**Probability Ratio or Difference?**

How Do People Perceive Risk?

*Harshada Vidwans, Rohini Kharate and Milind Wate*

In public health literature, the risk of death or disease associated with genetic, dietary, environmental, or behavioral factors is most commonly denoted by the odds ratio (OR), hazard ratio (HR), or risk ratio (RR). But how do people intuitively perceive risk? We conducted a small experiment in which respondents of different ages, sex, and education were asked to rank the risks associated with four different habits based on the incidence data of an imaginary deadly disease. Results showed that people judge risk by probability difference rather than probability ratio. Even individuals formally trained to use OR and HR as risk indicators preferred using probability differences over ratios to judge their own risk. This preliminary inquiry into intuitive statistical perception suggests that designing statistical indices based on people’s innate perception may be a better strategy than trying to educate people to understand the indices designed by expert statisticians.

**Introduction**

Thinking quantitatively, sampling their environment, and analyzing it in some way or the other to make behavioral decisions is an innate ability demonstrated by many species of animals in multiple studies [1–3]. Therefore, it is a fair assumption that humans too have some evolved innate algorithms for sampling their surroundings and drawing useful inferences [4, 5]. There have been a few studies addressing how people informally analyze data and whether and how it differs from textbook statistics.

*Vol 26, No.11, DOI: https://doi.org/10.1007/s12045-021-1262-5*
Bayesian statistics is the closest to how people make inferences. The way people perceive the relationship between two variables differs from commonly used least square regression [6]. But barring a few exceptions, how people make inferences from their own sampling is largely an unexplored question. One important function of people’s innate statistical ability is to assess risks of various kinds to avoid them or to optimize behavioral strategies accordingly. It can be easily imagined how and why this ability would have evolved under strong selective pressure. But whether the evolved algorithm for risk perception is similar or different from risk assessment in formal statistics is an unanswered question. Answering this question is relevant for the effective use of statistics in public awareness and education.

Risk assessment is extensively used in formal statistics, particularly in public health literature. The commonly used indices include risk ratio (RR), odds ratio (OR), or hazard ratio (HR) [7]. Although differing in some details, they depict the fold increase in the chance of an adverse event. For example, if the probability of cancer in non-smokers is 0.001 and in smokers is 0.01, the risk ratio of cancer among smokers compared to non-smokers is 10. This means that smokers are 10 times more likely to get cancer than non-smokers. This is likely to be perceived as high risk. However, the absolute increase in probability is only 0.009, which may be perceived as a small increase. Thus the probability ratio and difference may lead to very different perceptions. Some indices based on probability difference are also described in the literature [8] but are not commonly used in public health practice.

We ask a single relevant question here: Do people have an innate tendency of preference for probability ratios or difference while judging risk.

We gave an imaginary task to respondents in which they had to rank the risk of an imaginary disease from four hypothetical habits. Data on the percentage of people contracting the disease (such as cancer) with and without the hypothetical habit (such as smoking) was provided. Among the four habits in question, two had the same probability ratio but different probability difference,
while two had the same probability difference but different ratios. Maintaining these features, the characteristics of habits A to D were reshuffled, and three different versions of the questionnaire were used. Potential respondents were contacted through personal online and offline contacts. So it was convenience sampling rather than random sampling. Nevertheless, the samples represented a wide variety of age groups, educational and professional backgrounds. The minimum limit for education was secondary school since it was important to understand the question by reading it. Any queries about the clarity of the question were appropriately replied to. After receiving the responses, the shuffled versions were realigned to the master questionnaire (given in the supplementary information). Respondents were asked to rank the risks. Tied ranking was permitted. We request interested readers to look at the questionnaire and make their own judgments before reading the results to make further reading more interesting.

The data provided in the questionnaire and the corresponding response summary are given in Table 1.

The mean of the ranks given was used to calculate the average perceived risk (see supplementary material for details of the mean risk score calculation). It can be seen in Table 1 that habits A and B have the same ratio but different probability difference. A and C, as well as, B and D, have the same difference but different ratios. If the responses were random, we would have expected equal mean scores for all four habits. The 98 completely filled ques-

<table>
<thead>
<tr>
<th>Habit</th>
<th>Disease prevalence in people without the habit</th>
<th>Disease prevalence in people with the habit</th>
<th>Probability ratio</th>
<th>Probability difference</th>
<th>Mean risk score (s.e.) (lowest=1, highest =4)</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>0.1 %</td>
<td>0.3%</td>
<td>3</td>
<td>0.2</td>
<td>1.70 (0.11)</td>
</tr>
<tr>
<td>B</td>
<td>20%</td>
<td>60%</td>
<td>3</td>
<td>40</td>
<td>3.23 (0.08)</td>
</tr>
<tr>
<td>C</td>
<td>20.1%</td>
<td>20.3%</td>
<td>1.01</td>
<td>0.2</td>
<td>1.96 (0.08)</td>
</tr>
<tr>
<td>D</td>
<td>0.1 %</td>
<td>40.1</td>
<td>401</td>
<td>40</td>
<td>3.10 (0.1)</td>
</tr>
</tbody>
</table>

Table 1. The data given is represented in columns 1 to 3. The ratios and differences respectively in columns 4 and 5 and the average ranks in column 6.
tionnaires revealed that the responses were non-random and concordant. The mean risk perception scores differed significantly, with B and D being perceived as high-risk habits and A and C as low-risk habits (One way ANOVA $F = 77.53, p = 2.6 \times 10^{-39}$).

