Variability in Crash and Near-Crash Risk among Novice Teenage Drivers: A Naturalistic Study

Feng Guo, PhD[1,2], Bruce G. Simons-Morton, EdD, MPH[3], Sheila E. Klauer, PhD[2], Marie Claude Ouimet, PhD[4], Thomas A. Dingus, PhD[2,5], and Suzanne E. Lee, PhD[2]

Objective Using video monitoring technologies, we investigated teenage driving risk variation during the first 18 months of independent driving.

Study design Driving data were collected on 42 teenagers whose vehicles were instrumented with sophisticated video and data recording devices. Surveys on demographic and personality characteristics were administered at baseline. Drivers were classified into 3 risk groups using a K-mean clustering method based on crash and near-crash (CNC) rate. The change in CNC rates over time was evaluated by mixed-effect Poisson models.

Results Compared with the first 3 months after licensure (first quarter), the CNC rate for participants during the third, fourth, and fifth quarters decreased significantly to 59%, 62%, and 48%, respectively. Three distinct risk groups were identified with CNC rates of 21.8 (high-risk), 8.3 (moderate-risk), and 2.1 (low-risk) per 10 000 km traveled. High- and low-risk drivers showed no significant change in CNC rates throughout the 18-month study period. CNC rates for moderate-risk drivers decreased substantially from 8.8 per 10 000 km in the first quarter to 0.8 and 3.2 in the fourth and fifth quarters, respectively. The 3 groups were not distinguishable with respect to personality characteristics.

Conclusion Teenage CNC rates varied substantially, with distinct high-, moderate-, and low-risk groups. Risk declined over time only in the moderate-risk group. The high-risk drivers appeared to be insensitive to experience, with CNC rates consistently high throughout the 18-month study period, and the moderate-risk group appeared to learn from experience. *(J Pediatr 2013;163:1670-6)*.

M otor vehicle crash risk is higher for teenage drivers than for older age groups. In the US in 2009, the rate of fatal crashes per 100 000 licensed drivers was 39.1 for 15- to 20-year olds, compared with 32.7 for age group 21-24 and 23.7 for age group 25-34 years.[1] Previous studies have shown that the first few months of driving after licensure are the most dangerous.[2] Crash rates decline during the first years of licensed driving but remain much higher compared with experienced adults older than 25 years.[3] However, driving risk may not be evenly distributed in the population of young drivers. Previous studies of young drivers have shown the existence of subgroups of drivers at different risk levels, with some drivers having a much higher risk than others.[4,5]

Teenage driving risk could be affected by various factors. Simulation, test track, and observational studies have shown that novice and young drivers have poorly developed driving skills and are prone to distraction.[6-8] Studies using epidemiologic and observational methods indicate that crash rates and risky behavior among teenage drivers are higher for males than for females, at night, and in the presence of passengers.[3,9,10] Crash risk has been associated with certain personality characteristics, such as sensation seeking and some of the Neuroticism-Extroversion-Openness-Five Factor Inventory (NEO-FFI) personality factors (eg, extraversion, agreeableness, and conscientiousness) in systematic literature review,[11,12] which may partially differentiate risk groups.[4,13]

The Naturalistic Teenage Driving Study (NTDS) used continuous video and data recoding techniques to evaluate driving risk of novice teenage drivers.[14] The high resolution data collected through naturalistic studies provide objective information about crash events and risk factors.[15] The method is particularly valuable for studying novice teenage drivers whose driving performance can be expected to change substantially during the initial months after licensing. Previous findings from the NTDS showed that despite significant declines in crash and near-crash (CNC) rates over the 18-month study period, risk among novice teenage drivers was significantly higher than adults over the entire study period.[16,17]

The purpose of the present study is to examine the variation in risk among novice teenage drivers, specifically the change in CNC rate over time by risk

10K kmt	10 000 km traveled
CNC	Crash and near-crash
NEO-FFI	Neuroticism-Extroversion-Openness-Five Factor Inventory
NTDS	Naturalistic Teenage Driving Study

group and to evaluate the demographic, driving conditions, and personality characteristics of risk groups.

Methods

Forty-two teenage drivers from the New River Valley and Roanoke areas of the Commonwealth of Virginia in the US were recruited within 3 weeks after licensure and followed for 18 months using methods described by Lee et al.[14] Study participants included 22 females and 20 males with a mean age of 16.4 years (SD = 0.33). Teenage participant assent and parent consent were obtained, and teens were provided incentives according to procedures approved by the Virginia Tech Institutional Review Board. A Certificate of Confidentiality was also obtained from the US Department of Health and Human Services to further protect the participants' privacy and data.

