



Short Communication

Missed opportunities to advance knowledge on traffic safety: Accessibility of driver licensing and crash data for scientific research

Meghan E. Carey^{a,*}, Evan D. Anderson^b, Rania Mansour^a, Jason Sloan^{c,d}, Allison E. Curry^{a,d}^a Center for Injury Research and Prevention, Children's Hospital of Philadelphia, 2716 South Street, Philadelphia, PA 19146, USA^b University of Pennsylvania School of Nursing, 418 Curie Blvd., Philadelphia, PA 19104, USA^c University of Pennsylvania Law School, 3501 Sansom Street, Philadelphia, PA 19104, USA^d Perelman School of Medicine, University of Pennsylvania, 3400 Civic Center Blvd., Philadelphia, PA 19104, USA

ARTICLE INFO

Keywords:

Public health law
 Legal mapping
 Traffic safety
 Data collection
 Data linkage

ABSTRACT

Objective: Identifiable individual-level driver licensing and motor vehicle crash data are essential to advancing transportation safety research. However, epidemiologic studies using such data are rare, which may reflect their inaccessibility. We conducted a legal mapping study to evaluate US state laws regulating access to driver licensing and motor vehicle crash data for use in scientific research.

Methods: Legal statutes regulating the release of driver licensing and motor vehicle crash data for all 50 US states and the District of Columbia (D.C.) were retrieved. Legal text was evaluated to determine whether these jurisdictions authorize release of identifiable individual-level licensing and crash data for use in non-governmental research.

Results: Thirty-six states and D.C. explicitly authorize release of identifiable individual-level licensing data to researchers. Only five states and D.C. authorize release of identifiable individual-level crash records. No states explicitly prohibit the release of individual-level data about licensing records and only three states prohibit release of individual-level crash record data, meaning that in many states it is ambiguous whether and when releasing such data to researchers is permitted.

Conclusions: It is important to understand why licensing data are not used more frequently in transportation safety research given that many state laws permit access for non-governmental researchers. Reforming state laws to clarify and increase access to identifiable individual-level crash report data is an important priority for transportation safety advocates and researchers.

1. Introduction

Reductions in crash-related morbidity and mortality represent one of the ten most important public health achievements of the last hundred years (Centers for Disease Control and Prevention National Center for Injury Prevention and Control, 1999). Travel on roads and highways is vastly safer today in large part because of access to high quality traffic safety data (Burriss and Anderson, 2013). Such data have reliably measured a multitude of risk factors for crash-related injuries, enabling researchers to identify and evaluate transportation safety interventions. The development and refinement of transportation safety laws is considered one of the best examples of policy learning in which ongoing evaluation of state innovations guides continuous reform (Burriss and Anderson, 2013). Despite access to data and a record of persistent improvement in safety that is enviable among public health domains,

motor vehicle crashes remain one of the most common causes of injury-related mortality. Moreover, there is some troubling evidence that reductions in crash-related mortality may be stalling or even reversing in the last few years (National Highway Traffic Safety Administration, 2015).

Preserving and extending transportation safety gains requires rigorous and comprehensive analyses to identify both underlying and proximate factors that increase the risk of crashes and crash-related morbidity. Currently, the traffic safety field is largely precluded from conducting such analyses, as crash data provided to researchers by the National Highway Traffic Safety Administration (e.g., Fatality Analysis Reporting System) and certain state jurisdictions is de-identified (National Highway Traffic Safety Administration, 2014). Without access to identifiable individual-level data, crash events can only be studied in isolation. Thus, multiple events experienced by an individual

* Corresponding author at: Center for Injury Research and Prevention, Children's Hospital of Philadelphia, 2716 South Street, 13th floor, Philadelphia, PA 19146, USA.

E-mail address: careym3@email.chop.edu (M.E. Carey).

