INJURY BIOMECHANICS RESEARCH
Proceedings of the Sixth International Workshop

PROBLEMS IN ANALYZING IMPACT INJURY DATA

by

ARNOLD K. JOHNSON

INTRODUCTION:

This paper centers on the problem of correlating injury severity with impact test parameters, such as forces accelerations, velocities, test subject's size, etc.

Biomechanical impact tests of such severity as to produce injury are regularly conducted with primates and cadavers. Injury severity can be measured in several ways, one of which is the number of rib fractures caused by thoracic loading.

This paper will focus on the problem of correlations of injury data when used in routine curve fitting by the method of least squares. For convenience we will restrict our attention ot the assessment of injury severity in terms of the number of rib fractures to cadavers when restrained by lap and torso belts.

AN EXAMPLE OF A CORRELATION:

One can correlate or predict the number of rib fractures as a function of the three independent variables (1) maximum tensile force in the torso belt, (2) the cadaver's age and (3) the cadaver's weight. Such a correlation is shown in Figure 1 of data from Reference 1.* The predicted values shown along the vertical axis are based on the equation derived by least squares shown at the bottom of Figure 1. The observed values are shown along the horizontal axis. If there were 100 percent accuracy, all points would lie along the 45° line. In addition to the expected random scatter, there is also bias or skew. The predicted values are too high at the low end and too low at the high end. The dashed line fitted

* References are given on page 103

by least squares to the data points plotted on this graph shows the bias clearly. However, consider the high predictions at the low end. There is no such thing as a negative number of rib fractures. The bias is in the direction of positive numbers, which is reasonable. A simple, partial explanation for this bias at low numbers of rib fractures will be given later.

Figure 2 is a similar least squares plot for a second order polynomial. Its general form is shown at the bottom of the figure. The same general bias is displayed. The standard deviation for the equation of Figure 1 is 4.72, whereas for the equation of Figure 2 it is 4.68. The difference is not significant. Both the bias and the standard deviation seem independent of whatever curve is fitted to the data. A partial explanation of these observed facts can be made by reviewing in basic terms what is involved in correlation of impact injury.

REVIEW OF BASICS:

For the purposes of illustration, Figure 3 shows a simplification relating to the data we have just reviewed. The figure shows a correlation only between the number of rib fractures and the maximum tensile force in the torso belt. For simplicity the true underlying correlation we want to estimate is shown in the figure to be linear with two parameters F_o and θ. To estimate this straight line, we could conduct tests involving injury producing forces greater than F_o and observe the number of rib fractures. Figure 4 shows a fictitious plot of such data, and a least squares derived dashed line is shown fitted to the data points. The dashed line's intercept with the horizontal axis is an estimate of F_o and the line's slope is an estimate of the angle θ. Suppose that in generating this injury data there had also been generated data for which no injury had occurred. Could these noninjurious data points be added to this graff and included in a least square fit. The answer is negative.

All noninjurious data points would lie exactly on the horizontal axis as shown in Figure 5. Their effect if included in a least squares fit would be to reduce the dashed lines slope and its intercept with the horizontal axis. In summary, Figure 5 shows two distinct sets of data points: (1) those which truly help to estimate the parameters F_o and θ, and (2) those which lie on the horizontal axis and provide on such a plot no information whatsoever in estimating F_o and θ. Moreover, their use in a least square fit would contribute to a miscorrelation.

We have just reviewed a miscorrelation related to zero injury data for least squares curve fitting. At this point it needs to be noted that the predicted values of Figure 1 did not change much when the zero data were removed. However, the zero data points were only 5 percent of the total data sample.

Should we reject zero injury, data? The answer is negative provided we use a more appropriate method, one not involving least squares. The method that will now be proposed will center on the probability distribution for rib fractures.

PROBABILITY DISTRIBUTION FOR RIB FRACTURE:

For a specific thoracic impact, we want to know the probability of no rib fractures, the probability of one rib fracture, the probability of two rib fractures, etc. By using a probability distribution, we are setting up a mathematical procedure which can use all data, including that of zero. I am not aware of any data which provides information on the probability distribution of rib fractures. To decide the attributes that such a distribution should have, we appeal to a theorem of statistics and to common sense. We consider the attributes that such a distribution should have at the two extremes, that of severe impact and that of mild impact. We first look at severe.

Figure 6 shows a hypothetical frequency plot of rib fractures for repeated tests for which almost identical cadavers are impacted under almost identical test conditions. For a severe impact the probability of zero impact is essentially zero. Figure 6 shows the frequency plot approximating a bell shaped normal distribution. Many experimentally observed outcomes are normally distributed, and the central-limit theorem of statistics* helps to explain this fact.

