Female vs. Male Relative Fatality Risk in Fatal Crashes: 1975-2018

M. Z. Abrams[1] and C. R. Bass[1]
[1] Duke University

ABSTRACT

Known sex-based physiological and biomechanical differences do not explain the much larger fatality risk to young women in car crashes compared with men under matched conditions, suggesting a need for expanded test methodologies and research strategies to address as-yet unexplored sex contributing to crash outcomes. Vehicle occupants who died in crashes with at least two occupants between 1975 and 2018 were examined. National Highway Traffic Safety Administration (NHTSA) Fatality Analysis Reporting System (FARS) data was used to calculate relative risk using a modified Double Pair Comparison method, with a bootstrap estimation for determining confidence intervals. This method matches subject occupants exhibiting one of two characteristics based on the characteristics of a common control occupant in the same vehicle. Participants were matched using vehicle seating location, vehicle type, airbag deployment, seatbelt usage, age and a control occupant. Male (n=205,325) and female (n=126,075) vehicle occupants sustaining fatal injury in passenger car or light truck crashes between 1975 and 2018 were examined. This population was binned by crash year (1975-2018, n=331,400; 1999-2018, n=112,032; 2010-2018, n=36,481; 2015-2018, n=14,796). Across all conditions, 25-year-old female vehicle occupants in crashes from 2010-2018 exhibit a relative fatality risk of 1.107 (95% CI 1.097 – 1.116) compared to 25-year-old male occupants. There is an age-dependent risk to females in vehicle crashes, with younger females at higher risk of death than younger males. This difference is robust across many variables, and is not attributable to vehicle weight, belt usage, or airbag deployment. Female vehicle occupants aged 20–30 years are 20–25% more likely to die as a result of a fatal crash than males in the same age range. Known sex-based differences do not explain this large risk differential, and should be further examined to ensure all occupants are well-protected.

INTRODUCTION

Motor vehicle accidents are one of the leading causes of mortality worldwide, as the 8th leading cause of death for people of all ages, and the #1 cause of death for children and young adults 5-29 years old (World Health Organization, 2018). In the United States, mortality from motor vehicle accidents is the 2nd leading nationwide cause of unintentional injury death, resulting in 1.4 million years of life lost annually (Webb, 2020). Motor vehicle fatalities rank in the top 3 causes of death for individuals under the age of 34 (Webb, 2020).

**This paper has not been peer-reviewed*

The National Highway Traffic Safety Administration (NHTSA) maintains the Fatality Analysis Reporting System (FARS) (National Highway Traffic Safety Administration, 2018, 2019), which tracks all traffic crashes in the USA since 1975 that involve at least one fatality. FARS data are used to inform safety decisions at the local, state and national levels, and provide key insights into the efficacy of changing vehicle and roadway safety standards (National Highway Traffic Safety Administration, 2019). To be included in FARS, a crash must occur on a public road and must result in at least one death within 30 days of the crash. Road fatalities in the US continue to decrease as better advanced safety technologies emerge and become standard features across the board. Occupant fatalities involving vehicles manufactured in the last five or 10 years have decreased steadily, down significantly since 1975 (Figure 1). Despite improvements, when compared to 15 peer nations, the United States ranks last in reducing the rate of annual vehicle fatalities (Ahangari et. al., 2016).

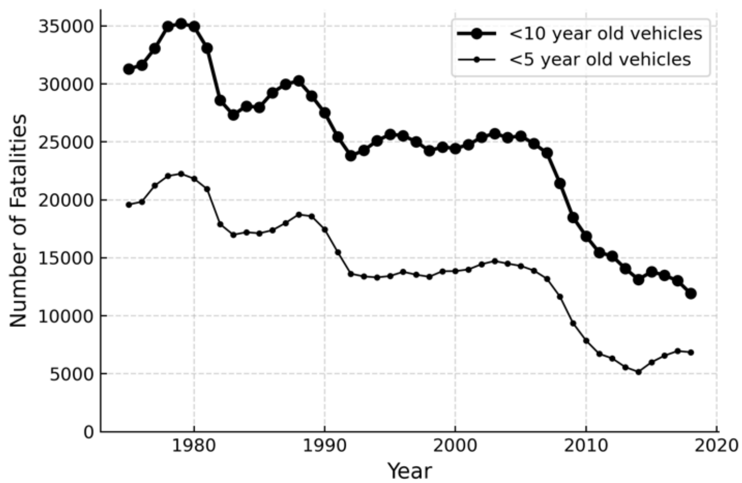

Figure 1: Annual occupant fatalities with vehicles manufactured in the prior five or 10 years.

