

# Dose-Injury Relation as a Model for Uncertainty Propagation from Input Dose to Target Dose

Hongyun Wang<sup>1</sup>, Wesley A. Burgei<sup>2</sup>, Hong Zhou<sup>3\*</sup>

<sup>1</sup>Department of Applied Mathematics, University of California, Santa Cruz, CA, USA

<sup>2</sup>U.S. Department of Defense, Joint Non-Lethal Weapons Directorate, Quantico, VA, USA

<sup>3</sup>Department of Applied Mathematics, Naval Postgraduate School, Monterey, CA, USA

Email: \*hzhou@nps.edu

**How to cite this paper:** Wang, H, Burgei, W.A. and Zhou, H. (2018) Dose-Injury Relation as a Model for Uncertainty Propagation from Input Dose to Target Dose. *American Journal of Operations Research*, 8, 360-385.

<https://doi.org/10.4236/ajor.2018.85021>

**Received:** July 8, 2018

**Accepted:** September 15, 2018

**Published:** September 18, 2018

Copyright © 2018 by authors and Scientific Research Publishing Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

## Abstract

We study a general framework for assessing the injury probability corresponding to an input dose quantity. In many applications, the true value of input dose may not be directly measurable. Instead, the input dose is estimated from measurable/controllable quantities via numerical simulations using assumed representative parameter values. We aim at developing a simple modeling framework for accommodating all uncertainties, including the discrepancy between the estimated input dose and the true input dose. We first interpret the widely used logistic dose-injury model as the result of dose propagation uncertainty from input dose to target dose at the active site for injury where the binary outcome is completely determined by the target dose. We specify the symmetric logistic dose-injury function using two shape parameters: the median injury dose and the 10 - 90 percentile width. We relate the two shape parameters of injury function to the mean and standard deviation of the dose propagation uncertainty. We find 1) a larger total uncertainty will spread more the dose-response function, increasing the 10 - 90 percentile width and 2) a systematic over-estimate of the input dose will shift the injury probability toward the right along the estimated input dose. This framework provides a way of revising an established injury model for a particular test population to predict the injury model for a new population with different distributions of parameters that affect the dose propagation and dose estimation. In addition to modeling dose propagation uncertainty, we propose a new 3-parameter model to include the skewness of injury function. The proposed 3-parameter function form is based on shifted log-normal distribution of dose propagation uncertainty and is approximately invariant when other uncertainties are added. The proposed 3-parameter function form provides a framework for extending skewed injury model from a test population to a target population in application.

## Keywords

Dose Injury Relation, Dose Propagation Uncertainty, Median Injury Dose, 10 - 90 Percentile Width, Skewness, Mapping Injury Model from One Population to Another

## 1. Introduction

In many injury assessment situations, injury status of a subject is simply characterized in the form of binary outcome. For example, in a study of skull fracture injury related to highway traffic safety [1] and in a study of rib fracture injury caused by blunt-impact non-lethal weapons [2], in each situation subjects tested are classified as either fractured or not fractured. Mathematically, occurrences of binary injury outcomes are statistically described by injury probability (also called injury risk). Let

- $(v_1, v_2, \dots, v_k)$  be a list of input factors that affect the injury outcome,
- $I$  be the binary injury outcome (random variable), and
- $p$  be the corresponding injury probability:  $p = \Pr(I = \text{“injured”})$

Here the binary injury outcome  $I$  is a random variable even when all input factors  $(v_1, v_2, \dots, v_k)$  are given and fixed. One approach of building a simple and practical model for assessing the injury risk is to use a single metric  $x$  to capture the overall effects of all input variables  $(v_1, v_2, \dots, v_k)$  [3] [4]. Quantity  $x$  is called the input dose, serving as the single metric best predictor of the injury probability. Input dose  $x$  may be one of the input variables  $(v_1, v_2, \dots, v_k)$  or a combination of these input variables. Depending on the application situations, input dose  $x$  is also called the determinant of injury, the risk factor, the exposure level, or the predictor variable [3] [4].

When the input dose  $x$  is directly controllable and measurable, an experimental data set consists of  $m$  entries, each containing a measured value of input dose and the corresponding binary injury outcome in an independent trial:

$$\text{Data} = \{(x_j, I_j), j = 1, 2, \dots, m\}. \quad (1)$$

Injury models are constructed in the general form of injury probability vs input dose.

$$p(x) = \text{injury probability at input dose } x$$

In many application situations, however, the input dose is not directly measurable. For example, for bone fracture injuries, we may use the stress at the impact site as the input dose. But it is difficult to measure directly the stress at impact site. In a study of behind-armor blunt trauma (BABT) [5] and a study of human body response to blunt impacts using advanced total body model (ATBM) [6], an estimated value of stress caused by the impact is calculated via computer simulations. The estimation is based on the measured mass and

velocity of projectile and using representative median material properties of the projectile, the subject body and the armor. When the true input dose is not directly measurable, an experimental data set contains pairs of estimated input dose and the corresponding binary injury outcome:

$$\text{Data} = \left\{ \left( x_j^{(est)}, I_j \right), j = 1, 2, \dots, m \right\}. \quad (2)$$

In these situations, practical injury models are constructed in the form of injury probability vs the *estimated* input dose

$$p\left(x^{(est)}\right) = \text{injury probability at estimated input dose } x^{(est)}$$

The estimated input dose, in general, is different from the true input dose, and the discrepancy between the two is population dependent since the actual material properties of individual subjects are different from the selected representative material properties and are population dependent. In addition, the relation of injury probability vs *true* input dose is also population dependent because the material properties of subjects significantly affect the injury outcome even when the true input dose is fixed. For example, at a fixed impact force, the injury probability varies considerably among groups of different ages, among groups of different body types, body sizes and body compositions. The experimentally established relation of injury probability vs estimated input dose is heavily influenced by the particular population tested. As a result, applying the injury model established for one population, straightforwardly without modification, to assess the injury risk of a different population will inevitably lead to large errors. In many applications, however, we face exactly this task: we are given an injury model established on a particular test population and we need to predict the injury risk of a different population. For example, a data set for human forearm fracture was assembled in [7] from drop test results conducted on PMHS forearms from cadaver donors of average age 55. The purpose of assembling the data set, however, is to build an injury model for assessing the risk of forearm fracture among a population of live human subjects with an age distribution significantly different from that of cadaver donors. In this study, we develop a simple mathematical framework for this task. The key idea is based on interpreting the probabilistic injury model as the consequence of dose propagation uncertainty from input dose to target dose at the active site for injury where the binary outcome is uniquely determined by the target dose. The framework of dose propagation uncertainty makes it mathematically convenient to accommodate different uncertainties associated with different populations. The formulation developed provides a mechanism of mapping injury function from one population to another by simply updating the model parameters.

## 2. Mathematical Formulation

We first review the logistic model for binary outcomes [8]. Note that the injury

probability  $p$  is not directly observable in experiments unless we repeat the experiment a large number of times at each fixed value of input dose. The logistic regression model was designed by working with the hidden injury probability  $p$  and considering the logit function of  $p$ , which is defined as the logarithm of the injury odds:  $\text{logit}(p) \equiv \log\left(\frac{p}{1-p}\right)$ . In the logistic model,  $\text{logit}(p)$  is postulated to be a linear function of input dose  $x$ ,

$$\text{logit}(p) = \alpha(x - D_{50}) \quad (3)$$

Writing probability  $p$  as a function of  $x$ , we obtain the logistic dose response relation

$$p = \text{logit}^{-1}(\alpha(x - D_{50})) = \frac{1}{1 + \exp(-\alpha(x - D_{50}))} \equiv f_{\text{logistic}}(x) \quad (4)$$

We write the linear function in (3) as  $\alpha(x - D_{50})$  so that constant  $D_{50}$  has the meaning of the median injury dose, at which the injury probability is 50%:  $p(D_{50}) = 50\%$  [9]. In general, we introduce  $D_{\eta}$  to denote the dose value at which  $p(D_{\eta}) = \eta\%$ . For example,  $D_{90}$  is the dose with  $p(D_{90}) = 90\%$ . Coefficient  $\alpha$  controls the steepness of transition (*i.e.*, the sensitivity of injury probability with respect to dose change). We define the width of injury function as

$$W \equiv D_{90} - D_{10}$$

Conceptually, the width  $W$  is not the 10 - 90 percentile range of  $x$  since dose  $x$  is not a random output of an experiment; it is the controlled input. However, if we view the injury function as the cumulative distribution function (CDF) for  $x$  and draw random samples of  $x$  based on the CDF, then the width  $W$  is indeed the 10 - 90 percentile range of random samples drawn. For simplicity, we shall call  $W$  the 10 - 90 percentile width even though  $x$  is not a random variable. In the logistic model, the width  $W$  is inversely proportional to coefficient  $\alpha$ .

$$W \equiv D_{90} - D_{10} = \frac{2 \ln(9)}{\alpha} \quad (5)$$

We point out that the steepness coefficient  $\alpha$  exists only in the logistic model. In contrast, the width of injury function is universally defined and meaningful for all injury models. To facilitate the comparison of various models, we shall use the width ( $W$ ) instead of coefficient  $\alpha$  whenever it is appropriate to do so. The logistic model in terms of shape parameters ( $D_{50}, W$ ) has the expression.

$$p = \frac{1}{1 + \exp\left(-\frac{2 \ln(9)}{W}(x - D_{50})\right)} \equiv f_{\text{logistic}}(x; D_{50}, W) \quad (6)$$

Logistic model is widely used as a phenomenological model for binary outcomes [3] [4] [10]. In this study, we interpret it and approximate it in the framework of dose propagation uncertainty from input dose to target dose. The

key assumption in our interpretation is that there is an active site at which target dose  $Z$  uniquely determines the binary outcome  $I$ . Mathematically, target dose  $Z$  at the active site has the features described below:

- Binary outcome  $I$  is the indicator function of  $Z > z^{(c)}$

$$I = \begin{cases} 1, & \text{if } Z > z^{(c)} \\ 0, & \text{otherwise} \end{cases} \quad (7)$$

where  $z^{(c)}$  is the critical threshold for target dose in transition from non-injury to injury. The transition is a discontinuous jump with respect to target dose  $Z$  at the active site. However, with respect to the input dose  $x$  that is away from the active site, the injury probability vs  $x$  generally is a smooth and gradual transition.