Figure 7 shows the probability distribution for an extremely mild impact for which the probability of zero fractures is essentially one. All the probability is concentrated at zero.

---

*See Reference 2, Section 7.6.

Figure 8 shows another specific probability distribution, but for a less mild impact. The figure shows a 0.5 probability of zero rib fractures. Since all the probabilities of any distribution have to total one, there is only a 0.5 probability remaining to distribute among the non-zero rib fractures. One possible distribution is shown in Figure 8, and the distribution does not approximate the bell shaped curve of a normal distribution. Furthermore, there is no possible way to distribute the remaining 0.5 probability so as to obtain a bell shaped probability curve.

In summary, the probability distribution is assumed to be normal for severe impact, and it is clearly not normal for mild impact. In this latter case, the distribution is skewed.

Shortly two candidate distributions which have the attributes shown in Figures 6, 7, and 8 will be presented. But we need to review first an important limitation on the method of least squares curves fitting.

A LIMITATION ON LEAST SQUARES CURVE FITTING:

Least squares curve fitting is strictly valid for data from a normal distribution. When this requirement is satisfied, then each point on the least squares curve is an estimate of the mean of a normal distribution. From page 58 of Reference 3 there is the statement that the method of least squares "rests upon the assumption that the scatter of the individual observations around a fitted line will approximate a normal distribution." One must consider that miscorrelations by the method of least squares curve fitting can occur when a data sample contains not only zero injury data, but also mild injury. Mild injury data cannot be expected to be observations from a normal distribution.

TWO CANDIDATE PROBABILITY DISTRIBUTIONS:

The two probability distributions of Equations 1 and 2 are offered as candidates for predicting the probability of rib fracture. Equation 1

$$P(k) = e^{-(k/\beta)^{\alpha}} - e^{-\left(\frac{k+1}{\beta}\right)^{\alpha}}, \quad k = 0, 1, 2, 3, \ldots \quad (Eq. 1)$$

$$Q(k) = \frac{\lambda^{k} e^{-\lambda}}{k!}, \quad k = 0, 1, 2, 3, \ldots \quad (Eq. 2)$$

is related to a Weibull distribution, and Equation 2 is the Poisson Distribution. Each equation by proper choice of parameteric values can simulate the probability distributions of Figures 6, 7, and 8. Equation 2 is clearly the simpler, because it contains only the one parameter λ. Equation 1 has the two parameters α and β.

FITTING DISTRIBUTIONS TO DATA:

The parameters of the distributions of Equations 1 and 2 depend, of course, on the independent variables of an experiment. An equation establishing this dependence for each parameter has to be assumed. In respect to the single parameter λ in Equation 2, consider

$$\lambda = A + BF \quad (Eq. 3)$$

as an example. We are still thinking in terms of rib fractures, and F stands for the maximum tensile force in the torso belt. The letters A and B are constants whose values have to be adjusted to fit a given set of data. The fitting of parameters in a distribution to a given set of data takes us away from the area of probability and into the area of statistics. The statistical method of maximum likelihood is perhaps the strongest procedure available for fitting a distribution's parameters to a set of data. Section 8.3 of Reference 2 explains in detail the method of maximum likelihood.

BENEFITS OF A PROBABILITY DISTRIBUTION:

The benefits of a probability distribution are as follows:

(1) Least squares curve fitting and it miscorrelation problems are avoided.

(2) It permits using all injury data, including that of zero injury.

(3) It permits the implementation of the statistical principle of maximum likelihood.

(4) It is more informative. (The distribution provides a probability of observing exactly k rib fractures, k or more rib fractures, etc. One can compute the distribution's mean, median, etc.)

PARTIAL EXPLANATION OF BIASED PREDICTED VALUES:

Reference has been made earlier to the bias of predicted values in Figure 1. We are now in a position to provide a partial explanation for this bias at least for small numbers of rib fractures.

Reference is made to Figure 9 which shows a hypothetical plot of the computed mean of a probability distribution for rib fracture vs the observed value. At the observed value of zero all the means are shown as positive. This has to be. When a probability distribution has only zero or positive outcomes, as would always be the case for rib fractures, the distribution mean has to be zero or positive. However, as shown, at the observed value of one it is possible to have a mean less than one. What is shown in Figure 9 is a partial explanation of the bias toward positive numbers seen in Figures 1 and 2.

RECOMMENDATIONS:

The following recommendations are made:

(1) The probability distributions of Equations 1 and 2 be studied further.

(2) Based on the results of Item (1), make rational decisions as to how best to model the distribution parameters.

(3) Implement the method of maximum likelihood.

