Sex-Related Vehicle and Crash Differences and their Potential to Confound Relative Injury Risk Analyses

Matthew L. Brumbelow

Abstract Many studies have found that females are at higher risk of injury than males in similar crashes. However, vehicle selection and crash characteristics differ by sex. This study was designed to investigate vehicle and crash differences between males and females and the extent to which they may confound estimates for relative injury and fatality risk. Results indicate that crash-involved female drivers were more often driving cars, SUVs and minivans, while males were more often driving pickups. In crashes involving one female and one male driver, the female was more often in side- and rear-struck vehicles and had a median curb weight disadvantage of 104 kg across all crash types. Vehicle differences represented a 75% increase in relative female fatality odds in head-to-head crashes. In terms of injury risk estimation, differences in crush-based delta-V were found to bias results towards overreporting relative risk for females, although this bias has decreased over time. Results demonstrate the importance of accounting for vehicle and crash differences between women and men when seeking to identify relative injury risk. They also stress the need to address vehicle incompatibility in the U.S. vehicle fleet as part of improving outcomes for females.

Keywords Compatibility, delta-V, females, injury risk, males.

I. INTRODUCTION

In 2020, females accounted for 28% of U.S. passenger-vehicle driver fatalities, the lowest proportion since 1991 [1]. Much of the overrepresentation of males among driver fatalities is due to exposure, with males driving farther on average each year [2] and engaging in more risk-taking behaviours [1]. Therefore, to evaluate the effectiveness of existing crashworthiness countermeasures or the need for improvements, studies of different injury or fatality outcomes must account for exposure-related factors. Differences in crash rate can be controlled by restricting analyses to a set of crashes that meet certain inclusion criteria. However, even within a crash dataset, sex-related vehicle and crash differences have the potential to confound injury risk analyses by sex.

Different strategies can be used to adjust for crash severity in studies of relative risk between females and males. Perhaps the simplest approach is to compare outcomes for a set of crashes that involve one female and one male driver, while controlling for other characteristics of the vehicles or occupants. Kahane [3] compared head-on fatal crashes involving two belted drivers to evaluate the effect of different measures taken from the U.S. New Car Assessment Program (NCAP) for 1979–1991 model years. While his study was not focused on the driver sex difference, Kahane found that female drivers had 44% higher odds of fatality than male drivers while controlling for vehicle curb weight and NCAP dummy readings of head injury criterion, chest acceleration and femur loads. Only 3% of the study vehicles were equipped with front airbags. Later, Farmer [4] conducted a similar analysis of vehicles evaluated in the Insurance Institute for Highway Safety (IIHS) moderate overlap test. In head-on crashes involving two vehicles of the same type (e.g. car-to-car, pickup-to-pickup), he found a non-significant 24% lower odds of fatality for females than males, while controlling for vehicle weight and IIHS rating. While all vehicles were equipped with front airbags, there were too few crashes involving two belted drivers in rated vehicles for analysis, so fatality odds ratios included both belted and unbelted drivers. Using Swedish crash data, Kullgren *et al.* [5] found that female drivers had higher non-fatal injury risks than male counterparts in the same crash. Since crash configuration was not controlled, the results would be affected if, for example, males more often drove the front-striking vehicle in front-to-side or front-to-rear configurations.

The double-pair comparison method has been used to expand analyses to unpaired crashes while using other

M. L. Brumbelow (e-mail: mbrumbelow@iihs.org; tel: +1-434-985-4600) is a Senior Research Engineer at the Insurance Institute for Highway Safety in Virginia, USA.

occupants in each vehicle as controls. Several studies using this method to analyse the Fatality Analysis Reporting System (FARS) have reported increased fatality risks for females relative to males in the range of 17–28% at least for younger ages [6-8]. Most recently, a double-pair study by Noh *et al.* [9] found relative fatality risk for females was lowest for belted occupants in the newest vehicles, with only a 0.3% higher risk for female drivers in 2010–2020 models. Across seating positions in the newest vehicles, relative female risks were lowest in far-side (1.5%), followed by front (7.4%) and near-side (26.9%) crashes.

The double-pair method greatly increases sample sizes over assessments of individual driver pairs in the same crash, but it has a few limitations. First, it assumes that the sex of the control occupants does not affect the fatality outcome of the case occupant [6]. While this may generally be true in front crashes, typical mass and size differences between females and males may influence outcomes in side impacts and rollovers, an effect that could be amplified when unbelted occupants are used as controls, as they typically have been. This may help explain the elevated relative risk for females in near-side crashes reported by Noh *et al.* [9] or the results of Abrams and Bass [8], which were not stratified by crash type. Studies that do not control for crash type are subject to the additional prospect that the different crash distributions for female and male drivers [10] may affect the relative injury risk between case and control occupants. Other limitations of all fatal crash data analyses arise from their dependence on police-reported values for belt use and airbag deployment, which often are used as grouping variables in double-pair studies. Furthermore, as FARS does not include occupant height and mass information, these studies cannot identify whether dissimilar outcomes are due to anthropometric differences between women and men or other limits of crashworthiness technologies in providing equal levels of protection. Finally, specific injury types or crash configurations cannot be evaluated using fatality data alone.

An alternative to the double-pair approach is to stratify outcomes by some measure of crash severity. Malczyk and Kröling [11] created a binary variable from maximum vehicle exterior deformation to represent severity in analyses of German insurers' crash data. They reported that sex-related differences in the risk of moderate injury (Abbreviated Injury Scale (AIS) ≥2) were not significant. However, their findings that crash rurality did have a significant effect on injury outcome and that rurality differed systematically by driver sex suggest the presence of uncontrolled crash severity differences between females and males. (The data also were insufficient to control for the possibility that belt use differed by sex.) As an alternative to a single crush measurement, the velocity change (delta-V) in a collision is a commonly used severity measure and is a coded variable in the National Automotive Sampling System-Crashworthiness Data System (NASS-CDS) and its successor, the Crash Investigation Sampling System (CISS). In these databases, delta-V is calculated by the WinSMASH algorithm using vehicle damage profile measurements along with other characteristics of the vehicle and impact. Relative to FARS, the detailed anthropometric and injury information in NASS-CDS and CISS also enable more granular evaluation of the unique injury risks faced by females and males. In one such study of frontal crashes, Forman *et al.* [12] found that females had 73% higher odds of serious injury (AIS≥3) and 142% higher odds of moderate injury (AIS≥2) than males. Among the body regions studied, the relative odds for females was greatest for ankle injuries (281%) and smallest for skull fractures (-53%).