Overall, evolving crash testing and vehicle standards are estimated to have prevented ~60% of all potential vehicle fatalities in the USA (Kahane, 2015; Viano and Parenteau, 2016). However, the bulk of the fatality reduction is in males (Figure 2). Qualitatively, female fatality rates have been relatively stable when compared to the overall trend. While historically, males have driven more miles per licensed driver than females, that gap is steadily closing, with decreasing differences in licensure rates and driving exposure (Fucci, 2018; Mayhew et. al., 2003; Sivak, 2013). Driving exposure for females has increased, and female drivers have been reported to display similar behavioral risk profiles to males when behind the wheel (Fucci, 2018; Romano et. al., 2008; Tsai et. al., 2008). Previous research has shown that female drivers and vehicle occupants are more likely than males to suffer severe or fatal injuries when involved in a fatal crash (Bose et. al., 2011; Evans, 1988, 2001; Romano et al., 2008). This potential disparity between men and women in automobile crashes is a major public health issue with implications for automobile design, personnel protection, and governmental regulation. The goal of this study is to determine relative risk of fatality by age and sex in a large dataset by matching crash conditions among different sex and age cohorts.

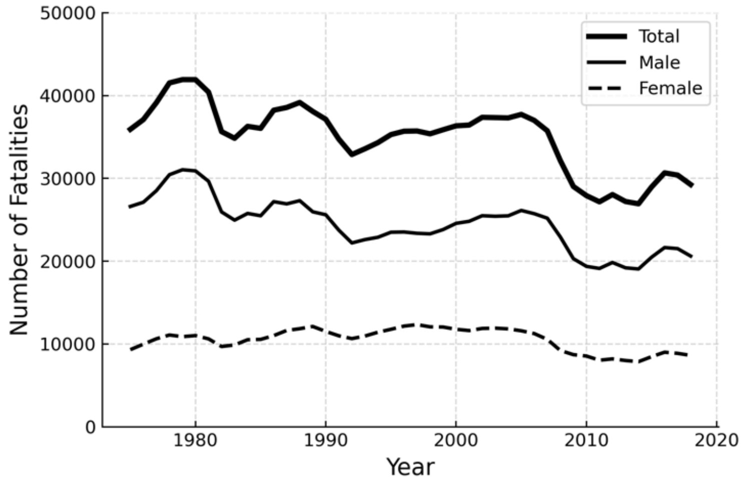

Figure 2: Annual vehicle occupant fatalities have decreased overall since 1975, yet fatalities among females have remained largely the same year-to-year.

METHODS

Data Pre-Processing

Data were downloaded from the US Department of Transportation – National Highway Traffic Safety Administration Fatality Analysis Reporting System (NHTSA FARS) FTP directory. Data analyses were performed using Python v.3.7.7, with packages installed and managed using Anaconda v4.8.2 on MacOS 10.15.3. Files were downloaded using pooch (Uieda et. al., 2020), and processed using tools in SciPy (Virtanen et. al., 2020), namely Pandas, primary package used for data analysis (McKinney, 2010, 2011), Dask for parallelised processing (Virtanen et al., 2020), Matplotlib (Hunter, 2007) and Seaborn (Waskom, 2021) for visualisations, pyjanitor for cleanup and recoding of fields (Ma et. al., 2019), and missingno (Bilogur, 2018) for preliminary visualisation of missingness in the dataset.

Data preprocessing was performed using the fars-cleaner package, produced by the authors for this work. Preprocessing merged changes in the FARS dataset over the last 50 years, adjusting outdated and modified codes using the FARS Analytical User Manual as reference (National Highway Traffic Safety Administration, 2019). The bulk of the data analysis was performed using parametrized (with Papermill) Jupyter notebooks (Kluyver et. al., 2016; Perez and Granger, 2007).

Double Pair Comparison

The double pair comparison method developed by Evans (Evans, 1986) isolates specific features of fatality risk in a crash on a comparable basis. One group of vehicle occupants is selected as the subject occupants, and another is selected as the control. The assessment of fatality risk is performed across the control group. To illustrate the use of the double pair method to determine the relative fatality risk of female vs. male drivers, two sets of crashes are selected with a consistent control occupant (for example, male passengers seated in the front right seat of the vehicle) so that

A = Number of female drivers killed in vehicles with a control occupant.
B = Number of control occupants killed in vehicles with a female driver.
C = Number of male drivers killed in vehicles with a control occupant.
D = Number of control occupants killed in vehicles with a male driver.

