Jun 29th, 12:00 AM

# Driving Consistency Errors Overestimate Crash Risk from Cellular Conversation in Two Case-Crossover Studies

Richard Young

*Wayne State University School of Medicine, Detroit, MI*

Follow this and additional works at: https://ir.uiowa.edu/drivingassessment

# DRIVING CONSISTENCY ERRORS OVERESTIMATE CRASH RISK FROM CELLULAR CONVERSATION IN TWO CASE-CROSSOVER STUDIES

Richard Young
Wayne State University School of Medicine
Detroit, Michigan, USA
Email: ryoun@med.wayne.edu

**Summary**: The goal of this study is to help resolve the discrepancy in relative risk estimates between recent and early epidemiological studies of call-crash association. Recent epidemiological studies estimate a crash risk for cellular conversation near that of baseline driving – a relative risk of about one. In contrast, two early case-crossover studies estimated a relative crash risk of about four for cellular conversation while driving. One hypothesis to explain this four-fold discrepancy is that the early studies had less driving time in the control window on a day before the crash, than in the crash window just before the crash. This bias in driving exposure translated into relatively lower exposure to cellular conversation during control windows than during crash windows, thereby introducing an overestimate of the relative risk for cellular conversation while driving. To test this hypothesis, the present study developed a new *driving consistency index* (DCI), which measures the percentage overlap in driving times from one day to the next. The mean DCI for 240 vehicles in a Chicago GPS study with known driving times for two successive days was a surprisingly low 14.8%, substantially below the driving consistency estimates in the early case-crossover studies. After adjustment by the mean DCI, the relative risk estimates for cellular conversation while driving in the early case-crossover studies are about one, resolving the discrepancy with the more recent epidemiological studies.

## INTRODUCTION

This study attempts to provide at least a partial resolution to the discrepancy between different real-world epidemiological studies about the relative risk of cellular conversations while driving when compared to baseline driving. Recent studies estimate little change in baseline crash risk from cellular conversations while driving (a relative risk of about one), while early studies estimate about a four times increase in crash risk compared to baseline driving for cellular conversations. The present study advances the hypothesis that the early studies of calls and crashes had less driving time in a "control window" on a day before the crash, than in a "crash window" just before the crash. I show how this hypothesis, if validated, can explain much if not all of the elevated risk ratios in the early studies. I then develop a new *driving consistency index*, and apply it to objective GPS data of day-to-day driving, in order to test the hypothesis. The hypothesis was validated, and I then go on to adjust the relative risks in the early studies by the new estimates of day-to-day driving consistency. After adjustment by the new index, the relative risk estimates for cellular conversation while driving in the early studies are about one, consistent with the more recent epidemiological studies. These results indicate that the early studies likely overestimated the driving time in control windows on days before a crash, leading to overestimates of the relative risk in those studies.

**Epidemiological Studies of Call-Crash Association**

Four recent epidemiological studies (Klauer et al., 2006; Young and Schreiner, 2009; Olson et al., 2009; Klauer et al., 2010) indicate that the relative risk of a crash while driving and engaging in a cellular conversation is near one, where "one" indicates risk parity with baseline driving. Klauer et al. (2010) did not give an explicit relative risk estimate for calls in their case-crossover study. However, during calls on a hand-held phone, the drivers' eyes were off the road about 10% of the time (Klauer et al., 2010, Fig. 2, p. vi), which in turn was associated with a relative risk estimate near one (Klauer et al., 2010, Fig. 13). [Note: In the present paper, the term *call* refers solely to the *conversation* portion of a cellular call (i.e., talking and listening), and does not include dialing, holding a phone, or any other visual-manual aspect of cellular phone usage.]

Two early case-crossover studies (Redelmeier and Tibshirani, 1997; McEvoy et al., 2005) estimated a relative risk for calls while driving of about four, a substantial discrepancy from a relative risk near one as estimated by the four recent epidemiological studies. Redelmeier and Tibshirani (1997) investigated the call behavior of 699 drivers in Toronto, Canada, who were involved in crashes. The authors compared calls in the ten-minute period prior to the estimated time of the crash (here called the *crash window*) to calls by the same individual during a comparable period on a prior day (the *control window*). For each driver, the call and crash categories of interest were: (a) the driver had calls during the crash window but not during the control window; and (b) the driver had calls during the control window but not during the crash window. In the case-crossover method, the ratio of the number of drivers in category (a) to (b) estimates the relative risk associated with calls while driving. If more drivers fall into category (a) than (b), then calls are associated with a crash risk greater than baseline driving (i.e., a relative risk greater than one). The study concluded that calls were "associated with a quadrupling of the risk of a collision." A similar study by McEvoy et al. (2005) in Australia, with more controls for driving exposure, reached the same conclusion.