<https://doi.org/10.1016/j.aap.2020.105500>

Received 2 October 2019; Received in revised form 17 January 2020; Accepted 11 March 2020

0001-4575/© 2020 Elsevier Ltd. All rights reserved.

driver cannot be connected, limiting our ability to understand context and characteristics of our most frequently crash-involved drivers. Additionally, lack of access to identifiable data precludes linkage to other valuable existing data sources—including those relating to driver licensure, traffic citations, emergency medical services, hospital treatment, and long-term care. This severely limits our ability to situate crash events in the context of a driver's previous driving history, a powerful predictor of ongoing risk for crash involvement, to follow individual drivers over the course of licensure, and to more fully understand the effect of crashes on short- and long-term injury and health outcomes (Das et al., 2015; Curry et al., 2019).

Like other important public health data, identifiable driver licensing and crash databases have become less widely available to researchers due in part to privacy concerns (Wartenberg and Thompson, 2010). In particular, state-level driver license and crash data have become increasingly inaccessible over recent decades since the 1994 federal Driver's Privacy Protection Act (DPPA) implemented specific legal restrictions on the release of license data. DPPA was adopted not because of researcher malfeasance or because of any issues related to bulk access to records; rather, DPPA was adopted because a few individuals used license data from state motor vehicle departments to locate—and in one instance murder—other citizens (Cosgrove, 2000). Although DPPA specifically exempted the release of data for research from its prohibitions, state implementation of DPPA has seemingly limited access to such data for scientific purposes.

This study aimed to systematically evaluate state laws regulating the release of driver licensing and motor vehicle crash data for scientific research. To do this, we conducted a legal mapping study, which is a type of systematic study used in the field of public health law to elucidate elements of laws pertaining to a particular issue and determine patterns or trends in these laws across states or jurisdictions (Burris, 2018). To our knowledge, no prior study has systematically measured state laws regulating access to identifiable individual-level crash or license data for non-governmental researchers.

2. Methods

This study followed recognized best-practices for systematic legal research (Tremper et al., 2010). The study retrieved legal text related to the release of driver licensing and motor vehicle crash records from law.justia.com—an extensive collection of legislation from the 50 US states and D.C.—and secondarily from LexisNexis, a subscription-based computer-assisted legal research tool used by legal practitioners. For comparison purposes, we also gathered legal statutes pertaining to the release of vital statistics, long considered the 'backbone' of linked administrative data (Jutte et al., 2011). To identify the most recent legal statutes regulating data release, the search terms and syntax in Table 1 were iteratively refined, and then deployed in searches in late 2017 and

2018.

Two co-authors (MC, RM) were trained to abstract legal text by another co-author with expertise in public health law (EA). An *a priori* codebook was drafted to guide the initial abstraction of legal text from five randomly chosen states. After examining the resulting data, the research team revised and finalized the codebook. The finalized codebook contained the following questions for licensing data: (1) is individual-level license information available to researchers; (2) is the information provided de-identified, or identifiable; and (3) if identifiable, are there any additional limitations or conditions placed on breadth of identifiable data? Codebook questions for crash and vital statistics records were complementary to those for license data. To assess the reliability of the coding schema, legal text from an additional ten randomly chosen states was double coded by a separate trained abstractor (JS). Divergences in coding were observed in less than 22 % of the codes pertaining to licensing and in less than 14 % of crash and vital statistics after initial coding. All discrepancies were then discussed by the research team until agreement was reached. Any discrepancies were re-coded and clarifying updates were made to the codebook, before coding proceeded. We characterized a codebook item as "unknown availability" if the legal text provided no clear answer. We conducted descriptive analyses using SAS Version 9.4 (SAS Institute, Inc.).

3. Results

Thirty-seven states and D.C. explicitly authorize the release of individual-level licensing records to researchers (Fig. 1A). We could not discern whether release to researchers is authorized for the remaining 13 states. Among the 38 jurisdictions that authorize release, all but two specified that they could provide identifiable licensing data. Some states placed additional conditions on requests for identifiable data, including (1) publication and re-disclosure of personal information, (2) contacting individuals, (3) release of Social Security Number, and (4) release of license photos. Only five states specified that data could be released in bulk.