The mean scores of B and D that had the same difference (40%) but different ratios (3 versus 401) did not differ significantly ($F = 1.26, p = 0.26$). A and C also had the same difference (40%) but different ratios (3 versus 1.01), but here C was perceived to have a marginally greater risk ($F = 4.18, p = 0.04$). The difference in ratios between B and D was much greater than that of A and C, but risk perception difference in the former was non-significant whereas in the latter was statistically significant. Therefore ratios do not seem to explain the perception of risk significantly. In contrast, between A and B that had the same ratio (3), but different difference (0.2 versus 40), the risk perception changed substantially ($F = 56.01, p < 0.0001$). Taken together, people seemed to judge the risk primarily by probability difference rather than by probability ratio.

When included as a second factor in two-way Anova analyses age (66 in the <25 age group and 32 above, $F = 0.01$, NS), sex (female 50 male 48, $F = 0$, NS), formal education in health science (with 33 and without 65, $F = 0$, NS), formal education in statistics (with 39 and without 59, $F = 0$, NS), did not show any significant effects on risk perception. The last one is particularly interesting because the use of ratio-based indices such as OR and HR is commonly taught in public health statistics to reflect risk, whereas probability difference based indices may not receive even a mention. It is interesting that even after undergoing training with ratio-based indices when one makes a decision for oneself, probability difference is used unanimously and not probability ratio.

In reality, we frequently see that people tend to ignore several serious health warnings despite continuous attempts at making them aware. The health warnings on all tobacco products, the use of masks, and social distancing during the Covid-19 pandemic are well-known examples. Although tobacco may be increasing
the probability of cancer ten times, if the probability difference is not very large, tobacco may not be treated as a high-risk factor by people. The Covid-19 example is recent and illustrates important principles. Despite continued warnings and awareness campaigns, people violated the rules on several occasions including festivals, elections, and riots. They engaged in crowding without adequate precaution. It is not that people were not aware that they were increasing the risk of infection by crowding. But if the perceived difference in the probability of death with and without Covid-19 was not large, they are unlikely to take the health advice seriously. The concept that people use their own statistical inference, and their methods could be different from those used by epidemiologists is conspicuously absent from public health thinking. A large mismatch between people’s innate statistics and that of the health authorities may be hindering the effective implementation of public health policies.

Although our experiment was simple and the results could be treated only as preliminary and exploratory, the finding that human-intuitive statistics may work differently than mainstream academic statistics, suggests a rethinking of the risk assessment indices. The ratio-based and difference-based indices have different strengths. For example, the OR or HR index has better mathematical properties and can be expressed on a time-independent scale. The probability differences may change with the follow-up time, but they resonate better with people’s perceptions. Accordingly, the use of the two types of indices should be purpose-driven and audience-specific. Ratios may serve better for research purposes, but while communicating research findings to clinicians, patients, and general readers, it would be advisable to use the methods, tools, and indices that resonate better with people’s innate perceptions. This is particularly relevant to public health statistics and may have important implications for policymaking in the interest of people.

Even more important is the relevance of this experiment to education. The human mind has substantial innate capacities. Education needs to be designed to resonate with them. Inculcating alternative methods from outside is most likely to be ineffective.
Currently, we are deficient in our understanding of innate human quantitative abilities. Serious research into this field is likely to reshape mathematics, statistics, and science education of the future.

Box 1. Supplementary Information: The Questionnaire Used For the Study

You have four habits, which doctors say are bad for health. According to reliable studies, all four habits increase the chance of developing one or more serious health problems, as tabulated below, but different habits have different levels of risk. Consider all the health problems to be equally serious, but the probabilities of developing them are different for different habits.

If it is difficult for you to give up all four habits, which ones will you be ready to give up on priority? Give your order of preference.

Tabulated below are % individuals developing a serious health problem.

<table>
<thead>
<tr>
<th></th>
<th>People without the habit</th>
<th>People with the habit</th>
</tr>
</thead>
<tbody>
<tr>
<td>Habit A</td>
<td>0.1 %</td>
<td>0.3%</td>
</tr>
<tr>
<td>Habit B</td>
<td>20%</td>
<td>60%</td>
</tr>
<tr>
<td>Habit C</td>
<td>20.1%</td>
<td>20.3%</td>
</tr>
<tr>
<td>Habit D</td>
<td>0.1%</td>
<td>40.1%</td>
</tr>
</tbody>
</table>

Your preference order for giving up the habit: If you think two habits need to be given the same preference, enter them at the same level.

1 (to be given up first)

2

3

4 (to be given up last/may not be given up at all)

Contd.
Box 1. Contd.

Information about the respondent:
Age: (Years)  Sex: Male/Female
Education:  
Profession:  
Field of work: Medicine/Engineering/Administrative/Academic/Office service/Business/Household work/Artist/Journalist/Other (specify)
Received any formal education in health science? Yes/No
Received any formal education in statistics? Yes/No

Note: For the question “Which habit will you give up first?” the highest perceived risk gets the rank of
1. However for the perception of the reader, we have inverted the scale for calculating mean risk score in
which 4 is the highest perceived risk and 1 is the lowest.

Suggested Reading


Address for Correspondence
Milind Watve
E-1-8, Girija Shankar Vihar
Karve Nagar
Pune 411 052, India
Email: milin.watve@gmail.com