Recreational Boating Fatality Statistics: an exploration of fatality rate calculations and comparisons

Background

In 2015, a charge team of NASBLA’s Engineering, Reporting & Analysis Committee (ERAC) continued digging into the data and findings generated by the U.S. Coast Guard’s 2012 National Recreational Boating Survey (NRBS). On the heels of its 2014 quest to understand what had gone into the production of the survey’s exposure hour estimates and then convey that information to NASBLA members, the team identified three other priority topics to study in the new charge cycle.

One of the research priorities was to examine the pros and cons associated with using survey-generated exposure hours versus state-reported registered boats as the denominator in boating risk equations. The examination led the team to also consider approaches that could be used to compensate for the reality that survey-generated exposure hours data—generally regarded as the preferred measure for calculating risk—will not be produced every year. Later in this research brief, we take a look at the results of that preliminary investigation into other variables identified for their potential, meaningful relationship with exposure hours; that are collected more frequently; and that—for the interim between survey collections—might eventually be used to “predict” or serve as “indicators” for exposure hours (see pp. 13-16).

But as is often the case in such explorations, the outcomes spawned more questions, prompting the team to take a step back to evaluate the reasons why it was even worthwhile to identify potential approaches for developing interim indicators of exposure hours. As such, the bulk of this research brief discusses work initiated in the 2015 charge cycle to address some related and fundamental questions about risk calculations and what it takes to standardize—or what will more accurately be referred to here as “normalize”—the data in order to make valid casualty rate comparisons between recreational boating and other activities, and even between states.

Before proceeding, however, two caveats must be noted. The first is that this brief focuses on calculations involving recreational boating fatality data as these are typically considered to be more accurate than non-fatal injury or property damage data. The second is that the calculations illustrated in this brief are the result of exploratory data analysis—that is, examining various

---

1 Brief prepared by Dr. L. Daniel Maxim, U.S. Coast Guard Auxiliary, and Dr. Deborah A. Gona, Research Consultant to NASBLA. Dr. Maxim is a member of the ERAC committee, and Dr. Gona serves as staff to ERAC.
3 1) Explore the pros/cons of using exposure hours v. registered boats as the denominator in risk equations; 2) Consider the methods to address ongoing issues regarding the calculation of casualty and fatality rates; and 3) As part of the post processing and evaluation of the NRBS, compare its estimates and findings with estimates from other surveys.
4 “Normalize” means to adjust by some mathematical manipulation so as to provide a more useful statistic.
ways of looking at data—as opposed to adjudicatory data analysis, the production of finished analysis. As we learn more and have the opportunity to further check all sources of data and replicate the computations, we will either be able to make them more rigorous or we will determine that they cannot be made sufficiently accurate for valid use. For this reason, the contents of this brief should be viewed more as a “work in progress” rather than a completed study. Nonetheless, its content should spur critical thinking and discussion about these topics.

In this brief

There are a number of options for normalizing fatality statistics data, particularly as they relate to recreational boating. This brief illustrates several methods for doing so, and depending on the research question being asked, each method, with one exception, may be appropriate.

As described in the following pages, one particularly useful basis for making comparisons across transportation modes or various activities is the fatality rate measured in units of fatalities per million exposure hours. The calculated rate for recreational boating, based on 2012 NRBS estimates, is approximately 0.44 fatalities per million boat exposure hours. To put this figure into perspective, this brief also presents calculations for comparable rates associated with other activities ranging from “low risk,” such as working in an office, to “high risk,” such as sky diving or BASE jumping. These preliminary calculations suggest that the fatality rate for operators or passengers in recreational boats is approximately equal to that for operators or passengers in motor vehicles. Activities carrying greater risk on a per exposure hour basis include BASE jumping, sky diving, hang gliding, operating or being a passenger on a motor cycle, and bicycling. Lower risk activities include bus and train travel and many occupations.

Finally, because exposure hours offer a convenient basis for normalizing recreational boating fatality rates, this brief also includes the results of the ERAC charge team’s consideration of other variables that could be used to estimate exposure hours for those years where the NRBS-generated data are not available.

---

5 See, for example, the description of exploratory data analysis at [http://www.itl.nist.gov/div898/handbook/eda/section1/eda11.htm](http://www.itl.nist.gov/div898/handbook/eda/section1/eda11.htm)

6 The U. S. Coast Guard Office of Auxiliary and Boating Safety has indicated that it fully supports additional research using data derived from the NRBS. However, it does not specifically endorse the results of any exploratory data analysis until results are reviewed and made final. The ERAC charge team and the primary authors of this brief appreciate the constructive input of Dr. Philippe Gwet, U.S. Coast Guard.

7 The only exception is in the case of fatalities per vehicle mile, often used for comparing fatality rates for various transportation modes.

8 Recreational boating exposure rates can be calculated and presented in different ways, including boat hours and person hours. For purposes of the calculations intended for comparisons between recreational boating and other activities (most of which do not engage multiple participants), boat exposure hour estimates are used. Moreover, the calculations for motor vehicles, presented later, are on this same basis.

9 BASE is an acronym for Buildings, Antennas, Span (bridges), and Earth (cliffs), and is one of many “extreme sports” growing in popularity. It differs from skydiving in several respects. Typically the BASE jumpers wear a wingsuit, enabling them to glide some distance before opening a parachute. BASE jumpers also typically jump from lower altitudes, which are more dangerous. To qualify for a “BASE number,” the jumper must have jumped from all four launch points (BASE) and to qualify for a “night BASE number” this feat must have been done at night.
Why “normalize” fatality data or rates?

The intent of normalization is to convert fatality data to common units so that meaningful comparisons or conclusions can be drawn.