Different epidemiological study designs should produce similar effect sizes if the risk is small and no bias is present (Rothman et al., 2008). The four-fold discrepancy in effect sizes between the recent and early epidemiological studies indicates that one or both sets of studies contain substantial bias. An analysis of possible sources of such bias is essential for guiding future research on the issue of calls while driving, a topic of major societal and public policy interest. The current study examines whether the early case-crossover studies fully controlled for errors in driving times during control windows that could have introduced bias in their results.

**Controls for Driving Consistency in Early Case-Crossover Studies**

Redelmeier and Tibshirani (1997) recognized that bias is introduced if someone who crashed did not drive during a control window (what they termed driving *intermittency*). Control windows form the denominator of the relative risk, so underestimation of calls during a control window results in a biased overestimation of relative risk. For example, assume people have a higher calling rate when driving than when not driving. If someone did not drive during a control window (but did drive during a crash window), there would be fewer calls during the control window than during the crash window. It would therefore appear (erroneously) that calls were more likely during a crash window than during a control window, leading to the (false)

conclusion that calls while driving are associated with a relative risk greater than one. The reality is simply that there are two associations that confound the estimate of the relative risk of calls while driving: (1) calls are associated with driving (i.e., people have more calls per minute when driving than when not driving), and (2) driving is associated with crashes (i.e., people who drive have an elevated risk of being involved in a crash). Association (2) almost guarantees that there will be more driving during a crash window than during a control window, because if there is a crash it is highly probable that someone was driving. More driving during the crash window means more calls during the crash window because of association (1). In turn, a greater number of calls during a crash window than during a control window gives rise to an increase in the relative crash risk estimate of calls while driving. However, this increase is entirely due to a third confounding variable, the difference in driving exposure between control and crash windows.

Redelmeier and Tibshirani (1997) attempted to control for driving intermittency bias by conducting a survey of a different group of 100 subjects who were not in the case-crossover study, 35% of whom did not recall driving during a "selected period." Subtracting the 35% intermittency from 100% yields an estimate of 65% for driving consistency. I prefer the newly coined term *driving consistency* rather than its converse driving intermittency, and adopt it for the rest of this paper. To correct for driving consistency bias, Redelmeier and Tibshirani multiplied their crude risk ratio and confidence limits [6.5, with a 95% confidence interval (C.I.) from 4.5 to 9.9], by their 65% driving consistency estimate, reducing their "crude" relative risk estimate to an adjusted value of 4.3 (95% C.I. 3 to 6.5), which is their final published estimate.

McEvoy et al. (2005) used three control windows for each of 744 drivers who crashed, creating 2,232 control windows. Of those, they kept only those 801 control windows where "the driver reported driving." That yields a consistency estimate of 36% (801 ÷ 2,232), or about half the 65% consistency estimate of Redelmeier and Tibshirani. Because McEvoy and colleagues used only those control windows in which drivers recalled driving, they believed their estimated odds ratio of 4.1 (95% C.I. 2.2 to 7.7) did not have to be further adjusted for driving consistency.

However, neither early case-crossover study accounted for bias introduced by part-time driving in control windows. Part-time driving is more likely in control windows than in crash windows because driving before a crash is highly probable. For this reason, driving exposure in control windows relative to crash windows is likely biased low in the early case-crossover studies. Assuming a higher call rate when driving than not driving, relative risk would then be overestimated. A closer examination of driving consistency in the two early case-crossover studies seems warranted. Ideally, one would like to have objective measures of driving times from the original studies, but unfortunately, these do not exist. The present study therefore estimated driving consistency from a GPS study with objective driving times, and used these results to adjust the relative risk estimates in the early case-crossover studies.

**METHODS**

**GPS Database**

The Chicago "Travel Tracker" database compiled travel behavior for a large sample of households in the Chicago region during 2007-2008 (http://www.cmap.illinois.gov/travel-

tracker-survey). Most data were from surveys, but one sub-study collected GPS vehicle data over a 7-day period. The GPS sub-study focused on households with at least one heavy traveler, lastly defined as "making at least 10 trips per day" (Bricka and Wolf, 2008, p. 7). After signing a confidentiality agreement, I received access to an anonymous version of this GPS database, agreed to for research purposes by the subject consent form. The first two days were on average the heaviest travel days of the 7-day GPS sample (Bricka and Wolf, 2008, Table 4), and so those days were conservatively chosen to maximize consistency. There were 240 vehicles, of which 223 had GPS data for both days, and 17 had GPS data for day 2 but not day 1 (which the survey verified was a non-travel day). Separate controls (not shown) produced similar results using only: (1) the 223 vehicles; or (2) the 72% of vehicles where the first two days were weekdays.

