Published in final edited form as:

Traffic Inj Prev. 2016 May 18; 17(4): 346–351. doi:10.1080/15389588.2015.1091072.

Validation of quasi-induced exposure representativeness assumption among young drivers

Allison E. Curry, PhD, MPH^{a,b}, Melissa R. Pfeiffer, MPH^a, and Michael R. Elliott, PhD^{c,d}

Melissa R. Pfeiffer: m.riley@pfeiffer@gmail.com; Michael R. Elliott: mreliott@umich.edu

^aCenter for Injury Research and Prevention, The Children's Hospital of Philadelphia, 3535 Market Street, Suite 1150, Philadelphia, PA 19104 USA

^bCenter for Clinical Epidemiology and Biostatistics, Perelman School of Medicine at the University of Pennsylvania, 423 Guardian Drive, Philadelphia, PA 19104 USA

^cDepartment of Biostatistics, School of Public Health, University of Michigan, 1415 Washington Heights, Ann Arbor, MI 48109 USA

^dSurvey Methodology Program, Institute for Social Research, University of Michigan, Rm. 4068, 426 Thompson Street, Ann Arbor, MI 48109

Abstract

Objective—Young driver studies have applied quasi-induced exposure (QIE) methods to assess relationships between demographic and behavioral factors and at-fault crash involvement, but QIE's primary assumption of representativeness has not yet been validated among young drivers. Determining whether non-responsible young drivers in clean (i.e., only one driver is responsible) two-vehicle crashes are reasonably representative of the general young driving population is an important step toward ensuring valid QIE use in young driver studies. We applied previously established validation methods to conduct the first study, to our knowledge, focused on validating the QIE representativeness assumption in a young driver population.

Methods—We utilized New Jersey's (NJ) state crash and licensing databases (2008–2012) to examine the representativeness assumption among 17- to 20-year-old non-responsible drivers involved in clean multi-vehicle crashes. It has been hypothesized that if not-at-fault drivers in clean two-vehicle crashes are a true representation of the driving population, it would be expected that non-responsible drivers in clean three-or-more-vehicle crashes also represent this same driving population (Jiang and Lyles 2010). Thus, we compared distributions of age, gender, and vehicle type among: (1) non-responsible young drivers in clean two-vehicle crashes and (2) the first non-responsible young driver in clean crashes involving three or more vehicles to (3) all other non-responsible young drivers in clean crashes involving three or more vehicles. Distributions were compared using chi-square tests and conditional logistic regression; analyses were conducted for all young drivers and stratified by license status (intermediate vs. fully licensed drivers), crash location, and time of day of the crash.

Results—There were 41,323 non-responsible drivers in clean two-vehicle crashes and 6,464 non-responsible drivers in clean three-or-more-vehicle crashes. Overall, we found that the distributions of age, gender, and vehicle type were not statistically significantly different between the three groups; in each group, approximately one-fourth of drivers were of each year of age 17 through 20, half were males, and approximately 80% were driving a car/station wagon/minivan. In general, conclusions held when we evaluated the assumption within intermediate and fully licensed young drivers separately and by crash location and time.

Conclusions—It appears that the representativeness assumption holds among the population of NJ young drivers. We encourage young driver studies utilizing QIE methods to conduct internal validation studies to ensure appropriate application of these methods and we propose utilization of QIE methods to address broader foundational and applied questions in young driver safety.

Keywords

at-fault crash; crash culpability; crash propensity; crash responsibility; induced exposure; teen driver

INTRODUCTION

Valid estimation of motor vehicle crash rates requires specification of a denominator that accurately reflects a groups' collective time "at risk" for experiencing a crash. Although the choice of denominator should ideally be motivated by the specific research question of interest, it instead is usually determined by the accessibility to or feasibility of collecting required data (Cummings et al. 1995). Thus, even in cases in which the denominator of interest is "at-risk" miles driven, studies are often limited to counts of population or licensed drivers given the difficulty of measuring miles driven via diaries or direct observations—in particular when disaggregated estimates are needed to compare crash rates for driver or vehicle subgroups or under different driving conditions.