Normalizing the data might not always be necessary. Whether or not data need to be transformed depends on the research question to be answered, and for some questions, “untreated” or raw data will be sufficient. For example, if our intent is to learn whether diabetes or major cardiovascular disease is the “bigger” public health problem in the United States, the raw annual fatality numbers would be appropriate: for the year 2013, there were 75,587 deaths due to diabetes compared to 796,494 deaths due to major cardiovascular diseases; heart disease killed more than 10 times as many people as diabetes.\(^\text{10}\)

Sometimes, too, there is interest in fatality data within some restricted group of activities, such as comparing fatalities among categories of transportation. To illustrate, in February 2015, the National Transportation Safety Board (NTSB) assembled and presented preliminary 2013 fatality data for various transportation modes, including marine transportation\(^\text{11}\) (see Figure 1). According to these raw data, recreational boating fatalities were far greater than those of any other marine activity—they accounted for 560 (91.1 percent) of the 615 marine transportation fatalities that year. Of course, the relative proportion of total fatalities depends upon which other activities are included. Motor vehicle fatalities, for example, are substantially greater than boating fatalities, so including other transportation modes would alter the raw data comparisons.

**FIGURE 1. Marine transportation fatalities in 2013 (source: NTSB)**

<table>
<thead>
<tr>
<th>Activity</th>
<th>Fatalities</th>
</tr>
</thead>
<tbody>
<tr>
<td>Recreational boating</td>
<td>560</td>
</tr>
<tr>
<td>Commercial fishing</td>
<td>24</td>
</tr>
<tr>
<td>Commercial passengers</td>
<td>18</td>
</tr>
<tr>
<td>Cargo transport</td>
<td>13</td>
</tr>
</tbody>
</table>

\(^{10}\) See [http://www.cdc.gov/nchs/data/nvsr/nvsr64/nvsr64_02.pdf](http://www.cdc.gov/nchs/data/nvsr/nvsr64/nvsr64_02.pdf)

Normalization by participants

To answer other types of research questions, though, normalizing the data in some way will produce sharper or more relevant comparisons.

Suppose we are interested in comparing the relative likelihood of dying in a motor vehicle accident to dying in a recreational boating accident. Using just the raw fatality data, as described on page 3, would provide one basis for comparison. In 2010, for example, 35,332 persons reportedly died in motor vehicle accidents, \(52.6\) times greater than the number (672) who died in recreational boating accidents in the same year. Thus, in 2010, motor vehicle fatalities were far more numerous than recreational boating fatalities. The problem is that this offers a limited comparison—it fails to take into account the number of persons who use motor vehicles compared to those who use boats and the annual hours of operation (one measure of exposure).

An improved basis for comparing automobile accidents to boating accidents, then, might be to compute the annual risks of dying among those who participate in each activity:

- Data from the Insurance Information Institute for 2010 indicate (as noted previously) that 35,332 persons died in motor vehicle accidents in that year; that works out to annual odds of about one in 8,763. To calculate this figure, you need to divide the number of persons who rode in a motor vehicle in 2010, which is approximately equal to the entire U.S. population of 309.3 million in that year, by the number of persons who died in motor vehicle accidents in that same year. But this is an aggregate figure that includes deaths of men and women of all ages, drunk or sober, driving or riding in all types of motor vehicles, on all types of roads, during the day and at night, in all weather conditions, and so on. It also includes some persons who occupied the vehicle for just a short time and others who had very much greater exposure. As such, it might not be an accurate estimate of any particular individual’s likelihood of dying in a motor vehicle accident.

- Data from Recreational Boating Statistics 2010 indicate that there were 672 boating deaths that year. Using the 2012 NRBS estimate of the number of boaters who participated in recreational boating as a stand-in for the same quantity in 2010, some 74,537,000 persons participate in recreational boating annually; this works out to annual fatality odds of approximately one in 110,918. Similar to the motor vehicle statistics, this aggregate figure includes occupants of all types of recreational boats under all weather and visibility conditions, day and night, drunk or sober, in all types of waters, and so on. Risks for any single individual could be substantially different than this average risk, especially considering the amount of time spent on board the vessel.

So, by this second measure, based on participation estimates, the annual odds of dying in a boating accident also are substantially lower than those for dying in a motor vehicle accident—lower by a factor of 110,918/8,763 or 12.6. This ratio is arguably a more relevant measure of relative odds since the number of persons who engage in each activity annually is explicitly taken into account. But, while this “improved” calculation is technically correct, it neglects the fact that most people spend many more hours in their cars annually than they spend in boats, and does not account for other risk factors mentioned above.

---

12 See [http://www.iii.org/fact-statistic/mortality-risk](http://www.iii.org/fact-statistic/mortality-risk)
14 This calculation is 74,537,000/672 = 110,918.
Normalization by exposure hours

A third potentially relevant comparison of transportation modes, then, would be to calculate fatalities per hour of exposure. To do this, we need additional data on exposure hours for each.

- The National Highway Traffic Safety Administration (NHTSA) publishes annual data on highway fatalities, but not exposure hour estimates. Rather, NHTSA uses vehicle miles as the exposure measure. While this might be appropriate for automobiles, buses, motorcycles, and similar modes, it is not the best measure for boats. So, we need to convert vehicle miles to vehicle hours to attempt a comparison. To convert the motor vehicle fatality rate from fatalities per vehicle mile to fatalities per vehicle exposure hour, we use an average speed (in MPH). See Table 1. To smooth possible year-to-year variability, data are included for 2002 to 2012.\textsuperscript{15} \textbf{The result? The average fatality rate for motor vehicles on an exposure hour basis over the period from 2002-2012 is equivalent to approximately 0.42 fatalities per million vehicle exposure hours.}\textsuperscript{16,17}

\begin{table}[h]
\centering
\begin{tabular}{|c|c|c|c|}
\hline
Year & Total Fatalities & Vehicle Miles Traveled (billions) & Source \\
\hline
2003 & 42,884 & 2,890 & \\
2004 & 42,836 & 2,965 & \\
2005 & 43,510 & 2,989 & \\
2006 & 42,708 & 3,014 & \\
2007 & 41,259 & 3,031 & \\
2008 & 37,423 & 2,977 & \\
2009 & 33,883 & 2,957 & \\
2010 & 32,999 & 2,967 & \\
2011 & 32,367 & 2,946 & \\
\hline
Total & 426,435 & 32,538 & \\
\hline
\end{tabular}
\end{table}

\begin{table}[h]
\centering
\begin{tabular}{|c|c|c|c|}
\hline
Assumed average speed (MPH) & Exposure hours (million) & Fatalities per million hours & Notes \\
\hline
20 & 1,626,907 & 0.262 & By some estimates (see http://www.ridetowork.org/transportation-fact-sheet) the average vehicle speed is 32 MPH, which would make the average fatality rate approximately 0.42 per million vehicle hours. \\
25 & 1,301,525 & 0.328 & \\
30 & 1,084,604 & 0.393 & \\
\textbf{32} & \textbf{1,016,817} & \textbf{0.419} & Failure Analysis Associates estimate is 0.47 per million year unspecified. http://cyclehelmets.org/1026.html. \\
35 & 929,661 & 0.459 & \\
\hline
\end{tabular}
\end{table}