Assessment of relative injury risk using WinSMASH delta-V carries its own limitations. First, because crush measurements are required for each case vehicle, data collection is more challenging and resulting sample sizes often are smaller than studies of FARS. Second, even when crush measures are available, the accuracy of the WinSMASH estimate can be affected by several factors. A previous study found that differences between delta-V as estimated by WinSMASH and values downloaded from event data recorders (EDRs) were highly dependent on the degree of lateral and vertical overlap in front crashes [13]. These and other variables have the potential to confound the analysis of sex differences if men and women drive different vehicle types or tend to be involved in different crash types. Brumbelow and Jermakian [14] reported that, while controlling for WinSMASH delta-V, all estimated injury odds for females relative to males decreased when some additional constraints were placed on crash and vehicle differences. Furthermore, WinSMASH values in NASS-CDS and CISS can be missing for different reasons, some of which may be due to the type of crash. For example, when a case vehicle strikes a yielding object other than another passenger vehicle, the object absorbs an unknown portion of the crash energy, and the delta-V cannot be estimated. If there are sex-related differences in vehicles and crash type, imputation of missing delta-V is another potential source of biased injury risk estimates. Finally, even when delta-V is reported accurately, it does not capture other information about the crash pulse that differs by vehicle type [15] and that may also affect injury [16].

It is important to identify how non-physiological risk factors may affect injury risk estimates for females and males in order to encourage the most robust and effective countermeasures. For example, resolving injury risk disparities due primarily to vehicle mass and size will require a different approach than those resulting from restraint systems overly optimised to specific crash test configurations. This study was designed to investigate vehicle and crash differences between men and women and the extent to which they may confound estimates for relative injury and fatality risk.

II. METHODS

Vehicle and Crash Type Differences

The Crash Report Sampling System (CRSS) is a survey-weighted sample of police-reported crashes that occur in the USA. It is maintained by the National Highway Traffic Safety Administration (NHTSA) and data collection began in 2016. For this study, years 2016–2020 were queried for crashes involving a passenger vehicle with known values for the driver sex and sufficient characters from the Vehicle Identification Number (VIN) to decode the vehicle year, make and model. VIN decoding was performed by the Highway Loss Data Institute (HLDI) using its VINDICATOR software [17] and vehicle curb weights were taken from the HLDI database. Where VIN data were insufficient to determine the trim level and associated curb weight, or where curb weights were unknown for a certain trim level, the median curb weight for other vehicles of the same make, model and model year was used. Vehicle type, make/model, age (crash year minus model year) and curb weight differences were quantified by driver sex. To investigate differences in crash role and partner-vehicle curb weight, CRSS data were further filtered to identify two-vehicle crashes involving one female and one male driver in which the struck vehicle was towed due to damage. Struck vehicles were defined as side- or rear-impacted in front-to-side or front-to-rear crashes, and either vehicle in front-to-front crashes. Front impacts were defined as initial impact locations of 11, 12, or 1, rear impacts as 5–7, and side impacts as 2–4, 8–10, 61–63 or 81–83. Crashes in which either vehicle rolled over were excluded, as were crashes in which the most harmful event was coded as anything other than the impact with the partner vehicle (this includes non-collision events such as fires). Since analyses did not depend on injury outcome, no belt use criteria were applied. The “survey” package in the R programming language was used to estimate all proportions and curb weight distributions while accounting for sampling weight [18].

Fatal head-to-head crashes were evaluated for the potential of curb weight and vehicle type to influence estimated differences by driver sex. The Fatality Analysis Reporting System (FARS) is a census of all police-reported fatal crashes in the USA. FARS years 1995–2020 were queried for two-vehicle, front-to-front crashes (11-, 12-, or 1-o’clock impact locations) where both drivers were reported belted and one driver was killed. Model years were restricted to 1995 and later, after most vehicles were equipped with driver front airbags [19]. Only passenger vehicles were included, and the vehicle type and curb weight were taken from the HLDI VIN-decoded values described above. Missing curb weight values were filled in with the median from vehicles of the same make, model and model year (or within five years when there were no exact matches). Vehicles that were still missing curb weight values were excluded. Conditional logistic regression was used to estimate the effect of driver sex while controlling for differences in the ages of the two drivers, the model years of the two vehicles and airbag deployment status. Results from this baseline model were compared with results from three additional models. In the first, the difference in the logs of each vehicle’s curb weight was included as a covariate. In the second, the curb weight covariate was maintained, and crashes were restricted to those involving vehicles of the same type (i.e. both vehicles were either cars, SUVs, minivans, full-sized vans, or pickups). Finally, crashes were additionally restricted to those where the difference in model years of the two vehicles was five years or less. All conditional regression models were fit using the “survival” package in R [20].

Potential Bias in Injury Odds Estimates from Crash Severity Errors

Delta-V is often used to permit comparisons of injury outcome across different vehicle types and sizes. NASS-CDS and CISS include estimates of longitudinal and lateral delta-V for many case vehicles. Like CRSS, CISS (and NASS-CDS previously) is a survey-weighted sample of police-reported crashes in the USA maintained by NHTSA. However, it is focused only on tow-away crashes involving passenger vehicles. Trained investigators gather additional information about sampled cases, including direct measurements of vehicle damage. These crush measurements are used as inputs in the WinSMASH algorithm, along with other characteristics of the vehicle and impact, to produce delta-V estimates.

Beginning in 2004, NASS-CDS investigators began downloading crash pulses from event data recorders (EDRs) when available, allowing a direct comparison of delta-V measured from vehicle instrumentation with the WinSMASH estimates. In front crash tests, Tsoi *et al.* [21] found that EDR-reported delta-V was an average of 6% lower than measurements taken with laboratory accelerometers, in line with previous studies [22-23]. The accuracy of WinSMASH delta-V relative to EDR-reported values has been studied, and the WinSMASH algorithm has been updated over the years in efforts to improve its accuracy. Hampton and Gabler [24] reported that the 2008 update resulted in values that were much closer to the EDR values than the previous version of WinSMASH. NHTSA reported additional changes as part of the introduction of CISS in 2016 [25]. Any WinSMASH errors that do exist will not bias relative injury risk as long as they are similar for females and males. However, previous studies have found evidence that WinSMASH accuracy varies by crash and vehicle type [13][24], potentially affecting sex effect estimates [14].