- The target dose  $Z$  is caused by the input dose  $x$ . While in most experiments the input dose  $x$  can be controlled, at least to some extent, the target dose  $Z$  is neither directly observable nor directly controllable.
- For a given input dose  $x$ , the corresponding target dose  $Z$  is a random variable, reflecting the uncertainty in the propagation from input dose to target dose.

We use an example to illustrate the propagation from input dose to target dose.

Example: Passing exam vs amount of study time

In this example, the input dose  $x$  is the amount of study time. Note that although the target dose  $Z$  is caused by the input dose  $x$ , quantities  $Z$  and  $x$  may have different physical dimensions. For passing an exam, the target dose  $Z$  is the effective fraction of actual exam contents correctly completed in the exam by the student. We use a flow chart to show a possible propagation from input dose to target dose.

- $x$  = the nominal amount of study time invested
- $Z_1$  = effective amount of study time  
*affected by the student's attentiveness, efficiency, and overall load*
- $Z_2$  = amount of course contents learned  
*affected by the student's prior preparation and ability of memorizing key items*
- $Z_3$  = fraction of actual exam contents learned  
*affected by the exam scope and weighting of components in exam*
- $Z$  = effective fraction of actual exam contents correctly completed  
*affected by the student's general health condition on exam day, and ability of working under time pressure and in presence of noise/disturbance* (8)

Mathematically, we write the target dose explicitly as  $Z(x, \omega)$ , emphasizing that  $Z$  is a random variable depending on the input dose  $x$  and depending on the random factor  $\omega$  in the dose propagation. The probability that a given input dose  $x$  leads to injury is

$$\Pr(I = \text{"injured"}) = \Pr(Z(x, \omega) > z^{(c)}) \quad (9)$$

We consider two models for uncertainty in dose propagation: 1) target dose  $Z(x, \omega)$  has a normal distribution; and 2) target dose  $Z(x, \omega)$  is expressed in terms of a normally distributed intermediate variable. For example, intermediate variable  $Y(x, \omega) \equiv \ln(Z(x, \omega) - x_0)$  has a normal distribution, and target dose  $Z(x, \omega)$  is a shifted log normal distribution, expressed in terms of intermediate variable  $Y(x, \omega)$  as  $Z(x, \omega) = \exp(Y(x, \omega)) + x_0$ .

### 3. Logistic Dose-Injury Relation Interpreted as Normally Distributed Target Dose

We model the target dose as proportional to the sum of the input dose and an additive Gaussian noise.

$$Z(x, \omega) = r \times (x - \mu + \sigma \varepsilon)$$

where  $\varepsilon \sim N(0,1)$ , a standard normal random variable. We scale target dose  $Z$  and the associated critical threshold  $z^{(c)}$  to make  $r=1$  by changing the physical unit for measuring  $z$ -values, or equivalently by changing the physical unit for measuring  $x$ -values. Thus, we set  $r=1$  and proceed with

$$Z(x, \omega) = x - \mu + \sigma \varepsilon \quad (10)$$

In this section, we first examine the dose-response relation for normally distributed dose uncertainty, which is the probit model [11]. Then we discuss how to accommodate different uncertainties corresponding to different populations, including how to incorporate additional uncertainties into the dose-response relation.

#### 3.1. Dose-Response Relation

The binary injury outcome is governed by the sign of random variable

$$Z(x, \omega) - z^{(c)} = x - z^{(c)} - \mu + \sigma \varepsilon \quad (11)$$

The injury probability ( $p$ ) corresponding to input dose  $x$  is

$$p = \Pr\left(\left(x - z^{(c)} - \mu + \sigma \varepsilon\right) > 0\right) = \Pr\left(-\varepsilon < \frac{x - z^{(c)} - \mu}{\sigma}\right)$$

Recall that the cumulative distribution function (CDF) of standard normal is given by the error function,  $\text{erf}(u)$ , which is defined as

$$\text{erf}(u) \equiv \frac{2}{\sqrt{\pi}} \int_0^u \exp(-s^2) ds$$

The dose response relation for normally distributed target dose  $Z$  has the expression:

$$p = \frac{1}{2} + \frac{1}{2} \text{erf}\left(\frac{x - z^{(c)} - \mu}{\sqrt{2}\sigma}\right) \equiv f_{\text{normal}}(x) \quad (12)$$

We approximate dose-response relation (12) using the logistic function form (4) with tunable parameters  $D_{50}$  and  $\alpha$ . First, we match the two functions at

$p = 50\%$  to obtain  $D_{50} = z^{(c)} + \mu$ . To simplify the search for optimal  $\alpha$ , we apply the transformation

$$x_{\text{new}} = \frac{x_{\text{old}} - D_{50}}{\sigma}$$

After the transformation, (4) and (12) as functions of  $x_{\text{new}}$  have standard forms:

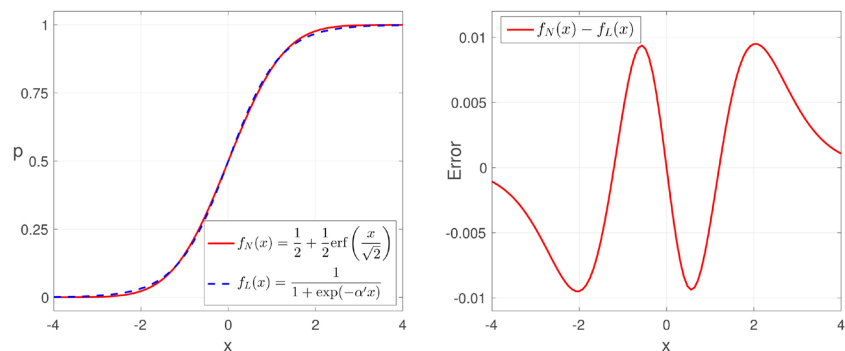
$$f_{\text{logistic}}(x_{\text{old}}) \mapsto f_L(x) \equiv \frac{1}{1 + \exp(-\alpha'x)} \tag{13}$$

$$f_{\text{normal}}(x_{\text{old}}) \mapsto f_N(x) \equiv \frac{1}{2} + \frac{1}{2} \operatorname{erf}\left(\frac{x}{\sqrt{2}}\right) \tag{14}$$

where the scaled coefficient  $\alpha'$  is related to  $\alpha$  by  $\alpha' = \sigma\alpha$ . For conciseness, we denote  $x_{\text{new}}$  simply as  $x$ . The task of approximating (12) with (4) is reduced to finding an optimal value of  $\alpha'$  such that the distance between  $f_N(x)$  and  $f_L(x)$  is minimized. Using numerical optimization, we find that the best approximation is achieved at  $\alpha'_{\text{opt}} = 1.701$ .

**Figure 1** compares functions (14) and (13) at  $\alpha'_{\text{opt}} = 1.701$ . It is clear that the two functions are very good approximations of each other. The maximum difference is bounded by 0.01 (*i.e.*, difference in predicted injury probability is less than 1%). With that error tolerance, the logistic model and the normal distribution model can practically substitute each other. In other words, the widely used logistic model can be viewed as a very good approximation of the normal distribution model, which was derived based on normally distributed dose propagation uncertainty from input dose to target dose.

Models (13) and (14) are nevertheless mathematically different. When the data set of binary injury outcomes ( $J$ ) is sufficiently large, eventually, the two models will be distinguishable. Let  $m$  be the number of samples in the data set. We look into the question of how large  $m$  needs to be in order to statistically distinguish the two models. We consider a collection of independent data sets, each of the form



**Figure 1.** Comparison of  $f_N(x) \equiv \frac{1}{2} + \frac{1}{2} \operatorname{erf}\left(\frac{x}{\sqrt{2}}\right)$  and  $f_L(x) \equiv \frac{1}{1 + \exp(-\alpha'x)}$  at  $\alpha'_{\text{opt}} = 1.701$ . Left panel: plots of the two functions. Right panel: plot of the difference between the two functions. The results shown demonstrate that the two functions are very good approximations to each other.

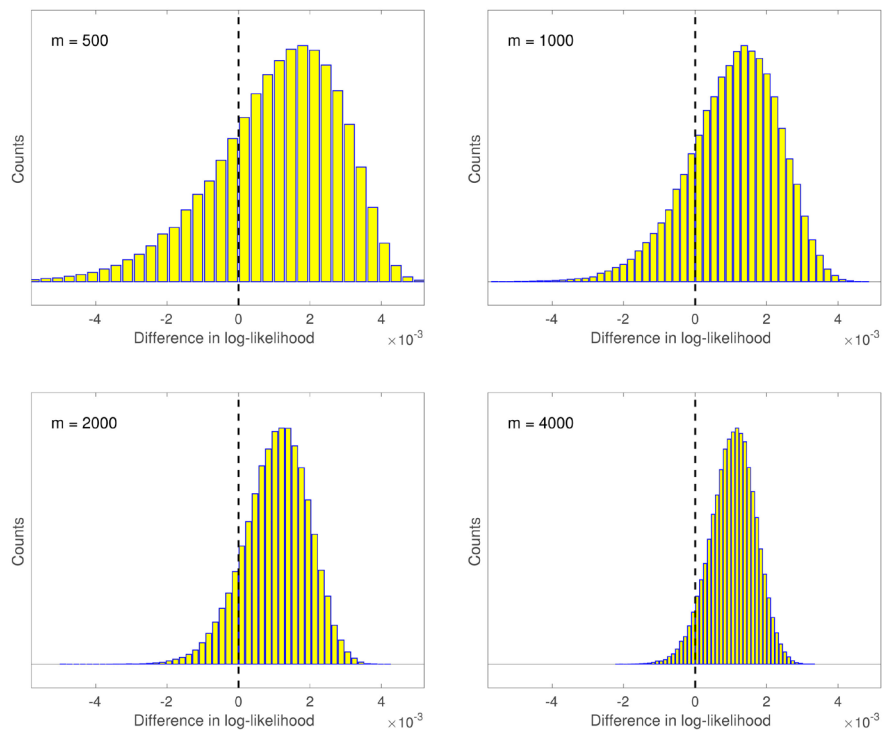
$$D = \{(x_j, I_j), j = 1, 2, \dots, m\}$$

where  $x_j$  is the input dose of the  $j$ -th experiment and  $I_j$  the corresponding binary injury outcome. To test if the two models are statistically distinguishable, we generate data sets according to the normal distribution model  $f_N(x)$  in (14). In all data sets, values of input dose  $\{x_j\}$  are uniformly distributed in  $[-3, 3]$ , and for each input dose  $x_j$  the corresponding binary injury outcome  $I_j$  is sampled using injury probability  $f_N(x_j)$ .