(4) Evaluate the entire method by applying it to sets of impact injury data.

CONCLUSIONS:

Although this study is not yet complete, it is felt that the following conclusions are valid based on the analysis presented.

(1) Least square curve fitting should be restricted to injury data believed to be approximately normal.

(2) Least squares curve fitting should not use zero injury data.

ACKNOWLEDGEMENT:

The author wishes to thank Dr. George V. Waldo of the Naval Ship Research and Development Center for suggesting the probability distribution of Equation 1.

REFERENCES

(1) Rolf H. Eppinger, "Prediction of Thoracic Injury Using Measurable Experimental Parameters." Report on the Sixth International Technical Conference of Experimental Safety Vehicles, October 1976.

(2) Alexander M. Mood, "Introduction to the Theory of Statistics," McGraw-Hill Book Company, Inc., New York, 1950.

(3) Mordecai Ezekial, "Methods of Correlating Analysis," John Wiley and Sons, Inc., New York, 1930.

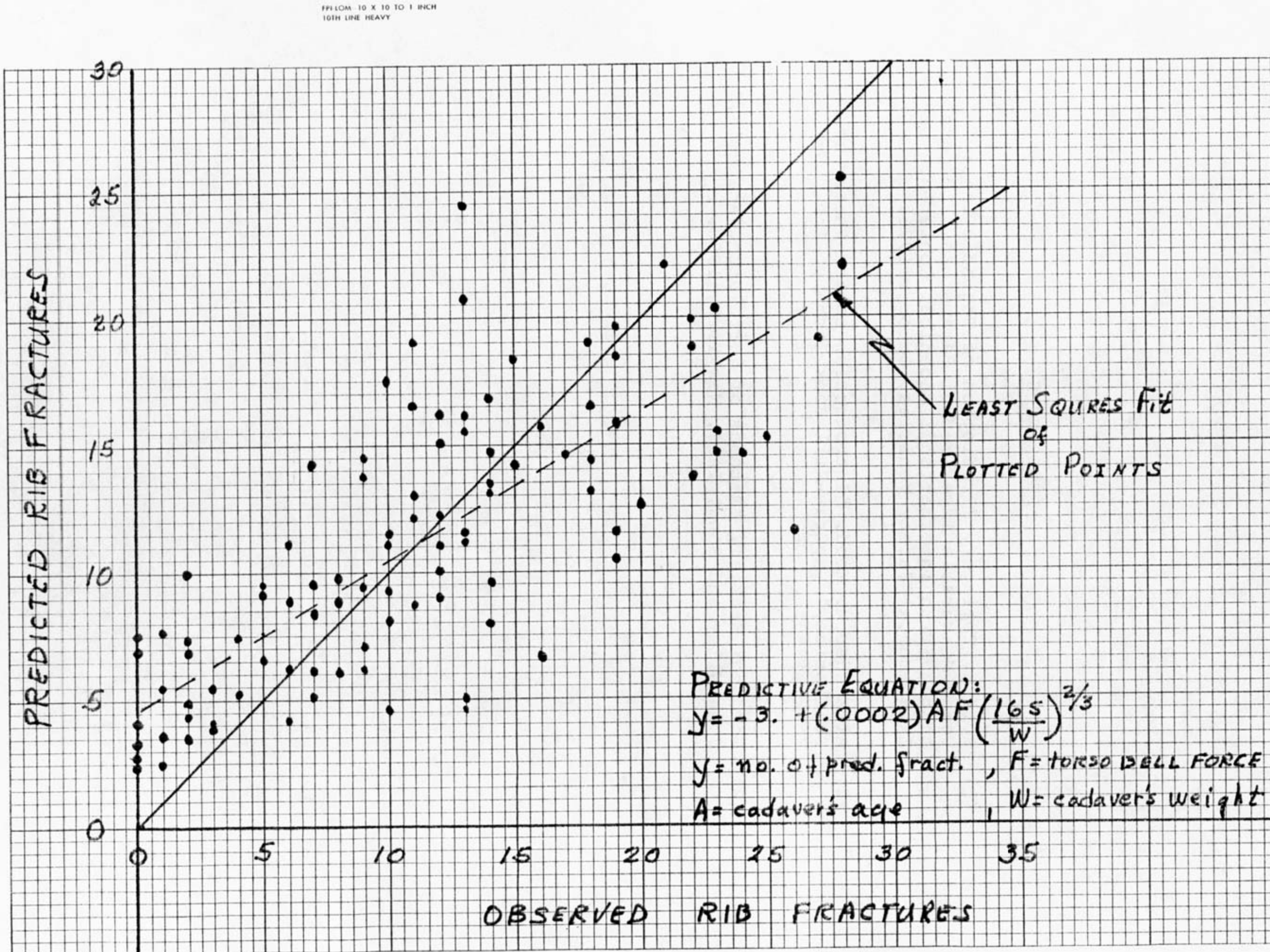

Figure 1 - Predicted Number of Rib Fractures versus the Observed Number

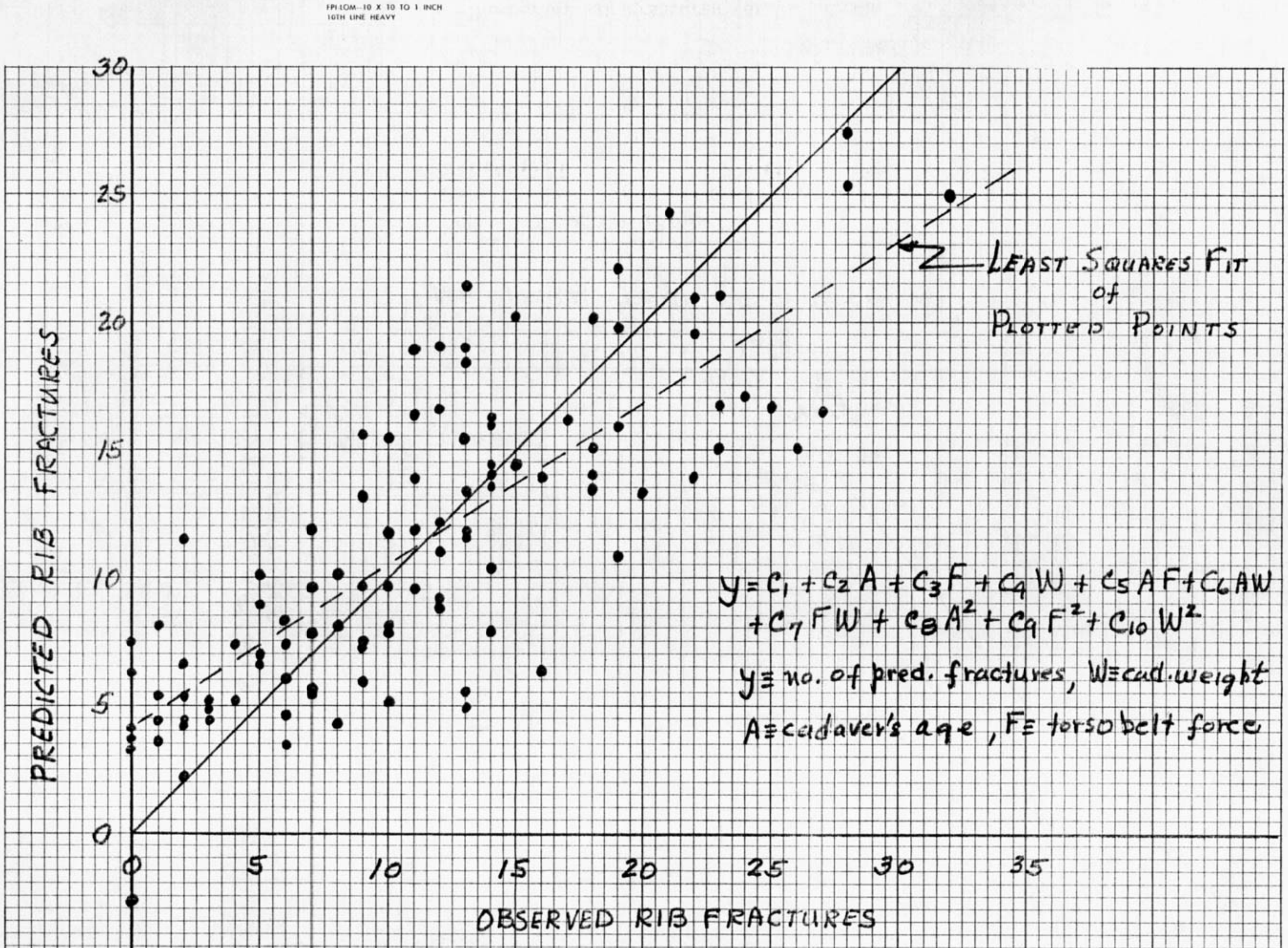

Figure 2 - Predicted Number of Rib Fractures versus the Observed Number (Based on the Equation shown).