For this study, front crashes from 2004 to 2015 NASS-CDS and from 2017 to 2021 CISS were queried for those involving a vehicle with EDR crash-pulse data. (NHTSA investigated a small number of CISS pilot cases in 2016 but did not produce datasets for this year.) While NASS-CDS did not contain dedicated tables for EDR data, EDR information is accessible through the web-based case viewer tool. The “xml2” package in R was used to scrape EDR data from the web pages for each NASS-CDS vehicle [26]. For both NASS-CDS and CISS, the following criteria were used to identify pulses from which the total longitudinal delta-V could be calculated: the minimum delta-V was between -5 km/h and -120 km/h, the slope of the delta-V curve reached a value that was greater than -5 g, and there were non-zero values recorded prior to the minimum delta-V.

Table I outlines the process used to estimate the bias in injury odds estimates due to crash severity errors. Based on changes to the WinSMASH algorithm over time, three distinct calendar year ranges were studied: 2004–

TABLE I

PROCESS FOR ESTIMATING POTENTIAL INJURY ODDS BIAS DUE TO CRASH SEVERITY ERRORS

Excluding effect of WinSMASH imputation

Item	Front crash data	Analysis
1	Vehicles with EDR + WS	Linear regression: effect of driver sex on WS dV, controlling for EDR dV
2	Vehicles with WS, drivers restrained by belt + airbag	Weighted logistic regression: effect of WS dV on injury, controlling for driver age Separate estimates for AIS≥2 and AIS≥3 injury
3	Simulated	MC: simulate 1,000,000 female and male driver pairs with WS difference for each pair sampled from distribution reported in #1
4	Simulated, results of #3	MC: for each driver pair, convert WS difference to estimated AIS≥2 and AIS≥3 injury odds ratio from distributions reported in #2 Report final odds ratio as mean of all samples and 95% CI as 2.5 and 97.5 percentile values

Including effect of WinSMASH imputation

Item	Front crash data	Analysis
5	Vehicles with EDR	Linear regression: effect of driver sex, WS missing status, and interaction on EDR dV
6	Drivers restrained by belt + airbag	Proportion of missing WS by driver sex
7	Simulated	MC: simulate 1,000,000 female and male driver pairs with WS difference for each pair sampled from distribution reported in #1 For proportions of each sex from #6, first adjust EDR dV difference by amount sampled from distributions reported in #5
8	Simulated, results of #7	MC: for each driver pair, convert WS difference to estimated AIS≥2 and AIS≥3 injury odds ratio from distributions reported in #2 Report final odds ratio as mean of all samples and 95% CI as 2.5 and 97.5 percentile values

Note: AIS = Abbreviated Injury Scale; CI = confidence interval; dV = delta-V; EDR = event data recorder; MC = Monte Carlo; WS = WinSMASH.

2007 NASS-CDS; 2008–2015 NASS–CDS; and 2017–2021 CISS. For each range of years, linear regression was used to construct a parallel slopes model for WinSMASH delta-V based on EDR delta-V and driver sex (Table I, Item 1). The estimates for driver sex reflect the difference in WinSMASH estimates due to sex-related vehicle or crash type differences when EDR delta-V is the same. Using logistic regression (Item 2) and Monte Carlo simulation (Items 3 and 4), these estimates were then translated to equivalent injury odds ratios that would be attributed to driver sex in analyses stratified by WinSMASH delta-V.

Additional steps simulated the effect of imputing unknown values of WinSMASH delta-V. First, the effect of driver sex and WinSMASH missing status on EDR delta-V was modeled using linear regression (Item 5). The differences by sex were applied to a proportion of the Monte Carlo simulation runs (Item 7) equal to the proportion of cases in the crash databases with missing WinSMASH values (Item 6). Simulation of non-missing WinSMASH values was performed as before (Item 1).

All simulated values were sampled from the Gaussian distributions represented by the regression model estimates. As the simulations were used to estimate differences in delta-V, rather than delta-V itself, boundary conditions were unnecessary. The linear regressions were unweighted, since NASS-CDS and CISS case selection factors should not influence the relationship between EDR and WinSMASH delta-V. However, case weights were included in the logistic regression models for injury outcome using the “survey” package in R [18].

III. RESULTS

Vehicle and Crash Differences

Analysis of 353,097 vehicles in 2016–2020 CRSS showed systematic differences in the crash-involved vehicles driven by females and males. Relative to males, females were more likely to be driving cars, SUVs and minivans, and less likely to be driving pickups (Fig. 1). Figure 2 shows the vehicle models involved in the largest number of crashes (weighted values) for drivers of either sex. Among these 20 models, the biggest disparity in driver sex was observed in the Ford F-250 (93% male). The Toyota RAV4 was the model with the highest proportion of female drivers (64%). Vehicle age differed more by vehicle type than by driver sex but females tended to drive newer vehicles, especially SUVs (Fig. 3).

There were 12,964 two-vehicle crashes that involved one female and one male driver and resulted in tow-away damage to the struck vehicle, or to either vehicle in front-to-front crashes. Of these, 41% were front-to-side, 38% were front-to-front, and 21% were front-to-rear crashes. Females more often were driving the struck vehicle in front-to-side (53.1% weighted proportion; 95% confidence interval [CI]: 51.5–54.8%) and in front-to-rear crashes (61.8%; 95% CI: 59.6–64.0%). The weighted distribution of curb weight differences for the two vehicles in each crash are shown in Fig. 4 by crash type and sex of the driver in the striking vehicle. In front-to-front crashes, female drivers had a median curb weight disadvantage of 78 kg (95% CI: 52–97 kg) relative to the male driver’s partner vehicle. When struck in the rear or side by the front of a vehicle with a driver of the opposite sex, females had a median curb weight disadvantage of 173 kg (130–210 kg) and 216 kg (186–253 kg), respectively, while males had an advantage of 12 kg (-25–54 kg) and 35 kg (3–65 kg), respectively. Across all crash types, the median difference was a 104 kg (95% CI: 91–117 kg) lower curb weight for the female-driven vehicle, regardless of its role in the crash.