The vehicles of the participants were instrumented with an advanced data acquisition system developed by the Virginia Tech Transportation Institute with four video cameras, a global positioning system, radars, a 3-dimensional accelerometer, a speedometer, and other sensors.[14] Participants were instructed to drive as they would normally. The teenage participants and their parents did not have access to the data, and no feedback was provided. The camera views monitored the driver's face and driver side of the vehicle, the forward view, the rear view, and an over-the-shoulder view of the driver's hands and surrounding area. The data acquisition system also recorded radar, global positioning system, vehicle network, and kinematic data, including speed and 3-dimensional acceleration.[14] The video and driving performance data were continuously recorded at 10 Hz (10 data points per second). Data were stored on the onboard hard drive and later downloaded to a secure central database for analyses. Audio during driving was not recorded. All participants were aware of the recording equipment in the vehicles. Previous studies have demonstrated that drivers quickly forget about the recording equipment, and exhibit very natural driving and nondriving behaviors, often within the first hour.[15]

Coders reviewed the video data for each trip and noted a wide spectrum of trip-based information including driver identity and passenger presence. The data coding followed a standardized and rigorous protocol developed in previous studies to ensure the accuracy and reliability of information.[14] Trip length and duration, including total distance traveled, percentage of distance traveled at night, length of individual trip, and percentage of trips traveled by passenger status, were calculated using automated computer algorithms.

A baseline questionnaire survey was administered including demographic information, vehicle sharing status, and two measures of personality. The NEO-FFI questionnaire form S is a 60-item questionnaire with 5 subscales: neuroticism, extraversion, openness to experience, agreeableness, and conscientiousness.[18] Each subscale includes 12 questions with response choices ranging from "strongly disagree" to "strongly agree" and can have a score ranging from 0-48. Reported internal consistency of the NEO-FFI ranges from alpha = 0.68 for agreeableness to 0.86 for neuroticism, test-retest reliability is reported as r = 0.79 to r = 0.83, and acceptable convergent and discriminant validity have been demonstrated.[18] Predictive validity has been established with different outcomes, including college student grade point average.[19] Sensation seeking was assessed with the 40-item Sensation Seeking Scale form V.[20] The measure provided a global score and 4 subscales: thrill and adventure seeking, experience seeking, disinhibition, and boredom susceptibility. Each subscale consists of 10 questions each with a binary choice and can have a score between 0 and 10. The internal consistency alpha ranges from 0.55 for boredom susceptibility and 0.80 for the global score.[20] Studies have demonstrated the association between sensation seeking scale and different behaviors, including risk driving. [21]

Two types of safety critical events, CNCs, were assessed. A crash was defined by Lee et al[14] as "any contact with an object, either moving or fixed, at any speed in which kinetic energy is measurably transferred or dissipated. Includes other vehicles, roadside barriers, objects on or off of the roadway, pedestrians, cyclists, or animals." The definition of near-crash was as follows: "Any circumstance that requires a rapid, evasive maneuver by the subject vehicle, or any other vehicle, pedestrian, cyclist, or animal to avoid a crash. A rapid, evasive maneuver is defined as steering, braking, accelerating, or any combination of control inputs that approaches the limits of vehicle capabilities." CNCs were identified by kinematic triggers (eg, harsh braking ≤ -0.65 g; Lee et al[14] for the full list) followed by visual verification. By definition, near-crashes are safety critical events that threaten safety and should be avoided. Given the generally small number of participants in naturalistic driving studies, combining CNCs provides considerable analytic advantages and is widely adopted in naturalistic driving study analyses.[22,23]

Statistical Analyses

Participants were classified into risk groups based on overall CNC rates during the entire study period. A K-mean cluster method was used to provide an objective classification by partitioning participants into predefined K-clusters by minimizing the within-cluster variations.

Depending on specific metrics, the demographic, driving, and personality characteristics of the risk groups were compared using an ANOVA method or the Cochran-Mantel-Haenszel test.

A mixed-effect Poisson regression approach was used to model CNC rates. The model assumption is that CNC counts are generated from a Poisson process, and the subject-specific effects are incorporated through a random effect term.

Results

During the 18 months of data collection teenage study participants drove more than 71 000 trips involving approximately