Implications of using ED-50 and probit analysis in comparing retinal injury threshold data

David H. Sliney, John Mellerio and Karl Schulmeister

Please register to receive our Laser, LED & Lamp Safety NEWSLETTER
(about 4 times a year) with information on new downloads:
http://laser-led-lamp-safety.seibersdorf-laboratories.at/newsletter

This paper can be downloaded from http://laser-led-lamp-safety.seibersdorf-laboratories.at


Copyright 2001 Society of Photo-Optical Instrumentation Engineers. One print or electronic copy may be made for personal use only. Systematic electronic or print reproduction and distribution, duplication of any material in this paper for a fee or for commercial purposes, or modification of the content of the paper are prohibited.
Laser and Noncoherent Light Ocular Effects: Epidemiology, Prevention, and Treatment

Ed. Bruce E. Stuck
Michael Belkin
Chairs/Editors

22 January 2001
San Jose, USA

Sponsored by
U.S. Air Force Office of Scientific Research
SPIE – The International Society for Optical Engineering

Proceedings of SPIE
Volume 4246
Implications of using ED-50 and probit analysis in comparing retinal injury threshold data

David H. Slineya*, John Melleriob, and Karl Schulmeisterc

aU.S. Army Center for Health Promotion and Preventive Medicine, Aberdeen Proving Ground, MD 21010-5403, USA
bUniversity of Westminster, School of Biosciences, W1M 8JS London, U.K.
cAustrian Research Centers Seibersdorf, A-2444 Seibersdorf, Austria

ABSTRACT

An indication of the level of uncertainty in laser injury studies relates to the slope of the transformed dose-response curve, or the "probit plot" of the data. The most cited threshold in a laser injury experiment is the point on the probit plot that represents a 50 % probability of injury: the ED-50. This value is frequently referred to as the "threshold," even though some experimental damage points exist below this "threshold." An analysis of any number of example data sets reveals that the slope in most experiments could not be explained by biological variation alone. The optical, thermophysical and biological factors influencing the probit plot are critically analyzed. By theoretically modelling an experiment, small errors in focus are shown to produce a substantial change in the ED-50 and the slope of the probit plot. The implications of plotting spot-size dependence with ED-50 values are shown to be significant, and can lead to erroneous conclusions regarding the apparent spot-size dependence.

1. INTRODUCTION

1.1 The Dose Response Curve

In studies of laser-induced injury of biological tissue, a dose response curve is frequently prepared to present the result. Any statistical analysis of uncertainties in retinal threshold studies has traditionally centered on an examination of the slope of the dose response curve expressed as a probit plot of experimental data. It frequently had been thought in some circles that the most important reference point on the probit plot is the exposure that represents a 50 % probability of injury: the ED-50. Indeed, this value is frequently referred to as the "threshold," even though some experimental damage points exist below this "threshold." The steepness of the curve is not only related to the type of damage mechanism and variation among individual animals, but also indicates problems in conducting the experiment. The techniques of probit analysis come from toxicology, and certain inherent assumptions relating to the expected distribution of toxicological threshold data, where each exposure is presumed to be independent, have been carried over to laser safety studies, however the assumptions underlying this technique require some discussion. We define the slope as the ratio of ED-84/ED-50.

The thresholds for injury are normally obtained by exposing laboratory animals under controlled conditions to simulate "worst-case" human exposure conditions and threshold is quoted as a single ED-50 value. Mush21 and Wolbarsht and Sliney29 argue that this approach is unsatisfactory and that far too much emphasis is placed on the ED-50 value, rather than also reporting other points on the curve. Nevertheless, quoting the ED-50 is the accepted convention. The value of the ED-50 will be influenced by the choice of experimental endpoint and the delay until examination after exposure. Although the biological change is most frequently determined by direct observation, more sensitive endpoints based upon microscopy or chemical analytical techniques are also employed. In the determination of thresholds of laser-induced injury, direct observation by ophthalmic instruments is most frequently used. For retinal studies, the experimentalist, to observe the retinal location where the laser exposure takes place, uses the ophthalmoscope or slit-lamp microscope. In some cases the visibility of the laser-induced lesion can be improved by a technique known as fluorescein angiography, and some workers have even used light and electron microscopy to examine the exposed tissue. These measures claim to have greater sensitivity, but are more costly and less practical. Although fluorescein angiography is frequently used, the detailed histological studies with microscopic examination of thin slices of exposed tissue have only been carried out for specified laser wavelengths and exposure durations to quantify the reduction in the ophthalmoscopically determined ED-50 by the more elaborate techniques. Normally this reduction factor is approximately 2, and the committees deriving the total "safety factor" required to arrive at an MPE take this factor into account.
Figure 1 shows an example of an early, hand-drawn retinal damage probability curve where the actual bars indicate the absence or presence of damage are recorded above and below the curve. It also shows the lowering of the threshold values when histological criteria for damage are applied. The slope “S” of the dose response curve reflects not only natural biological variation, which may have a lognormal distribution, but also the impact of experimental errors, which may show a different statistical distribution. The class of damage mechanism will also alter the steepness of the curve. When the slope is shallower, this indicates an increased standard deviation, which in turn may be due to increased experimental uncertainties. This illustrates that any derivation of human exposure limits for laser-induced injury requires one to estimate the true biological variation and separate this from the added experimental errors that reduce the steepness of the slope in the probit plot.