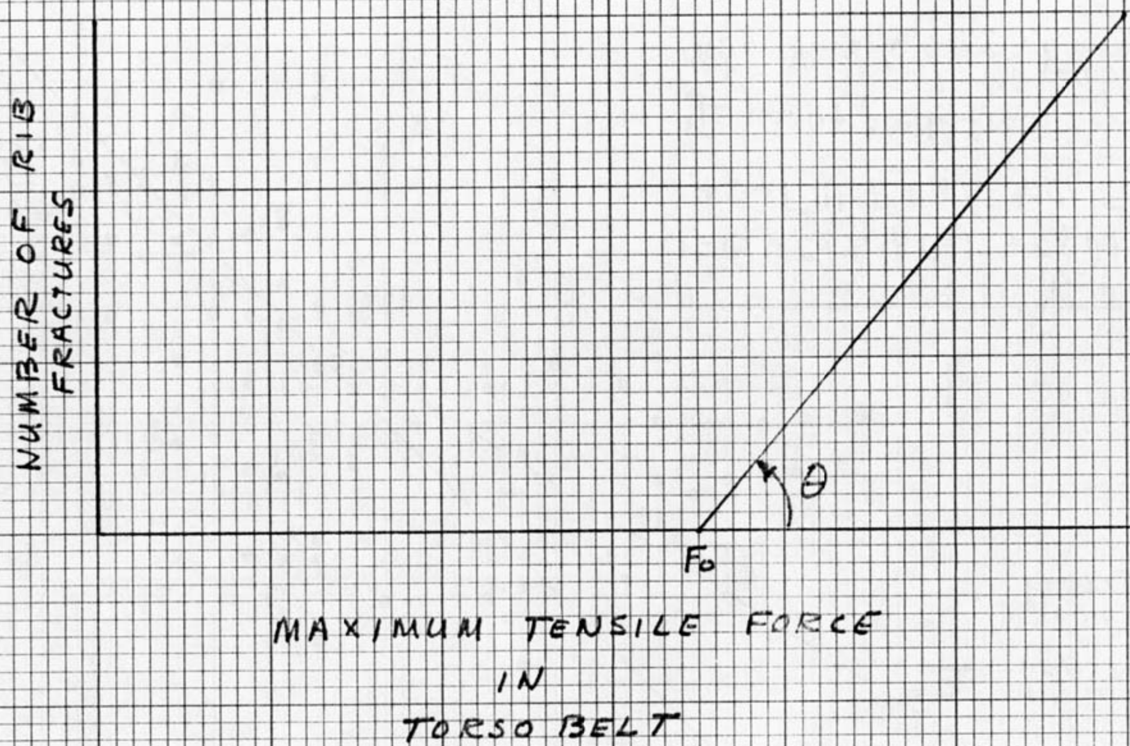

Figure 3 - Number of Rib Fractures versus Maximum Force in Torso Belt.