\textsuperscript{15} Reported 2010 fatalities are slightly different than used above, a minor discrepancy among reporting sources. 
\textsuperscript{16} Calculated from the total vehicle miles traveled over the period divided by average speed (32 MPH) to get exposure hours in millions. The fatality rate is total fatalities divided by total exposure hours over this same period. Rates for other average speed assumptions illustrate the sensitivity of the estimate to the assumed average MPH. 
\textsuperscript{17} If exposure is measured in person hours, rather than boat hours, the figure would be approximately 0.18 fatalities per million person exposure hours. The difference is the average number of occupants per boat.
Table 2 on page 8 shows the same type of calculation for recreational boating using data derived from the 2012 NRBS. The result? For all boat types pooled, the corresponding value is approximately 0.44 fatalities per million boat hours.

In terms of fatalities per vehicle exposure hour, then, the average risk of driving or riding in a motor vehicle is almost the same as being an operator or passenger in a recreational boat. By this measure, the risks of boating or driving are about equal. The reasons why many more people die annually in motor vehicles compared to boats are that (1) more people use motor vehicles, and (2) users are exposed to risks for a greater amount of time per year.

Dr. Chauncey Starr, a former Dean of the UCLA School of Engineering and Applied Science and founder of the Electric Power Research Institute, offered the following comments on the choice of exposure hours in a pioneering paper on risk analysis.\(^{18}\)

“The hour-of-exposure unit was chosen because it was deemed more closely related to the individual’s intuitive process in choosing an activity that a year of exposure would be, and gave substantially similar results. Another possible alternative, the risk per activity [e.g., the risk per voyage or per trip], involved a comparison of too many dissimilar units of measure: thus, in comparing the risk for various modes of transportation, one could use risk per hour, per mile, or per trip.”

And, indeed, a primary goal of the U.S. Coast Guard in conducting the NRBS was to gather data on exposure hours as the risk per exposure hour was viewed as a significant improvement over measures such as the annual fatalities per registered boat.

**Normalization by vehicle miles traveled**

A final example of a common basis for normalizing fatality rate data is vehicle miles traveled. It is a common measure of exposure used to compare accident or fatality rates for various transportation modes\(^ {19}\) and is one of the measures used by the U.S. Department of Transportation in the annual publication National Transportation Statistics.

However, vehicle miles are not likely to be a useful basis for normalizing boating fatality rates because the objective of most boating trips is to enjoy a recreational experience rather than to travel to a particular destination. Moreover, doubling the boat miles may not translate to doubling the risk—the boat miles traveled will be a function of usage patterns and the fleet size and composition. Taken to its logical extreme, if all boats were to remain at the dock at all times during a year and there was as much as one fatality associated with a fire or explosion that had resulted from the vessel or vessel equipment (thereby making it a reportable accident), the calculated fatality rate per boat mile would be infinite because the denominator would be zero.

---


There is also a practical problem in using vehicle miles to normalize fatalities for recreational boating—the lack of data on the average speed of boats to compute vehicle miles traveled. We can presume that the average speed is relatively low for rowboats, sailboats, and most paddle craft and probably higher for many powerboats. But even for high speed powerboats, such as bass boats, the average speed might be quite low, because much of a typical voyage will involve drifting, slow trolling powered by an electric motor, or even being anchored for swimming or fishing. All things considered, the average boat speed for all boats will certainly be lower than the average automobile speed.

Now because,

\[
\text{Fatalities/million vehicle miles} = \text{Fatalities/million hours} \times \text{hours/vehicle mile},
\]

the computed fatality rate per vehicle mile will be the fatality rate per hour divided by the average vehicle speed. As recreational boats and motor vehicles have about the same fatality risk per hour, and the average speed of boats—though not known with precision—is less than the average motor vehicle speed, if the fatality rate is expressed on a per vehicle mile basis, the fatality rate for boating will be higher than that for vehicles.

*The bottom line is that it is important to specify the basis for normalizing the data, because, as has been illustrated here, depending upon what basis is used, the fatality rates for recreational boats might be much lower or even higher than those for motor vehicles.*

All that said, in this exploratory analysis, based on available data and our assumptions about recreational boating, using exposure hours as the basis for normalizing the boating fatality data and making rate comparisons seems most appropriate.
TABLE 2. Days and Exposure Hours of Recreational Boating Participation by Boat Type in the United States - 2012.