Given data set  $D$ , the log-likelihood for a general probability function  $f(x)$  is

$$\mathcal{L}(f(\cdot) | D) \equiv \frac{1}{m} \sum_{j=1}^m (I_j \log(f(x_j)) + (1 - I_j) \log(1 - f(x_j))) \quad (15)$$

We use log-likelihood (15) to compare models  $f_N(x)$  and  $f_L(x)$ . Since  $f_N(x)$  is the exact probability model for the data set while  $f_L(x)$  is a slightly incorrect model, the difference in log-likelihood  $\mathcal{L}(f_N(\cdot) | D) - \mathcal{L}(f_L(\cdot) | D)$  is expected to be positive. However, due to randomness of data sets, the difference in log-likelihood between two models fluctuates from one data set to another. We examine the sample distribution of differences in log-likelihood based on  $N = 100000$  independent data sets. **Figure 2** plots the histograms of  $\mathcal{L}(f_N(\cdot) | D) - \mathcal{L}(f_L(\cdot) | D)$  for various values of  $m$ , the size of each data set.



**Figure 2.** Histograms of  $\mathcal{L}(f_N(\cdot) | D) - \mathcal{L}(f_L(\cdot) | D)$  for various values of  $m$ , the size of individual data sets, each yielding a sample for the histogram. Top left panel: histogram based on  $N = 100000$  independent data sets, each containing  $m = 500$  samples; top right panel:  $m = 1000$ ; bottom left panel:  $m = 2000$ ; and bottom right panel:  $m = 4000$ .



To clarify, here  $N$  is the number of data sets used in each histogram and  $m$  is the number of binary outcomes in each data set. In **Figure 2**, each sample of difference in log-likelihood requires one data set. That is why we use  $N = 100000$  independent data sets to plot each histogram.

Suppose we use the sign of  $\mathcal{L}(f_N(\cdot)|D) - \mathcal{L}(f_L(\cdot)|D)$  to classify data sets as the normal distribution model (positive sign) or as the logistic model (negative sign). All data sets examined in **Figure 2** are generated based on the normal distribution model. Thus, data sets with  $\mathcal{L}(f_N(\cdot)|D) - \mathcal{L}(f_L(\cdot)|D) < 0$  will be falsely identified as the logistic model (false negative). In **Figure 2**, all counts to the left of the dashed black line in each histogram correspond to false negative identification. For data sets of  $m = 500$  samples each (top left panel), the false negative rate is 25.26%. For  $m = 1000$  (top right panel), the false negative rate decreases to 19.44%. When the sample size is increased to  $m = 2000$  (bottom left panel), the false negative rate falls to 12.49%. Finally, when the sample size is doubled again to  $m = 4000$  (bottom right panel), the false negative rate drops down to 5.63%. Based on the simulation results, we see that to reduce the false negative rate to less than 20%, for example, we need to work with data sets, each consisting of  $m = 1000$  samples. This is above the typical sample size of data sets for injury models. Thus, in real applications, the normal distribution model (14) and logistic model (13) are practically the same unless we work with injury data sets of very large sample size.

We go back to the pre-transformation logistic model, function (4) specified by steepness coefficient  $\alpha$ , and function (6) specified by width  $W$ . The corresponding optimal values for  $\alpha$  and for  $W$  are respectively

$$\begin{aligned}\alpha_{\text{opt}} &= \frac{\alpha'_{\text{opt}}}{\sigma} = \frac{1.701}{\sigma} \\ W_{\text{opt}} &= \frac{2 \ln(9)}{\alpha_{\text{opt}}} = \frac{2 \ln(9)}{1.701} \sigma = 2.584\sigma\end{aligned}\quad (16)$$

Since the 10 - 90 percentile width is well defined for all injury functions, we choose to specify the logistic model using width  $W$  instead of coefficient  $\alpha$ . We conclude that normal distribution model (12) based on dose propagation uncertainty is practically equivalent to logistic model (6) with shape parameters  $(D_{50}, W)$  given by

$$W = 2.584\sigma, \quad D_{50} = z^{(c)} + \mu \quad (17)$$

(17) describes the best approximation to the normal distribution model (12) from the logistic model family (6). The best approximation is obtained numerically by minimizing the distance between the two functions (**Figure 1**). Alternatively, a straightforward approximation can be written out by simply matching the widths of two injury functions. The width of normal distribution model is given by the inverse error function

$$W_{\text{normal}} = 2\sqrt{2} \operatorname{erf}^{-1}(0.8)\sigma = 2.563\sigma$$

Notice that the two widths, the width of normal distribution model  $W_{\text{normal}}$

and the width of its best logistic model approximation  $W_{\text{opt}}$ , are indeed very close to each other. We will use these two interchangeably.

Similar to the situation of logistic model, the normal distribution model is also completely specified by the shape parameters  $(D_{50}, W)$ . It has the form

$$p = \frac{1}{2} + \frac{1}{2} \operatorname{erf} \left( \frac{2 \operatorname{erf}^{-1}(0.8) \cdot (x - D_{50})}{W} \right) \equiv f_{\text{normal}}(x; D_{50}, W) \quad (18)$$

where shape parameters  $(D_{50}, W)$  are related to parameters of dose propagation uncertainty in (17). It should be pointed out that in general, the target dose  $Z$  is hidden, not observable or controllable; none of parameters  $z^{(c)}$ ,  $\mu$  or  $\sigma$  is directly observable. These are internal quantities in the mathematical model, explaining why the injury probability follows the normal distribution model (12). In an idealized situation, the input dose  $x$  should be a controllable/measurable variable, and shape parameters  $(D_{50}, W)$  may be determined from experimental measurements. In realistic applications, however, the true input dose  $x$  may not be directly measurable, which we will discuss in next subsection. At the end of this subsection, we summarize the normal distribution model for dose propagation uncertainty, and its connection to the widely used logistic model.

#### Summary of the injury model based on dose propagation uncertainty

- We select the physical unit for measuring the target dose  $Z$  such that in the absence of dose propagation uncertainty, target dose  $Z$  is the same as input dose  $x$ :

$$Z|_{(\text{zero uncertainty})} = x.$$

- In the normal distribution model, the difference between target dose and input dose is an additive Gaussian noise:

$$Z(x, \omega) - x = -\mu + \sigma N(0, 1).$$

- The binary injury outcome is completely determined by the condition  $Z(x, \omega) > z^{(c)}$  where  $z^{(c)}$  is the critical threshold for target dose  $Z$ .
- The probability of injury caused by the input dose  $x$  is described by the CDF of normal distribution. Practically the injury probability is very well approximated by the widely used logistic dose-response relation.
- As given in (17), the median injury dose of injury function is the critical threshold for the target dose, shifted by the bias in the dose propagation:

$$D_{50} = z^{(c)} + \mu,$$

and the width of injury function is proportional to the uncertainty in dose propagation (standard deviation of the Gaussian noise):

$$W = 2.584\sigma$$

The larger the uncertainty, the more spread out the injury function is.

- In terms of shape parameters  $(D_{50}, W)$ , the logistic model is expressed in (6); the normal distribution model is given in (18).

Next, we study how to incorporate additional uncertainties in the framework of dose-response relation, and how to model a new population with different uncertainty.

### 3.2. Effects of Additional Uncertainties

In the previous subsection, we interpreted the dose-response relation as a consequence of dose propagation uncertainty. In this subsection we study how to incorporate additional uncertainties by changing the shape parameters  $(D_{50}, W)$  in logistic model (6) or in normal distribution model (18).

We start by considering a homogeneous population consisting of statistically identical subjects, which means quantities  $z^{(c)}$ ,  $\mu$  and  $\sigma$  are fixed and stay the same for all subjects in the population. In a homogeneous population, the dose propagation uncertainty is statistically the same for all subjects. Its effect is already reflected in the dose response relation specified by shape parameters  $(D_{50}, W)$ , which are related to internal parameters  $(z^{(c)}, \mu, \sigma)$  in (17). In particular, the width  $W$  is proportional to the standard deviation of uncertainty. If there is no uncertainty present in the dose propagation, the dose-response relation would be a sharp transition (a step function).

Now we consider a more realistic situation: a heterogeneous population consisting of subjects with variable critical threshold  $z^{(c)}$ , denoted here in the new setting as  $Z^{(c)}$ , following the convention of using uppercase letters for random variables. In addition to the uncertainty in  $Z^{(c)}$ , the input dose  $x$  may not be directly measurable. In some situations, the input dose  $x$  is not directly measured; instead, input dose  $x$  is derived from a controllable/measurable variable  $y$ . In these situations, the value of input dose  $x$  is calculated via computer simulations from measurable quantities using idealized representative properties of subjects, such as the 50-percentile properties of the general population [5] [6]. We use the example below to illustrate the situation of controllable variable  $y$  vs true input dose  $X^{(true)}$  vs estimated input dose  $X^{(est)}$ . Consider the experiment in which we test the shatter resistance of a product by dropping it from a specified height. In this example, the various quantities in the model are described as follows:

- The height  $y$  is the controllable/measurable variable.
- The estimated input dose  $X^{(est)}$  is the impact force calculated in a computer simulation from height  $y$  using the representative median properties, such as the weight of the product, the aerodynamic properties, the mechanical properties of the product and the ground surface, and the orientation angle of the product at impact.
- The true input dose  $X^{(true)}$  is the actual impact force, which in general is different from the estimated input dose  $X^{(est)}$ . The difference  $(X^{(true)} - X^{(est)})$  depends on how much the true properties deviate from the selected representative properties. The distribution of difference varies from one population to another.