From these counts, the relative risk of fatality for a female driver vs. the control occupant, $r_1 = \frac{A}{B}$, and the relative risk of fatality for a male driver vs. the control occupant, $r_2 = \frac{C}{D}$, were used to derive the female vs. male relative risk, $R = \frac{r_1}{r_2}$. In this technique, the control occupants (**B** and **D**) are eliminated from the calculation of relative risk, so male and female subject occupants can be compared. The original double pair method determines standard error in the estimates of R (ΔR):

$$\Delta R = R\sqrt{\sigma_\mu^2 + \frac{1}{A} + \frac{1}{B} + \frac{1}{C} + \frac{1}{D}}$$

where σ_μ is an estimate of "intrinsic uncertainty," set to either 0.05 or 0.1 as in (Leonard Evans, 1986; Evans, 1988). The introduction of this constant term forces all variance estimates to be similar, regardless of the pointwise variance in each estimate. Since these are used in the weighted summaries, results may be biased by granting larger weight to samples with more uncertainty. To mitigate this bias, an alternative method was used for describing the variance of the risk ratios (Cummings et. al., 2003). To correct this deficiency, the counts were further stratified using:

A = Number of female drivers killed in vehicles with a control occupant (also killed).
B = Number of female drivers killed in vehicles with a control occupant (not killed).
C = Number of control occupants killed in vehicles with a female driver (not killed).
E = Number of male drivers killed in vehicles with a control occupant (also killed).
F = Number of male drivers killed in vehicles with a control occupant (not killed).
G = Number of control occupants killed in vehicles with a female driver (not killed).

Note that the original variables **A**, **B**, **C**, and **D** can be derived as the sum of the new counts detailed above (**A-G)**. Now, the relative risk ratio is

$$R = \frac{A + B/_{A + C}}{E + F/_{E + G}}$$

As in the work by Cummings variance for the log of the relative risk ratio is

$$\Delta R = \frac{\left[\left(A \times (A + B + C) + (B \times C)\right) \times (F + G)\right] + \left[\left(E \times (E + F + G) + (F \times G)\right) \times (B + C)\right]}{(A + B) \times (A + C) \times (E + F) \times (E + G)}$$

Weighted risk ratios ($\bar{R}$) and weighted estimates of variance ($\Delta\bar{R}$) are

$$\bar{R} = \exp\left(\frac{\sum(\ln R \times 1/\Delta R)}{\sum 1/\Delta R}\right)$$

$$\Delta\bar{R} = \frac{1}{\sum 1/\Delta R}$$

Confidence intervals (95%) are derived with a bootstrap method, sampling with replacement many times and calculating values for $\bar{R}$ and $\Delta\bar{R}$ for each new sample set (Cummings et al., 2003). This procedure was replicated 5000 times for each weighted average. Each bootstrap run produces a distribution of results, from which the 95% confidence interval is derived.

Cases were selected with at least two occupants and with at least one fatality in the vehicle. Fatality was determined as coded within FARS and includes those declared dead at the crash and within 30 days from crash-related causes. Age ranges were examined in five-year periods for subject occupants, while control occupants were grouped as in previous analyses: ages 16–24, 25–34, 35–54, and 55+ (Evans, 2001). Table 1 shows the subject breakdown for 1975–2018 without airbag deployment (*n*=321,320).

Table 1: Distribution of 321,320 Fatally Injured Subject Occupants, 1975–2018, no airbag deployment

Vehicle	**Subject Occupant**	**Restraint Use**	**Female Fatalities**	**Male Fatalities**	**Total**
Car	Driver	Unbelted	16,994	54,196	71,190
Car	Right-front passenger	Unbelted	38,478	43,043	81,521
Car	Driver	Belted	8,060	15,424	23,484
Car	Right-front passenger	Belted	19,296	11,270	30,566
Car	Left-rear passenger	Unbelted	4,395	6,298	10,693
Car	Right-rear passenger	Unbelted	5,012	7,270	12,282
Light truck	Driver	Unbelted	3,994	24,882	28,876
Light truck	Right-front passenger	Unbelted	11,015	18,889	29,904
Light truck	Driver	Belted	2,071	6,338	8,409
Light truck	Right-front passenger	Belted	5,148	4,166	9,314
Light truck	Left-rear passenger	Unbelted	696	1,092	1,788
Light truck	Right-rear passenger	Unbelted	790	1,153	1,943
Motorcycle	Passenger	Helmeted	3,687	1,211	4,898
Motorcycle	Passenger	Unhelmeted	4,496	1,956	6,452
Totals			**124,132**	**197,188**	**321,320**

Fatal crash cases were grouped by vehicle type (car, truck, motorcycle), passenger seating position (front right seat, rear left seat, etc.), seat-belt use, and number of vehicles involved in the crash. In each analysis, a weighted average of the value for R was taken within each driver age subset to find the overall risk for a driver. Further, driver age was grouped in five-year chunks to increase sample size for additional analyses.

