COMMENTARY

Lethal misconceptions: interpretation and bias in studies of traffic deaths

Donald A. Redelmeiera,b,c,d,*, Christopher J. Yarnellb,d

a Department of Medicine, University of Toronto, Toronto, ON, Canada
b Clinical Epidemiology Program, Sunnybrook Health Sciences Centre, University of Toronto, Toronto, ON, Canada
c Institute for Clinical Evaluative Sciences, Toronto, ON, Canada
d Injury Prevention Program, Sunnybrook Research Institute, University of Toronto, Toronto, ON, Canada

Accepted 15 September 2011; Published online xxxx

Abstract

Clinical epidemiology studies are vulnerable to subtle confounding, leading skeptics to claim that an odds ratio below three rarely indicates a clinically important finding. We argue that such a high threshold is inappropriate when interpreting traffic death studies in clinical epidemiology research. We review 10 concepts that emphasize the value of modest effect sizes by taking into account the baseline frequency, nonfatal disability, numbers needed to treat, shared responsibility, event diversity, behavioral offsets, measurement error, indirect reinforcement, delayed progression, and economic affordability. An awareness of these concepts may help when interpreting effect sizes in studies of traffic deaths. © 2011 Elsevier Inc. All rights reserved.

Keywords: Road injury; Major trauma; Clinical reasoning; Evidence-based medicine; Effect size interpretation; Clinical significance

1. Introduction

A complacent attitude toward traffic deaths might be reasonable if crashes were difficult to predict, hard to prevent, or costly to avoid. In a series of articles, however, our group has shown that seemingly minor differences in behavior can lead to measurable differences in traffic deaths. For example, U.S. presidential elections lead to a 19% increase in traffic deaths during the hours of polling and the average Super Bowl leads to a 41% increase in traffic deaths during the hours following the broadcast [1,2]. These and other studies [3–11] call into question community attitudes that characterize accidents as random acts, divine planning, or unavoidable fates. Instead, scientific studies prove how easily the risks of a traffic death can be increased or decreased throughout society.

Studies of traffic deaths, however, often fail to change clinical practice for at least three reasons. First, the traditional practice of medicine focuses more on relieving suffering in the aftermath of an event rather than primary prevention in otherwise asymptomatic patients [12]. Second, the failures to prevent traffic deaths can stay invisible to clinicians if reports are not relayed back to health care providers [13,14]. Third, effective methods for preventing traffic deaths rarely take the form of prescribing a medication or a surgical procedure; instead, most interventions involve added inconvenience to the patient and clinician [15]. As a consequence, formal medical training and ongoing clinical practice can accentuate rather than mitigate complacent attitudes toward traffic deaths.

We wondered whether the inbuilt structure of traffic death studies also thwarts progress for more thoughtful attention and better patient outcomes. In accord with scientific standards, effect sizes in traffic death studies are usually quantified with numerical units such as an odds ratio or relative risk reduction [16]. In turn, the relative risk reductions almost never reach 100% and sometimes range to single-digit integers (Table 1). Well-intentioned readers, therefore, may have difficulty distinguishing trivial from substantial levels of effectiveness. The purpose of this article is to explain why small differences in observed risks...
Table 1. Effect size in studies of traffic deaths

<table>
<thead>
<tr>
<th>Concept</th>
<th>Example</th>
<th>Risk ratio (95% CI)</th>
<th>Explanation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Baseline frequency</td>
<td></td>
<td>2.33 (1.43–4.03)</td>
<td>Although modest effect size, absolute increase over 10-year study was 12,954 deaths</td>
</tr>
</tbody>
</table>
| Nonfatal disability | | 1.39 (1.04–1.59) | Every teen crash fatality associated with 15–40 times as many 15–40 years old compared with those aged 16–20 | Concept Example Risk ratio (95% CI) Explanation
| Misleading NNT | | 2.32 (1.43–3.7) | Every teen crash fatality associated with 15–40 times as many 15–40 years old compared with those aged 16–20 |
| Shared responsibility | | 1.26 (1.03–1.54) | Increased crash frequency understates degree of age-related decline in ability |
| Event diversity | | 2.6 (1.03–6.66) | Distinguishing crashes because of glycemic complications would likely increase effect size estimate |
| Behavioral offset | | 2.2 (1.02–1.04) | Although modest effect size, absolute increase over 10-year study was 12,545 deaths |
| Measurement error | | 2.00 (1.89–2.12) | Alcohol dependence notoriously difficult to diagnose and detect, leading to measurement error |
| Indirect reinforcement | | 2.43 (1.24–4.89) | Dangers of sleeping behind wheel become more appreciated by commercial truckers over time |
| Delayed progression | | 2.20 (1.89–2.34) | Once prevented, a fatal crash on one day is not rescheduled to another day |
| Economic affordability | | 2.18 (1.92–2.56) | Front seatbelts are cost saving to society, despite the modest risk reduction effect |