As shown in Fig. 1B, more than two-thirds (n = 38) of jurisdictions did not specify whether release of individual-level crash records to researchers is authorized. Only nine states and D.C. legally authorize the release of crash records to researchers; five states and D.C. explicitly stated they would release identifiable data. Only two jurisdictions mentioned specific conditions for releasing identifiable data: D.C. does not allow publishing or re-disclosing personal information and Michigan will not release a person's name unless the person authorizes release in writing. Three states (HI, KY, and NM) explicitly prohibit release of individual-level crash data to researchers. Conversely, 47 jurisdictions authorize release of individual-level vital statistics data for researchers (Fig. 1C); the release of identifiable data was only specified in seven states and D.C.

Table 1
Search terms utilized to find relevant legal text.

<i>A priori</i> ^a	Search Term 1	Search Term 2	Search Term 3
Driver's license	State name AND 'driving records confidentiality'	State name AND 'driving records disclosure'	State name AND 'driving records availability'
Crash records	State name AND 'crash records confidentiality'	State name AND 'crash records disclosure'	State name AND 'crash records availability'
Vital statistics	State name AND 'vital statistics confidentiality'	State name AND 'vital statistics disclosure'	State name AND 'vital statistics availability'
Based on legal text review ^b	Search Term 1	Search Term 2	Search Term 3
Driver's license	State name AND 'record' AND 'vehic*'	State name AND 'license' AND 'record'	State name AND ('record' OR 'license') AND ('vehic*' OR 'driver')
Crash records	State name AND ('record' OR 'report') AND ('crash!' OR 'accident') AND 'vehic!'	-	-
Vital statistics	State name AND ('vital statistics' OR 'birth records' OR 'death records')	-	-

^a These search terms did not solicit all necessary legal text.

^b Search terms used after line by line reading of the legal text returned from the *a priori* search.