<table>
<thead>
<tr>
<th>Boat Type</th>
<th>Total boats</th>
<th>Boats used (%)</th>
<th>Average number of days use per year</th>
<th>Average number of hours use per day</th>
<th>Average number of persons aboard use per day</th>
<th>Boating person-hours per year</th>
<th>Boating boat-hours per year</th>
<th>2012 fatalities</th>
<th>Fatalities per million person hours</th>
<th>Fatalities per million boat hours</th>
<th>Annual Fatalities per million boats</th>
</tr>
</thead>
<tbody>
<tr>
<td>All boats</td>
<td>21,611,000</td>
<td>NA</td>
<td>11.3</td>
<td>5.7</td>
<td>2.4</td>
<td>3,584,000,000</td>
<td>1,493,333,333</td>
<td>651</td>
<td>0.1816</td>
<td>0.4359</td>
<td>30.12</td>
</tr>
<tr>
<td>Power Boats</td>
<td>10,147,000</td>
<td>NA</td>
<td>12.0</td>
<td>6.0</td>
<td>2.7</td>
<td>2,035,000,000</td>
<td>753,703,704</td>
<td>346</td>
<td>0.1700</td>
<td>0.4591</td>
<td>34.10</td>
</tr>
<tr>
<td>Sail Boats</td>
<td>735,800</td>
<td>NA</td>
<td>11.1</td>
<td>7.8</td>
<td>2.4</td>
<td>154,000,000</td>
<td>64,166,667</td>
<td>27</td>
<td>0.1753</td>
<td>0.4208</td>
<td>36.69</td>
</tr>
<tr>
<td>PWCs</td>
<td>1,704,000</td>
<td>NA</td>
<td>11.0</td>
<td>4.7</td>
<td>2.3</td>
<td>212,000,000</td>
<td>92,173,913</td>
<td>58</td>
<td>0.2736</td>
<td>0.6292</td>
<td>34.04</td>
</tr>
<tr>
<td>Canoes</td>
<td>2,508,000</td>
<td>NA</td>
<td>8.6</td>
<td>6.3</td>
<td>2.3</td>
<td>362,000,000</td>
<td>157,391,304</td>
<td>52</td>
<td>0.1436</td>
<td>0.3304</td>
<td>20.73</td>
</tr>
<tr>
<td>Kayaks</td>
<td>3,916,000</td>
<td>NA</td>
<td>11.2</td>
<td>4.6</td>
<td>1.3</td>
<td>280,000,000</td>
<td>215,384,615</td>
<td>50</td>
<td>0.1786</td>
<td>0.2321</td>
<td>12.77</td>
</tr>
<tr>
<td>Pontoon</td>
<td>854,000</td>
<td>NA</td>
<td>14.9</td>
<td>4.1</td>
<td>3.8</td>
<td>220,000,000</td>
<td>57,894,737</td>
<td>44</td>
<td>0.2000</td>
<td>0.7600</td>
<td>51.52</td>
</tr>
<tr>
<td>Row/Inflatable/</td>
<td>1,747,000</td>
<td>NA</td>
<td>10.0</td>
<td>6.8</td>
<td>2.4</td>
<td>322,000,000</td>
<td>134,166,667</td>
<td>64</td>
<td>0.1988</td>
<td>0.4770</td>
<td>36.63</td>
</tr>
<tr>
<td>Other boats</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Unknown</td>
<td>21,611,800</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>10</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>651</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Why make comparisons between fatality rates?

If “fatalities per exposure hour” is likely to be the most appropriate risk measure, then we need to figure out how to interpret this statistic. It is difficult to appreciate the significance of a fatality rate of 0.44 fatalities per million boat hours if this number is considered in isolation. The statistic fails to answer the question, “Is this a risky activity or not?” and invites the rejoinder, “Risky, compared to what?” So, to help put this statistic into perspective, it might be useful to compare the risks per exposure hour for boating to those for other activities. Driving or being a passenger in a motor vehicle serves as one obvious “risk benchmark.” Additionally, it might be useful to have benchmarks for other activities (both voluntary and involuntary) for comparison. For some activities (for example, general aviation or occupational risks), exposure hour data are directly available; for others, some conversions are required to put risks into common units.

Making similar comparisons between boating fatality rates per exposure hour and several other activities is not a trivial task, for several reasons, not the least of which include the need to: cover comparable time periods (to capture the effects of trends); consider the possible uncertainty in the risk estimates; and to carefully specify the effects of various assumptions in the calculated value.

Building a “risk ladder”

One way to illustrate these comparative risk calculations is in the form of a “risk ladder” where each rung on the ladder corresponds to a different risk level. The ladder is typically constructed on a logarithmic scale, meaning that successive steps up the ladder increase risks by a factor of 10. The lower rungs on the ladder include activities such as working in an office, whereas the upper rungs on the ladder would correspond to activities with greater risk, such as driving a motorcycle, flying a private plane, or BASE jumping.

Table 3 on the next page shows a sample of some rungs on the ladder. Appendix A (page 17) and Table A-1 (Excel file, ERAC 2015 C1 NRBS_Research brief_Rate Normalization Table A-1_July 2015) present references and risk calculations in support of the others listed in this table.

---

20 The distinction between voluntary and involuntary risks is potentially important because people have different tolerance for these risks, see Starr, C., (1969). Social benefit versus technological risk, Science, 165: 1232-1238. Only voluntary activities are included in Table 3, but we may want to include some involuntary risks in the future.  
21 See, for example, http://www.psandman.com/articles/cma-0.htm.  
22 The uncertainty of the calculated average risks for boating and motor vehicles is quite low—for example, exposure hours for boating are estimated to be within +/- 5% and for fatalities +/-1%—so the conclusion that boating and roadway fatality rates are comparable is quite robust. This is not necessarily the case for some of the other activities, such as swimming or BASE jumping.
### TABLE 3. Fatality rates per million exposure hours for boating compared to other activities and relative risks (preliminary calculations)

<table>
<thead>
<tr>
<th>Activity</th>
<th>Fatalities per million hours</th>
<th>Relative risk</th>
</tr>
</thead>
<tbody>
<tr>
<td>BASE Jumping</td>
<td>12,948.00</td>
<td>29,704.06</td>
</tr>
<tr>
<td>Sky diving</td>
<td>75.89</td>
<td>174.10</td>
</tr>
<tr>
<td>Hang gliding</td>
<td>25.86</td>
<td>59.33</td>
</tr>
<tr>
<td>General aviation</td>
<td>22.44</td>
<td>51.00</td>
</tr>
<tr>
<td>Motorcycles</td>
<td>8.40</td>
<td>19.26</td>
</tr>
<tr>
<td>Scuba diving</td>
<td>6.29</td>
<td>14.43</td>
</tr>
<tr>
<td>Pedestrian</td>
<td>1.26</td>
<td>2.88</td>
</tr>
<tr>
<td>Swimming</td>
<td>1.07</td>
<td>2.45</td>
</tr>
<tr>
<td>Snowmobiling</td>
<td>0.88</td>
<td>2.02</td>
</tr>
<tr>
<td>Bicycling</td>
<td>0.57</td>
<td>1.31</td>
</tr>
<tr>
<td>Commercial fishing</td>
<td>0.51</td>
<td>1.16</td>
</tr>
<tr>
<td>Recreational boats</td>
<td>0.44</td>
<td>1.00</td>
</tr>
<tr>
<td>Motor vehicles</td>
<td>0.42</td>
<td>0.96</td>
</tr>
<tr>
<td>ATVs</td>
<td>0.38</td>
<td>0.87</td>
</tr>
<tr>
<td>Skiing</td>
<td>0.14</td>
<td>0.32</td>
</tr>
<tr>
<td>Working agriculture</td>
<td>0.128</td>
<td>0.29</td>
</tr>
<tr>
<td>Train travel</td>
<td>0.055</td>
<td>0.13</td>
</tr>
<tr>
<td>Bus travel</td>
<td>0.035</td>
<td>0.08</td>
</tr>
<tr>
<td>Wholesale trade</td>
<td>0.025</td>
<td>0.06</td>
</tr>
<tr>
<td>Government</td>
<td>0.014</td>
<td>0.03</td>
</tr>
<tr>
<td>Retail trade</td>
<td>0.010</td>
<td>0.02</td>
</tr>
</tbody>
</table>