- The target dose  $Z(X^{(true)}, \omega)$  is the maximum stress at the most vulnerable part of the product.

The bottom line is that the true input dose  $X^{(true)}$  is a random variable when the controllable variable  $y$  is specified. We model the difference  $(X^{(true)} - X^{(est)})$ , the dose propagation uncertainty  $(Z(X^{(true)}, \omega) - X^{(true)})$ , and the critical threshold  $Z^{(c)}$  as additive Gaussian noises. Mathematically, we formulate the problem as

$$(Z(X^{(true)}, \omega) - X^{(true)}) = -\mu_1 + \sigma_1 \varepsilon_1 \quad (19)$$

$$(X^{(true)} - X^{(est)}) = -\mu_2 + \sigma_2 \varepsilon_2 \quad (20)$$

$$Z^{(c)} = z^{(c)} + \mu_3 + \sigma_3 \varepsilon_3 \quad (21)$$

where  $\{\varepsilon_1, \varepsilon_2, \varepsilon_3\}$  are *i.i.d.* samples of  $N(0,1)$ . The binary injury outcome is governed by the sign of random variable

$$Z(X^{(true)}, \omega) - Z^{(c)} = X^{(est)} - z^{(c)} - (\mu_1 + \mu_2 + \mu_3) + \sqrt{\sigma_1^2 + \sigma_2^2 + \sigma_3^2} \varepsilon \quad (22)$$

At a given value of  $X^{(est)}$ , random variable  $(Z(X^{(true)}, \omega) - Z^{(c)})$  has the same mathematical form as random variable  $(Z(x, \omega) - z^{(c)})$  in (11). As a result, the injury probability vs the estimated input dose has the expression

$$\begin{aligned} p|_{X^{(est)}=x} &= \Pr\left[\left[x - z^{(c)} - (\mu_1 + \mu_2 + \mu_3) + \sqrt{\sigma_1^2 + \sigma_2^2 + \sigma_3^2} \varepsilon\right] > 0\right] \\ &= \frac{1}{2} + \frac{1}{2} \operatorname{erf}\left(\frac{x - z^{(c)} - (\mu_1 + \mu_2 + \mu_3)}{\sqrt{2} \sqrt{\sigma_1^2 + \sigma_2^2 + \sigma_3^2}}\right) \end{aligned} \quad (23)$$

Injury function (23) has the same form as (12). Thus,  $p|_{X^{(est)}=x}$  is described by the normal distribution model with shape parameters  $(D_{50}, W)$  given as follows.

$$p|_{X^{(est)}=x} = f_{\text{normal}}(x; D_{50}, W) \quad (24)$$

$$W = 2.584 \sqrt{\sigma_1^2 + \sigma_2^2 + \sigma_3^2}$$

$$D_{50} = z^{(c)} + (\mu_1 + \mu_2 + \mu_3)$$

In a well controlled lab setting, the true input dose  $X^{(true)}$  is measurable. For example, in experiments of male forearm fracture [12], a cylinder of specified mass is dropped from a specified height along a vertical track onto the PMHS forearm sample. Both the forearm sample and the cylinder are connected to accelerometers, allowing accurate measurements of the dynamic impactor load and the support loads. In addition, in situ strain gauges are used to record time series of strains at various locations during the loading. In this idealized setting, there is no measurement error in  $X^{(true)}$ . The injury probability as a function of the true input dose,  $p|_{X^{(true)}=x}$ , can be determined from the observed binary injury outcomes vs measured values of  $X^{(true)}$ . Injury function  $p|_{X^{(true)}=x}$  follows the normal distribution model with shape parameters  $(D_{50}^{(0)}, W^{(0)})$  given below.

$$\begin{aligned}
 p|_{X^{(true)}=x} &= f_{\text{normal}}\left(x; D_{50}^{(0)}, W^{(0)}\right) \\
 W^{(0)} &= 2.584\sqrt{\sigma_1^2 + \sigma_3^2} \\
 D_{50}^{(0)} &= z^{(c)} + (\mu_1 + \mu_3)
 \end{aligned}
 \tag{25}$$

With this formulation, we can map back and forth between injury functions  $p|_{X^{(true)}=x}$  and  $p|_{X^{(est)}=x}$ . We can also revise the injury function  $p|_{X^{(est)}=x}$  measured on one population to construct the injury function for a different population. We now discuss these two problems.

**Problem 1:**

Suppose we are given an injury model  $p|_{X^{(true)}=x}$ , specified by shape parameters  $(D_{50}^{(0)}, W^{(0)})$ . The given injury function is for an idealized setting where the true input dose is directly measured. Our goal is to extend the given injury function  $p|_{X^{(true)}=x}$  to predict the injury probability,  $p|_{X^{(est)}=x}$ , as a function of estimated input dose for the same population when the true input dose is not measurable.

**Solution:**

Injury function  $p|_{X^{(true)}=x}$  is specified by shape parameters  $(D_{50}^{(0)}, W^{(0)})$  given in (25) while injury function  $p|_{X^{(est)}=x}$  is specified by shape parameters  $(D_{50}, W)$  given in (24). Combining (25) with (24), we write  $(D_{50}, W)$  as an update on  $(D_{50}^{(0)}, W^{(0)})$ .

$$\begin{aligned}
 W^2 &= W_0^2 + (2.584)^2 \sigma_2^2 \\
 D_{50} &= D_{50}^{(0)} + \mu_2
 \end{aligned}
 \tag{26}$$

**Problem 2:**

Suppose we are given an injury model  $p|_{X^{(est)}=x}$ , specified by shape parameters  $(D_{50}, W)$ . The given injury function is established based on measurements of a heterogeneous population, labeled population 1. Population 1 is characterized by uncertainties in the input dose estimation and in the critical threshold, as described in (20) and (21)

$$\begin{aligned}
 (X^{(true)} - X^{(est)}) &= N(-\mu_2, \sigma_2^2) \\
 Z^{(c)} &= z^{(c)} + N(\mu_3, \sigma_3^2)
 \end{aligned}$$

Now consider a different heterogeneous population, labeled population 2, with uncertainties described by

$$\begin{aligned}
 (X^{(true)} - X^{(est)}) &= N(-\tilde{\mu}_2, \tilde{\sigma}_2^2) \\
 Z^{(c)} &= z^{(c)} + N(\tilde{\mu}_3, \tilde{\sigma}_3^2)
 \end{aligned}$$

Here we assume that the propagation uncertainty from true input dose to target dose  $(Z(X^{(true)}, \omega) - X^{(true)}) = N(-\mu_1, \sigma_1^2)$  is statistically the same for the two populations. Our goal is to predict the injury function  $(p|_{X^{(est)}=x})_2$  for population 2 based on the given injury function  $p|_{X^{(est)}=x}$  for population 1.

**Solution:**

Injury function  $p|_{X^{(ext)}=x}$  for population 1 is specified by shape parameters  $(D_{50}, W)$  while injury function  $(p|_{X^{(ext)}=x})_2$  for population 2 is specified by shape parameters  $(\tilde{D}_{50}, \tilde{W})$ . We write  $(\tilde{D}_{50}, \tilde{W})$  as an update on  $(D_{50}, W)$  to take into account the differences in uncertainties between the two populations.

$$\begin{aligned}\tilde{W}^2 &= W^2 + (2.584)^2 (\tilde{\sigma}_2^2 - \sigma_2^2 + \tilde{\sigma}_3^2 - \sigma_3^2) \\ \tilde{D}_{50} &= D_{50} + (\tilde{\mu}_2 - \mu_2) + (\tilde{\mu}_3 - \mu_3)\end{aligned}\quad (27)$$

#### 4. Dose-Injury Function for Target Doze of Log-Normal Distribution

For the discussion below, we adopt the normal-distribution model as the base formulation, switching away from the logistic model. There are several reasons behind the switching.

- The normal-distribution model is based on 1) viewing the binary injury outcome as completely determined by the target dose at the active site, 2) explaining the randomness in injury outcome as the consequence of uncertainty in dose propagation from input dose to target dose, and 3) modeling the dose propagation uncertainty as an additive Gaussian noise. This interpretation is both theoretically and operationally appealing.
- Mathematically, the injury function form of normal-distribution model is exactly invariant when additional normally distributed noise/uncertainty is incorporated into the model.
- We will study dose-injury models based on normally distributed *intermediate variable*. Mathematically, such an injury model is conveniently treated as a transformation of the normal-distribution model since the target doze is expressed as a function of the normally distributed intermediate variable.
- As we demonstrated in the previous section, the logistic model is practically equivalent to the normal-distribution model with the same shape parameters  $(D_{50}, W)$ .