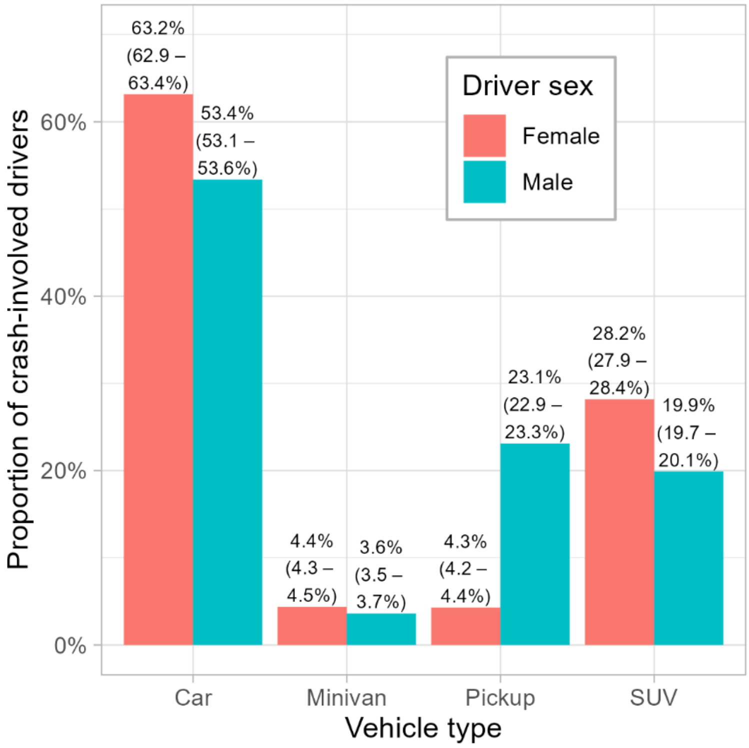

Fig. 1. Distribution of CRSS vehicle types by driver sex, with 95% confidence intervals.

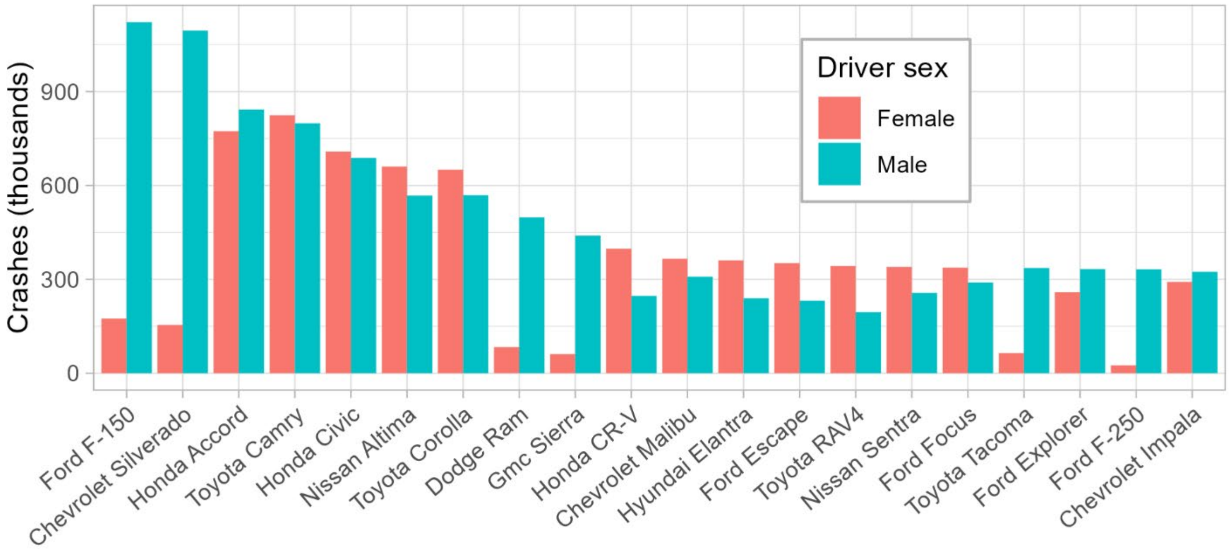

Fig. 2. CRSS vehicle models with most crashes for drivers of either sex.

Results of conditional logistic regression models for 1995–2020 FARS head-to-head crashes are shown in Table II. Relative to male drivers, the estimated fatality odds ratio for females was 1.75 (95% CI: 1.64–2.86) when neither curb weight nor vehicle type were accounted for. When curb weight differences were controlled, the female fatality odds ratio was reduced to 1.14 (95% CI: 1.03–1.26). When additionally limiting to head-to-head crashes of the same vehicle type, the estimate for driver sex was no longer significant at α = 0.05 (female odds ratio [OR]: 1.04; 95% CI: 0.89–1.21), and when the crashed vehicles were further restricted to a model year difference of 5 or less, the fatality odds ratio was effectively 1.

When included, the curb weight difference between the two vehicles always had a strong effect on fatality risk. This likely helps explain the airbag effect estimated in the first model. In the absence of a curb weight covariate, lack of airbag deployment likely was a surrogate for a smaller momentum change and greater mass. In the second model, with curb weight as a covariate, lack of an airbag deployment indicated an increase in fatality likelihood, although this estimate was not statistically significant.