Apart from presenting calculations for fatality rates using exposure hours as their basis, Table 3 also illustrates another method for normalizing fatality rates—the calculation of “relative risks.” The relative risk is the calculated fatality rate per exposure hour for the activity shown divided by the comparable fatality rate for recreational boating. The preliminary results in Table 3 indicate that on an exposure hour basis, general aviation is approximately 50 times more risky than boating and BASE jumping is nearly 30 thousand times more risky! The calculated risks presented in Table 3, with a few exceptions, are consistent with intuition. Among those exceptions, for example, are the risks to pedestrians which seem higher than might be expected; however, these risks arise not because of slips, trips, falls, and heart attacks (which we might expect and are not considered in the statistic), but rather because of pedestrians being killed by automobiles, busses, or trains.

---

23 This is one estimate (among several) that needs to be checked and possibly updated. There are data on drowning statistics, but it is much more difficult to estimate exposure hours. To illustrate, consider drowning fatalities in swimming pools. Is the relevant exposure only the actual time spent swimming or should the time spent walking around the pool (which might lead to drowning if the person falls in) be included?
Making comparisons among states

Officials responsible for boating safety in the states are likely to be interested in state-by-state comparisons of boating fatality rates for a variety of reasons. In principle, normalizing the fatality statistics to compare fatality rates on an exposure hour basis is preferable to making these comparisons based on either the raw total of boating fatalities\(^{24}\) or on fatalities per registered boat.\(^{25}\) However, even the exposure hour-based statistic is not a perfect measure of the efficacy of safe boating initiatives.

For example, the statistic does not consider the state where the boat is registered, only the state where the fatality occurred. Many states with attractive boating opportunities are “magnets” for boaters from other states. But fatalities involving those “out-of-state” boaters are typically included in the calculated fatality rates for the state where the fatal accident occurred rather than the state where the boat was registered. Out-of-state boaters may lack local knowledge and, in any event, are at least partially a product of the safety culture prevailing in their respective “home states.” For these reasons, apart from examining the implications of using exposure hours as the basis for the denominator in calculating risk equations (and in pursuit of its second, priority research topic for 2015), the ERAC charge team also is investigating the feasibility and implications of recalculating the numerator for fatality rates based on the state of origin of the boater, rather than on the state where the fatal accident occurred.

The idea of “correcting” for in-state versus out-of-state fatalities requires that the total exposure hours in the state of interest be partitioned into those from in-state boaters and those from out-of-state boaters. Dr. Philippe Gwet of the U.S. Coast Guard, and an active participant on the ERAC charge team, was able to do this with data collected in the 2012 NRBS.\(^{26}\) For example:

- For New Jersey, out-of-state exposure hours were estimated to be 3,343,000 exposure hours, which accounted for 7.2 percent of the total boating exposure hours in that state.

- The available data also indicate that 31 out-of-state operators were involved in a boating accident in New Jersey in that year, which amounted to 26.5 percent of the total (both in-state and out-of-state) operators involved in boating accidents in New Jersey.

Now, if out-of-state boaters were to have the same accident rate per hour as in-state boaters, the percentage of the operators involved in accidents would be equal to the corresponding percentage in exposure hours, 7.2 percent. However, this figure is 26.5 percent, which means that out-of-state boaters, by these 2012 estimates, accounted for a significantly greater percentage of the accidents in New Jersey. Across the entire United States, it is estimated that out-of-state boating accounted for 15.4 percent of exposure hours and a slightly greater proportion of accidents, 18.4 percent.

---

\(^{24}\) As a statistic, total fatalities neglect boating participation, the length of the boating season, and other factors.

\(^{25}\) This comparison is potentially a problem because not all boats (e.g., paddle craft) are registered in each state and because not all boats are used to the same extent.

\(^{26}\) This is documented in a “Fact Sheet” from the U.S. Coast Guard Office of Auxiliary and Boating Safety (CG-BSX) titled “Recreational Boating Across State Lines,” 15 Feb. 2015. Specifically, Table 2 provides--by state--the “out-of-state” exposure hours and the percentage of the total hours. This same table contains the “out-of-state” operators involved in a boating accident.
Figure 2 plots the fraction of out-of-state exposure hours versus the fraction of accidents accounted for by these same boaters in 2012. Each point on this plot represents data from a different state.

**FIGURE 2. Accidents and relative exposure of various states, 2012**

If accidents matched relative exposure exactly, the points on this plot would appear as a straight line, denoted the “parity line.” In this figure, the points above the line represent states in which accidents of out-of-state boaters accounted for a larger percentage of total accidents in that state than exposure hours. Much like New Jersey, New Hampshire (specifically labeled on Fig. 2) is one such state—47.5 percent of the accidents occurred with out-of-state boaters, whereas these boaters only accounted for 7.5 percent of the exposure hours in 2012. Points below the parity line are those states where out-of-state boaters contributed less than proportionally to accidents—New Mexico and Rhode Island (also labeled on Fig. 2) fall into this category. Of interest is that many states lie on or close to the parity line. In Wyoming, for example, out-of-state boaters accounted for 24.9 percent of the exposure hours and 25 percent of the accidents. Likewise, in Arizona, out-of-state boaters accounted for 41 percent of the exposure hours and 41.8 percent of the accidents in that state in 2012.