**New Driving Consistency Index and Adjustment of Relative Risk**

This study develops a new driving consistency index that can be used to correct for inconsistencies in driving between control and crash days in epidemiological studies. To accomplish this goal, two data arrays were created, each with 1,440 bins corresponding to the 1,440 minutes in each of two 24-hour days. A "1" in a 1-min bin indicated that driving took place during that bin. Day 2 is analogous to the "crash day" in the case-crossover studies, and day 1 to the "control day." For every vehicle, each bin for day 1 was compared to its corresponding bin in day 2. The number of bins with overlapping ones in the two days, divided by the sum of the driving minutes in day 2, defines the *driving consistency index* (DCI) (Eq. 1). In Eq. 1, $day_1(i)$ is the bin entry for the control day (1 if driving, 0 if not), and $day_2(i)$ is the corresponding value for the subsequent day.

$$DCI = 100 \times \frac{\sum_{i=1}^{1440} day_1(i) \times day_2(i)}{\sum_{i=1}^{1440} day_2(i)} \tag{1}$$

The DCI is simply the amount of overlap in the driving times in two successive days, divided by the total driving time for the second day. (To demonstrate robustness, I also reversed the day order, designating day 2 as the "control day" and day 1 as the "crash day," which simply changes the sum of minutes in the denominator from day 2 to day 1, and produced similar results.)

The mean DCI across all the vehicles is the expected driving percentage during an arbitrary control window of any duration and at any time of day (given driving during a corresponding crash window the next day). Consider that the mean DCI divided by 100 gives the probability $p$ of driving in any arbitrary minute on day 1, given a matching minute of driving on day 2. Assume that a person had been driving for $k$ minutes on day 2 and a crash occurs at clock time $t$. The expected minutes that the person would have driven in a $k$-minute control window on day 1 (preceding the corresponding clock time $t$) would be $k*p$. For example, with a mean DCI of 15%, and a 10-min crash window, the amount of expected driving during the corresponding 10-min control window on the previous day would be only 1.5 min (15% of the control window).

I next used Redelmeier and Tibshirani's (1997) out-of-sample adjustment method to adjust their crude risk ratio and confidence limits by multiplying by the mean DCI across all 230 vehicles.

301

However, McEvoy et al. (2005) had *a priori* rejected control windows where there was no recalled driving (0% consistency), so I adjusted their odds ratio and confidence limits by multiplying by the mean DCI across only vehicles with driving on control day (i.e., > 0% DCI).

## RESULTS

### Driving Consistency Index for Chicago GPS Study

The mean number of trips per day for the 240 vehicles was 10, with mean driving time of 68.1 minutes on day 1, and 70.8 minutes on day 2. As expected, the heavier the driving on day 1, the higher the DCI ($r = 0.3$, $t = 4.8$, df = 238, $p < 0.00001$). Fig. 1 shows that 110 of the 240 vehicles (46%) had no driving minutes on day 1 that overlapped with those on day 2 (0% DCI). The 46% "no driving" amount is within the 36% to 65% range for control windows without driving in the early case-crossover studies. The other 54% of the vehicles had at least some driving minutes in the control windows which overlapped those in the crash windows (i.e., a DCI above 0%), but most (93%) had less than a 50% DCI. Only 25 of the 240 vehicles (10%) had a DCI of 50% or more, and only two vehicles (1%) had complete overlap (100% DCI). The mean DCI across all 230 vehicles was a surprisingly low 14.8%.
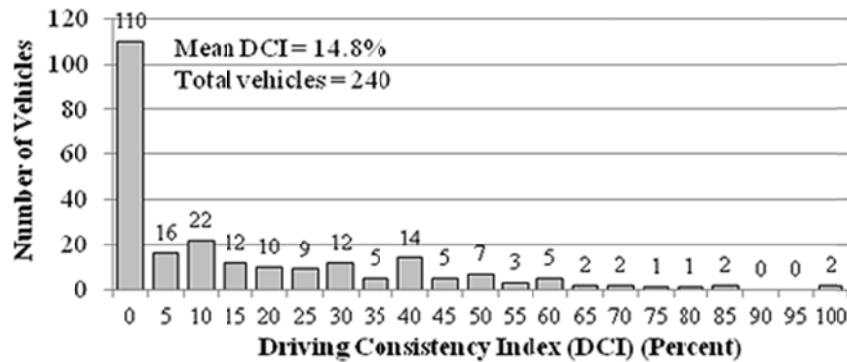