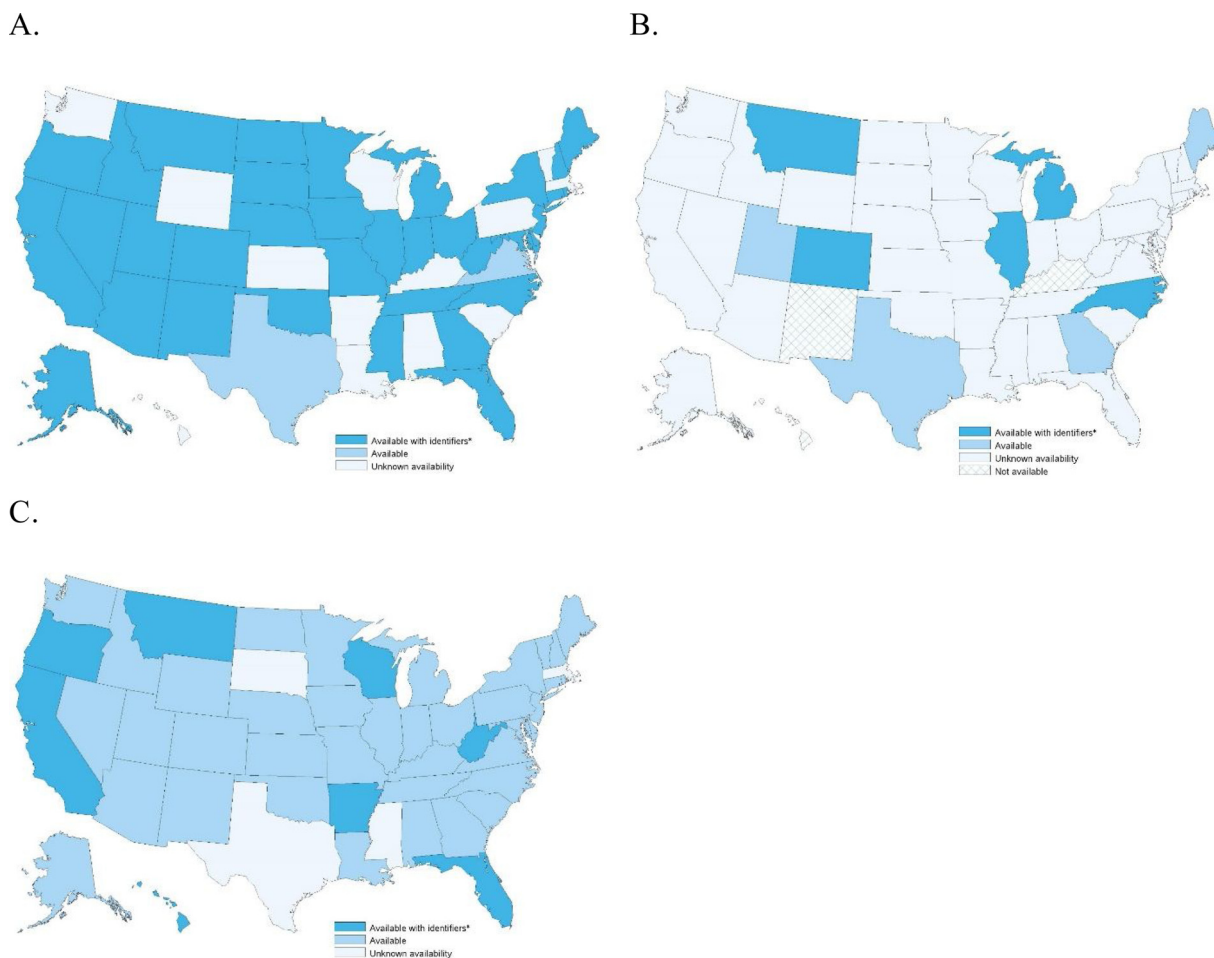


Fig. 1. Availability of individual-level (A) driver licensing data, (B) crash report data, and (C) vital statistics data for researchers according to the legal statutes of each US state and the District of Columbia. * Legal statutes did not specify which identifiers were available for researchers, only that data would include identifiers.

4. Discussion

This study examined legal statutes regulating the release of identifiable individual-level driver licensing and crash data in the US. Results indicate that the majority of US jurisdictions authorize release of identifiable driver licensing data to researchers. In contrast, it is largely unclear under existing statutes whether the release of identifiable crash data is permitted for non-governmental research.

As highlighted in CDC's 2019 publication *Linking Information for Nonfatal Crash Surveillance*, there is a critical gap in the availability of comprehensive and novel data resources that can propel scientific advancements in the prevention of motor vehicle crashes (Centers for Disease Control and Prevention National Center for Injury Prevention and Control, 2019). Access to these data resources would certainly catalyze our ability to address novel and important questions in traffic safety. Linkages between licensing and crash databases would enable researchers to monitor drivers longitudinally over time and determine the extent to which adverse driving events (e.g., crashes, citations, DUIs, suspensions) predict future crash risk. Further, linkage of crash data with hospital outcome databases can deepen our understanding of the risk of injuries from crashes and the contribution of relevant factors (e.g., seatbelt use) to injury severity; this foundational data can guide policy development and ultimately improve outcomes among road users. There are numerous examples of international traffic safety researchers accessing and linking existing identifiable crash data to hospital data or trauma registries, including in Australia, Denmark, New Zealand, and France (Watson et al., 2015; Alsop and Langley, 2001; Amoros et al., 2006; Janstrup et al., 2014). However, to our knowledge

only a handful of US-based research groups and programs have been able to access and link identifiable individual-level crash and license data (Centers for Disease Control and Prevention National Center for Injury Prevention and Control, 2019).

The current study was designed to identify whether governmental agencies *can* provide licensing and crash data for researchers; notably, these laws do not specify whether such agencies *must* or *actually do* release data to researchers. On the actual release of licensing data, two previous studies directly surveyed representatives at the government entity responsible for maintaining driver licensing data in each US state (Walsh et al., 2011; Littenberg and Lubetkin, 2016). Walsh et al. aimed to utilize driver licensing data to recruit controls for a hypothetical case-control study on breast cancer; in this study, government officials from fourteen states indicated they would release individual-level data for their study (Walsh et al., 2011). Littenberg and Lubetkin aimed to collect specific variables—including height, weight, and address—from driver records to examine the relationship between body mass index and the built environment (Littenberg and Lubetkin, 2016). Eighteen states provided them with data, though only fourteen states provided the full extent of data requested (Littenberg and Lubetkin, 2016).