To compare with Evans (Evans, 2001) original study, initial analyses used cases with no airbag deployment. The various analysis combinations are described in Table 2. Additional analyses were performed with matched airbag deployment, or in cases where both the subject and the control occupant experienced the same airbag deployment at their seating position (deployment or no deployment). Cases with unknown airbag deployment were excluded. Additional analyses were performed to investigate the robustness of the results, including rural vs. urban cases, cases with and without alcohol or drug involvement, and cases involving only late-model vehicles.

Table 2: Analyses performed (Matched condition indicates airbag deployment for both subject and control, or for neither occupant)

Years	**Vehicle Type**	**Subject Occupants**	**Air Bag Deployment**
1999–2018	Car, Light Truck, Motorcycle	Driver, front-right, rear-left/right passengers	None
2010–2018	Car, Light Truck, Motorcycle	Driver, front-right, rear-left/right passengers	None
1975–2018	Car, Light Truck, Motorcycle	Driver, front-right, rear-left/right passengers	None
1975–2018	Car, Light Truck	Driver, front-right, rear-left/right passengers	Deployed
1999–2018	Car, Light Truck	Driver, front-right, rear-left/right passengers	Deployed
2010–2018	Car, Light Truck	Driver, front-right, rear-left/right passengers	Deployed
2015–2018	Car, Light Truck	Driver, front-right, rear-left/right passengers	Deployed
1975–2018	Car, Light Truck	Driver, front-right passengers	Matched
1999–2018	Car, Light Truck	Driver, front-right passengers	Matched
2010–2018	Car, Light Truck	Driver, front-right passengers	Matched
2015–2018	Car, Light Truck	Driver, front-right passengers	Matched

RESULTS

To best illustrate the results produced by the double pair method, consider the case of belted drivers, aged 23–27 years (25-year-old (yo) drivers), in passenger cars, using belted front-right seat passengers as a control, matching airbag deployment, in crashes between 2010 and 2018. These results are given in Table 3 for each sex/age cohort under these conditions. Taking the weighted average across the controls presented in Table 3 for each set of subject occupants (belted/unbelted drivers and front-right passengers, with and without airbag deployment, in cars and light trucks), we produce Table 4, for 25yo vehicle occupants.

From Table 3, for example, belted 25yo female drivers with airbag deployment were 8.1% (95% CI -8.6 – 21.9%) more likely to die in a fatal crash than male drivers under the same conditions from 2010 to 2018. Taken with all other cases for 25yo female occupants, the overall risk (as shown in Table 4), is 10.7% (95% CI 9.7 – 11.6%) higher risk for females. This value is plotted as the corresponding value in Figure 3. Note that the value is plotted at 23yo, not directly at 25yo. In different analyses, slightly different age bins were used to better capture the distribution of subject and control occupants. Bins were combined into five-year chunks, using the midpoint of the original groupings to determine membership in the final bin. Data were then plotted at the midpoint of each bin. Younger female drivers have a higher fatality risk than younger males when driving or sitting in the front passenger seat, regardless of seat-belt use. Figure 4 shows the results of repeating this analysis for all crashes from 1975 to 2018. Figure 5 is representative of the confidence interval estimation utilizing the bootstrap method, and shows the distribution of R values calculated with 5000 repetitions for passenger car occupants.

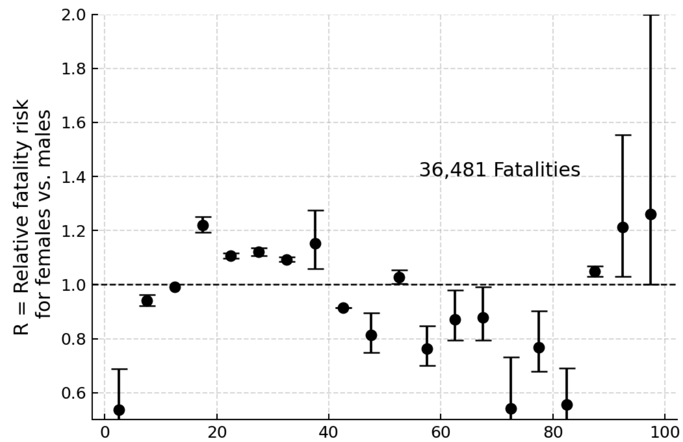

Figure 3: All passenger car and light truck fatalities, 2010-2018, matched airbag deployment condition.