**Figure 1. Consistency of Driving Times on Successive Days**

### Adjustment of Relative Risks of Early Case-Crossover Studies

Multiplying by the mean DCI of 14.8%, Redelmeier and Tibshirani's (1997) crude risk ratio and confidence interval of 6.5 (95% C.I. 4.5 to 9.9) were reduced to 0.96 (95% C.I. 0.67 to 1.5). This adjusted risk ratio of 0.96 is about four times smaller than Redelmeier and Tibshirani's adjusted risk ratio of 4.3, because the current 14.8% consistency estimate is about four times smaller than the 65% estimate in that study. The adjusted risk ratio of 0.96 is in line with the results of recent epidemiological studies, which estimate a relative risk near one for calls while driving.

Before estimating the odds ratio, McEvoy et al. (2005) correctly discarded the 36% of control windows in their study with no recalled driving, but they did not correct for partial driving in the remaining control windows. In the current study, the mean DCI for only those vehicles actually driven on day 1 (i.e., a DCI above 0) was 27.4%. After multiplying by this 27.4% consistency, the odds ratio and confidence interval of 4.1 (95% C.I. 2.2 to 7.7) of McEvoy et al. were reduced to 1.1 (95% C.I. 0.60 to 2.1), again in line with the results of recent epidemiological studies.

## DISCUSSION

The present study developed a new *driving consistency index* (DCI) to measure day-to-day driving time repeatability. The DCI is the percentage of driving time in a control window on a given day, which overlaps with the driving time in a corresponding window on the following day. It accounts for any duration of driving at any time of day, repetitive trips such as daily commutes to work or school, or any other possible driving pattern. The index provides a more accurate and objective estimate of the amount of driving that occurred in control windows on days before a crash, as in the Redelmeier and Tibshirani (1997) and McEvoy et al. (2005) studies. These studies properly removed from their analysis any control windows in which the subject did not recall driving, because people may have more calls per minute when driving than when not driving. However, these studies did not control for part-time driving during control windows, which could reduce the call exposure there, leading to overestimates of relative risk.

In the present study, the mean DCI in driving times measured with objective GPS data is only 14.8%, several times lower than consistency estimates in the two early case-crossover studies based on subjective recall. Hence, the call exposure in the control windows in those early studies is likely biased downward by having less driving there than was assumed. Using the methods of Redelmeier and Tibshirani (1997) to adjust for driving consistency bias, but with accurate and objective GPS data for driving times, the adjusted relative risk estimates in both early case-crossover studies are near one, with a 95% confidence interval from slightly below to slightly above one. These adjusted estimates are in line with those reported in recent epidemiological studies of calls and crashes, including the case-crossover study by Klauer et al. (2010) that had objective measurements of driving and call times in the crash and control windows.

### Limitations of Present Study

A limitation of the present study is that it used GPS data which included at least one heavy traveler in each household (see Methods), and heavy travelers have more consistency (higher DCI values) than average travelers as these results show. If the early case-crossover studies had average travelers, the mean DCI for those travelers would likely be lower than 14.8%, indicating that the relative risks in those studies should be adjusted to even lower values than shown here.

In addition, the present study suffers from the same limitation as the study of Redelmeier and Tibshirani (1997) in that it did not use the same drivers as in the case-crossover analysis to estimate driving consistency, and so loses the benefits of the case-crossover method (Maclure and Mittleman, 1997). Indeed, the current data is from a different sample, in different areas (and different countries), and at different points in time from the case-crossover studies. There may be differences between driving styles in Chicago and in the Canadian and Australian locales where the case-crossover studies were conducted. There is no data on cell phone ownership or usage rates in the Chicago GPS database, so it likely included drivers who did not own or use cell phones while driving. The early case-crossover subjects were all cellular phone users, who tend to drive more per day than non-users (Laberge-Nadeau et al., 2003). Those who drive more per day have more consistent day-to-day driving patterns, as shown by the present results. Hence, the present study may have underestimated the DCI relative to that of a sample of all cellular phone users, and therefore may have adjusted the relative risk estimates too far downward.