Collectively, findings from the current study and previous research indicate that the proportion of state agencies that are legally authorized to release identifiable licensing data is much higher than the proportion that reported a willingness to release these data. This suggests that the primary barriers affecting researchers' ability to obtain identifiable individual-level licensing data are factors other than legal availability. Future research is critically needed to understand why identifiable individual-level licensing data are not being utilized in traffic safety

research, given they are largely legally available to researchers. Specifically, future studies should determine the extent to which factors such as data infrastructure, informal departmental policies, concern around inappropriate use and redisclosure of protected data, and competing priorities might preclude agencies from disclosing individual-level data for research purposes. Facilitators should also be identified so these data can be leveraged to the greatest extent possible (Walsh et al., 2011; Littenberg and Lubetkin, 2016).

One notable limitation of this study is its reliance on LexisNexis, a subscription-based tool, to identify relevant legal text. Our LexisNexis search algorithm was superior to that on law.justia.com, a free service, as the same search terms in law.justia.com prompted a host of irrelevant results (e.g., court cases, bills). As a descriptive study, this project did not explore the ethical implications of the current legal requirements. In recent years, there have been troubling reports that state Departments of Motor Vehicles have sold individual driver information to corporations and private investigators (Landen, 2019). It is essential to note that we only explored – and would only propose increasing – availability of identifiable individual-level data elements required to perform a high quality data linkage for the purpose of scientific research, which is subject to federal, state, and institutional privacy requirements.

4.1. Conclusions

According to legal statutes, thirty-five states and D.C. authorize the release of identifiable individual-level licensing data to external non-governmental researchers. Providing public health researchers with identifiable data from drivers' license and motor vehicle crash records would allow linkage to other data sources (e.g., crash records, hospital discharge records), exponentially increasing the value of each linked dataset. The disconnect between legal authorization of releasing driver licensing data and its scarce use in applied studies highlights the need to further understand existing barriers to researcher access. Additionally, there is a need to improve the clarity of state laws surrounding the release of crash records.

Funding statement

This work was supported by the Eunice Kennedy Shriver National Institute of Child Health and Human Development at the National Institutes of Health [grant number R21HD092850] and the Nicholas Crognale Chair for Emergency Medicine, Children's Hospital of Philadelphia, PA, USA. The funders of this study had no involvement in study design or conduct; collection, analysis and interpretation of data; writing the report; or deciding to submit the article for publication.

CRedit authorship contribution statement

Meghan E. Carey: Conceptualization, Formal analysis, Writing - original draft, Project administration. **Evan D. Anderson:** Methodology, Writing - original draft, Writing - review & editing. **Rania Mansour:** Validation, Investigation, Writing - review & editing. **Jason Sloan:** Validation, Writing - review & editing. **Allison E. Curry:** Conceptualization, Writing - review & editing, Supervision, Funding acquisition.

Declaration of Competing Interest

None.

Acknowledgments

The authors would like to acknowledge Jacqueline Teed, BS, and Lauren D'Arinzo, BS, for their assistance with the initial collection of legal statutes and assistance developing initial iterations of the codebook. We would also like to thank Kristi Metzger, PhD, MPH, for her help with visualization of study results.