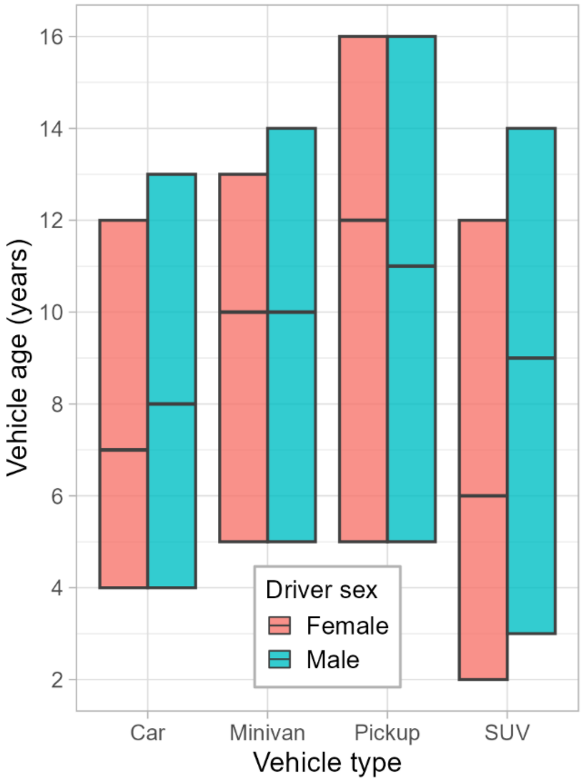

Fig. 3. First, second, and third quartiles of CRSS vehicle age by driver sex and vehicle type.

Potential Bias in Injury Odds Estimates

Table III shows front-crash case count data for the three versions of WinSMASH that were evaluated. Each version had more EDR cases available for analysis than the previous version. Vehicles driven by males had a higher rate of missing WinSMASH values than those driven by females, both when considering cases with available EDR data and all crashes with a driver restrained by a seat belt and airbag.

Table IV shows the results of the linear and logistic regression models that were used as inputs in the Monte Carlo simulations. The linear regression models of WinSMASH delta-V showed that given the same EDR delta-V, vehicles driven by males consistently had higher estimated WinSMASH values than vehicles driven by females. However, the magnitude of the difference declined with each version of WinSMASH and was not significant at the α = 0.05 level in the second and third versions. The linear regression model of EDR delta-V showed that vehicles usually had lower estimated delta-V in cases without WinSMASH estimates. The estimated differences for vehicles driven by males were significant

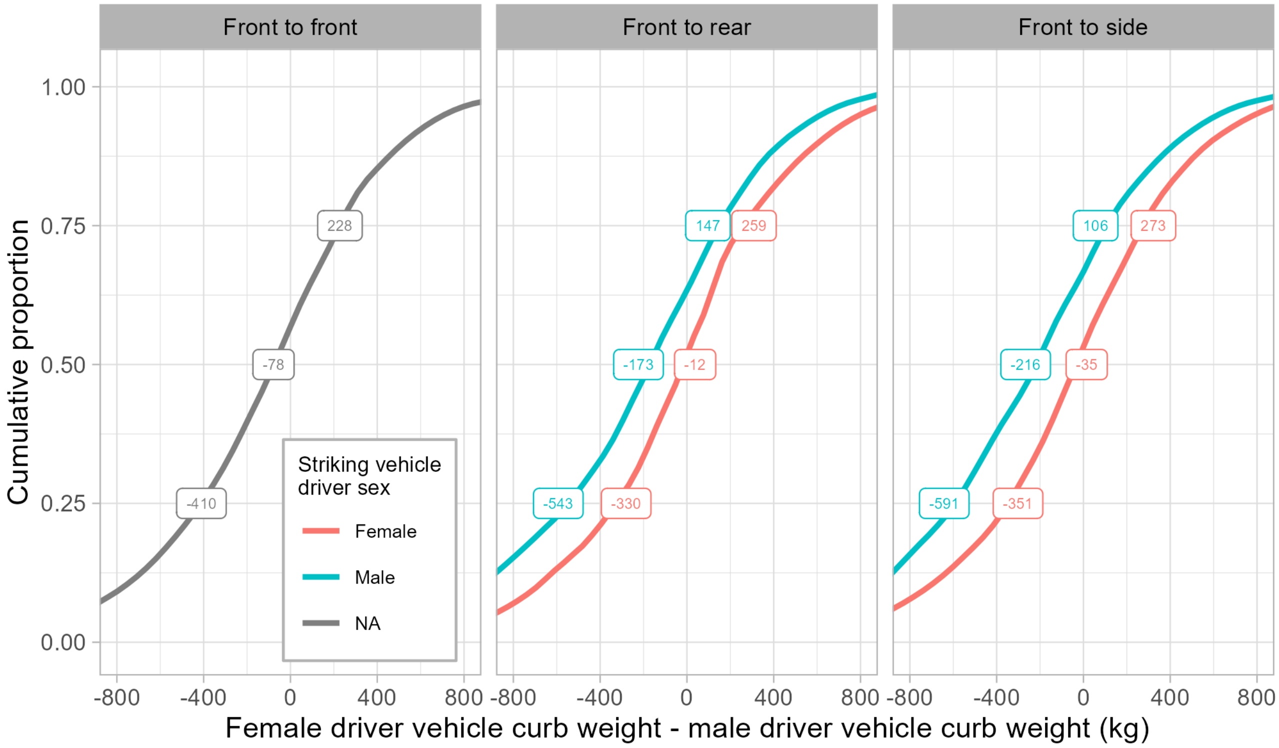

Fig. 4. Weighted distribution of vehicle curb weight differences for two-vehicle crashes involving one female and one male driver, by sex of driver in striking vehicle. Quartile values are labelled.

TABLE II

RESULTS OF CONDITIONAL LOGISTIC REGRESSION FATALITY MODELS IN HEAD-TO-HEAD CRASHES

Model	All crashes	Different sex	Covariate	Fatality odds ratio	95% Confidence interval
Baseline	11,007	5,041	Age (+ 1)	1.04	(1.04, 1.04)
			Airbag not deployed (ref: deployed)	0.87	(0.77, 0.98)
			Model year (+1)	0.92	(0.92, 0.93)
			Female (ref: male)	1.75	(1.64, 1.86)
Curb weight covariate	11,007	5,041	Age (+ 1)	1.06	(1.06, 1.07)
			Airbag not deployed (ref: deployed)	1.17	(0.97, 1.41)
			Curb weight (log; +0.1)	0.46	(0.44, 0.47)
			Model year (+1)	0.93	(0.92, 0.94)
			Female (ref: male)	1.14	(1.03, 1.26)
Curb weight covariate Same vehicle type	3,570	1,570	Age (+ 1)	1.07	(1.06, 1.07)
			Airbag not deployed (ref: deployed)	1.05	(0.78, 1.42)
			Curb weight (log; +0.1)	0.50	(0.47, 0.53)
			Model year (+1)	0.91	(0.90, 0.93)
			Female (ref: male)	1.04	(0.89, 1.21)
Curb weight covariate Same vehicle type Model year difference ≤5	2,069	915	Age (+ 1)	1.07	(1.06, 1.07)
			Airbag not deployed (ref: deployed)	0.95	(0.64, 1.41)
			Curb weight (log; +0.1)	0.49	(0.46, 0.53)
			Model year (+1)	0.91	(0.87, 0.95)
			Female (ref: male)	0.98	(0.81, 1.19)