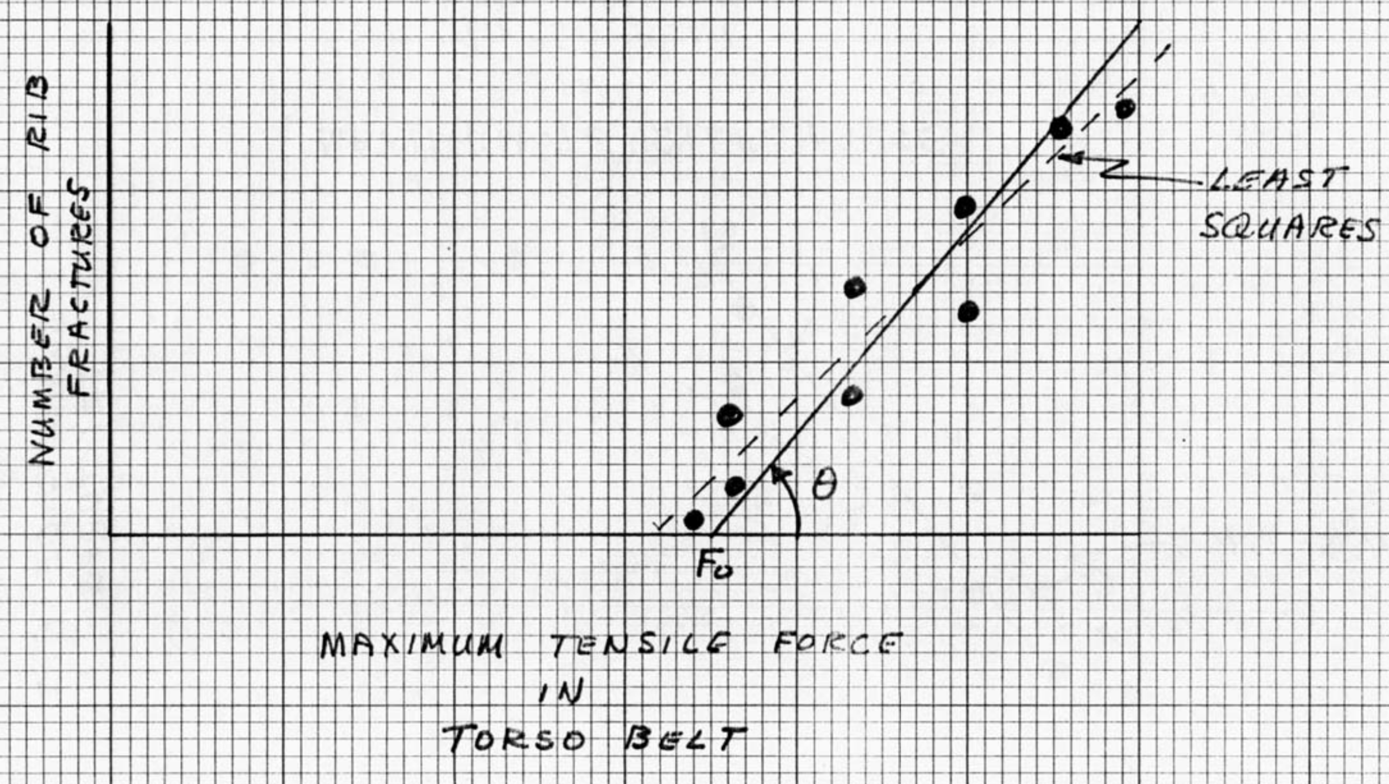

Figure 4 - Data Points to Estimate by Least Squares Curve Fitting the Number of Rib Fractures vs. Maximum Force in Torso Belt.