In response, the quasi-induced exposure (QIE) method was proposed to allow estimation of relative driving exposure for specified groups using only the crash database itself and subsequently to allow for adjusted crash involvement rates (Stamatiadis and Deacon 1997). The fundamental assumption of QIE is that the responsible driver in a "clean" (i.e., only one driver is responsible) two-vehicle crash "chooses his or her innocent victim at random from all drivers present" at the same time and location (Lyles et al. 1991). In other words, non-responsible drivers in clean two-vehicle crashes are a random sample of the driving population at the time and place of the crash (Stamatiadis and Deacon 1997), and distribution of the factor of interest (e.g., driver gender) in that sample serves as a representative measure of the relative driving exposure with respect to that factor among all drivers.

Several studies have validated QIE's fundamental assumption of representativeness, primarily by using crash databases given the difficulty of collecting "gold standard" exposure data via diaries or direct observation. Lyles et al. argued that if non-responsible drivers are indeed randomly selected by drivers responsible for crashes, then the gender distribution of responsible and non-responsible drivers should be independent—that is, male

and female responsible drivers should crash with the same proportion of male and female non-responsible drivers (Lyles et al. 1991). By comparing matrices of responsible/non-responsible driver genders in clean two-vehicle crashes in Michigan in the late 1980's, the authors found that, for a given roadway type, non-responsible drivers appeared to be a random sample of the driving population. Stamatiadis and Deacon reported similar results with respect to driver age and vehicle type (Stamatiadis and Deacon 1997). An enhanced approach taken by two recent studies involved comparing the distribution of age, gender, and vehicle type among groups of non-responsible drivers in two-vehicle crashes and non-responsible drivers in crashes involving three or more vehicles (Chandraratna and Stamatiadis 2009; Jiang and Lyles 2010). These studies were guided by the notion that the randomness of crash involvement among the second or later non-responsible drivers in three-or-more-vehicle crashes is greater than among the first non-responsible driver. Thus, if non-responsible drivers in two-vehicle crashes are a true representation of the driving population (at the time and place of the crash), it would be expected that this group would be similar to later non-responsible drivers in three-or-more vehicle crashes. Taken together, these studies found these samples to be statistically similar with respect to driver gender, crash location, and environmental factors and concluded that QIE yielded reasonable estimates of exposure with respect to these factors.

Notably, the above-described validation studies were focused on the overall driving population. With respect to young drivers, the studies either grouped them altogether (e.g., 15- to 19-year-olds) or with older drivers (e.g., <25 years old) and investigated other factors such as gender without regard to driver age or experience. Meanwhile, a number of applied young driver studies have employed QIE methods to assess relationships between demographic and behavioral factors (e.g., gender, nighttime driving, cell phone use) and at-fault crash involvement (Keall and Newstead 2009; Kirk and Stamatiadis 2001; Preusser et al. 1998; Redondo-Calderon et al. 2001; Rice et al. 2003). None of these studies appear to have validated (or at least did not explicitly describe having validated) QIE's primary assumption in their particular (or any) young driver sample; several rely on references to these larger validation studies that include driver populations of all ages. However, young drivers may differ from older drivers in important ways, including exposure patterns (e.g., less highway driving) and different gender-specific risk profiles (i.e., young males as a particularly high-risk group). Thus, validating the assumption specifically among young drivers—that is, determining whether non-responsible crash-involved young drivers reasonably represent the general young driving population—is an important step toward ensuring valid QIE use in studies that estimate driving exposure and crash rates among different subgroups of young drivers (e.g., male vs. female young drivers).

To this end, we applied methods previously developed by Chandraratna and Stamatiadis (2009) and Jiang and Lyles (2010) to conduct the first study, to our knowledge, focused on validating the QIE representativeness assumption among the young driver population. Specifically, we utilized New Jersey's state crash database (2008–2012) to examine whether non-responsible 17- to 20-year-old New Jersey (NJ) drivers in clean multi-vehicle crashes can be considered reasonably representative of NJ's overall young driving population. Given our particular interest in applying these methods within young driver groups of varying driving experience, we examined the assumption's validity both overall and separately for

intermediate drivers (for whom Graduated Driver Licensing [GDL] restrictions apply) and fully licensed drivers. Finally, inherent in this assumption is that non-responsible drivers are representative of the road user population *at the time and space of the crash*. Thus, we also examined the representative assumption separately by time of day and road type (highway/interstate vs. other road) of crash occurrence.