TABLE III

NASS-CDS AND CISS FRONT-CRASH CASE COUNTS FOR ANALYSIS OF INJURY ODDS BIAS

	Table I	2004–2007 NASS-CDS			2008–2015 NASS-CDS			2017–2021 CISS		
	Item	Female	Male	Total	Female	Male	Total	Female	Male	Total
Vehicles with EDR	5	236	204	440	563	601	1,164	1,870	2,041	3,911
Above, with WS	1	203	170	373	479	486	965	1,448	1,450	2,898
Percentage missing WS		14.0	16.7	15.2	14.9	19.1	17.1	22.6	29.0	25.9
Drivers with belt+AB		2,667	2,987	5,654	3,070	3,258	6,328	2,531	2,716	5,247
Above, with WS	2	2,087	2,154	4,241	2,246	2,277	4,523	1,821	1,795	3,616
Percentage missing WS	6	21.7	27.9	25.0	26.8	30.1	28.5	28.1	33.9	31.1

Note: EDR = event data recorder; WS = WinSMASH.

TABLE IV

RESULTS OF REGRESSION MODELS USED AS INPUTS FOR MONTE CARLO SIMULATIONS

	Table I		2004–2007 NASS-CDS			2008–2015 NASS-CDS			2017–2021 CISS		
Outcome	Item	Covariate	Est.	SE	*p* value	Est.	SE	*p* value	Est.	SE	*p* value
WS dV	1	(Intercept)	7.205	0.993	<0.001	6.342	0.845	<0.001	5.597	0.382	<0.001
		EDR dV (+1 km/h)	0.548	0.033	<0.001	0.655	0.025	<0.001	0.640	0.011	<0.001
		Males (ref: females)	2.143	0.756	0.005	0.518	0.685	0.45	0.309	0.344	0.37
AIS≥2 injury	2	WS dV (+1 km/h)	0.081	0.008	<0.001	0.070	0.007	<0.001	0.068	0.008	<0.001
		Age (+1 year)	0.022	0.005	<0.001	0.025	0.006	<0.001	0.036	0.007	<0.001
AIS≥3 injury	2	WS dV (+1 km/h)	0.115	0.009	<0.001	0.097	0.009	<0.001	0.070	0.010	<0.001
		Age (+1 year)	0.031	0.006	<0.001	0.055	0.010	<0.001	0.056	0.006	<0.001
EDR dV	5	WS missing (ref: not missing), females	−0.872	2.352	0.71	−3.221	1.616	0.05	−1.369	0.887	0.12
		WS missing (ref: not missing), males	0.450	2.354	0.85	−5.541	1.417	<0.001	−2.302	0.783	0.003
		WS missing:male interaction	1.322	3.328	0.69	−2.320	2.149	0.28	−0.934	1.183	0.43

Note: AIS = Abbreviated Injury Scale; dV = delta-V; EDR = event data recorder; Est. = estimate; SE = standard error; WS = WinSMASH.

at the α = 0.05 level for the second and third versions of WinSMASH, as was the estimate for vehicles driven by females included in the second version of WinSMASH. None of the interaction terms with driver sex was significant at the α = 0.05 level.

The Monte Carlo simulation results are presented in Table V and Fig. 5. While other factors may affect the true injury risk difference between the sexes, the simulations show injury odds ratios that could be attributed to females based on their sex but which are actually due to typical differences in WinSMASH delta-V for vehicles driven by females and males. Odds ratios were greatest for cases from 2004 to 2007 NASS-CDS, which represent the first version of WinSMASH evaluated in this study. When cases with missing WinSMASH delta-V were excluded, the simulation results indicated female odds ratios of 1.19 (95% CI: 1.05–1.36) for AIS≥2 injuries and 1.29 (1.08–1.53) for AIS≥3 injuries. Estimated female injury odds ratios decreased with each successive version of WinSMASH, and were only 1.02 for both injury outcomes in 2017–2021 CISS cases (AIS≥2 95% CI: 0.98–1.07; AIS≥2 95% CI: 0.98–1.07). For the second and third versions, estimated odds ratios were larger when the effect of imputing missing WinSMASH delta-V was accounted for, reflecting the lower EDR delta-Vs typical for vehicles driven by men in cases without WinSMASH values (Table IV, EDR dV interaction term). However, due to the large amount of uncertainty around this difference (standard errors in last row of Table IV), the 95% confidence intervals on the female odds ratios were all wider and included 1 when the effect of imputation was simulated.

TABLE V

MONTE CARLO SIMULATION RESULTS FOR APPARENT FEMALE INJURY ODDS RATIOS
DUE TO DIFFERENCES IN WINSMASH DELTA-V ESTIMATION;
FRONT CRASHES WITH BELTED DRIVERS AND DEPLOYED AIRBAG

		2004–2007 NASS-CDS		2008–2015 NASS-CDS		2017–2021 CISS	
Missing WS	Outcome	Female OR	95% CI	Female OR	95% CI	Female OR	95% CI
Excluded	AIS≥2	1.19	(1.05, 1.36)	1.04	(0.94, 1.14)	1.02	(0.98, 1.07)
Excluded	AIS≥3	1.29	(1.08, 1.53)	1.05	(0.92, 1.20)	1.02	(0.97, 1.07)
Included	AIS≥2	1.17	(0.86, 1.58)	1.13	(0.70, 1.72)	1.05	(0.86, 1.28)
Included	AIS≥3	1.26	(0.81, 1.90)	1.19	(0.61, 2.12)	1.06	(0.86, 1.29)

Note: AIS = Abbreviated Injury Scale; CI = confidence interval ; OR = odds ratio; WS = WinSMASH.