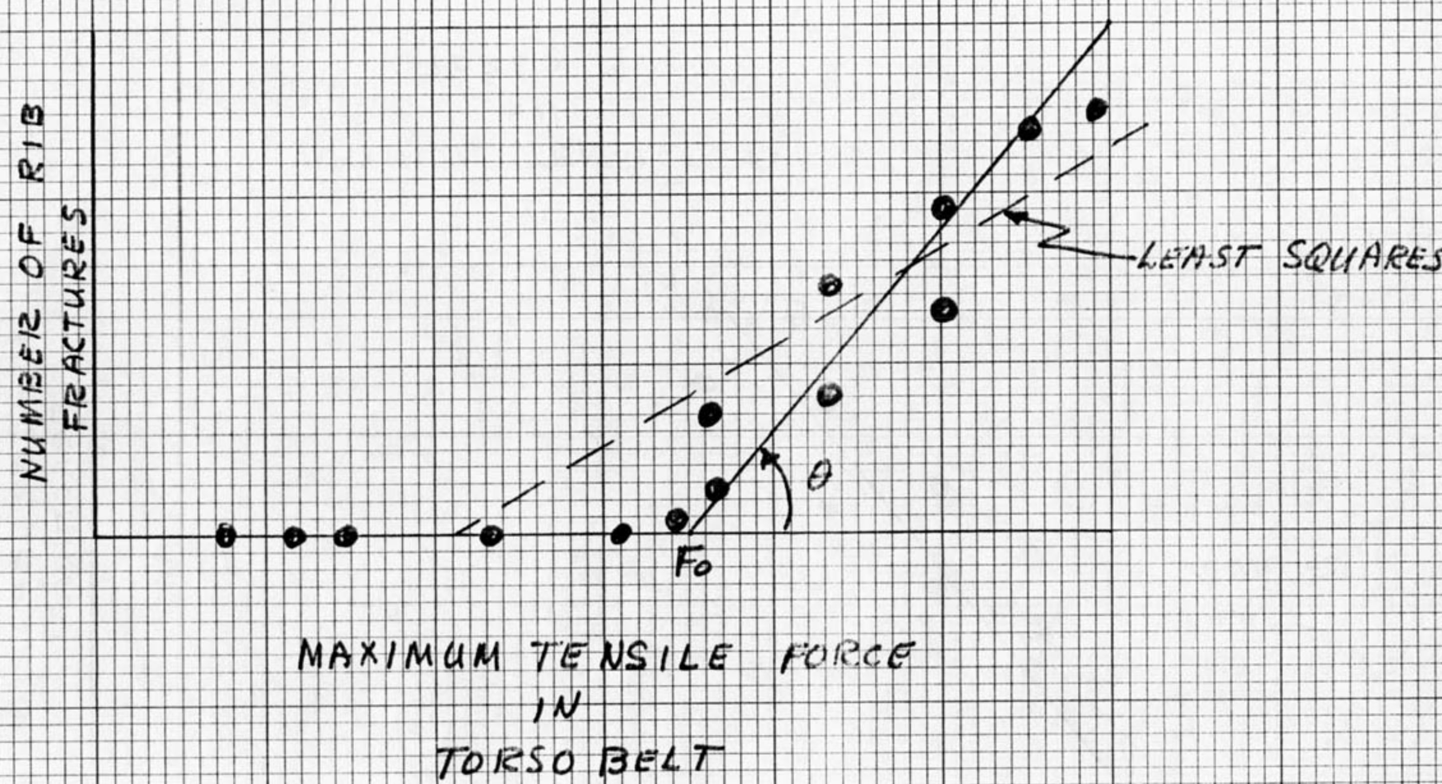

Figure 5 - The Change in the Dashed Line Derived by Least Squares (Compared to Figure 4) When Zero Data Points (i.e., Data Points for No Injury) are Used.

FPI-LOM-10 X 10 TO 1 INCH
10TH LINE HEAVY

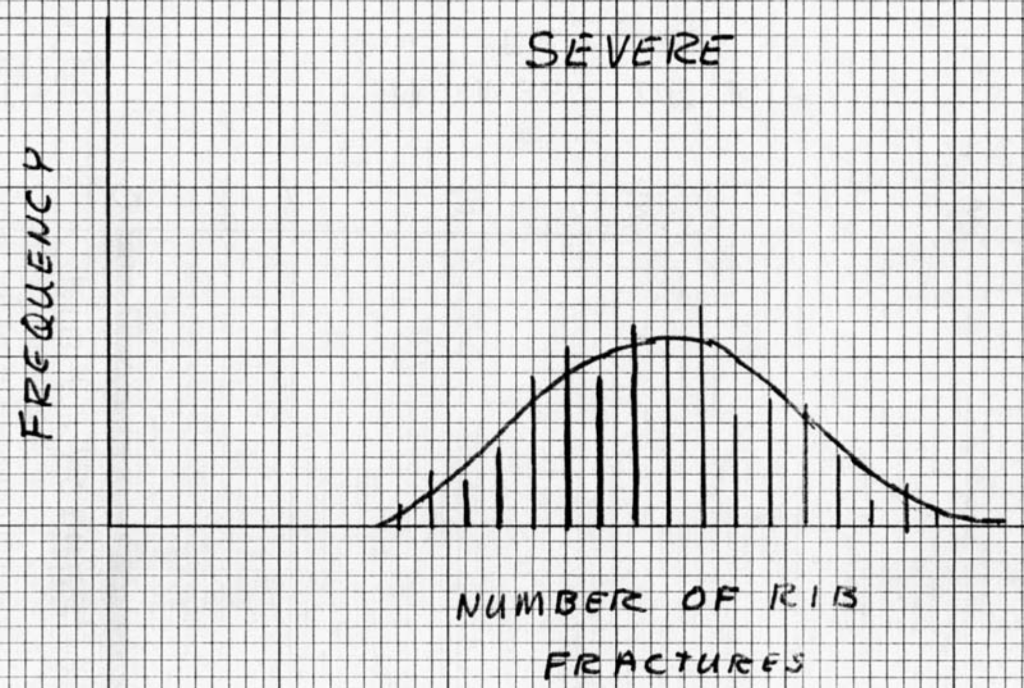

Figure 6 - Hypothetical Frequency Plot of Rib Fractures vs. the Number of Rib Fractures for Severe Impacts.