27 It is necessary to consider the precision of the exposure hour estimates, which differ from state-to-state because the same number of respondents did not respond to the NRBS. The exposure estimate for New Mexico, in particular, had substantially greater uncertainty. In future analyses, the precision of exposure estimates for each state will be explicitly considered.
Of note is that the majority of states, 34 of 50 (68 percent), lie above the parity line, suggesting that the hypothesis that out-of-state boaters are more likely to be involved in accidents may have merit.

Because fatalities are subject to less error than accidents, the calculations for Figure 2 will need to be redone in terms of boating fatalities,\(^2\) rather than boating accidents, as part of the charge team’s ongoing exploratory data analysis effort.

It is important to note in any such comparisons, however, that state boating fatality rates are also likely to be affected by factors other than the efficacy of state safety programs. For example, the composition of the “fleet” (for example, relative fractions of various types of craft) or the types of waters involved may differ from state to state.

**What happens when the basis for normalizing the data isn’t available every year?**

The results of the exploration to this point suggest that using recreational boating exposure hours to normalize boating fatality data is likely the most appropriate basis for making risk comparisons. The NRBS generated valid exposure hour estimates for 2012, but budget and other resource realities preclude conducting the survey and producing that data each year. That being the case, the ERAC charge team took up the question of whether—for use in the non-survey years—it might be meaningful and feasible for another variable or two to “stand-in” for or serve as a “surrogate” of sorts for exposure hour data in years when survey data are not available. That would require identifying variables with individually or collectively high correlations (that is, relationships) to boating exposure hours and with more frequent data collection—preferably on an annual basis—for each of the states.

The balance of this research brief describes a test of correlations between the 2012 NRBS-generated boating exposure hour estimates (denoted as EXP in million person hours) and several other variables for which 2012 data were also available and for which data are generally available on a more frequent basis. Six variables were identified by the team as potentially having some meaningful relationship to exposure hours—and, from a statistical standpoint—several do. The variables (with abbreviated names and data sources) that were included in the examination and some of the caveats associated with each are:

- **REG:** Number of registered boats by state (in thousands) 2012 from the U. S. Coast Guard. It is widely recognized that boat registration requirements differ from state-to-state. For example, according to the 2012 *Recreational Boating Statistics*, all watercraft were required to be registered in the state of Ohio, whereas in Delaware only motorboats were required to be registered. Thus, the actual number of boats is likely to be larger than just those registered, at least for certain states.

\(^2\)To avoid problems arising from the variability in year-to-year fatalities, a five- or 10-year average will be used. Additionally, statistical confidence bounds will be calculated to show the possible uncertainty of the estimates. Because the numbers of operators involved in accidents in some states are quite small, the confidence bounds are likely to be very broad.
• **VAL:** Economic value of recreational boating by state ($ billions) 2012 from the National Marine Manufacturers Association (NMMA) (Table 2.11, source: InfoGroup, Recreational Marine Research Center, MSU). Economic value components include “direct,” “indirect,” and “induced” economic value. Other things being equal, the economic value of boating is likely to be related to the economic activity associated with boating and the exposure hours.

• **DPI:** Disposable personal income per capita ($ thousands) 2012 from NJ Dept. of Labor ([http://lwd.dol.state.nj.us/labor/lpa/industry/incpov/dpci.htm](http://lwd.dol.state.nj.us/labor/lpa/industry/incpov/dpci.htm)). Disposable personal income per capita is the amount of discretionary income (personal income minus personal taxes) an individual has to purchase goods and services. A higher disposable income generally means a higher amount of consumer spending. Recreational boating is properly described as a “superior good.” Similar to a normal good, it is an item for which demand (consumption) increases as a person’s income increases. Unlike a normal good, its income elasticity of demand is always greater than one. As income increases people tend to spend proportionally more on superior goods.

• **POP:** State populations (millions) 2012 from U. S. Census Bureau. It is plausible that, other things being equal, states with larger populations will have more persons who spend more time on boating activity.

• **Marinas:** Number of marinas (NAICS code 71393) in 2012 by state from Quarterly Census of Employment and Wages, Bureau of Labor Statistics. Recreational boaters (at least those with larger boats) often base their boats at marinas, so it is plausible that states with more marinas will have greater boating activity. Of course, many boaters are able to engage in boating activity without use of a marina (paddle craft and trailerable boats), so this variable might be an imperfect surrogate for boating activity. In future explorations, and if a viable data source can be accessed, the number of launch ramps or lanes might help account for activity associated with smaller boats.

• **FIHU:** Fishing and hunting expenditures by state residents 2011 ($ Billions) from National Survey of Fishing, Hunting and Wildlife Associated Recreation from U. S. Fish and Wildlife Service and U. S. Census Bureau (conducted every five years). This variable excludes certain recreational expenditures (e.g., wildlife observation) that do not require the use of a boat. Additionally, people use boats for activities other than fishing and hunting. However, it is plausible that this variable might be correlated with boating activity.

A few jurisdictions were excluded from this part of the exploratory data analysis (i.e., New Mexico, North Dakota, and District of Columbia) because their data were either missing or not reported in the data source.

**The correlations with exposure**

Figure 3 on the next page shows a picture of how each variable correlates with EXP. The small pictures in the overall frame show the scattergrams of the variables and the marginal distribution of each variable as histograms (shown as the top picture in each column).

Inspection of the first column (excluding the top picture for EXP) shows that most of these variables are indeed correlated with exposure hours. A scattergram that is more tightly packed—

---


30 A scattergram is a plot of data points for each of the two variables included. For example, excluding the top left box, the one directly below shows a plot of the exposure hours for each state on the x-axis compared to the number of registered boats in each state.
i.e., closer to a straight line—suggests a closer relationship between that variable and exposure hours. One surprise is that DPI appears to be negatively correlated with exposure hours. Apparently as disposable personal income increases (one state compared to another state) people do not necessarily use their boats more. Perhaps they go to the Opera, read books, or drink fine wines; but based on this exploration it does not appear that they spend more time on boating or boating-related activities.