References

- Alsop, J., Langley, J., 2001. Under-reporting of motor vehicle traffic crash victims in New Zealand. *Accid. Anal. Prev.* 33 (3), 353–359.
- Amoros, E., Martin, J.L., Laumon, B., 2006. Under-reporting of road crash casualties in France. *Accid. Anal. Prev.* 38 (4), 627–635. <https://doi.org/10.1016/j.aap.2005.11.006>.
- Burris, S.C., 2018. How to Write a Legal Mapping Paper. *Temple Univ Leg Stud Res Pap No 2018-10* <https://doi.org/10.2139/ssrn.3133065>.
- Burris, S., Anderson, E.D., 2013. Legal regulation of health-related behavior: a half century of public health law research. *Annu Rev Law Soc Sci.* 9, 95–117. <https://doi.org/10.1146/annurev-lawsocsci-102612-134011>.
- Centers for Disease Control and Prevention National Center for Injury Prevention and Control, 1999. Achievements in Public Health, 1900-1999 Motor-Vehicle Safety: A 20th Century Public Health Achievement Vol 48. <http://www.cdc.gov/mmwr/preview/mmwrhtml/mm4818a1.htm>.
- Centers for Disease Control and Prevention National Center for Injury Prevention and Control, 2019. *Linking Information for Nonfatal Crash Surveillance*. Atlanta, GA.
- Cosgrove, R.T., 2000. *Reno v. Condon: The Supreme Court Takes a Right Turn in Its Tenth Amendment Jurisprudence by Upholding the Constitutionality of the Driver's Privacy Protection Act*. <http://ir.lawnet.fordham.edu/flr/vol68/iss6/13>.
- Curry, A.E., Pfeiffer, M.R., Carey, M.E., Cook, L.J., 2019. Catalyzing traffic safety advancements via data linkage: development of the New Jersey Safety and Health Outcomes (NJ-SHO) data warehouse. *Traffic Inj. Prev.* 0 (0), 1–4. <https://doi.org/10.1080/15389588.2019.1679552>.
- Das, S., Sun, X., Wang, F., Leboeuf, C., 2015. Estimating likelihood of future crashes for crash-prone drivers. *J. Traffic Transp. Eng. (English Ed.)* 2 (3), 145–157. <https://doi.org/10.1016/j.jtte.2015.03.003>.
- Janstrup, K.H., Hels, T., Kaplan, S., Lauritsen, J., Prato, C.G., 2014. Understanding traffic crash under-reporting: linking police and medical records to individual and crash characteristics. *Traffic Inj. Prev.* <http://trid.trb.org/view/2014/C/1316934>.
- Jutte, D.P., Roos, L.L., Brownell, M.D., 2011. Administrative record linkage as a tool for public health research. In: In: Fielding, J.E., Brownson, R.C., Green, L.W. (Eds.), *Annual Review of Public Health Vol 32*. pp. 91–108. <https://doi.org/10.1146/annurev-publhealth-031210-100700>. Vol 32.
- Landen, X., 2019. How Vermont's DMV Makes Millions of Dollars Selling Personal Information. Published November 29 Accessed December 14, 2019. <https://www.vnews.com/How-Vermont-s-DMV-makes-millions-of-dollars-selling-personal-information-30912137>.
- Littenberg, B., Lubetkin, D., 2016. Availability, strengths and limitations of US state driver's license data for obesity research. *Cureus* 8 (3), e518. <https://doi.org/10.7759/cureus.518>.
- National Highway Traffic Safety Administration, 2014. *Fatality Analysis Reporting System (FARS)*. Published. <https://www.nhtsa.gov/research-data/fatality-analysis-reporting-system-fars>.
- National Highway Traffic Safety Administration, 2015. *Research Note: 2015 Motor Vehicle Crashes: Overview (August)*. <https://crashstats.nhtsa.dot.gov/Api/Public/ViewPublication/812318>.
- Tremper, C., Thomas, S., Wagenaar, A.C., 2010. Measuring law for evaluation research. *Eval. Rev.* 34 (3), 242–266.
- Walsh, M.C., Trentham-Dietz, A., Palta, M., 2011. Availability of driver's license master lists for use in government-sponsored public health research. *Am. J. Epidemiol.* 173 (12), 1414–1418. <https://doi.org/10.1093/aje/kwr018>.
- Wartenberg, D., Thompson, W.D., 2010. Privacy versus public health: the impact of current confidentiality rules. *Am. J. Public Health* 100 (3), 407–412. <https://doi.org/10.2105/AJPH.2009.166249>.
- Watson, A., Watson, B., Vallmuur, K., 2015. Estimating under-reporting of road crash injuries to police using multiple linked data collections. *Accid. Anal. Prev.* 83, 18–25. <https://doi.org/10.1016/j.aap.2015.06.011>.