Abbreviations: CI, confidence interval; NNT, number needed to treat.
increase the risk of mortality, yet often do little to increase disability. Moreover, patients with small aneurysms often have no symptoms unless they rupture. In contrast, those factors that contribute to traffic collisions are distinct because they can increase both the risks of mortality and disability. Traffic studies often express results as counts of short-term deaths, thereby leading readers to overlook the wider spectrum of harm or benefit that might be altered by the seemingly modest effect size.

2.3. Misleading numbers needed to treat

The number needed to treat (NNT) has gained popularity in medical journals as an accepted metric for gauging the clinical importance of effective interventions. An estimated NNT, however, is meaningless without stating the duration of the effect [33]. Because of a limited duration of follow-up, traffic death studies generally underestimate a person’s lifetime risk of involvement in a road crash and thereby lead to an overly pessimistic estimate of the NNT. The average person in North America has a lifetime risk of about 99% for being involved in a serious police-reported motor vehicle collision (and many are involved in more than one such crash over their lifetime) [10]. As a consequence, a seemingly minor odds ratio of 1.06 might equate to an impressive NNT of about 16 if extrapolated to a patient’s full destiny.

Other medical disorders are often treated with interventions whose published NNT overestimates the anticipated benefit for patients [34,35]. In outpatients with depression, for example, a double-blinded randomized controlled trial suggested that the addition of pindolol to fluoxetine led to a significant reduction in depression severity that corresponded to an estimated NNT of about 4.8 (where the referent treatment was fluoxetine without pindolol) [36]. This NNT, however, is somewhat misleading because the study design entailed a pre-study run-in during which researchers screened out resistant patients as well as placebo responders [37]. In contrast, most traffic death studies are field research that does not engineer results that overstate an intervention’s efficacy in the community.

2.4. Shared responsibility

A distinct bias arises in traffic death studies that does not appear in most other clinical research because of the shared nature of most road crashes [38]. Consider that a particular driver might be at-fault in half of his or her crashes and not at-fault in the remaining crashes. If so, a risk factor that leads to a 100% increase in the frequency of error for the particular driver might result in only a 50% increase in the overall crash risks for the particular driver. Similarly, a 100% increase in observed crash risk might actually signify a 200% worsening in the driver’s ability. The same shared ecological effects also mean that directly decreasing the traffic death risk for one patient in the community may indirectly decrease the risk for many others in the community as well.

Other medical disorders usually do not exhibit such shared responsibility, aside from cases of communicable infections or environmental toxins. The risk of coronary ischemia in a middle-aged woman, for example, is mostly a reflection of her genetics, age, and lifestyle (such as whether she smokes) [39]. Her risk is only a partial reflection of whether there are others at home who also smoke [40]. If the patient stops smoking, therefore, her own risk of coronary ischemia attenuates greatly regardless of whether others at home continue to smoke. For most medical diseases, a patient’s risk is direct and not buffered to a large degree by surrounding circumstances.

2.5. Event diversity

Most science generally acknowledges the complexity of patient outcomes and the distinction between all-cause mortality and disease-specific mortality. A perfect cure for lung cancer, for example, would be heralded as a triumph even if the treatment did nothing to lessen colon cancer or prostate cancer. By analogy, therefore, traffic death studies might set an overly lofty target when they include overall crash risks regardless of whether the incident was a head-on crash, an alcohol-involved event, or a roadway washout. A more meaningful approach might be to study a specific crash risk difference where expected and the absence of a different specific crash risk difference where not expected.

Other medical disorders are often narrowed to highly selected subgroups that become the sole focus of inquiry. Lung cancer studies, for example, usually distinguish small-cell lung cancer from non–small cell lung cancer (and confine the comments to just one subtype). In a recent issue of *Lancet*, each of these lung cancers was reviewed by different authors with only occasional mention of the other subtype in each article [41,42]. Although traffic death studies dedicated to specific types of crash events exist [43], they are rare. In traffic death studies, the focus on global event types can dilute the effect sizes of risk reduction ratios and obscure the full complexity of different types of traffic deaths.

