Accident Analysis and Prevention 82 (2015) 257–262

Contents lists available at ScienceDirect

Accident Analysis and Prevention

journal homepage: www.elsevier.com/locate/aap

Evaluating adverse rural crash outcomes using the NHTSA State Data System

Christine Peura[a], Joseph A. Kilch[a], David E. Clark[a,b,*]

[a] Center for Outcomes Research and Evaluation, Maine Medical Center, Portland, ME, USA
[b] Department of Surgery, Maine Medical Center, Portland, ME, USA

ARTICLE INFO

Article history:
Received 17 March 2015
Received in revised form 12 June 2015
Accepted 13 June 2015
Available online 24 June 2015

Keywords:
Crash
Injury
Mortality
Rural
NHTSA
State Data System

ABSTRACT

Introduction: The population-based rate of motor vehicle crash mortality is consistently higher in rural locations, but it is unclear how much of this disparity might be due to geographic barriers or deficiencies in emergency medical services (EMS). We sought to analyze separately factors associated with the occurrence of a severe injury and those associated with death after injury had occurred.
Methods: Data from all police-reported crashes in 11 states from 2005–2007 were obtained through the National Highway Traffic Safety Administration (NHTSA) State Data System (SDS). Logistic regression was used to estimate factors associated with (1) death; (2) severe (incapacitating or fatal) injury; and (3) death given severe injury. Models included covariates related to the person, vehicle, and event; county location was specified using Rural–Urban Continuum Codes (RUCC).
Results: Older age, not wearing a belt, ejection, alcohol involvement, high speed, and early morning times were associated with increased risk of both severe injury and death. Controlling for these factors, and restricting analysis to persons who had suffered a severe injury, the adjusted odds ratio (aOR) associated with death was higher for counties classified rural (RUCC 6–7, aOR 1.23, 95% CI 1.16–1.31) or very rural (RUCC 8–9, aOR 1.31, 95% CI 1.18–1.46).
Conclusions: Persons severely injured in crashes are more likely to die if they are in rural locations, possibly due to EMS constraints. As NHTSA-SDS data become more available and more uniform, they may be useful to explore specific factors contributing to this increased risk.

1. Introduction

The increased risk for rural residents to die from a motor vehicle crash has been recognized for many years (Baker et al., 1987; Brodsky and Hakkert, 1983). While this disparity may be partly due to an increased incidence of severe crashes, it appears to be attributable more to the difference in outcome for persons who have been injured (Goldstein et al., 2011; Muelleman et al., 2007). This disparity in outcomes may raise questions about the quality of care delivered by emergency medical services (EMS) and emergency departments (ED), as well as the obvious problems of communication, transportation, and scarce resources in more remote locations (Cummings and O'Keefe, 2000). We sought to explore these issues in order to help identify any factors that might be modified to improve outcomes.

The National Highway Traffic Safety Administration (NHTSA) has developed several crash databases and made them available to researchers at no cost. The best known is the Fatality Analysis Reporting System (http://www.nhtsa.gov/FARS), a census of all crashes since 1975 in which at least one person died. A stratified random sample of similar (but less detailed) information about nonfatal crashes has been provided since 1988 by the National Automotive Sampling System (http://www.nhtsa.gov/NASS). These databases have been used extensively by traffic and automotive engineers, and occasionally for epidemiologic or health services research.

A less frequently used NHTSA database is the State Data System (http://www.nhtsa.gov/Data/State+Data+Program+&+CODES), which is a compilation of state based police accident reports from participating states, including information about the event, vehicles, and persons similar to that available in the National Automotive Sampling System (NASS). Studies using the NHTSA State Data System (NHTSA-SDS) have been infrequently published outside of NHTSA (Cheung and McCartt, 2011; Eisenberg and

* Corresponding author at: Maine Medical Center, Center for Outcomes Research and Evaluation, 509 Forest Avenue, Suite 200, Portland, ME 04101, USA. Fax: +1 207 774 0459.
E-mail addresses: peurac@mmc.org (C. Peura), Joseph.Kilch@med.uvm.edu (J.A. Kilch), clarkd@mmc.org (D.E. Clark).

http://dx.doi.org/10.1016/j.aap.2015.06.005
0001-4575/

Warner, 2005; Karaca-Mandic and Ridgeway, 2010; Lyon et al., 2012), but it records data from a much larger number of rural counties than the few sampled by the NASS. NHTSA-SDS is therefore a potentially valuable database for EMS research, which could help overcome some of the limitations and complement the findings from the other NHTSA databases.

A primary goal of this study was to investigate further the rural/urban outcome disparities in traffic crashes and the effects of post-crash factors on outcomes. We were most interested in the person, vehicle, and location factors associated with mortality among persons who had been severely injured in a crash, since this would be the outcome most likely to be affected by EMS or trauma care systems. We also intended to compare findings using NHTSA-SDS to published results based on estimates from the NASS General Estimates System (GES).

2. Methods

NHTSA-SDS data for 2005–2007 were obtained at nominal cost through the NHTSA Office of Data Acquisitions. Access to the data from each state required specific approval of an official in that state, and in some cases there were additional state-specific requirements. At the time, 33 states were participating in NHTSA-SDS, and we attempted to obtain data from 20 of them; seven explicitly denied access except to internal NHTSA researchers, and two others did not respond to repeated requests. The 11 states that agreed to provide data were Arkansas, California, Florida, Kentucky, Maryland, Michigan, Minnesota, New Mexico, South Carolina, Washington, and Wyoming. An institutional review board exempted this study from review because it contained only preexisting and de-identified data.