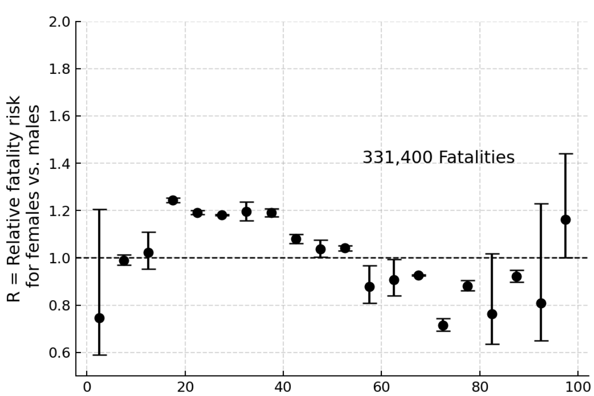

Figure 4: All passenger car and light truck fatalities, 1975–2018, matched airbag deployment condition.

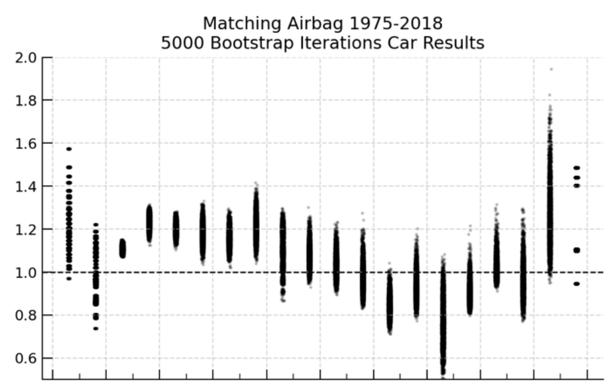

Figure 5: Bootstrap distributions, passenger car fatalities 1975-2018, matched airbag deployment condition.

In line with previous findings (Evans, 1988, 2001), the relative risk for female drivers is higher than that for males until around age 60. The highest difference in risk is between ages 20 and 40, with females ~20% more likely to die in a crash. This general trend is apparent when isolating vehicle type, seat-belt use, number of vehicles involved, urban vs. rural road type, and airbag deployment.

Table 3: Female vs. Male Fatality Risk, Belted 25YO Car Drivers, Airbag Deployment, 2010–2018 (Note: Control occupants are belted, front-right passengers)

Control Occupant, Age	Fatalities				Ratios			
	A	**B**	**C**	**D**	**r_1 = A/B**	**r_2 = C/D**	**R = r_1/r_2**	**ln ΔR (Eq. 13)**
Male Passenger, 16–24yo	24	26	132	118	0.923	1.119	0.825	0.059
Male Passenger, 25–34yo	54	48	83	87	1.125	0.954	1.179	0.040
Male Passenger, 35–54yo	12	11	21	27	1.091	0.778	1.403	0.129
Male Passenger, 55+yo	1	6	2	4	0.167	0.500	0.333	1.917
Female Passenger, 16–24yo	30	24	110	111	1.250	0.991	1.261	0.075
Female Passenger, 25–34yo	24	28	59	77	0.857	0.766	1.119	0.065
Female Passenger, 35–54yo	12	25	12	20	0.480	0.600	0.800	0.190
					Weighted average		1.081	
					95% Confidence Interval		[0.914, 1.219]	

Table 4: Values of R, Fatality Risk to 25yo Females Compared to 25YO Males, All Subject Occupants, 2010-2018, Matched Airbag Deployment

Vehicle	**Subject Occupant**	**Female Fatalities**	**Male Fatalities**	**Total Fatalities**	**R**	**95% CI**
Car	Unbelted drivers, airbag deployed	482	1,446	1,928	0.945	[0.835, 1.079]
Car	Unbelted right front passengers, airbags	938	1,137	2,075	1.379	[1.211, 1.519]
Car	Belted drivers, airbags	1,747	3,330	5,077	1.081	[0.902, 1.223]
Car	Belted right front passengers, airbags	3,781	2,391	6,172	0.985	[0.891, 1.164]
Car	Unbelted drivers, no airbags	300	852	1,152	1.254	[0.932, 1.960]
Car	Unbelted right front passengers, no airbags	477	678	1,155	1.051	[0.640, 1.354]
Car	Belted drivers, no airbags	653	1,484	2,137	1.280	[1.011, 2.453]

Car	Belted right front passengers, no airbags	1,386	1,079	2,465	1.169	[0.775, 1.450]
Light truck	Unbelted drivers, airbags	211	666	877	1.292	[0.856, 1.823]
Light truck	Unbelted right front passengers, airbags	433	549	982	1.058	[0.854, 1.380]
Light truck	Belted drivers, airbags	661	1,754	2,415	0.857	[0.675, 1.182]
Light truck	Belted right front passengers, airbags	1,643	960	2,603	1.075	[0.830, 1.441]
Light truck	Unbelted drivers, no airbags	436	1,502	1,938	0.976	[0.671, 1.441]
Light truck	Unbelted right front passengers, no airbags	759	1,081	1,840	1.476	[1.238, 2.013]
Light truck	Belted drivers, no airbags	489	1,349	1,838	1.213	[0.961, 1.549]
Light truck	Belted right front passengers, no airbags	964	863	1,827	0.843	[0.727, 1.203]
			Weighted Average		1.107	[1.097, 1.116]
Totals		**15,360**	**21,121**	**36,481**		

DISCUSSION

Despite significant advances in vehicle safety since 1975, female vehicle occupants involved in fatal crashes have a higher risk of death compared to males in matched circumstances. These results show that when in a fatal crash, a younger female occupant is approximately 20% more likely to suffer a fatal injury than a male occupant of the same age, regardless of seating position, airbag deployment, or seat-belt usage.