We first recall the function form of the normal-distribution model. In terms of internal variables  $(\sigma, \mu, z^{(c)})$ , it is given by (12). In terms of shape parameters  $(D_{50}, W)$ , it is expressed in (18). Geometric quantities  $D_{50}$ ,  $D_{10}$ ,  $D_{90}$  and  $W$  of the injury function are related to internal variables  $(\sigma, \mu, z^{(c)})$  as

$$\begin{aligned}D_{50} &= z^{(c)} + \mu \\ D_{10} &= D_{50} - \text{erf}^{-1}(0.8)\sqrt{2}\sigma \\ D_{90} &= D_{50} + \text{erf}^{-1}(0.8)\sqrt{2}\sigma \\ W &\equiv D_{90} - D_{10} = 2\sqrt{2}\text{erf}^{-1}(0.8)\sigma\end{aligned}\quad (28)$$

Because of the symmetry of error function  $\text{erf}(z)$ , the normal distribution model (18) is symmetric around the median injury dose  $D_{50}$ :

$$\text{Symmetry: } f_{\text{normal}}(D_{50} + x) - \frac{1}{2} = \frac{1}{2} - f_{\text{normal}}(D_{50} - x) \quad (29)$$

We now study a skewed injury function that breaks this symmetry. Consider the situation where the target dose  $Z(x, \omega)$  has a log-normal distribution

$$Z(x, \omega) = x \cdot \exp(-\mu + \sigma\varepsilon)$$

Again  $\varepsilon \sim N(0,1)$  is a standard normal random variable. In this case,  $\ln(Z(x, \omega))$  and  $\ln(x)$  are simply related by an additive Gaussian noise.

$$\ln(Z(x, \omega)) = \ln(x) - \mu + \sigma\varepsilon \quad (30)$$

If we use  $\ln(x)$  and  $\ln(Z(x, \omega))$  to measure, respectively, the input dose and the target dose, then the injury probability vs  $\ln(x)$  follows the same function form as (12) with  $(x; z^{(c)})$  replaced by  $(\ln(x); \ln(z^{(c)}))$ :

$$p(x) = f_{\text{normal}}(\ln(x); \ln(z^{(c)})) = \frac{1}{2} + \frac{1}{2} \operatorname{erf}\left(\frac{\ln(x) - \ln(z^{(c)}) - \mu}{\sqrt{2}\sigma}\right) \quad (31)$$

We examine the injury probability as a function of the original input dose  $x$ . The purpose is to investigate 1) under what condition the injury probability vs  $x$  can be approximated by the symmetric normal-distribution model, and 2) when the normal distribution approximation is invalid, what additional parameter we need to introduce to describe the injury function for the original input dose  $x$ .

Since the injury probability vs  $\ln(x)$  follows the normal distribution model (12), we use results (28) for (12) to write out  $D_{50}^{(\ln)}$ ,  $D_{10}^{(\ln)}$  and  $D_{90}^{(\ln)}$  for quantity  $\ln(x)$ .

$$\begin{aligned} D_{50}^{(\ln)} &= \ln(z^{(c)}) + \mu \\ D_{10}^{(\ln)} &= D_{50}^{(\ln)} - \operatorname{erf}^{-1}(0.8)\sqrt{2}\sigma \\ D_{90}^{(\ln)} &= D_{50}^{(\ln)} + \operatorname{erf}^{-1}(0.8)\sqrt{2}\sigma \end{aligned} \quad (32)$$

The corresponding  $D_{50}$ ,  $D_{10}$  and  $D_{90}$  for quantity  $x$  are

$$\begin{aligned} D_{50} &= \exp(D_{50}^{(\ln)}) = z^{(c)} \cdot e^{\mu} \\ D_{10} &= \exp(D_{10}^{(\ln)}) = D_{50} \cdot \exp(-\operatorname{erf}^{-1}(0.8)\sqrt{2}\sigma) \\ D_{90} &= \exp(D_{90}^{(\ln)}) = D_{50} \cdot \exp(\operatorname{erf}^{-1}(0.8)\sqrt{2}\sigma) \end{aligned} \quad (33)$$

In this case, it is clear that  $(D_{90} - D_{50}) > (D_{50} - D_{10})$ . The injury probability vs quantity  $x$  is not exactly symmetric around  $D_{50}$ . We introduce a measure of skewness to represent the asymmetry of injury probability vs quantity  $x$ .

$$\gamma \equiv \ln\left(\frac{D_{90} - D_{50}}{D_{50} - D_{10}}\right) \quad (34)$$

Specifically,  $\gamma$  defined above measures the skewness of interval  $[D_{10}, D_{90}]$  around  $D_{50}$ .

- When  $\gamma = 0$ , interval  $[D_{10}, D_{90}]$  is symmetric around  $D_{50}$ .
- When  $\gamma > 0$ , we have  $(D_{90} - D_{50}) > (D_{50} - D_{10})$ , which implies that the

upper half (above  $D_{50}$ ) of injury function is flatter than the lower half (below  $D_{50}$ ).

- When  $\gamma < 0$ , we have  $(D_{90} - D_{50}) < (D_{50} - D_{10})$ , and that the upper half of injury function is steeper than the lower half.

Skewness  $\gamma$  is an indicator of how well the injury function for  $x$  can be approximated by the symmetric normal distribution model. For a target dose of log-normal distribution, the skewness is  $\gamma = \text{erf}^{-1}(0.8)\sqrt{2}\sigma$ . When  $\sigma$  is small, the skewness  $\gamma \approx 0$ , and the injury function is nearly symmetric around  $D_{50}$ . When  $\sigma^2 > 0$ , the skewness  $\gamma$  is positive, and in (31) the injury probability as a function of  $x$  is not symmetric. In this case, the injury function is characterized by three shape parameters:  $(D_{50}, W, \gamma)$ .

$$\begin{aligned} D_{50} &= z^{(c)} \cdot e^{\mu} \\ W &\equiv D_{90} - D_{10} = D_{50} \cdot 2 \sinh\left(\text{erf}^{-1}(0.8)\sqrt{2}\sigma\right) \\ \gamma &\equiv \ln\left(\frac{D_{90} - D_{50}}{D_{50} - D_{10}}\right) = \text{erf}^{-1}(0.8)\sqrt{2}\sigma \end{aligned} \quad (35)$$

Notice that even though expressions of  $(D_{50}, W, \gamma)$  in (35) contain three variables  $(\mu, \sigma, z^{(c)})$ , two variables  $\mu$  and  $z^{(c)}$  appear only as a combination  $(z^{(c)} \cdot e^{\mu})$  in  $D_{50}$ . Mathematically, the three shape parameters  $(D_{50}, W, \gamma)$  are completely specified by  $(D_{50}, \sigma)$ , and thus, have only two degrees of freedom. As a result, the three shape parameters  $(D_{50}, W, \gamma)$  cannot be set independently of each other. For example, in (35) when  $\gamma$  is small, the width  $W$  will be small unless the median dose  $D_{50}$  is large. Formulation (35), based on target dose of log-normal distribution (30), cannot accommodate any negative skewness ( $\gamma < 0$ ). It cannot even accommodate the simple symmetric case of  $\gamma = 0$  with finite  $W > 0$  and  $D_{50} < +\infty$ . We like to revise the formulation and construct an injury model in which the three shape parameters  $(D_{50}, W, \gamma)$  can be set independently of each other.

## 5. A Dose-Injury Model with Skewness Based on a Normally Distributed Intermediate Variable

We construct a model that accommodates the median injury dose ( $D_{50}$ ), the width ( $W$ ) and the skewness ( $\gamma$ ) as 3 independent parameters. In previous section, we studied the formulation based on target dose of log-normal distribution, in which the skewness is always positive and the 3 shape parameters  $(D_{50}, W, \gamma)$  are not independent of each other. A log-normal random variable can be viewed as the exponential of normal random variable. To accommodate negative skewness and to make  $(D_{50}, W, \gamma)$  independent of each other, we extend the formulation to the case of target dose being a more general function of normal random variable.

We consider the situation where the dose propagation uncertainty is an additive Gaussian noise in quantity  $\ln|x - x_0|$  with  $x_0$  as a new tunable parameter. The target dose  $Z(x, \omega)$  and the input dose  $x$  are related by



$$Z(x, \omega) - x_0 = (x - x_0) \exp(-\mu + \sigma\varepsilon)$$

In this setting,  $(Z(x, \omega) - x_0)$  has the same sign as  $(x - x_0)$ . The domain of  $x$  is divided by  $x_0$  into two regions:  $x > x_0$  and  $x < x_0$ . Only the region containing the critical threshold  $z^{(c)}$  will be relevant for the injury model. The other region of  $x$  produces target dose  $Z(x, \omega)$  always above or always below  $z^{(c)}$ . For example, when  $x_0 > z^{(c)}$ , only the region  $x < x_0$  is relevant for the injury model; the region  $x > x_0$  leads to target dose  $Z(x, \omega) > x_0 > z^{(c)}$  and thus, leads to an injury probability of 100%. We discuss separately the case of  $x_0 > z^{(c)}$  and the case of  $x_0 < z^{(c)}$ .

### 5.1. Case 1: $x_0 < z^{(c)}$

In this case, the region  $x < x_0$  yields target dose  $Z(x, \omega) < x_0 < z^{(c)}$  and an injury probability of 0%. We focus on the region  $x > x_0$ , the relevant region for the injury model. The logarithm of shifted target dose  $\ln(Z(x, \omega) - x_0)$  and logarithm of shifted input dose  $\ln(x - x_0)$  are related by an additive Gaussian noise.

$$\ln(Z(x, \omega) - x_0) = \ln(x - x_0) - \mu + \sigma\varepsilon \tag{36}$$

where  $\varepsilon \sim N(0,1)$ . We apply the shift  $x_{new} = x_{old} - x_0$  on all dose quantities (including  $D_{50}$  and  $z^{(c)}$ ). After the shift, problem (36) above is exactly the same as problem (30) in the previous section. It follows that the injury probability has the same function form as (12) with  $(x; z^{(c)})$  replaced by  $(\ln(x - x_0); \ln(z^{(c)} - x_0))$

$$p(x) = f_{\text{normal}}\left(\ln(x - x_0); \ln(z^{(c)} - x_0)\right) \text{ for } x_0 < z^{(c)}$$

$$= \frac{1}{2} + \frac{1}{2} \operatorname{erf}\left(\frac{\ln(x - x_0) - \ln(z^{(c)} - x_0) - \mu}{\sqrt{2}\sigma}\right) \tag{37}$$

Based on results (33) and (35), we write out  $(D_{50}, \gamma, W)$  for injury function (37).

$$D_{50} - x_0 = (z^{(c)} - x_0) \cdot e^{\mu}$$

$$D_{10} - x_0 = (D_{50} - x_0) \cdot \exp(-\operatorname{erf}^{-1}(0.8)\sqrt{2}\sigma)$$

$$D_{90} - x_0 = (D_{50} - x_0) \cdot \exp(+\operatorname{erf}^{-1}(0.8)\sqrt{2}\sigma) \tag{38}$$

$$W \equiv D_{90} - D_{10} = (D_{50} - x_0) \cdot 2 \sinh(\operatorname{erf}^{-1}(0.8)\sqrt{2}\sigma)$$

$$\gamma \equiv \ln\left(\frac{D_{90} - D_{50}}{D_{50} - D_{10}}\right) = \operatorname{erf}^{-1}(0.8)\sqrt{2}\sigma$$