Data maintenance and analysis was performed using Stata Version 12.1 (StataCorp, College Station, TX). States had independently collected and reported their data, resulting in variability in the completeness and possibly the accuracy of information. Variable names and coding schemes were also inconsistent, and these were renamed and redefined so that they would be uniform across each state database. Each state database included separate person, vehicle, and crash files that were merged to create a single dataset for each state. A multiple-state database was then created by merging all of the state datasets, which included selected person-, vehicle-, and crash-level variables from each of the 11 states.

Persons other than occupants of cars and light trucks were excluded. Person-related variables of interest included age, sex, seating position, safety belt usage, ejection, and police-reported injury severity (no injury, possible injury, non-incapacitating injury, incapacitating injury, fatal injury, injury of unknown severity). Age was categorized as under 15 years, 15–39, 40–64, 65–79, and 80 years or more. Persons were considered unbelted if they were not recorded as wearing a seat belt or if seat belts were not available in the vehicle.

NHTSA-SDS (like NASS-GES) does not contain measures of injury severity ascertained by hospital personnel or medical examiners. However, the definition of “incapacitating injury” used in police reports has been standardized to mean an injury “which prevents the injured person from walking, driving or normally continuing the activities the person was capable of performing before the injury occurred” (National_Safety_Council, 2007). The number of persons with incapacitating injuries determined by police reports has been previously used as a denominator for persons at risk of death after traffic crashes (Brodsky and Hakkert, 1983; Brown, 1979). For this study, we defined a “severe” injury category to include injuries categorized by police as either incapacitating or fatal.

Other variables were selected primarily with a view to comparison with the findings that Travis et al. (2012) had obtained using NASS-GES, but with further detail limited by the comparability of the SDS data from each state. Vehicle-related variables included speed limit, speed estimates, driver alcohol involvement, rollovers, vehicle deformation, and towing information. Most states were missing data on one or more of the variables that might be used to describe vehicle crash damage, so it was not possible to derive a standard categorization. For this study, we classified vehicles as traveling at high speed if the police estimated that it was traveling over 50 miles per hour at the time of the crash, or if its speed was not reported but the posted speed limit was 50 or above. Alcohol was considered involved if the driver had a positive blood alcohol test.

Crash-related variables included time of day, poor visibility, slippery conditions, number of vehicles involved, and occurrence on an interstate highway. The time of the crash was divided into 6-h blocks.

The location of each crash was categorized using county Rural–Urban Continuum Codes, as defined by the U.S. Department of Agriculture (RUCC, http://www.ers.usda.gov/data-products/rural-urban-continuum-codes.aspx). RUCC classifies US counties into nine groups, which we stratified as urban (RUCC 1–3), partly rural (RUCC 4–5), rural (RUCC 6–7), and very rural (RUCC 8–9). The location was specified as southern if the crash occurred in Arkansas, Florida, New Mexico or South Carolina.

Additional county-level demographic, socioeconomic, and medical resource information was obtained from the Area Resource File provided by the U.S. Health Resources and Services Administration (ARF, http://www.arf.hrsa.gov). ARF variables included median age, racial and ethnic proportions, median income, unemployment rates, high school graduation rates, proportion of the population with health insurance, the number of emergency physicians, and the number of neurosurgeons.

After initial data exploration, logistic regression models (limited to occupants of motor vehicles) were used to examine

Table 1
Rural–urban continuum codes—definitions and distributions.

RUCC	Definition	US counties $n = 3142$	US population $n = 281{,}422{,}000$	Sample counties $n = 654$	Sample persons $n = 11{,}008{,}057$
1	>1 million, metro area	414 (13.2%)	149,224,000 (53.0%)	85 (13.0%)	6,044,182 (54.9%)
2	250,000 to 1 million, metro area	325 (10.3%)	55,514,000 (19.7%)	83 (12.7%)	2,408,260 (21.9%)
3	<250,000, metro area	351 (11.2%)	27,842,000 (9.9%)	73 (11.2%)	1,097,909 (10.0%)
4	≥20,000, adjacent to metro area	218 (6.9%)	14,442,000 (5.1%)	43 (6.6%)	348,967 (3.2%)
5	≥20,000, not adjacent to metro area	105 (3.3%)	5,573,000 (2.0%)	24 (3.7%)	225,258 (2.1%)
6	2500 to 19,999, adjacent to metro area	609 (19.4%)	15,134,000 (5.4%)	128 (19.6%)	390,697 (3.6%)
7	2500 to 19,999, not adjacent to metro area	450 (14.3%)	8,464,000 (3.0%)	112 (17.1%)	356,686 (3.2%)
8	<2500, adjacent to metro area	235 (7.5%)	2,426,000 (0.9%)	42 (6.4%)	46,548 (0.4%)
9	<2500, not adjacent to metro area	435 (13.8%)	2,803,000 (1.0%)	64 (9.8%)	67,947 (0.6%)

21,603 persons in the sample were missing county information.