**FIGURE 3. Scattergram showing the correlation between the candidate variables and exposure hours.**

*Pearson correlation matrix*

Another way to look at these data is in the form of a correlation matrix, as shown in Table 4 on the next page.
Table 4. Correlation matrix for candidate variables related to boating exposure hours.

<table>
<thead>
<tr>
<th></th>
<th>EXP</th>
<th>REG</th>
<th>VAL</th>
<th>DPI</th>
<th>MARINA</th>
<th>POP</th>
<th>FIHU</th>
</tr>
</thead>
<tbody>
<tr>
<td>EXP</td>
<td>1.000</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>REG</td>
<td>0.907</td>
<td>1.000</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>VAL</td>
<td>0.939</td>
<td>0.960</td>
<td>1.000</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DPI</td>
<td>-0.169</td>
<td>-0.175</td>
<td>-0.120</td>
<td>1.000</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MARINA</td>
<td>0.765</td>
<td>0.626</td>
<td>0.765</td>
<td>0.102</td>
<td>1.000</td>
<td></td>
<td></td>
</tr>
<tr>
<td>POP</td>
<td>0.685</td>
<td>0.720</td>
<td>0.826</td>
<td>-0.008</td>
<td>0.612</td>
<td>1.000</td>
<td></td>
</tr>
<tr>
<td>FIHU</td>
<td>0.854</td>
<td>0.899</td>
<td>0.926</td>
<td>-0.178</td>
<td>0.709</td>
<td>0.786</td>
<td>1.000</td>
</tr>
</tbody>
</table>

Number of observations: 48

The numbers shown are the simple correlation coefficients between each of the variables. These range between zero (0) and one (1.0). A value of zero indicates that there is no correlation or relationship, whereas a value of one (1.0) indicates that there is a perfect one-to-one correspondence. A negative sign (for example, DPI) means that the variables have an inverse correlation—that is, as one goes up the other goes down.

Examining the correlations between each of the variables and exposure hours (EXP), it is apparent that boat registrations and the economic value of recreational boating have the highest correlation to exposure hours. But boat registrations and economic value also are highly correlated with each other.

So where does this lead?

There might be other variables, yet unexplored, that also correlate with recreational boating exposure hours. But based on this initial exploratory analysis, if we had to pick variables for prediction purposes, the most likely picks would be boat registrations and economic value. Apart from their strong correlation with exposure hours, both of these variables have an advantage over some of the others (for example, FIHU) in that they are calculated and made available annually.

Because these two variables also are highly correlated (r = 0.960), however, there would not be much gained by using both for prediction; only one should be retained for such purposes. Given the source and the known methodology for collecting and calculating the data over the years, boat registrations would likely be the preferred variable.

However, there is a caveat to all of this. Even though the current exploration shows that boat registrations and exposure hours are highly correlated for 2012, the real test of the variable’s utility for predictive purposes will not come until we have the experience of additional surveys and collections of exposure hours. For the time being, however, the variable might be used as an indicator.

31 With 48 points, a correlation coefficient as high as 0.9 would be highly significant (p<0.0001) in the statistical sense meaning that this could not have originated from chance alone (see e.g., http://vassarstats.net/rsig.html).
Appendix A – Resources and References (see Table A-1 in separate Excel file)

All website references here and in Table A-1 were verified as of 12 June 2015.

**General Aviation**


<table>
<thead>
<tr>
<th>Year</th>
<th>Total Fatalities</th>
<th>Flight hours (000)</th>
<th>Fatalities per Million flight hours</th>
</tr>
</thead>
<tbody>
<tr>
<td>2002</td>
<td>581</td>
<td>25,545</td>
<td>22.7442</td>
</tr>
<tr>
<td>2003</td>
<td>633</td>
<td>25,998</td>
<td>24.3480</td>
</tr>
<tr>
<td>2004</td>
<td>559</td>
<td>24,888</td>
<td>22.4606</td>
</tr>
<tr>
<td>2005</td>
<td>563</td>
<td>23,168</td>
<td>24.3008</td>
</tr>
<tr>
<td>2006</td>
<td>706</td>
<td>23,963</td>
<td>29.4621</td>
</tr>
<tr>
<td>2007</td>
<td>496</td>
<td>23,819</td>
<td>20.8237</td>
</tr>
<tr>
<td>2008</td>
<td>496</td>
<td>22,805</td>
<td>21.7496</td>
</tr>
<tr>
<td>2009</td>
<td>479</td>
<td>20,862</td>
<td>22.9604</td>
</tr>
<tr>
<td>2010</td>
<td>457</td>
<td>21,688</td>
<td>21.0716</td>
</tr>
<tr>
<td>2011</td>
<td>448</td>
<td>23,750</td>
<td>18.8632</td>
</tr>
<tr>
<td>2012</td>
<td>440</td>
<td>20,881</td>
<td>21.0718</td>
</tr>
<tr>
<td>2013</td>
<td>387</td>
<td>20,887</td>
<td>18.5283</td>
</tr>
<tr>
<td>Total</td>
<td>6,245</td>
<td>278,254</td>
<td>22.4435</td>
</tr>
</tbody>
</table>

**Motorcycles**

Average speed: 29.8 MPH

Risk per exposure hour: Alternative estimate given by Failure Analysis Associates (now Exponent) 8.8 fatalities per million hours, see http://cyclehelmets.org/1026.html.

**Bus travel**

Average speed: 12 MPH (rounded upward from 11.42 for year 2009 given in Table 27 for commuter bus)
ATVs

Fatalities and number of 4-wheel ATVs (Table 4)

Hours per year: 252.3 hours (Table 6)

Swimming

Fatality rate: 1.07 fatalities per million hours, from Failure Analysis Associates, http://cyclehelmets.org/1026.html. Note that neither methodology nor date specified in this source. This estimate needs to be verified.

Hang Gliding


Skydiving


Duration per jump: 7 minutes, http://www.skydivecsc.com/skydive-questions-faq. The USPA uses a fatality index defined as the number of fatalities per 100,000 jumps. The available data indicate that this normalized fatality rate has decreased over the years (see http://parachutistonline.com/feature/indexing-toward-safer-sport).
Note fatality rate estimate per Failure Analysis Associates is 128.71 fatalities per million hours (http://cyclehelmets.org/1026.html), based on older data.