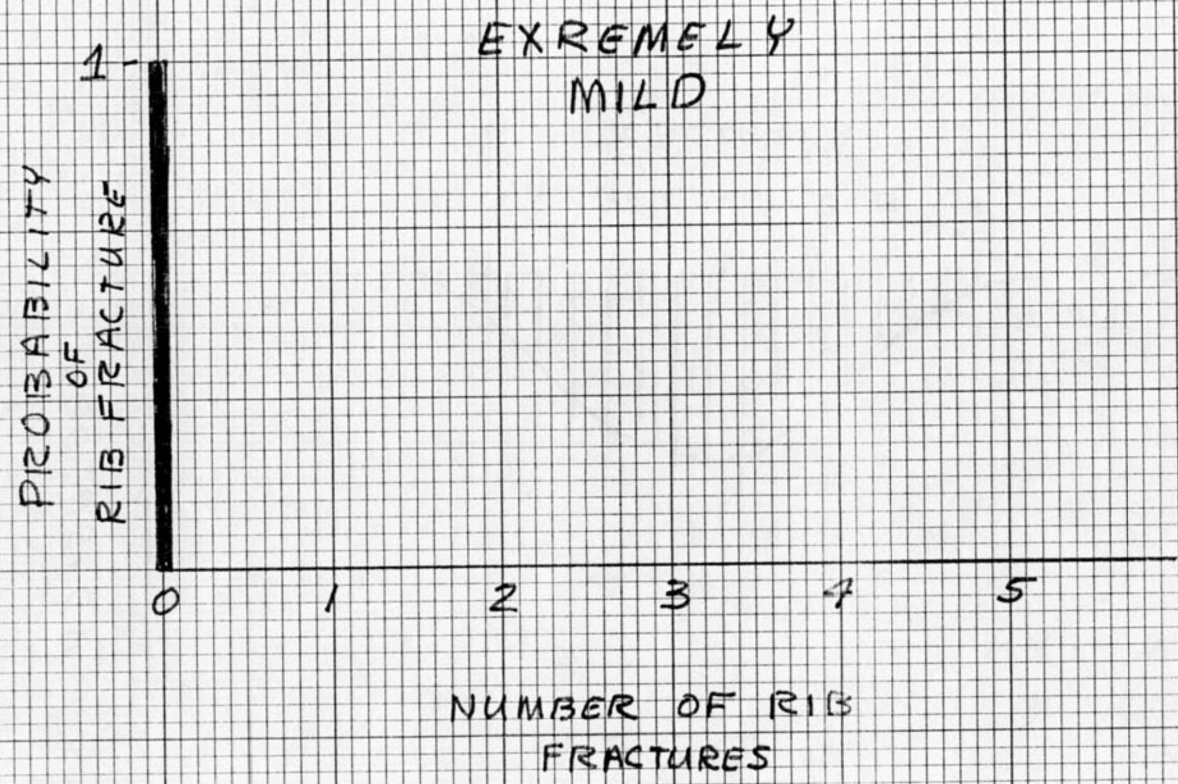

Figure 7 - Probability of Rib Fractures for an Extremely Mild Impact.

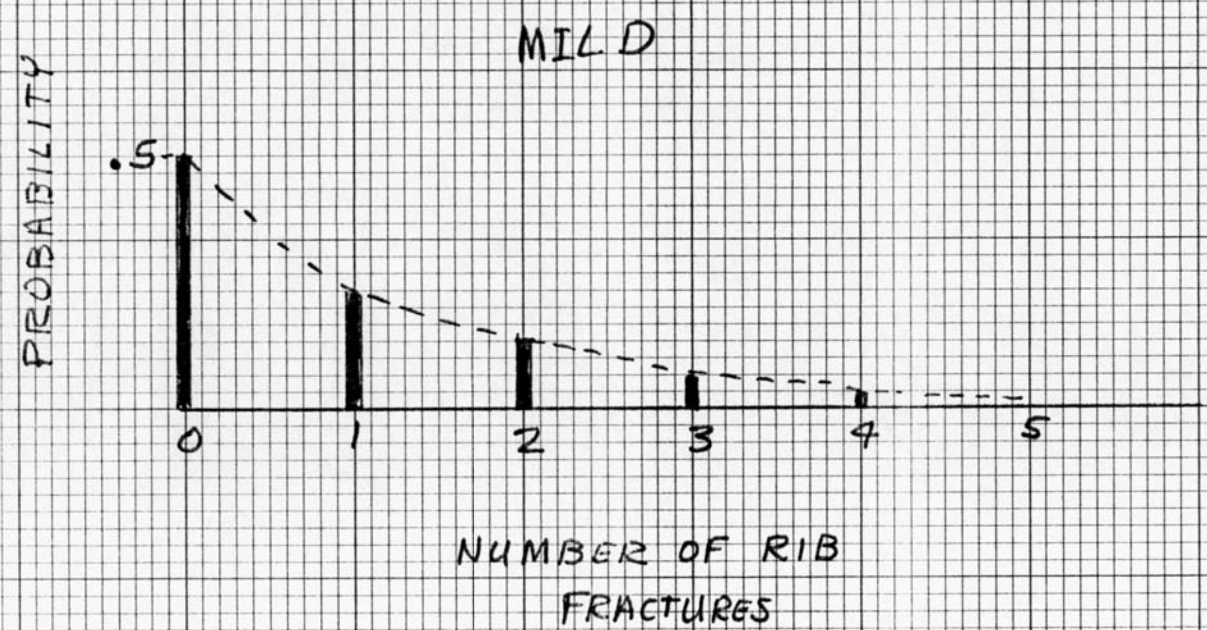

Figure 8 - Possible Probability Distribution for a Mild Impact (0.5 Probability of Zero Rib Fractures).