Note that both  $z^{(c)}$  and  $D_{50}$  are on the right side of  $x_0$  in the case of  $x_0 < z^{(c)}$ . As we will see,  $D_{50}$  and  $z^{(c)}$  are always on the same side of  $x_0$ . With Formulas (38) for the case of  $x_0 < z^{(c)}$ , we can accommodate shape parameters  $(D_{50}, W, \gamma)$  with positive skewness  $\gamma > 0$ . Specifically, at any fixed  $\mu$ , for each given set of  $(D_{50}, W, \gamma > 0)$  there is a unique corresponding set of

$$\begin{aligned}
 & (x_0, \sigma, z^{(c)}). \\
 & \sigma = \frac{\gamma}{\text{erf}^{-1}(0.8)\sqrt{2}} \\
 & (D_{50} - x_0) = \frac{W}{2 \sinh(\text{erf}^{-1}(0.8)\sqrt{2}\sigma)}, \quad x_0 = D_{50} - (D_{50} - x_0). \quad (39) \\
 & z^{(c)} = x_0 + (D_{50} - x_0)e^{-\mu}
 \end{aligned}$$

This works for any positive skewness  $\gamma > 0$ , corresponding to the situation where the injury probability has a flatter rise above the median injury dose  $D_{50}$  than below it.

To accommodate negative skewness  $\gamma < 0$ , however, we need  $x_0 > z^{(c)}$ .

**5.2. Case 2:  $x_0 > z^{(c)}$**

In this case, we focus on the region  $x < x_0$  since the region  $x > x_0$  yields target dose  $Z(x, \omega) > x_0 > z^{(c)}$  and an injury probability of 100%. The target dose and input dose are related by

$$-\ln(x_0 - Z(x, \omega)) = -\ln(x_0 - x) + \mu + \sigma\varepsilon \quad (40)$$

where  $\varepsilon \sim N(0,1)$ . Here we consider quantity  $-\ln(x_0 - Z(x, \omega))$  with the negative sign because it is an increasing function of  $Z(x, \omega)$ . Injury occurs when the target dose is above the critical threshold:  $Z(x, \omega) > z^{(c)}$ , which translates to

$$\text{Injury probability} = \Pr\left(-\ln(x_0 - Z(x, \omega)) > -\ln(x_0 - z^{(c)})\right)$$

The injury probability has the expression

$$\begin{aligned}
 p(x) &= f_{\text{normal}}\left(-\ln(x_0 - x); -\ln(x_0 - z^{(c)})\right) \text{ for } x_0 > z^{(c)} \\
 &= \frac{1}{2} + \frac{1}{2} \text{erf}\left(\frac{-\ln(x_0 - x) + \ln(x_0 - z^{(c)}) + \mu}{\sqrt{2}\sigma}\right) \quad (41)
 \end{aligned}$$

Notice that (31) with quantities denoted by (') and (41) are connected by transformation

$$(x - x_0) = \frac{-1}{x'}, \quad (z^{(c)} - x_0) = \frac{-1}{z^{(c)'}} \quad \mu = -\mu'$$

We use results (33) and (35) for injury function (31) to write out  $(D_{50}, \gamma, W)$  for (41).

$$\begin{aligned}
 D_{50} - x_0 &= -(x_0 - z^{(c)}) \cdot e^{\mu} < 0 \\
 D_{10} - x_0 &= (D_{50} - x_0) \cdot \exp(\text{erf}^{-1}(0.8)\sqrt{2}\sigma) < 0 \\
 D_{90} - x_0 &= (D_{50} - x_0) \cdot \exp(-\text{erf}^{-1}(0.8)\sqrt{2}\sigma) < 0 \quad (42) \\
 W &\equiv D_{90} - D_{10} = -(D_{50} - x_0) \cdot 2 \sinh(\text{erf}^{-1}(0.8)\sqrt{2}\sigma) > 0 \\
 \gamma &\equiv \ln\left(\frac{D_{90} - D_{50}}{D_{50} - D_{10}}\right) = -\text{erf}^{-1}(0.8)\sqrt{2}\sigma < 0
 \end{aligned}$$

In the case of  $x_0 > z^{(c)}$ , both  $z^{(c)}$  and  $D_{50}$  are on the left side of  $x_0$ . With Formulas (42) for the case of  $x_0 > z^{(c)}$ , we can accommodate shape parameters  $(D_{50}, W, \gamma)$  with negative skewness  $\gamma < 0$ . Specifically, at any fixed  $\mu$ , for each given set of  $(D_{50}, W, \gamma < 0)$  there is a unique corresponding set of  $(x_0, \sigma, z^{(c)})$ .

$$\begin{aligned} \sigma &= \frac{(-\gamma)}{\operatorname{erf}^{-1}(0.8)\sqrt{2}} \\ (D_{50} - x_0) &= \frac{-W}{2 \sinh(\operatorname{erf}^{-1}(0.8)\sqrt{2}\sigma)}, \quad x_0 = D_{50} - (D_{50} - x_0). \quad (43) \\ z^{(c)} &= x_0 + (D_{50} - x_0)e^{-\mu} \end{aligned}$$

This works for  $\gamma < 0$ , which indicates that the injury probability has a steeper rise above the median injury dose  $D_{50}$  than below it.

Next we combine the results of  $x_0 < z^{(c)}$  and  $x_0 > z^{(c)}$  to derive a unified formulation for accommodating shape parameters  $(D_{50}, W, \gamma)$  regardless of the sign of  $\gamma$ .

### 5.3. A Unified Formulation for All Values of Skewness

In the previous sub-section, we studied models based on target dose of shifted log normal distribution with shift as a parameter. We now synthesize the results obtained to develop a unified formulation of injury function in which the 3 shape parameters  $(D_{50}, W, \gamma)$  can be specified independently.

First, we show that at any fixed value of  $\mu$ , there is one-to-one correspondence between  $(x_0, \sigma, z^{(c)})$  and  $(D_{50}, W, \gamma)$ . For any given set of shape parameters  $(D_{50}, W, \gamma)$  regardless of the sign of  $\gamma$ , we combine results (39) and (43) to write out the corresponding  $(x_0, \sigma, z^{(c)})$ .

$$\begin{aligned} \sigma &= \frac{|\gamma|}{\operatorname{erf}^{-1}(0.8)\sqrt{2}} \\ x_0 &= D_{50} - \frac{W}{2 \sinh(\gamma)} \\ z^{(c)} &= x_0 + (D_{50} - x_0)e^{-\mu} \end{aligned} \quad (44)$$

Conversely, for any given set of  $(x_0, \sigma, z^{(c)})$ , we combine results (38) and (42) to write out the corresponding shape parameters  $(D_{50}, W, \gamma)$ .

$$\begin{aligned} D_{50} &= x_0 + (z^{(c)} - x_0) \cdot e^{\mu} \\ W &= |D_{50} - x_0| \cdot 2 \sinh(\operatorname{erf}^{-1}(0.8)\sqrt{2}\sigma) \\ \gamma &= \operatorname{sign}(z^{(c)} - x_0) \cdot \operatorname{erf}^{-1}(0.8)\sqrt{2}\sigma \end{aligned} \quad (45)$$

Again,  $D_{50}$  and  $z^{(c)}$  are always on the same side of  $x_0$ . Next we combine (37) for  $x_0 < z^{(c)}$  and (41) for  $x_0 > z^{(c)}$  to write out a unified injury probability vs  $x$ .

$$p(x) = \frac{1}{2} + \frac{1}{2} \operatorname{erf} \left( \frac{\operatorname{sign}(z^{(c)} - x_0)}{\sqrt{2}\sigma} \ln \left( \frac{x - x_0}{(z^{(c)} - x_0)e^{\mu}} \right) \right) \quad (46)$$

To specify the unified injury function in terms of shape parameters  $(D_{50}, W, \gamma)$ , we express all quantities in (46) using only  $(D_{50}, W, \gamma)$  and  $x$ .

$$\frac{\text{sign}(z^{(c)} - x_0)}{\sqrt{2}\sigma} = \frac{\text{erf}^{-1}(0.8)}{\gamma}$$

$$(z^{(c)} - x_0)e^{\mu} = (D_{50} - x_0) = \frac{W}{2\sinh(\gamma)}$$

$$x - x_0 = (x - D_{50}) + (D_{50} - x_0) = (x - D_{50}) + \frac{W}{2\sinh(\gamma)}$$

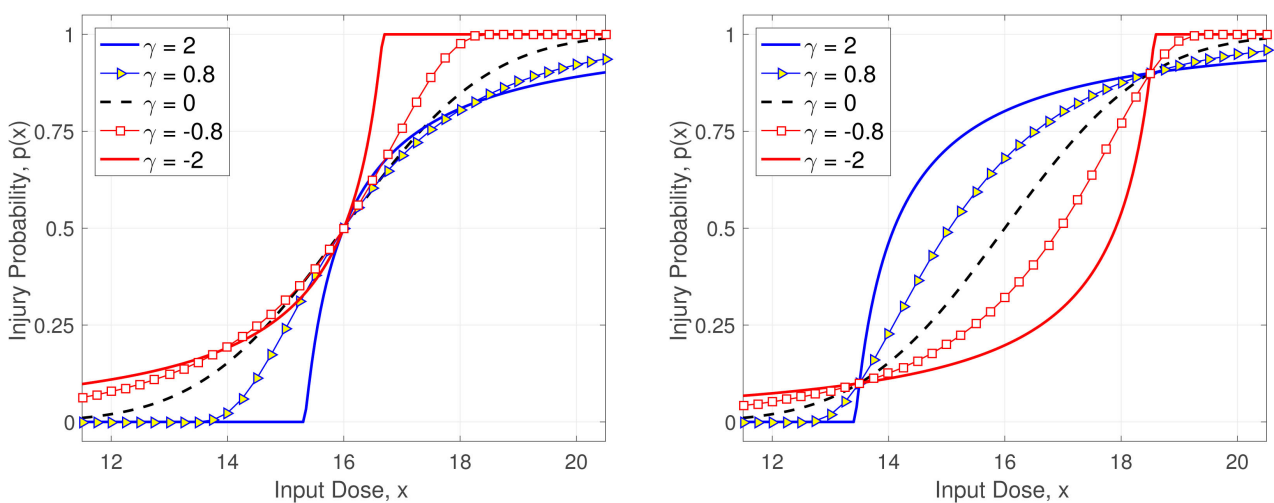
With these expressions, we write the unified injury function as

$$p(x) = \frac{1}{2} + \frac{1}{2} \text{erf}\left(\frac{\text{erf}^{-1}(0.8)}{\gamma} \ln\left(2\sinh(\gamma)\frac{(x - D_{50})}{W} + 1\right)\right) \equiv G(x; D_{50}, W, \gamma) \quad (47)$$