Base Jumping


Duration of jump: Estimates vary from a few seconds to 2 minutes. Two (2) minutes was used as a conservative figure. See http://www.stealingaltitude.com/BASE_JumpingFAQ.htm#13.
The choice of a shorter duration would increase the calculated risk per hour, perhaps substantially.
**SCUBA Diving**

**Fatality rate:** 1 per 211,864 dives, available at [http://scuba.about.com/od/divemedicinesafety/p/Is-Scuba-Diving-Safe-Or-Dangerous.htm](http://scuba.about.com/od/divemedicinesafety/p/Is-Scuba-Diving-Safe-Or-Dangerous.htm). Another reference (Denoble, P. J., Marroni, A., and Vann, R. D., (2011). Annual Fatality Rates and Associated Risk Factors for Recreational Scuba Diving, Recreational Diving Fatalities Workshop Proceedings: 73-85) is available at [http://www.diverbelow.it/attachments/article/52/DAN_Fatalities_8.pdf](http://www.diverbelow.it/attachments/article/52/DAN_Fatalities_8.pdf) or [http://archive.rubicon-foundation.org/xmlui/handle/123456789/9329](http://archive.rubicon-foundation.org/xmlui/handle/123456789/9329). This reference cites several studies with fatality rates ranging from 0.45 to 6 fatalities per 100,000 dives. The more recent rates (2000–2006) are more closely spaced and range from 0.45 to 0.57 fatalities per 100,000 dives, quite close to the value used in this analysis 0.472/100,000 dives.

**Skiing**


**Snowmobiling**

**Estimated fatality rate:** 0.88 per million hours, from Failure Analysis Associates, available at [http://cyclehelmets.org/1026.html](http://cyclehelmets.org/1026.html).

**Pedestrian**

**Pedestrian fatalities:** 4,594, average from 2002 to 2011 from [http://www-nrd.nhtsa.dot.gov/Pubs/811748.pdf](http://www-nrd.nhtsa.dot.gov/Pubs/811748.pdf). These are pedestrian traffic fatalities. Total pedestrian fatalities might be higher if additional sources (slips, trips, falls, and medical events) were included.


“Pedestrian data also indicate another difficulty with aggregate fatality rate calculations—these are not evenly distributed across the population. As one source (http://www.smartgrowthamerica.org/2014/05/20/dangerous-by-design-2014-highlights-preventable-pedestrian-fatalities/) notes:

“Pedestrian fatalities are disproportionately born by older adults, people of color, and children. While just 12.6 percent of the total population, those over the age of 65 years old account for nearly 21 percent of pedestrian fatalities nationwide. Among people of color, blacks and African Americans suffer a pedestrian fatality 60 percent higher than non-Hispanic whites, and Hispanics of any race have a rate nearly 43 percent higher. And, for the most recent years we have data, 4,394 children aged 15 and younger were killed while walking. Pedestrian injury is the third leading cause of death for this age group.”

**Bicycling**

**Fatality rate:** 0.573 from Beck et al., 2007 and Pucher et al., 2011.

Other estimates from web sites shown. Additional data from:


**Amtrak**

**Average Fatalities:** 6.46 per year, average over years 2000 to 2012, based on data from U.S. DOT (http://www.amtrak.com/ccurl/636/294/Amtrak-Sets-New-Ridership-Record-FY2012-ATK-12-092.pdf). Note that this includes just passenger fatalities. Total fatalities are much higher resulting from trespasser fatalities and accidents at grade crossings. In 2013, for example, the total number of rail fatalities was 891, but only six (6) of these were passenger fatalities. Nonetheless, train travel is relative safe; see Savage, I., Department of Economics and the Transportation Center Northwestern University 2001 Sheridan Road, Evanston IL 60208, USA (see http://faculty.wcas.northwestern.edu/~ipsavage/436.pdf). Savage computes an average annual fatality figure (876) then notes:

“The vast majority of the risk was suffered by people other than those involved in the production and consumption of rail service.

The breakdown of the average of 876 annual fatalities was:

490: Trespassers (primarily pedestrians) at places other than grade crossings,
281: Motorized highway users at grade crossings,
68: Pedestrians and non-motorized highway users at grade crossings,
26: On and off-duty employees and contractors working on the railroad,
7: Passengers on trains, and
4: Bystanders not on railroad property.”

The calculations given in Table A-1 include only passenger fatalities.
Average speed: 50 MPH. This is an assumption. The average train speed over any distance is a function of (among other things) the type of train, tracks, and number of stops. It is typically much lower than either the maximum train speed or the applicable speed limits. The average speed varies with the route and the type of train. Some examples;

- The average speed of the Acela Express is 64 MPH between Boston and Washington (see http://www.economist.com/blogs/gulliver/2013/12/amtrak-new-trains); another report indicates that the average speed is 70-80 MPH on this route (http://transportation.mit.edu/sites/default/files/documents/MIT_Amtrak_Report_2013.pdf),
- On the 423-mile stretch between Oakland and Los Angeles aboard Amtrak's daily Coast Starlight train, the scheduled travel time is more than 13 hours. Though the train reaches 79 miles per hour on some segments, the average speed on this trip ranges between 25 to 35 miles per hour (http://usatoday30.usatoday.com/travel/columnist/grossman/2008-08-22-amtrak-business-travelers_n.htm),
- Between Pittsburgh and Harrisburg PA, the average speed of the Keystone West train is 45 MPH (http://www.planthekeystone.com/keystonewest.html),
- The South East Corridor trains now run at an average speed of 46-48 MPH (http://www.sehsr.org/faq.html),
- Keystone service between New York, NY, and Harrisburg, PA completes the 195 mile journey in 3 hours 50 minutes, for an average speed of 51MPH,
- The Silver Star service between New York and Miami traverses a distance of 1,480 miles at a scheduled time of 32 hours, for an average speed of 46 MPH (http://www.railpassengerusa.com/routes/silverstarroute.php). The Silver Meteor traverses a route of 1,389 miles in 28 hours for an average speed of 50 MPH.

Occupational