We have investigated several potential covariates that might explain these findings, including rural vs. urban crashes, vehicle mass differences by sex, drug and alcohol use by drivers, number of passengers, and number of vehicles involved. Figure 6 shows the distribution of some of these covariates. The distribution of vehicles driven by females is similar to that for males, with a slightly higher proportion of female drivers using cars with vehicle masses of ~2,500 lb, but both sexes use heavier vehicles at the same rate. While male drivers are more frequently involved in single-car crashes compared to female drivers, the proportion of crashes involving multiple vehicles is similar across sex. The distribution by sex of the number of occupants in a vehicle are similar as well. The lack of a large qualitative difference in these covariates would imply limited effect on the relative risk to drivers.

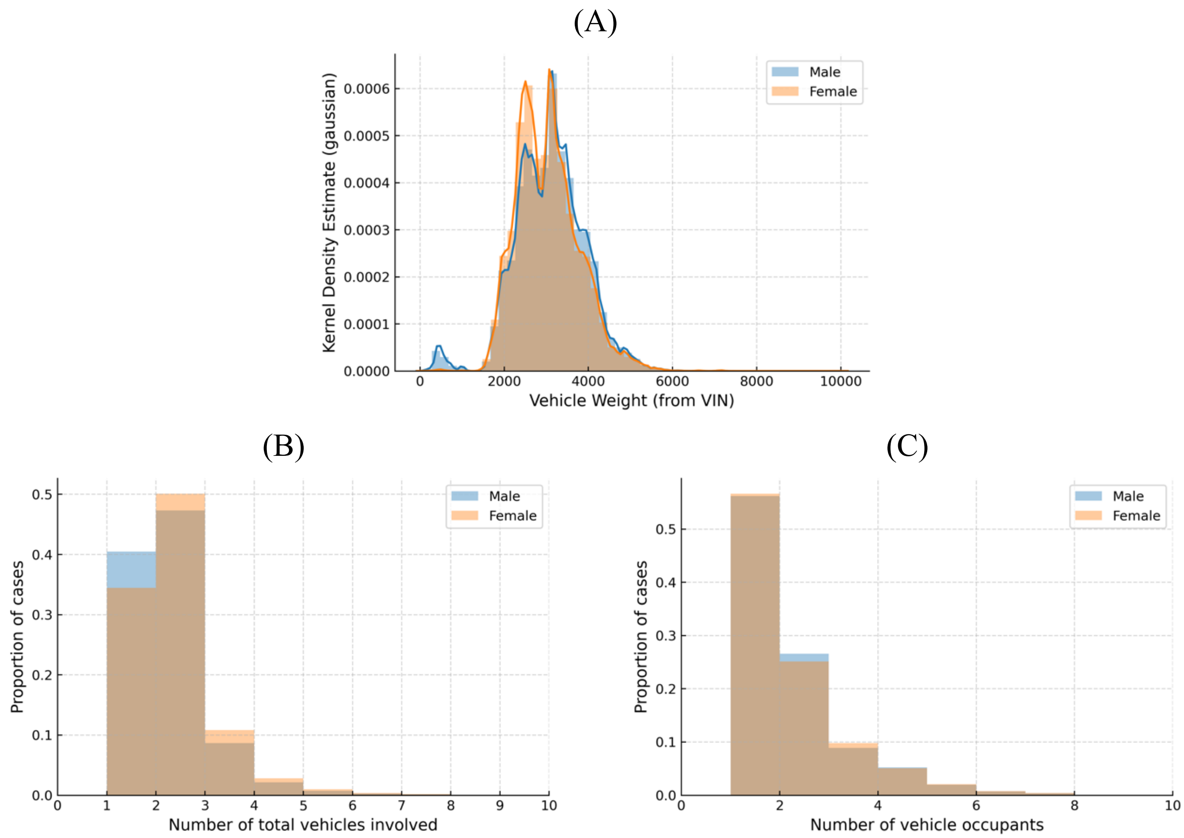

Figure 6: Breakdown of potential covariates for male vs. female drivers: (A) Gaussian kernel density estimate of vehicle weight; (B) normalized histogram of vehicles involved in a crash; (C) normalized histogram of number of vehicle occupants.