In injury model (47), the 3 shape parameters  $(D_{50}, W, \gamma)$  can be specified independently of each other. In particular, for small skewness  $\gamma \ll 1$ , expanding (47) in terms of  $\gamma$  reduces it to the symmetric normal-distribution model (18)

$$G(x; D_{50}, W, \gamma) = \frac{1}{2} + \frac{1}{2} \text{erf}\left(\frac{2\text{erf}^{-1}(0.8) \cdot (x - D_{50})}{W} + O(\gamma)\right)$$

**Figure 3** illustrates several injury functions of the form (47), respectively, for positive, zero and negative skewness. All injury functions shown have the same width  $W = 5$ . In the left panel of **Figure 3**, injury functions are aligned at  $D_{50} = 16$ . This alignment demonstrates that for  $\gamma > 0$  the left half of injury function is steeper than the right half; for  $\gamma < 0$  the left half of injury function is flatter than the right half; and for  $\gamma = 0$  the injury function is symmetric. In the right panel, injury functions are shifted to be aligned at  $D_{10} = 13.5$  and thus also aligned at  $D_{90} = D_{10} + W = 18.5$  because they all have the same width



**Figure 3.** Injury functions with positive, zero and negative values of skewness. All injury functions have the same width  $W = 5$ . Left panel: injury functions are aligned at the median injury dose  $D_{50} = 16$ . Right panel: injury functions are shifted to have the same interval  $[D_{10}, D_{90}] = [13.5, 18.5]$ .

$W = 5$ . With  $D_{10} = 13.5$  fixed, the median dose  $D_{50}$  varies with skewness  $\gamma$  from  $D_{50} = 14.1$  at  $\gamma = 2$ , to  $D_{50} = 16$  at  $\gamma = 0$ , and to  $D_{50} = 17.9$  at  $\gamma = -2$ . The alignment of interval  $[D_{10}, D_{90}]$  highlights that as  $\gamma$  increases from negative to zero to positive, the injury function becomes more concave down.

### 6. Effect of Input Dose Uncertainty on the Injury Function with Skewness

We study the effect of input dose estimation uncertainty on the dose-injury function with skewness. We use the term “composite injury function” to denote the injury model after the input dose uncertainty has been incorporated into the model. In general, the composite injury function will be somewhat different from the 3-parameter function form (47) we derived in the previous section. We calculate the three shape parameters  $(D_{50}, W, \gamma)$  of the composite injury function. Then we explore approximating the composite injury function using function form (47). We examine the difference between the composite injury function and model (47) with the same shape parameters  $(D_{50}, W, \gamma)$ . If the approximation error is small, then the 3-parameter function form (47) is approximately invariant with respect to input dose uncertainty, and it serves as an adequate framework for accommodating uncertainty in estimating the input dose. Furthermore, framework (47) provides a mechanism of mapping the injury function for one particular dose propagation uncertainty to that for a different uncertainty. Using this mechanism, we can construct an injury model for a target population in application, based on measured injury data for a test population in experiments.

We start with a function of injury probability vs true input dose that is exactly of form (47) specified by 3 shape parameters  $(D_{50}, W, \gamma)$ :

$$p_0(x^{(true)}) = G(x^{(true)}; D_{50}, W, \gamma)$$

We consider the situation where the true input dose  $x^{(true)}$  is not measurable. Instead, an estimated input dose,  $x$ , is obtained as an approximation for  $x^{(true)}$ . We assume

- the difference  $(x^{(true)} - x)$  is a normal random variable, and
- the difference  $(x^{(true)} - x)$  is independent of  $x$ .

We assess the injury probability as a function of the *estimated input dose*  $x$ . For each fixed value of  $x$ , the corresponding  $x^{(true)}$  is a normal random variable:  $x^{(true)} = x + \mu + \sigma\varepsilon$  where  $\varepsilon \sim N(0, 1)$ . The composite injury function,  $p_\sigma(x)$ , representing the injury probability at estimated input dose  $x$ , is a Gaussian weighted average of  $p_0(x^{(true)})$ :

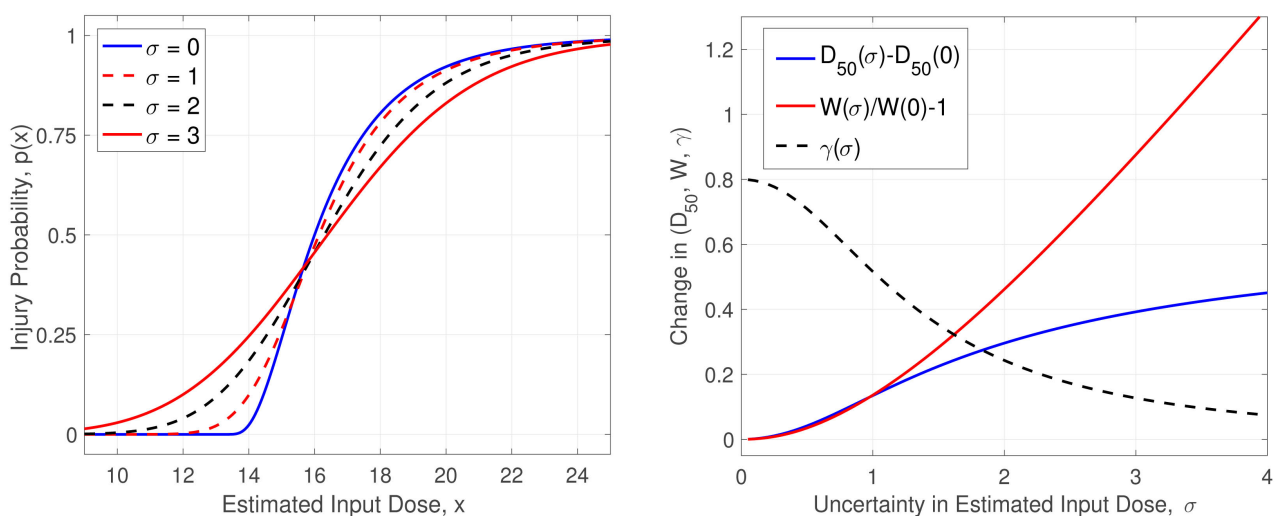
$$\begin{aligned}
 p_\sigma(x) &= E(G(x + \mu + \sigma\varepsilon; D_{50}, W, \gamma)) \\
 &= \int_{-\infty}^{\infty} G(x + \mu + s; D_{50}, W, \gamma) \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(\frac{-s^2}{2\sigma^2}\right) ds
 \end{aligned}
 \tag{48}$$

When injury function  $G(\cdot)$  has non-zero skewness, the Gaussian weighted

average of  $G(\cdot)$  on the right hand side of (48) does not have a simple analytical expression. We use numerical integration to calculate the composite injury function  $p_\sigma(x)$  and calculate its shape parameters  $(D_{50}(\sigma), W(\sigma), \gamma(\sigma))$ . We examine numerically if  $p_\sigma(x)$  is still well described by function form (47) with  $p_\sigma(x)$ 's shape parameters  $(D_{50}(\sigma), W(\sigma), \gamma(\sigma))$ .

In our numerical study,  $p_0(x^{true})$ , the injury probability vs the true input dose before input dose uncertainty is incorporated, has function form (47) and is specified by shape parameters  $D_{50} = 16$ ,  $W = 5$ , and  $\gamma = 0.8$ . We consider input dose uncertainty of normal distribution with  $\mu = 0$  (mean) and various values of  $\sigma$  (standard deviation). The composite injury function,  $p_\sigma(x)$ , contains the effect of input dose uncertainty, showing injury probability vs estimated input dose  $x$ . **Figure 4** examines the composite injury function  $p_\sigma(x)$  for  $\sigma$  between 0 and 3.