METHODS

We obtained detailed data on all drivers involved in a police-reported crash from January 1, 2008 through June 30, 2012 from the NJ Department of Transportation (Rutgers University Police Technical Assistance Program). Crashes in NJ are reportable if they result either in an injury to any person or property damage in excess of \$500. In order to specifically identify NJ drivers, we conducted a hierarchical deterministic linkage to link each driver's crash record to his or her corresponding record (if applicable) in the NJ Motor Vehicle Commission's licensing database, which contains information on each NJ driver's progression through the state's licensing process. Full details of this linkage are available elsewhere (Curry et al. 2013); in all, over 98% of crash-involved drivers with a licensing state of NJ (as noted on the crash report) matched to a unique record in the licensing database.

Consistent with QIE methods, we selected all crashes involving two or more passenger vehicles. We excluded all hit-and-run crashes and crashes involving a pedestrian, pedalcyclist, or vehicle other than a passenger vehicle (e.g., bus, truck, motorcycle). We determined the responsibility status of each crash-involved driver using the presence of a crash-contributing driver action, as noted on the crash report by the responding officer; New Jersey's *Police Guide for Preparing Reports of Motor Vehicle Crashes* instructs officers to determine the presence of these actions independently of whether a citation was issued (Rutgers University Police Technical Assistance Program). A driver was determined to be responsible if he/she committed one or more driver actions (e.g., inattention, unsafe speed, failure to obey traffic control device, failure to yield to vehicle/pedestrian) that the officer determined to have contributed to the crash. We and others have documented the relative advantages of using this method rather than legal culpability methods such as issuance of a citation or moving violation when determining crash responsibility (Curry et al. 2014; Jiang et al. 2012). For each crash, any number of drivers—including none—could have been determined to be responsible for the crash. For the current study, we identified all clean multi-vehicle crashes—that is, multi-vehicle crashes in which one and only one driver was deemed responsible.

We then selected all non-responsible NJ drivers involved in clean multi-vehicle crashes who were 17 to 20 years old on the day of the crash. We determined the license status of each young driver on the date of his or her crash (i.e., learner's permit, intermediate license, full license, suspended/unlicensed) using information on the start dates of the permit and intermediate license, license transactions (to ascertain the date of full licensure), and dates of license suspension and restoration; see Curry et al. (2015) for more information. Under NJ's GDL system, residents are eligible for an intermediate license at a minimum of age 17 and a full license at a minimum of age 18. Notably, NJ is the only state that applies full GDL rules

to all newly-licensed drivers under age 21; in most other states, newly-licensed drivers aged 18 and older are exempt from GDL restrictions (Insurance Institute for Highway Safety 2013). Thus, we defined intermediate licensed drivers as 17- to 20-year-old drivers who had a valid (i.e., not suspended or revoked) intermediate license at the time of the crash and fully licensed drivers as 18- to 20-year-old drivers who had a valid full license at the time of the crash. Gender was obtained from the licensing record and information on vehicle type (car/station wagon/minivan vs. other vehicle type), type of road the crash occurred on (state highway/interstate vs. other road), and time of crash (5:00 am – 4:00 pm [daytime], 4:01 pm – 11:00pm [evening], 11:01 pm – 4:59 am [hours restricted by NJ's GDL system]) was obtained from the crash report.

To test the validity of the representativeness assumption, we compared distributions of age, gender, and vehicle type among: (1) non-responsible drivers in clean two-vehicle crashes (D2) and (2) the first non-responsible driver in clean crashes involving three or more vehicles (D3_1) to (3) all other non-responsible drivers in clean crashes involving three or more vehicles (D3_other). The order of non-responsible drivers was determined based on the successive order of each driver's involvement in the sequence of harmful events (Rutgers University Police Technical Assistance Program). Chi-square tests were used to statistically compare distributions between the D2 and D3_other groups. Comparisons of D3_1 and D3_other were conducted using conditional logistic regression given that drivers were clustered by crash (and thus were not independent samples), with group (D3_1 vs D3_other) as the outcome and age, gender, and vehicle type separately as categorical independent variables. Given large sample sizes, inferential tests may indicate statistical significance at the $p=0.05$ level even when there is no practical difference between two populations. While no standard threshold for a meaningful practical difference exists, Jiang et al. previously defined a four percentage point difference based on empirical evidence (Jiang and Lyles 2010). We refrained from specifying a threshold given lack of parameters by which to base this judgment; instead, we present absolute percentage point differences and reference Jiang's selected threshold in our discussion of results.