To verify this conclusion, the above analyses were applied to subsets of vehicle occupants involved in one-, two- and multi-vehicle crashes, as well as cases with one, two, or several vehicle occupants. As there are systematically more crashes of higher severity in rural areas (Burgess, 2005), rural and urban crash locations were broken out into separate analyses as well, to serve as a surrogate for crash severity. The trend described above is robust across all of these cases.

The results presented here are similar to those presented by Evans in 2001, for FARS data through 1998. Notably, the inclusion of data from before 2010 in the analysis presented in Figure 3 mainly reduces the error assessment, without qualitatively altering the overall characteristics of the risk curve.

Physiological/anatomical differences are difficult to explore within the FARS dataset, but may provide an explanation. Current vehicle crashworthiness tests do not use a crash dummy representative of the average adult female (Linder and Svensson, 2019). The current US regulations require crash testing primarily with the 50th percentile Hybrid III (HIII) adult male ATD, with a few tests adding in the 5th percentile adult female as a passenger. Since the 5th percentile female ATD is primarily a dimensionally scaled version of the HIII male (Saul et. al.)

the physiological and anatomical differences between the sexes may not be completely reproduced in testing methods.

These results are not completely consistent with other well-described trends related to sex differences in injury. Females are more likely than males to suffer fractures past the age of 60 due to osteoporosis, and experience bone loss at an earlier age (Alswat, 2017), and females tend to have greater age related bone density loss than males (Hannan et. al., 2000). Increased fracture risk in elderly females, therefore, cannot explain the observed trends. Without detailed injury report data (not available within FARS), the cause of death cannot be determined for each case. This information would be valuable in parsing the differences between male and female crash survivability. We posit that there may be unobserved trends in the injury patterns, and therefore outcomes, between male and female occupants. These trends may be the result of unintentional vehicle design issues, or a potentially unexplained lack of biofidelity in the ATDs used in testing.

CONCLUSIONS

There is an age-dependent increased risk for females compared with males in automobile and other vehicle crashes with younger females at higher risk of death than males. This difference is robust across vehicle type, number of passengers, crash severity and has persisted since the mid 1970s. This increased risk peaks in the mid 30s age and is not attributable to vehicle weight, belt usage, or airbag deployment. Given recent advances in vehicle safety technologies, and a general trend towards fewer vehicle fatalities each year, the persistence of this trend is disturbing and requires further study. More work should be done to better understand the differences in female crash fatal and nonfatal outcomes compared to males. The results presented highlight a decades-long issue that has clearly not been fully addressed. Known underlying physiological differences are insufficient to explain the phenomenon described in this study and should be examined to ensure all occupants are well-protected in crash situations.

REFERENCES

AHANGARI, H., ATKINSON-PALOMBO, C., & GARRICK, N.W. (2016). Progress towards zero, an international comparison: Improvements in traffic fatality from 1990 to 2010 for different age groups in the USA and 15 of its peers. *J Safety Res, 57*, 61-70.

ALSWAT, K.A. (2017). Gender Disparities in Osteoporosis. *Journal of Clinical Medicine Research, 9*(5), 382-387.

BILOGUR, A. (2018). Missingno: a missing data visualization suite. *The Journal of Open Source Software, 3*(22).

BOSE, D., SEGUI-GOMEZ, M., & CRANDALL, J.R. (2011). Vulnerability of female drivers involved in motor vehicle crashes: an analysis of US population at risk. *American Journal of Public Health, 101*(12), 2368-2373.

BURGESS, M. (2005). Contrasting Rural and Urban Fatal Crashes 1994-2003. DOT HS 809 896. Retrieved from Washington, DC:

CUMMINGS, P., MCKNIGHT, B., & WEISS, N.S. (2003). Matched-pair cohort methods in traffic crash research. *Accident Analysis & Prevention, 35*(1), 131-141.

EVANS, L. (1986). Double pair comparison—A new method to determine how occupant characteristics affect fatality risk in traffic crashes. *Accident Analysis & Prevention, 18*(3), 217-227.

EVANS, L. (1988). Risk of fatality from physical trauma versus sex and age. *Journal of Trauma, 28*(3), 368-378.

EVANS, L. (2001). Female compared with male fatality risk from similar physical impacts. *Journal of Trauma, 50*(2), 281-288.

FUCCI, N.M.A. (2018). Summary of Travel Trends: 2017 National Household Travel Survey. FHWA-PL-18-019. Retrieved from Washington, DC:

HANNAN, M.T., FELSON, D.T., DAWSON-HUGHES, B., TUCKER, K.L., CUPPLES, L.A., WILSON, P.W., et al. (2000). Risk factors for longitudinal bone loss in elderly men and women: the Framingham Osteoporosis Study. *J Bone Miner Res, 15*(4), 710-720.