![Figure 1. Early, hand-drawn dose-response curve.](image)

We analysed a collection of experimental data reported in a variety of studies of corneal and retinal injury from pulsed laser radiation\(^\text{26}\). The analysis indicates that thermal and thermoacoustic damage mechanisms apparently have an intrinsic slope of approximately 1.05 to 1.15\(^\text{11,18,25}\). Often, however, experimental threshold data have shallower curves with slopes "S" that are numerically much greater, e.g., 1.5 - 1.7, and even up to 2.8\(^\text{19,27,28}\). This is really not surprising because the enormous difficulties of seeing a minimally visible lesion and focusing the laser beam to produce the nearly diffraction-limited image leads to this greater spread of data and shallower slopes. Thus, a probit plot applied during the derivation of exposure limits should have a slope adjusted to be steeper to account for these experimental contributions to the spread of data points.

The slope of the probit plot has generally been considered a good indication of the overall uncertainty and quality of the experimental data, but there has been some disagreement as to whether this slope could be considered an index of risk to a human population.

1.2 Retinal Injury Mechanisms

In an ideal experiment where the experimental errors introduced by the beam propagation through the ocular media were not present, the spread of retinal threshold data would be much smaller than that obtained by current experimental methods. The type of damage mechanism will affect the absolute spread of threshold data. Based upon the current understanding of the retinal injury mechanisms by short-pulsed lasers, thermal photocoagulation dominates for pulse durations of 10 ns to many milliseconds. Injury thresholds appear to vary little from 10 ns to about 18 – 50 µs. At longer durations, the thresholds increase because heat flow occurs during the exposure. At sub-nanosecond exposure durations, other, non-linear damage mechanisms come into play, such as self-focussing and laser-induced breakdown. The thermal and thermo-mechanical mechanisms have sharply defined thresholds in experiments where tissue is directly exposed without confounding factors, as
in CO$_2$ laser-induced corneal injury or in vitro studies of laser induced injury to RPE cell cultures. More stochastic effects appear in the sub-nanosecond regime and the dose-response curves become shallower.

2. METHODS AND ANALYSIS OF EXPERIMENTAL FACTORS
The calculated value of the ED-50 and the spread of data points can change by varying the experimental exposure or evaluation techniques. We analyzed the major sources of uncertainties in experimentally determined thresholds and dose-response curves for retinal thermal injury. There are a number of factors that influence the spread of data in laboratory experiments and each was evaluated in terms of the impact upon the reported ED-50 values and probit slopes. To explore these factors, we undertook a review of the extensive database from experimental animal threshold ocular injury studies. The methods of this analysis focused on the variation in the slopes as well as the ED-50 values in the dose-response plots, where available. All of the sources of experimental error - other than false-positive data - will increase the ED-50 value and the value of the slope. The published ED-50 data with slopes showed slopes $S$ ranging from about 1.04 to 2.5. As the slope values increase, corresponding to an increased spread of data, one implication is that experimental difficulties have also increased. For example, one early series of experiments to determine ED-50 values for a 30-ns q-switched ruby laser (694.3 nm) had particularly large slopes of 2.2 for the smallest spot sizes, and dropped to 1.3 for large (0.89 mm) retinal image sizes as shown in Figure 2. From a biophysical standpoint, one would not expect the dose response curve (and therefore probit slopes) to vary so much for different retinal image sizes. It is well known that ruby laser beams did not have clean Gaussian profiles and therefore, a consistent, minimal image diameter was difficult to achieve, and the spread of data were particularly great for these small spot sizes.