RESULTS

Over the four and a half year study period, there were a total of 47,787 non-responsible 17- to 20-year-old NJ drivers who were involved in clean multi-vehicle crashes. This includes 41,323 drivers in two-vehicle crashes and 6,464 drivers in crashes involving three or more vehicles. Of those in three-or-more-vehicle crashes, 3,081 were determined to be the first non-responsible driver (i.e., closest non-responsible driver to first harmful event) and the remaining 3,383 were the second or later non-responsible driver.

The distributions of age, gender, and vehicle type among non-responsible young drivers in two-vehicle crashes (D2), the first non-responsible young drivers in three-or-more-vehicle crashes (D3_1), and other non-responsible young drivers in three-or-more-vehicle crashes (D3_other) are shown in Table 1. Overall, there were no statistically significant differences—and one might argue no real practical differences—between the D2 and D3_other groups or between the two groups of non-responsible drivers in three-or-more-vehicle crashes (D3_1 and D3_other). In each sample, approximately a quarter of drivers are of each year of

age 17 through 20, half are males, and approximately 80% were driving a car/station wagon/minivan at the time of the crash. For comparisons of D2 and D3_other, there was at most a 1.5 percentage point difference for age (for 18-year-old drivers), 1.5 percentage point difference for gender, and 0.1 percentage point difference for vehicle type. Differences were similar for D3_1 and D3_other comparisons; the largest difference was for vehicle type (1.2 percentage points).

We further evaluated the validity of the representativeness assumption separately for intermediate and fully licensed young drivers. As shown in Table 2, the distributions of age, gender, and vehicle type for non-responsible intermediate drivers in clean two-vehicle crashes (D2) and the first non-responsible intermediate drivers in clean three-or-more-vehicle crashes (D3_1) were both similar to the distribution for later non-responsible drivers in three-or-more-vehicle crashes (D3_other). The distribution for all three groups of intermediate drivers was heavily weighted toward 17 year olds. Distributions for fully licensed drivers are shown in Table 3. Compared with later non-responsible drivers in three-or-more-vehicle crashes, non-responsible drivers in two-vehicle crashes are slightly older (38.8% vs. 36.1% are 20 years old); the p-value for the age comparison was 0.01 and represented at most a three percentage point difference.

Finally, we conducted analyses separately by crash location (highway/interstate, other road) and time of crash; there were too few non-responsible young driver crashes between 11:01 pm – 4:59 am (the time period restricted by NJ's GDL system) to power analyses for this time period. As shown in Figures 1 and 2, tests comparing age distributions were not statistically significant with the exception of the comparison of D_2 and D3_other on non-highway roads; the largest percentage point difference for this comparison was 4.1 for 20 year olds (24.5% for D_2 vs. 20.4% for D3_other). There were no statistically significant differences (at the 0.05 level) for gender (range of percentage point differences: 0.7–2.2) or vehicle type (range of percentage point differences: 0.1–2.1).

DISCUSSION

This study validated QIE's primary assumption of representativeness specifically in the young driver population in order to determine whether crash-involved young drivers who were not responsible for their crash may serve as a reasonably representative sample of the general population of young drivers on the road. Overall, we found that with respect to the distributions of age, gender, and vehicle type, both non-responsible young drivers in clean two-vehicle crashes and the first non-responsible young drivers in clean three-or-more-vehicle crashes are similar to the second or later non-responsible young drivers in clean crashes involving three or more vehicles—a group theorized to be an even closer representation of the characteristics of road users given their later involvement in the crash sequence (Chandraratna and Stamatiadis 2009). Thus, the groups appear to be from the same underlying population of young drivers on the road. In general, these conclusions also held when we evaluated the assumption within intermediate and fully licensed young drivers separately and by crash location and time. There were slight differences in the age distribution for fully licensed drivers and on non-highway roads when non-responsible drivers in two-vehicle crashes were compared with later non-responsible drivers in three-or-

more-vehicle crashes; for both comparisons, age distributions were weighted slightly toward 20-year-old drivers for D2 drivers compared with D3_other drivers. While statistically significant (likely in part due to large samples sizes), the two comparisons were within (in the case of age among fully licensed drivers) or nearly within (in the case of non-highway roads) the four percentage point difference previously suggested by Jiang et al. for defining practical significance (Jiang and Lyles 2010).