2.6. Behavioral offsets

Another rudimentary bias occurs in traffic death studies because the research is difficult to conduct in a blinded manner. As a consequence, a driver’s conscious and subconscious factors can influence final outcomes and distort the estimates of effectiveness. This is termed “risk compensation” in literature where observed benefits fall short of expected benefits [44]. A trend toward more aggressive speeds appears in vehicles equipped with anti-lock brakes [45], for example, and causes the inherent advantages of a positive improvement to become undermined by associated overconfidence of drivers. An intervention that might reduce a driver’s risk by 10% under controlled circumstances, yet induces a 4% increased subjective willingness to take risks under normal
circumstances, might thereby result only in an apparent 6% reduction in overall traffic risk.

Other medical disorders could be prone to behavioral offsets, yet such anomalies are rarely described in clinical trials. Pregnant women prescribed folate supplements, for example, decrease their risk of neural tube defects and do not seem to compensate by drinking increased amounts of alcohol or engaging in other risky exposures [46]. Indeed, positive feedback provided to patients generally leads to increases in adherence for many chronic disorders [47–49]. In contrast, people usually drive with less care when they see positive changes, such as sunny weather rather than stormy rain [50]. Some self-regulation is common among senior drivers [51], yet the path of least resistance in general driver behavior is often to cut corners and exchange safety for convenience.

### 2.7. Measurement error

Traffic death studies are usually conducted as a field research lacking the stringent controls of laboratory science. Even elementary data on a driver’s age and gender can be missing in some cases despite rigorous investigation by police and other authorities [52]. Highly nuanced data about the driver’s behavior, outcomes, and compliance are usually missing or extremely fallible [53]. In epidemiological science, this degree of measurement error is tantamount to some random misclassification that biases results toward the null. A true odds ratio of 2.00, for example, reduces to an observed odds ratio of only 1.86 if subjected to 5% random misclassification in both the exposed and not-exposed groups.

Other medical disorders more often have treatments that are tested in highly controlled settings with extremely precise baseline measures or other strategies that create a gap between efficacy and effectiveness [54]. A double-blinded randomized clinical trial comparing tiotropium and salmeterol for chronic obstructive lung disease patients, for example, featured 8 inclusion criteria and 22 exclusion criteria (including “significant diseases other than chronic obstructive pulmonary disease”) [55]. The trial found that tiotropium reduced the rate of severe exacerbations by 27% relative to salmeterol, with a respectable confidence interval of 18–34%. Traffic studies rarely apply such rigorous selection and outcome criteria, and thereby dilute both the estimated risk reduction ratios as well as the corresponding 95% confidence intervals.

### 2.8. Indirect reinforcement

Driver’s behavior is a far more socially interactive process than typical medical diseases that stay private within a patient or family. As a consequence, the full force of an intervention is rarely apparent in traffic death studies until widespread community adoption occurs. The effectiveness of seatbelt laws on preventing traffic deaths, for example, will generally be limited by the degree of compliance in the community. The surprising pattern, however, is that the compliance rates with seatbelt laws continue to steadily rise even in high-income countries that enacted laws many decades earlier [56]. The net result of delayed indirect community reinforcement is that early studies evaluating traffic death interventions may significantly underestimate their long-term steady-state effectiveness.

Other medical disorders rarely have such dramatic indirect effects aside from the recognized dangers of antibiotic resistance because of overusage. *Staphylococcus aureus*, for example, is notorious for rapidly developing resistance to antibiotics, including the emergence of vancomycin-resistant strains, because of recent widespread usage of vancomycin against methicillin-resistant *S. aureus* [57]. Although articles decrying the overuse of antibiotics are common [58], articles decrying the overuse of seatbelts are now difficult to find. A new antibiotic may demonstrate a more impressive initial effect size than a new traffic safety intervention, but the benefits of the traffic safety intervention may accentuate rather than attenuate with time.

### 2.9. Delayed progression

Traffic deaths are distinct from many other medical diseases because each death is entirely preventable. Death rates because of myocardial infarction, in contrast, can often be delayed—but not averted—with treatment of hyperlipidemia and other atherosclerotic risk factors. That is, medical care can yield a nominal decrease in deaths from myocardial infarction over a limited time horizon but the disease still progresses to cause eventual adverse outcomes. Hence, an observed relative risk reduction for many patients actually equates to changing the timing but not the actuality of an adverse event. In contrast, an observed relative risk reduction in a traffic death study often indicates that the adverse event was fully avoided rather than rescheduled to a later time.

