 2010.

[41] Liberman M, Mulder D, Lavoie A, Denis R, and Sampalis JS. Multicenter Canadian study of prehospital trauma care. *Annals of Surgery*, 237(2):153, 2003.

[42] Sukumaran S, Henry JM, Beard D, et al. Prehospital trauma management: a national study of paramedic activities. *Emergency Medicine Journal*, 22(1):60–63, 2005.

[43] Sasser S, Varghese M, Kellerman A, and Lormand JD. Prehospital trauma care systems. *Geneva: World Health Organization*.

[44] Sasser SM, Varghese M, Joshipura M, and Kellermann A. Preventing death and disability through the timely provision of prehospital trauma care. *Bulletin of the World Health Organization*, 84(7):507–507, 2006.

[45] Sanghavi P, Jena AB, Newhouse JP, and Zaslavsky AM. Outcomes after out-of-hospital cardiac arrest treated by basic vs advanced life support. *JAMA Internal Medicine*, 2014.

[46] Meredith JW, Evans G, Kilgo PD, et al. A comparison of the abilities of nine scoring algorithms in predicting mortality. *Journal of Trauma-Injury, Infection, and Critical Care*, 53(4):621–629, 2002.

[47] Stevenson M, Segui-Gomez M, Lescohier I, Di Scala C, and McDonald-Smith G. An overview of the injury severity score and the new injury severity score. *Injury Prevention*, 7(1):10–13, 2001.

[48] MacKenzie EJ, Rivara FP, Jurkovich GJ, Nathens AB, et al. A national evaluation of the effect of trauma-center care on mortality. *New Engl J of Med*, 354(4):366–378, 2006.

[49] Angrist JD, Imbens GW, and Rubin DB. Identification of causal effects using instrumental variables. *Journal of the American Statistical Association*, 91(434):444–455, 1996.

[50] Joshua D Angrist and Jörn-Steffen Pischke. *Mostly harmless econometrics: An empiricist's companion*. Princeton University Press, 2008.

[51] McClellan M, McNeil BJ, and Newhouse JP. Does more intensive treatment of acute myocardial infarction in the elderly reduce mortality?: analysis using instrumental variables. *JAMA*, 272(11):859–866, 1994.

[52] Prasad V and Jena AB. Prespecified falsification end points: can they validate true observational associations? *JAMA*, 309(3):241–242, 2013.

[53] Wang HE, Kupas DF, Hostler D, Cooney R, Yealy DM, and Lave JR. Procedural experience with out-of-hospital endotracheal intubation. *Critical Care Medicine*, 33(8):1718–1721, 2005.

[54] Lecky F, Bryden D, Little R, Tong N, and Moulton C. Emergency intubation for acutely ill and injured patients. *Cochrane Database Syst Rev*, 2, 2008.

[55] Wang HE, Peitzman AB, Cassidy LD, Adelson PD, and Yealy DM. Out-of-hospital endotracheal intubation and outcome after traumatic brain injury. *Annals of Emergency Medicine*, 44(5):439–450, 2004.

[56] Von Elm E, Schoettker P, Henzi I, Osterwalder J, and Walder B. Pre-hospital tracheal intubation in patients with traumatic brain injury: systematic review of current evidence. *British Journal of Anaesthesia*, 103(3):371–386, 2009.

[57] Stockinger ZT and McSwain Jr NE. Prehospital endotracheal intubation for trauma does not improve survival over bag-valve-mask ventilation. *The Journal of Trauma and Acute Care Surgery*, 56(3):531–536, 2004.

[58] Haut ER, Kalish BT, Cotton BA, et al. Prehospital intravenous fluid administration is associated with higher mortality in trauma patients: a National Trauma Data Bank analysis. *Annals of Surgery*, 253(2):371, 2011.

[59] Bickell WH, Wall Jr MJ, and Pepe PE others. Immediate versus delayed fluid resuscitation for hypotensive patients with penetrating torso injuries. *New Engl J of Med*, 331(17):1105–1109, 1994.

[60] Jauch EC, Saver JL, Adams HP, et al. Guidelines for the early management of patients with acute ischemic stroke a guideline for healthcare professionals from the American Heart Association/American Stroke Association. *Stroke*, 44(3):870–947, 2013.

[61] Sparrow MK. *License to steal: how fraud bleeds America's health care system*. Basic Books, 2000.

[62] US Department of Health and Human Services. Departments of justice and health and human services announce over $27.8 billion in returns from joint efforts to combat health care fraud, 2015. http://www.hhs.gov/news/press/2015pres/03/20150319a.html?utm_campaign=031915_hhs_news_hcfac&utm_medium=email&utm_source=hhs_news_releases&utm_content=031915_hhs_news_hcfac_titlelink.

[63] Berwick DM and Hackbarth AD. Eliminating waste in US health care. *JAMA*, 307(14):1513–1516, 2012.

[64] Medicare Payment Advisory Commission. Report to the congress, regional variation in medicare service use, 2011. http://www.medpac.gov/documents/reports/Jan11_RegionalVariation_report.pdf?sfvrsn=0.

[65] Osler T, Baker SP, and Long W. A modification of the injury severity score that both improves accuracy and simplifies scoring. *Journal of Trauma-Injury, Infection, and Critical Care*, 43(6):922–926, 1997.

[66] Brenneman FD, Boulanger BR, McLellan BA, and Redelmeier DA. Measuring injury severity: time for a change. *Journal of Trauma-Injury, Infection, and Critical Care*, 44(4):580–582, 1998.

[67] Sacco WJ, MacKenzie EJ, Champion HR, Davis EG, and Buckman RF. Comparison of alternative methods for assessing injury severity based on anatomic descriptors. *Journal of Trauma-Injury, Infection, and Critical Care*, 47(3):441–446, 1999.

[68] Clark DE, Osler TM, and Hahn DR. ICDPIC: Stata module to provide methods for translating International Classification of Disease (Ninth Revision) diagnosis codes into standard injury categories and/or scores. http://ideas.repec.org/c/boc/bocode/s457028.html, 2010.

[69] Sears JM, Blanar L, and Bowman SM. Predicting work-related disability and medical cost outcomes: A comparison of injury severity scoring methods. *Injury*, 45(1):16–22, 2014.

[70] National Association of State EMS Directors. National EMS Information System. http://www.nemsis.org.