With respect to application of these findings to young driver research, previous studies have used QIE methods to adjust associations between driver/vehicle factors and crash involvement for driving exposure or to identify a control group in epidemiologic case-control studies but have not conducted analyses to determine whether the representativeness assumption is valid among their population of young drivers; as illustrated in this study, analyses can and should be conducted to test this assumption. To our knowledge, none of the young driver studies that have applied QIE methods have conducted analyses to validate the assumption of representativeness in their sample. This should be considered a critical step, as validation analyses will provide researchers with the assurance that QIE methods can be appropriately applied in their specific population of young drivers. As demonstrated here, these analyses can be conducted relatively easily and cost-effectively using only the state-level crash database itself. Additionally, a recent study utilized QIE methods to compare the proportion of crashes involving non-responsible young drivers that occurred at night before and after implementation of a GDL night provision (Jiang and Lyles 2011). However, the authors of that study concluded that a “shortcoming of quasi-induced exposure is that it can only produce a *relative* value. It can’t answer a basic question such as, ‘what is the crash rate for [young drivers] under post-GDL’ or the like” (p.490). In contrast, we believe that if we are able to conclude—as supported by evidence in the current study—that the population of non-responsible crash-involved young drivers in clean two-vehicle crashes is a reasonably random sample (with respect to age, gender, vehicle type, etc.) of the population of young road users, then the rate of driving behaviors (night driving, peer passengers, cell phone use) in this group would be expected to estimate the rate among all young drivers on the road. In this way, we can address critical but difficult-to-answer foundational questions—e.g., “to what extent do males and females engage in high-risk driving behaviors and how does this differ by the time of day?”—and applied questions—e.g., “what is the effectiveness of a state-wide program that aims to increase intermediate drivers’ compliance with GDL restrictions?”

Some important considerations deserve discussion. The nature of two- and three-or-more vehicle crashes may differ in important ways. Three-or-more-vehicle crashes may be more likely to occur on certain roads (e.g., high-speed roads) and/or times of day (e.g., peak traffic hours), and distributions of age, gender, or vehicle type may vary by space and time (e.g., older young drivers are less likely to be restricted from driving at night). This may account for some of the observed differences in overall comparisons; previous research has suggested that further stratifying comparisons by road type and time of day—as we do in this study—may reduce bias when the representativeness assumption does not hold completely true (Stamatiadis and Deacon 1997). In addition, it is important to note that observed distributions are inherently designed to reflect distributions in the *driving* population, which is not necessarily equivalent to the population of *drivers*; that is, this

method inherently gives less weight to licensed drivers who drive less and in that way reflects driver trips more than the drivers themselves. Finally, we analyzed data only from one US state. There is likely to be variation among jurisdictions both in the distributions of relevant factors (e.g., age distribution, urbanicity) and in methods available to assign crash responsibility. Thus, applied young drivers studies should conduct QIE validation studies directly among the population under study.

CONCLUSION

In this study, we have directly applied established QIE methods to validate the representativeness assumption among a state-level population of young drivers. In doing so, we demonstrate that, with respect to age, gender, and vehicle type, the distribution of these factors among young crash-involved drivers who are not responsible for their crash appears to reasonably reflect distributions among NJ's young road users. We urge young driver studies that utilize QIE methods to conduct an internal validation study to ensure appropriate application of these methods and suggest utilization of QIE methods that extends beyond previously used applications to address broader foundational and applied questions in the field of traffic safety related to young drivers' engagement in relevant driving behaviors. Our future work will focus on utilizing QIE methods to directly address these questions and discussing implications to consider when doing so.

Acknowledgments

This work was supported by the Eunice Kennedy Shriver National Institute of Child Health and Human Development at the National Institutes of Health (Grant R03HD073248, PI: Curry). The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health. The sponsors had no role in the design and conduct of the study; in the collection, management, analysis, and interpretation of the data; or in the preparation, review, or approval of the manuscript. The authors thank Konny Kim, MPH for her help in data management and the NJ Motor Vehicle Commission, NJ Department of Transportation, and NJ Office of Information Technology for their assistance in providing data.