Fig. 5. Monte Carlo simulation results for apparent female injury odds ratios due to differences in WinSMASH delta-v estimation; front crashes with belted drivers and deployed airbag.
Note: AIS = Abbreviated Injury Scale; OR = odds ratio; WS = WinSMASH.

IV. DISCUSSION

Relative to males, females are more often driving the side- or rear-struck vehicle in front-to-side and front-to-rear crashes, and their vehicle typically has a lower curb weight than their male crash partner's vehicle in every type of two-vehicle crash. In head-to-head crashes involving one female and one male driver, vehicle differences represent a 75% higher fatality odds for the female driver on average (Table II, comparing baseline and final models). In their assessment of near-side crashes, Teoh and Arbelaez [27] found that increased vehicle mass reduced fatality risk in the struck vehicle. The size of the effect they reported would translate to a 24% higher female fatality odds at the median curb weight difference in front-to-side crashes found here. This does not account for other risk factors related to vehicle type, such as height incompatibility. In fatal front crashes, these produced additional risks for female drivers beyond curb weight differences alone (Table II).

Addressing these disparities without eliminating vehicle choice represents a challenge for the traffic safety field on several fronts. Improved education regarding the self-protection benefits of vehicle size may be helpful. United States consumer-information front-crash test scores assigned by IIHS and NHTSA represent impacts with similar-sized vehicles. While both organisations communicate this fact on their public websites, neither one features it prominently on the individual vehicle rating pages. Another challenge is the relationship between vehicle size and cost, whether in terms of the initial purchase, fuel, or to the environment. Consumers who prioritise economy in any combination of these aspects usually have been required to sacrifice a level of crash self-protection. While electric vehicles (EVs) have the potential to reduce operational and environmental costs, it may be some time before they offer parity in terms of purchase costs. Furthermore, it has not yet been established whether the additional battery mass will convey the same level of benefit that has been measured for additional mass in crashes of internal combustion engine (ICE) vehicles. Finally, more needs to be done to incentivise crash-partner protection for large vehicles, whether EV or ICE. The Euro NCAP mobile progressive deformable barrier test, introduced in 2020, was designed to do this [28]. Similar options should be explored in the USA, where the mean passenger-vehicle curb weight for 2021 models was around one-third higher (USA: 1,945 kg [29]; European Union and the United Kingdom: 1,481 kg [30]).

The fatality regression models for belted driver pairs in head-to-head crashes indicate that overall front-crash fatality risks are not significantly different for females and males after accounting for vehicle differences. This differs from the findings of Noh *et al.* [9], who reported a statistically significant 6.8% higher fatality risk for female drivers in front crashes of 2000–2020 model vehicles, the cohort that was closest to the 1995–2021 range analysed in this study. It is unknown whether the discrepancy is due to the assumptions inherent in the double-pair method, the inclusion of unbelted drivers in the earlier study, or differences between head-to-head crashes and front crashes in general.

Beyond fatality risk in head-to-head crashes, this study did not evaluate whether women and men have different risks after accounting for vehicle and crash factors. However, results of the Monte Carlo simulations based on EDR delta-V show the potential for vehicle and crash differences to indirectly influence injury odds ratios when WinSMASH delta-V is used as a crash severity control. This was especially the case for the version of WinSMASH used to estimate severity for 2004–2007 NASS-CDS cases. There is little indication that 2017–2021 CISS cases are susceptible to this type of bias, but future years should be similarly assessed. As studies continue to include NASS-CDS cases to supplement smaller CISS samples in the near future, researchers should consider the possible effects of the WinSMASH bias on their findings. For example, studies that include multiple versions of the WinSMASH algorithm would be expected to produce an apparently decreasing relative injury risk for females over time. It is possible that prior studies have been affected by this phenomenon. Properly accounting for this in a single regression model could require a three-way interaction term between delta-V, WinSMASH version and sex, but interpreting the resulting parameter estimates would be challenging. Simpler approaches would be to construct separate models for each version of WinSMASH or include tighter controls on vehicle and crash differences between females and males.

Additional caution should be used when imputing missing delta-Vs. The wide confidence intervals associated with the imputation process (right half of Fig. 5) represent the range of bias that was simulated for individual pairs of female and male drivers and should not be taken as evidence that no bias exists. For example, a study of AIS≥3 injuries in 2004–2015 NASS-CDS in which missing delta-Vs were imputed would be expected to produce a female odds ratio of 1.20–1.25 simply due to the difference in WinSMASH error by sex. Where missing delta-Vs need to be estimated, their accuracy could be improved by including vehicle type, curb weight and crash configuration as auxiliary variables during the imputation process. Ultimately, the best use of WinSMASH in the future may be its own use as an auxiliary variable when imputing missing EDR values for delta-V and other crash-pulse severity metrics.

That said, one limitation of the current study is the assumption that EDR delta-V is correct. Others have found EDR tends to underestimate crash test delta-V [21-23]. The inclusion criteria used for complete EDR pulses in this study likely increases the degree of underreporting. The possibility that there are sex-related differences in EDR accuracy due to vehicle or crash type could not be evaluated. A limitation of the FARS head-to-head crash regression models was the use of vehicle model year as an imperfect control for all crashworthiness differences. While fleet-wide crash test performance improves each year, there are differences by model year that could affect results.

V. CONCLUSIONS

Differences in the vehicles driven by females and males, as well as the crashes they are involved in, have the potential to confound relative injury risk analyses unless they are properly accounted for. In fatal head-to-head crashes, the increased odds of female fatality relative to males declined from 75% to -2% after controlling for vehicle differences. Using crush-based estimates of crash severity to allow comparison across unpaired crashes may be insufficient to remove the bias, and imputing missing values for these estimates may compound the problem. Better crash severity estimates, for example from the newest version of WinSMASH or from EDR pulses, may be necessary for comparing risk. When this is not possible, researchers should consider tighter controls of vehicle- and crash-based risk factors. More generally, these results demonstrate that vehicle compatibility improvements will be an important part of addressing the higher injury and fatality risks faced by females.