Given these limitations, the present study does not provide definitive proof that the high relative risk estimates in the early case-crossover studies are entirely attributable to driving consistency errors. One such proof would be a replication of the exact window methods of Redelmeier and Tibshirani (1997) and McEvoy et al. (2005) in an instrumented vehicle study, but with objective and direct measures of call times (while driving and not driving) for the same individuals. Unfortunately, no such direct objective yet exists (for example, none of the recent epidemiological studies of calls and crashes had information on call times when not driving). The DCI should also be calculated with GPS data from additional cities, with randomly-selected subjects. Ideally, a direct DCI comparison should be made between recalled and GPS-based driving times for two or more days of travel for the same subjects, but such data does not exist.

Despite these limitations, the present study establishes as plausible the hypothesis that the two early case-crossover studies overestimated relative risk for calls while driving due to an underestimation of driving exposure in the control windows used in their computations -- and, furthermore, demonstrates that such a hypothesis likely accounts for much of the discrepancy between the recent and early relative risk estimates associated with calls while driving. Thus, the new driving consistency hypothesis cannot be excluded as an explanation of the elevated risk ratios in these early case-crossover studies, until new objective data are collected.

**Cautions about Visual-Manual Aspects of Cellular Phone Use while Driving**

Cell phone "usage" is actually composed of two separate types of tasks – conversation (auditory-vocal), and visual-manual. This paper indicates that the risks associated with the conversation portion of a cellular call may have been overestimated in some early epidemiological studies. However, these results should not be interpreted to mean that all aspects of cellular phone usage during driving is therefore "safe." Indeed, the same recent epidemological studies that estimate a relative risk near one for conversations while driving, have estimated many-fold increases in the visual-manual aspects of cell phone usage – e.g., a 3-7 times increase in relative risk for manual 10-digit dialing or reaching for a cellular phone, and a 23 times increase for texting (Klauer et al., 2006; Olson et al., 2009). The on-road experimental results of Angell et al. (2006) are consistent with this dichotomy in finding less degradation in driving performance associated with auditory-vocal tasks during which drivers keep their eyes on the road (as during calls), compared with visual-manual tasks during which drivers take their eyes off the road (as during dialing).

**CONCLUSION**

Objective GPS data indicates that two early case-crossover studies likely overestimated the amount of driving during control windows, which translates into lower exposure for cellular conversation during control vs. crash windows, introducing overestimates of relative risk.

**ACKNOWLEDGMENTS**

## REFERENCES

Bricka, S., & Wolf, J. (2008). Chicago regional household travel inventory draft GPS final report. 1-30. Retrieved from http://www.cmap.illinois.gov/documents/20583/2af38a79-5084-426d-bcf9-867f9c50eab7, April 28, 2011.

Klauer, S. G., Dingus, T. A., Neale, V. L., Sudweeks, J. D., & Ramsey, D. J. (2006). The impact of driver inattention on near-crash/crash risk: An analysis using the 100-car naturalistic driving study data. (DOT HS 810 594). NHTSA, Washington, D.C.

Klauer, S. G., Guo, F., Sudweeks, J., & Dingus, T. A. (2010). An analysis of driver inattention using a case-crossover approach on 100-car data: Final report (DOT HS 811 334).

Laberge-Nadeau, C., Maag, U., Bellavance, F., Lapierre, S. D., Desjardins, D., Messier, S., & Saïdi, A. (2003). Wireless telephones and the risk of road crashes. *Accident Analysis & Prevention, 35*(5), 649-660.

Maclure, M., & Mittleman, M. A. (1997). Cautions about car telephones and collisions [editorial]. *The New England Journal of Medicine*, *336*(7), 501-502.

McEvoy, S. P., Stevenson, M. R., McCartt, A. T., Woodward, M., Haworth, C., Palamara, P., & Cercarelli, R. (2005). Role of mobile phones in motor vehicle crashes resulting in hospital attendance: A case-crossover study. *BMJ*, *331*, 428-430.

Olson, R. L., Hanowski, R. J., Hickman, J. S., & Bocanegra, J. (2009). Driver distraction in commercial vehicle operations. U.S. DOT, Report No. FMCSA-RRR-09-042, Wash., D.C.

Redelmeier, D., & Tibshirani, R. (1997). Association between cellular-telephone calls and motor vehicle collisions. *The New England Journal of Medicine*, *336*, 453-458.

Rothman, K., Greenland, S., & Lash, T. (2008). *Modern Epidemiology* (Third Ed.). Philadelphia: Lippincott Williams & Wilkins.

Young, R. A., & Schreiner, C. (2009). Real-world personal conversations using a hands-free embedded wireless device while driving: Effect on airbag-deployment crash rates. *Risk Analysis*, *29*(2), 187-204.