Other medical disorders are not usually eliminated with effective treatment, and sometimes therapy only delays the rate of progression. Patients with cerebrovascular disease who are treated with aspirin tend to have a reduced rate of stroke, but the absolute change in risk is about 1% per year. That is, such antiplatelet agents serve mostly to slow the rate of recurrence but not to reverse the fundamental underlying pathology [59]. Traffic safety interventions such as graduated licensing can completely prevent some traffic deaths because most risk reduction ratios in traffic studies represent outcomes that are eliminated, not just delayed [60]. The time horizons of most research can disguise this distinction and mislead well-intentioned readers who compare effect sizes across disciplines.

### 2.10. Economic affordability

A final nuance is that road crashes are remarkably costly events causing both human suffering and widespread
property damage. In the United States, for example, the average crash equates to 2–10 thousand dollars in property damage and the total annual costs of crashes accumulates to about 350 billion dollars in societal losses [61]. These costs are paid by members of society through higher taxes, increased insurance premiums, and raised consumer prices [62]. Hence, an intervention that could reduce road crashes by one-tenth of one percent could equate to savings on the order of 350 million dollars each year in the nation. Unlike the treatment of most medical disorders, effective interventions for traffic deaths can be cost saving to society.

Other medical diseases can rarely claim such large economic gains consequent to intervention. Programs for liver transplantation in patients with primary sclerosing cholangitis, for example, ultimately improve patient outcomes, yet cost about $41,000 per additional quality-adjusted life year when compared with no transplantation [63]. Hence, the medical intervention is effective to patients but not cost saving to society. Traffic safety interventions such as seatbelts and graduated licensing are both beneficial and cost saving if not taken to extremes [18]. In an era of worrisome economic deficits, interventions with dominating cost-benefit ratios should be welcomed even if their effect size is humble.

3. Summary

This article has summarized 10 reasons why small differences in observed risks may sometimes represent important findings in traffic death studies. Much of the material will be familiar to investigators skilled in statistical design; however, the framework may help guide those who are either inexperienced or confronted by an unenthused audience. Each of the 10 concepts can explain why a seemingly minor odds ratio can have a major implication to human health related to traffic risks. However, this article says nothing about the offsetting conflict of interest from industries that gain financially from deliberately downplaying the risk of traffic death [64]. In addition, the summary steers clear of the litigation strategies that cause dramatic claims around who is to be blamed for a particular crash [65].

The main limitation in the current review is the lack of scientific insight on emotional motivations that fundamentally underline the neglect of traffic deaths. Driving is such a common activity, for example, that readers generally have strong preconceptions before encountering any new study [66]. Furthermore, the physics of the situation seems so straightforward that nihilists might assume clinical epidemiology science has little to offer. Because so much quality of life depends on road travel [67,68], people may paradoxically be averse to learning more about traffic deaths. The shortfall in the funding of traffic death studies [69], as well, almost guarantees that studies can be additionally faulted because the science can rarely produce perfect data. These and other emotional factors can lead people to dismiss almost any odds ratio observed in a traffic death study.

Traffic deaths have fallen by about 30–40% in many industrialized countries over the past 40 years [70]. This reduction exceeds the corresponding reduction in cardiovascular deaths for the same countries during the same interval [71,72]. The reduction in traffic deaths has not been caused by any single intervention that had a published odds ratio typical of cardiovascular studies published in the medical literature. Instead, the advance has been the net result of interventions with remarkably modest odds ratios such as efforts to increase the use of seatbelts [73,74], stop drinking and driving [75,76], reduce excessive speed [77–79], and minimize distractions [3,5,80]. The future years will hopefully bring more scientific studies with modest effect sizes that mitigate the risk of traffic deaths.

Acknowledgments

This project was supported by the Canada Research Chair in Medical Decision Sciences and the Comprehensive Research Experience for Medical Students program of the University of Toronto Faculty of Medicine. The views expressed in this article are those of the authors and do not necessarily reflect the Ontario Ministry of Health. We wish to thank the following for helpful comments: William Chan and Neil Dattani.

References


[69] Nicholl JP. Health research funding: while injuries were ignored. *BMJ* 2006;332:1510.