References

- Chandraratna S, Stamatiadis N. Quasi-induced exposure method: Evaluation of not-at-fault assumption. *Accid Anal Prev*. 2009; 41(2):308–313. [PubMed: 19245890]
- Cummings P, Koepsell T, Mueller B. Methodological challenges in injury epidemiology and injury prevention research. *Annu Rev Public Health*. 1995; 16:381–400. [PubMed: 7639878]
- Curry AE, Pfeiffer MR, Durbin DR, Elliott MR. Young driver crash rates by licensing age, driving experience, and license phase. *Accid Anal Prev*. 2015; 80:243–50. [PubMed: 25939133]
- Curry AE, Pfeiffer MR, Localio R, Durbin DR. Graduated driver licensing decal law: Effect on young probationary drivers. *Am J Prev Med*. 2013; 44(1):1–7. [PubMed: 23253643]
- Curry AE, Pfeiffer MR, Myers RK, Durbin DR, Elliott MR. Statistical implications of using moving violations to determine crash responsibility in young driver crashes. *Accid Anal Prev*. 2014; 65C: 28–35. [PubMed: 24398139]
- Insurance Institute for Highway Safety. [Accessed June 2014] Graduated Driver Licensing state laws. Available at <http://www.iihs.org/iihs/topics/laws/graduatedlicenseintro?topicName=teenagers>
- Jiang X, Lyles RW. A review of the validity of the underlying assumptions of quasi-induced exposure. *Accid Anal Prev*. 2010; 42(4):1352–1358. [PubMed: 20441852]
- Jiang X, Qiu Y, Lyles RW, Zhang H. Issues with using police citations to assign responsibility in quasi-induced exposure. *Saf Sci*. 2012; 50(4):1133–1140.

- Jiang XG, Lyles RW. Exposure-based assessment of the effectiveness of Michigan's graduated driver licensing nighttime driving restriction. *Saf Sci*. 2011; 49(3):484–490.
- Keall M, Newstead S. Induced exposure estimates of rollover risk for different types of passenger vehicles. *Traffic Inj Prev*. 2009; 10(1):30–36. [PubMed: 19214875]
- Kirk A, Stamatiadis N. Crash rates and traffic maneuvers of younger drivers. *Transp Res Rec J Transp Res Board*. 2001; 1779(1):68–74.
- Lyles RW, Stamatiadis P, Lighthizer DR. Quasi-induced exposure revisited. *Accid Anal Prev*. 1991; 23(4):275–285. [PubMed: 1883467]
- Preusser DF, Ferguson SA, Williams AF. The effect of teenage passengers on the fatal crash risk of teenage drivers. *Accid Anal Prev*. 1998; 30(2):217–222. [PubMed: 9450125]
- Redondo-Calderon JL, de Luna-del-Castillo J, Jimenez-Moleon JJ, Garcia-Martin M, Lardelli-Claret P, Galvez-Vargas R. Application of the induced exposure method to compare risks of traffic crashes among different types of drivers under different environmental conditions. *Am J Epidemiol*. 2001; 153(9):882–891. [PubMed: 11323319]
- Rice TM, Peek-Asa C, Kraus JF. Nighttime driving, passenger transport, and injury crash rates of young drivers. *Inj Prev*. 2003; 9(3):245–250. [PubMed: 12966014]
- Rutgers University Police Technical Assistance Program. Police Guide for Preparing Reports of Motor Vehicle Crashes. http://www.state.nj.us/transportation/refdata/accident/pdf/NJTR-1Field_Manual.pdf
- Stamatiadis N, Deacon JA. Quasi-induced exposure: Methodology and insight. *Accid Anal Prev*. 1997; 29(1):37–52. [PubMed: 9110039]

Figure 1.

Comparison of the distribution of age among non-responsible young drivers in clean two-vehicle crashes (D2), the first non-responsible drivers in clean three-or-more-vehicle crashes (D3_1), and all other non-responsible drivers in clean three-or-more-vehicle crashes (D3_other), by crash location.