HUNTER, J.D. (2007). Matplotlib: A 2D Graphics Environment. *Computing in Science & Engineering, 9*(3), 90-95.

KAHANE, C.J. (2015). Lives Saved by Vehicle Safety Technologies and Associated Federal Motor Vehicle Safety Standards, 1960 to 2012 – Passenger Cars and LTVs – With Reviews of 26 FMVSS and the Effectiveness of Their Associated Safety Technologies in Reducing Fatalities, Injuries, and Crashes. DOT HS 812 069. Retrieved from Washington, DC:

KLUYVER, T., RAGAN-KELLEY, B., PÉREZ, F., GRANGER, B.E., BUSSONNIER, M., FREDERIC, J., et al. (2016). Jupyter Notebooks - a publishing format for reproducible computational workflows. ELPUB.

LINDER, A., & SVENSSON, M.Y. (2019). Road safety: the average male as a norm in vehicle occupant crash safety assessment. *Interdisciplinary Science Reviews, 44*(2), 140-153.

MA, E.J., BARRY, Z., ZUCKERMAN, S., & SAILER, Z. (2019). pyjanitor: A Cleaner API for Cleaning Data. Proceedings of the 18th Python in Science Conference.

MAYHEW, D.R., FERGUSON, S.A., DESMOND, K.J., & SIMPSON, H.M. (2003). Trends in fatal crashes involving female drivers, 1975-1998. *Accident Analysis & Prevention, 35*(3), 407-415.

MCKINNEY, W. (2010). Data structures for statistical computing in python. Proceedings of the 9th Python in Science Conference, Austin, TX.

MCKINNEY, W. (2011). pandas: a foundational Python library for data analysis and statistics. *Python for High Performance and Scientific Computing, 14*.

NATIONAL HIGHWAY TRAFFIC SAFETY ADMINISTRATION. (2018). Fatality Analysis Reporting System. In. Washington, DC.

NATIONAL HIGHWAY TRAFFIC SAFETY ADMINISTRATION. (2019). Fatality Analysis Reporting System (FARS) Analytical User's Manual, 1975-2018. DOT HS 812 827. Retrieved from Washington, DC:

PEREZ, F., & GRANGER, B.E. (2007). IPython: A System for Interactive Scientific Computing. *Computing in Science & Engineering, 9*(3), 21-29.

ROMANO, E., KELLEY-BAKER, T., & VOAS, R.B. (2008). Female involvement in fatal crashes: Increasingly riskier or increasingly exposed? *Accident Analysis & Prevention, 40*(5), 1781-1788.

SAUL, R.A., PRITZ, H.B., MCFADDEN, J., BACKAITIS, S.H., HALLENBECK, H., & RHULE, D. (2001). Description and performance of the Hybrid III three-year-old, six-year-old and small female test dummies in restraint system and out-of-position air bag environments. 17th Enhanced Safety of Vehicles Conference, Amsterdam, The Netherlands.

SIVAK, M. (2013). Female Drivers in the United States, 1963–2010: From a Minority to a Majority? *Traffic Injury Prevention, 14*(3), 259-260.

TSAI, V.W., ANDERSON, C.L., & VACA, F.E. (2008). Young female drivers in fatal crashes: recent trends, 1995-2004. *Traffic Injury Prevention, 9*(1), 65-69.

UIEDA, L., SOLER, S., RAMPIN, R., VAN KEMENADE, H., TURK, M., SHAPERO, D., et al. (2020). Pooch: A friend to fetch your data files. *Journal of Open Source Software, 5*(45).

VIANO, D.C., & PARENTEAU, C.S. (2016). Effectiveness of the revision to FMVSS 301: FARS and NASS-CDS analysis of fatalities and severe injuries in rear impacts. *Accident Analysis & Prevention, 89*, 1-8.

VIRTANEN, P., GOMMERS, R., OLIPHANT, T.E., HABERLAND, M., REDDY, T., COURNAPEAU, D., et al. (2020). SciPy 1.0: fundamental algorithms for scientific computing in Python. *Nat Methods, 17*(3), 261-272.

WASKOM, M.L. (2021). seaborn: statistical data visualization. *Journal of Open Source Software, 6*(60), 3021.

WEBB, C.N. (2020). Motor vehicle traffic crashes as a leading cause of death in the United States, 2016 and 2017. DOT HS 812 927. Retrieved from Washington, DC:

WORLD HEALTH ORGANIZATION. (2018). Global status report on road safety 2018. *Geneva, Switzerland*.