VI. REFERENCES

[1] Insurance Institute for Highway Safety (2023) Fatality Facts 2020: Males and Females. Available: https://www.iihs.org/topics/fatality-statistics/detail/males-and-females. Accessed 16 March 2023.

[2] McGuckin, N., Fucci, A. (2018) Summary of Travel Trends: 2017 National Household Travel Survey. Federal Highway Administration Report No. FHWA-PL-18-019. Available: https://nhts.ornl.gov/assets/2017_nhts_summary_travel_trends.pdf. Accessed 16 March 2023.

[3] Kahane, C. J. (1994) Correlation of NCAP performance with fatality risk in actual head-on collisions. National Highway Traffic Safety Administration, Report No. DOT HS 808 061.

[4] Farmer, C. M. (2005) Relationships of frontal offset crash test results to real-world driver fatality rates. *Traffic Injury Prevention*, **6**(1): pp.31–37.

[5] Kullgren, A., Stigson, H., Axelsson, A. (2020) Developments in car crash safety since the 1980s. *Proceedings of the 2020 IRCOBI Conference*, Online.

[6] Evans, L. (2001) Female compared with male fatality risk from similar physical impacts. *The Journal of Trauma: Injury, Infection, and Critical Care*, **50**(2): pp.281–288. doi:10.1097/00005373200102000-00014

[7] Kahane, C. J. (2013) Injury vulnerability and effectiveness of occupant protection technologies for older occupants and women. National Highway Traffic Safety Administration, Report No.: DOT HS 811 766.

[8] Abrams, M. Z., Bass, C. R. (2020) Female vs. male relative fatality risk in fatal crashes. *Proceedings of the 2020 IRCOBI Conference*, Online.

[9] Noh, E. Y., Atwood, J. R. E., Lee, E., Craig, M. J. (2022) Female crash fatality risk relative to males for similar physical impacts. National Highway Traffic Safety Administration Report No. DOT HS 813 358.

[10] Bingham, C. R., Ehsani, J. P. (2012) The Relative Odds of Involvement in Seven Crash Configurations by Driver Age and Sex. *Journal of Adolescent Health,* **51**(5): pp.484–490.

[11] Malczyk, A., Kröling, S. (2022) Occupant injury severity and gender: Evidence for a gender gap? German Insurance Association, Compact Accident Research No. 120. Available: https://www.udv.de/udv-en/publications.

[12] Forman, J., Poplin, G. S., *et al*. (2019) Automobile injury trends in the contemporary fleet: belted occupants in frontal collisions. *Traffic Injury Prevention,* **20**(6): pp.607–612.

[13] Brumbelow, M. L. (2019) Front crash injury risks for restrained drivers in good-rated vehicles by age, impact configuration, and EDR-based delta V. *Proceedings of the 2019 IRCOBI Conference*, Florence, Italy.

[14] Brumbelow, M. L., Jermakian, J. S. (2022) Injury risks and crashworthiness benefits for females and males: Which differences are physiological? *Traffic Injury Prevention*, **23**(1): pp.11–16.

[15] Wiacek, C., Nagabhushana, V., *et al*. (2015) Evaluation of frontal crash stiffness measures from the U.S. New Car Assessment Program. ESV Paper Number 15-0257.

[16] Hauschild, H., Halloway, D., Pintar, F. (2021) Delta-v slope as an indicator of injury. *Traffic Injury Prevention*, **22**(Sup1): pp.S165–S169.

[17] Highway Loss Data Institute (2023) VINDICATOR user's manual: VINDICATOR 2023.

[18] Lumley, T. (2020) "Survey: analysis of complex survey samples". R package version 4.0.

[19] Kahane, C. J. (2015) Lives Saved by Vehicle Safety Technologies and Associated Federal Motor Vehicle Safety Standards, 1960 to 2012. National Highway Traffic Safety Administration, Report No. DOT HS 812 069.

[20] Therneau, T. (2022) A Package for Survival Analysis in R. R package version 3.3-1, Available: https://CRAN.R-project.org/package=survival.

[21] Tsoi, A., Hinch, J., Ruth, R., Gabler, H. C. (2013) Validation of event data recorders in high severity full-frontal crash tests. *SAE International Journal of Transportation Safety*, **1**(1): pp.76–99.

[22] Niehoff, P., Gabler, H. C., *et al*. (2005) Evaluation of event data recorders in full systems crash tests. *Proceedings of 19th International Technical Conference on the Enhanced Safety of Vehicle*, 2005, Washington, DC, USA.

[23] Gabler, H. C., Thor, C. P., Hinch, J. (2007) Preliminary evaluation of advanced air bag field performance using event data recorders. U.S. Dept of Transportation Technical Report No: DOT HS 811 015.

[24] Hampton, C. E., Gabler, H. C. (2010) Evaluation of the accuracy of NASS/CDS delta-V estimates from the enhanced WINSMASH algorithm. *Annals of Advances in Automotive Medicine*, **54**: pp.241–252.

[25] Mynatt, M., Brophy, J. (2017) Improved field measurements in NHTSA's CISS program. *Proceedings of 25th International Technical Conference on the Enhanced Safety of Vehicles*, 2017, Detroit, USA.

[26] Wickham, H., Hester, J., Ooms, J. (2021) xml2: Parse XML. R package version 1.3.3, Available: https://CRAN.R-project.org/package=xml2.

[27] Teoh, E. R., Arbelaez, R. A. (2022) The association between data collected in IIHS side crash tests and real-world driver death risk. *Traffic Injury Prevention*, **23**(4): pp.198–202.

[28] Sandner, V., Ellway, J., van Ratingen, M. (2017) Euro NCAP frontal impact working group report. *Proceedings of 25th International Technical Conference on the Enhanced Safety of Vehicles*, 2017, Detroit, USA.

[29] Environmental Protection Agency (2022) The 2022 EPA automotive trends report: Greenhouse gas emissions, fuel economy, and technology since 1975. Report No: EPA-420-R-22-029.

[30] International Council on Clean Transportation (2022) European vehicle market statistics: Pocketbook 2022/23. Available: https://eupocketbook.org/wp-content/uploads/2023/01/Pocketbook_2022_23_Web.pdf.