Figure 2. Comparison of the distribution of age among non-responsible young drivers in clean two-vehicle crashes (D2), the first non-responsible drivers in clean three-or-more-vehicle crashes (D3_1), and all other non-responsible drivers in clean three-or-more-vehicle crashes (D3_other), by time of day.

Table 1

Comparison of the distribution of relevant factors among non-responsible young drivers in clean two-vehicle crashes (D2), the first non-responsible drivers in clean three-or-more-vehicle crashes (D3_1), and all other non-responsible drivers in clean three-or-more-vehicle crashes (D3_other).

Factor	Non-responsible driver, 2-vehicle crashes (D2) N=41,323	1st non-responsible driver, 3+ vehicle crashes (D3_1) N=3,081	All other non-responsible drivers, 3+ vehicle crashes (D3_other) N=3,383	Percentage point difference (D2–D3_other)	Percentage point difference (D3_1–D3_other)	p-value	
						D2 vs D3_other	D3_1 vs D3_other
Age (years)							
17	23.2	22.8	23.6	−0.4	−0.8	0.16	0.73
18	25.6	27.8	27.1	−1.5	0.7		
19	25.7	25.0	25.0	0.7	0.0		
20	25.5	24.3	24.3	1.2	0.0		
Gender							
Female	51.6	49.9	50.1	1.5	−0.2	0.10	0.68
Male	48.4	50.1	49.9	−1.5	0.2		
Type of passenger vehicle							
Car/station wagon/minivan	82.0	80.9	82.1	−0.1	−1.2	0.88	0.52
Other (SUV, pickup, passenger van)	18.0	19.1	17.9	0.1	1.2		

Table 2

Comparison of the distribution of relevant factors among non-responsible young intermediate drivers in clean two-vehicle crashes (D2), the first non-responsible drivers in clean three-or-more-vehicle crashes (D3_1), and all other non-responsible drivers in clean three-or-more-vehicle crashes (D3_other).

Factor	Non-responsible driver, 2-vehicle crashes (D2) N=20,219	1st non-responsible driver, 3+ vehicle crashes (D3_1) N=1,508	All other non-responsible drivers, 3+ vehicle crashes (D3_other) N=1,623	Percentage point difference (D2–D3_other)	Percentage point difference (D3_1–D3_other)	p-value	
						D2 vs D3_other	D3_1 vs D3_other
Age (years)							
17	46.2	45.8	48.6	–2.4	–2.8	0.17	0.45
18	26.1	28.9	26.0	0.1	2.9		
19	15.7	14.3	14.0	1.7	0.3		
20	12.0	11.1	11.4	0.6	–0.3		
Gender							
Female	53.8	51.5	52.1	1.7	–0.6	0.21	0.88
Male	46.2	48.5	47.9	–1.7	0.6		
Type of passenger vehicle							
Car/station wagon/minivan	81.5	79.6	80.7	0.8	–1.1	0.40	0.55
Other (SUV, pickup, passenger van)	18.5	20.4	19.3	–0.8	1.1		

Table 3

Comparison of the distribution of relevant factors among non-responsible young fully licensed drivers in clean two-vehicle crashes (D2), the first non-responsible drivers in clean three-or-more-vehicle crashes (D3_1), and all other non-responsible drivers in clean three-or-more-vehicle crashes (D3_other).

Factor	Non-responsible driver, 2-vehicle crashes (D2) N=20,270	1st non-responsible driver, 3+ vehicle crashes (D3_1) N=1,525	All other non-responsible drivers, 3+ vehicle crashes (D3_other) N=1,700	Percentage point difference (D2–D3_other)	Percentage point difference (D3_1–D3_other)	p-value	
						D2 vs D3_other	D3_1 vs D3_other
Age (years)							
18	25.5	27.3	28.5	–3.0	–1.2	0.01	0.58
19	35.7	35.8	35.4	0.3	0.4		
20	38.8	36.9	36.1	2.7	0.8		
Gender							
Female	50.0	48.4	48.8	1.2	–0.4	0.37	0.48
Male	50.0	51.6	51.2	–1.2	0.4		
Type of passenger vehicle							
Car/station wagon/minivan	82.6	82.4	83.4	–0.8	–1.0	0.41	0.40
Other (SUV, pickup, passenger van)	17.4	17.6	16.6	0.8	1.0		