The left panel of **Figure 4** shows the injury probability vs the estimated input dose  $x$ , respectively, for  $\sigma = 0, 1, 2, 3$ . The most pronounced effect of input dose uncertainty is to spread out the injury function and increase the width. We examine the trend of shape parameters  $(D_{50}(\sigma), W(\sigma), \gamma(\sigma))$  when the input dose uncertainty  $\sigma$  is added and increased. The right panel shows  $(D_{50}(\sigma), W(\sigma), \gamma(\sigma))$  vs  $\sigma$ , of the composite injury function. As the input dose uncertainty  $\sigma$  increases, both the median injury dose  $D_{50}(\sigma)$  and the width  $W(\sigma)$  increase monotonically, with  $W$  increasing more prominently than  $D_{50}$ . At the same time, when  $\sigma$  increases, the asymmetry of injury function is smoothed out by the Gaussian noise and as a result, the skewness  $\gamma(\sigma)$  decreases. The change in median injury dose  $D_{50}(\sigma)$  is attributed to the presence of skewness: the median injury dose increases (moves toward the right) when an injury function with positive skewness is smoothed out by a Gaussian noise. Conversely, the median injury dose decreases (moves toward the

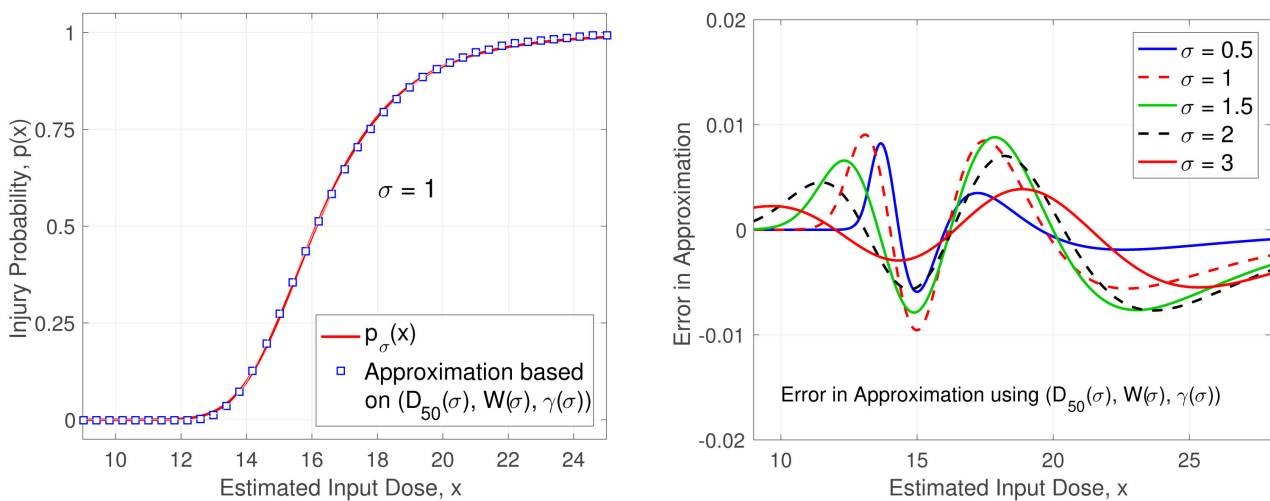


**Figure 4.** Effect of the input dose uncertainty  $\sigma$  on the injury function with skewness. Left panel: composite injury functions for several values of  $\sigma$ . Right panel: shape parameters  $(D_{50}(\sigma), W(\sigma), \gamma(\sigma))$  vs  $\sigma$  of the composite injury function.

left) when an injury function with negative skewness is smoothed out. The movement of the median injury dose is caused by smoothing an asymmetric function (see **Figure 3** for the general shape of injury functions with positive, zero, and negative skewness). For an injury function of zero skewness,  $D_{50}(\sigma)$  is invariant with respect to  $\sigma$  when the injury function is smoothed out by a Gaussian noise.

Next we examine whether or not the composite injury functions for  $\sigma^2 > 0$  shown in **Figure 4** are still approximately described by model (47). **Figure 5** compares the composite injury function and the approximation using function form (47) with shape parameters  $(D_{50}(\sigma), W(\sigma), \gamma(\sigma))$  of the composite injury function. The left panel of **Figure 5** compares the composite injury function for  $\sigma = 1$  and its approximation. The two functions are barely distinguishable from each other. To quantitatively examine the error of approximation, in the right panel we plot the difference between the composite injury function and its approximation. For all values of  $\sigma$  examined, the maximum error in approximation is less than 0.01 (1%). The results demonstrate that function form (47) specified by 3 independent shape parameters  $(D_{50}, W, \gamma)$  is an adequate model for quantitatively describing general injury functions with skewness.

With the framework of function form (47) and mapping transformation (48), we can filter out the effect of input dose uncertainty in measured injury data. Suppose we are given a measured injury function,  $p_{\sigma_1}(x)$ , of form (47) for a particular population with input dose uncertainty  $\sigma_1$ . We use transformation (48) to map it back to  $p_0(x^{(true)})$ , the injury function for the case of zero input dose uncertainty ( $\sigma = 0$ ). From there, we can apply the mapping transformation again to predict the injury model for another population with input dose uncertainty  $\sigma_2$ . There is no simple analytical expression for the mapping



**Figure 5.** Approximation of the composite injury function  $p_\sigma(x)$  using function form (47) with shape parameters  $(D_{50}(\sigma), W(\sigma), \gamma(\sigma))$ . Left panel: comparison of  $p_\sigma(x)$  and its approximation for  $\sigma = 1$ . Right panel: error of the approximation for several values of input dose uncertainty  $\sigma$ .

transformation. Both the forward and backward mappings need to be implemented numerically. The detailed numerical procedure will be discussed in a subsequent study.

## 7. Concluding Remarks

We considered injury models in the framework of dose propagation uncertainty. The mathematical formulation is based on that the binary injury outcome is completely determined by the target dose at the active site and the critical threshold. The randomness in the occurrence of injury at a given input dose is attributed to the dose propagation uncertainty from input dose to target dose. The normal distribution model describes the situation where the dose propagation uncertainty is normally distributed. We interpreted the widely used logistic model as a good approximation to the normal distribution model, and thus, interpreted it approximately as a consequence of normally distributed dose propagation uncertainty. In many applications, the input dose is not directly measurable. Instead, an estimated input dose is calculated via computer simulations from measured quantities using representative median parameter values of the general population. In many practical situations, injury models are constructed in the form of injury probability vs estimated input dose. The discrepancy between the estimated input dose and the true input dose can be viewed as an uncertainty in the input dose. With the interpretation of dose propagation uncertainty, the input dose uncertainty is conveniently incorporated into the injury model. The framework of dose propagation uncertainty provides a mechanism of extending an injury function established on a test population to predict the injury model for a different population in application. Both the logistic model and the normal distribution model are specified by two shape parameters: the median injury dose and the 10 - 90 percentile width. The mapping between the injury functions of two populations has a simple analytical form of updating the two shape parameters. Both the logistic model and the normal distribution model are symmetric around the median injury dose and have no skewness. To accommodate injury functions with skewness, we studied dose propagation uncertainties of shifted log normal distribution with shift as a parameter. Based on the shifted log normal model, we developed a function form for injury probability vs input dose that is specified by three shape parameters: median injury dose, the width, and the skewness. The proposed function form allows the three shape parameters to be set independent of each other. In particular, the proposed function form is capable of accommodating arbitrary skewness, positive or negative. In addition, we showed numerically that the proposed 3-parameter function form is approximately invariant with respect to additions or changes in input dose uncertainty. Therefore, the 3-parameter function form serves as a broad framework for modeling input dose uncertainty and modeling injury function skewness at the same time. This broad framework allows us to map injury function with



skewness from a test population to a different population in applications.

## Disclaimer and Acknowledgements

The authors thank C. Kramer and J. Swallow of Institute for Defense Analysis (IDA) for bringing the problem to their attention, and thank the Joint Non-Lethal Weapons Directorate of U.S. Department of Defense for supporting this work. The views expressed in this document are those of the authors and do not reflect the official policy or position of the Department of Defense or the U.S. Government.

## Conflicts of Interest

The authors declare no conflicts of interest regarding the publication of this paper.

## References

- [1] Vorst, M.V., Stuhmiller, J., Ho, K., Yoganandan, N. and Pintar, F. (2003) Statistically and Biomechanically Based Criterion for Impact-Induced Skull Fracture. *Annual Proceedings/Association for the Advancement of Automotive Medicine*, **47**, 363-381.
- [2] Cazares, S.M., Finnin, M.S., Holzer, J.R., King, A.L. and Kramer, C.M. (2017) Significance of Rib Fractures Potentially Caused by Blunt Impact Non Lethal Weapons. Institute for Defense Analyses (IDA), Alexandria, VA.
- [3] Chan, P., Ho, K. and Ryan, A.F. (2016) Impulse Noise Injury Model. *Military Medicine*, **181**, 59-69. <https://doi.org/10.7205/MILMED-D-15-00139>
- [4] Murphy, W.J., Khan, A. and Shaw, P.B. (2011) Analysis of Chinchilla Temporary and Permanent Threshold Shifts Following Impulsive Noise Exposure. Centers for Disease Control and Prevention, Cincinnati, OH.
- [5] Hampton, C.E. and Kleinberger, M. (2018) Material Models for the Human Torso Finite Element Model. US Army Research Laboratory (ARL), Aberdeen Proving Ground, MD.
- [6] Shen, W., Niu, E., Webber, C., Huang, J. and Bykanova, L. (2012) ATBM Analyst's Guide for Model Verification and Validation. L-3 Applied Technologies, San Diego, CA.
- [7] Kramer, C. and Swallow, J. (2018) Error Propagation in RSI Estimates of Blunt Impact NLWs Using the Advanced Total Body Model. Institute for Defense Analyses (IDA), Alexandria, VA.
- [8] Cox, D.R. (1958) The Regression Analysis of Binary Sequences. *Journal of the Royal Statistical Society. Series B*, **20**, 215-242.
- [9] Wang, H., Burgei, W.A. and Zhou, H. (2017) Interpreting Dose-Response Relation for Exposure to Multiple Sound Impulses in the Framework of Immunity. *Health*, **9**, 1817-1842. <https://doi.org/10.4236/health.2017.913132>
- [10] Sturdivan, L.M., Viano, D.C. and Champion, H.R. (2004) Analysis of Injury Criteria to Assess Chest and Abdominal Injury Risks in Blunt and Ballistic Impacts. *The Journal of Trauma: Injury, Infection, and Critical Care*, **56**, 651-663. <https://doi.org/10.1097/01.TA.0000074108.36517.D4>
- [11] Bliss, C.I. (1934) The Method of Probits. *Science*, **79**, 38-39.

<https://doi.org/10.1126/science.79.2037.38>

- [12] Duma, S.M., Schreiber, P.H., McMaster, J.D., Crandall, J.R. and Bass, C.R. (2002) Fracture Tolerance of the Male Forearm: The Effect of Pronation versus Supination. *Proceedings of the Institution of Mechanical Engineers, Part D: Journal of Automobile Engineering*, **216**, 649-654.