![Figure 2](image-url)

Figure 2. The variation of reported slope of the retinal threshold dose-response curve for one study.

2.1 Time of Examination
One of the most critical elements of the experimental protocol is the delay between laser exposure and time of examination, because biological changes undergo a time course following injury. Typical delays for examination range from 5 minutes to 48 hours, with 1 hour and 24 hours being the most frequently used. Depending upon damage mechanism (i.e., thermal, thermo-mechanical, photochemical, etc.), the optimum period for determining the ED-50 changes as the time-course varies for the biological response and amplification of the initial biophysical tissue insult. For thermal injury this appears to have a very small impact of 1.2 or less, and this was not studied further.

2.2 Animal Model
Another factor that affects the ED-50 value is the choice of experimental animal. One of the most frequently used animals in ophthalmic research is the rabbit, and this is used most often for corneal injury studies, and was used in many early retinal studies. However, the optical quality of the rabbit eye is poor, leading to distorted retinal images and inaccurate threshold determinations. There have been studies employing ophthalmic contact lenses and careful optical alignment of the anaesthetised rabbit to eliminate most corneal aberrations in the eye. These conditions produce nearly diffraction-limited retinal images leading to lower threshold values. However, There remains some uncertainty as to how close the contact-lens-rabbit model is to the awake, task-oriented human eye in everyday environments. Hence, the most favored animal model has
become the rhesus monkey for retinal studies for most investigators, and this has greatly increased the cost of the experiment, leading to fewer threshold determinations.

2.3 Retinal Pigmentation
Still another factor, the retinal pigmentation, will determine the fraction of incident energy absorbed at each wavelength and thereby affect the resulting ED-50. Pigmentation in the retinal pigment epithelium (RPE) and choroid varies with the individual subject, the species and the retinal location. Although the color of the rabbit retina appears to more closely resemble that of the human retina\textsuperscript{13}, it has been argued that since the RPE melanin pigment density is greater in the rhesus monkey compared to that of the human (and the retinal structure more similar), the rhesus monkey is a preferable experimental model for safety studies\textsuperscript{20, 25}. Gabel et al\textsuperscript{11, 12} showed that the RPE pigmentation had a limited variation of about two-fold across a human population although choroidal pigmentation varied. The macular pigmentation was the most uniform and consistent. To estimate the impact of localized pigment mottling, we scanned micrographs of RPE cells (Figure 3) to determine the pigmentation densities. The variation in relative absorption was approximately 1.3 for a 25 µm spot, 1.2 for a 40 µm spot and 1.15 for a 75-µm diameter spot. While this would impact the spread of data, it would not have a very great impact upon the slope of the dose-response curve.

![Figure 3. Photomicrograph of a flat-preparation of retinal pigment epithelium showing pigment mottling\textsuperscript{11}](image)

2.4 Refractive State
Also of concern is the refractive state of the experimental subject as this determines whether the minimal retinal image size of a collimated laser beam is achieved; this is critical\textsuperscript{10}. A one-diopter error in the refractive state can result in an increase in the achieved retinal blur circle from approximately 25 µm to about 120 µm -- a change in image area of 22 times. Investigators therefore normally attempt to correct the animal's state of refraction in the visible, but this does not take into account the chromatic aberrations -- also of the order of one diopter across the visible spectrum. Investigators are able to correct the refraction to within 0.25 D, which leads to an uncertainty in the actually achieved minimal retinal image diameter of approximately 20 µm to 50 µm. Furthermore, during an investigative session, the animal's refractive state will vary slightly from exposure to exposure, the degree depending upon the level of anaesthesia; this increases the spread of data.

2.5 Eye Movements
Eye movements during the experimental exposure will also affect the outcome. These eye-movements will only be of concern for exposures lasting more than 10 ms, particularly in anaesthetised animals\textsuperscript{22, 16}. Ness et al\textsuperscript{22} showed that laser energy was concentrated in a much smaller retinal area in the anaesthetized animal.

2.6 Retinal Location
The actual site of retinal exposure will influence the ED-50. This occurs for two reasons. First the optical quality of the eye for off-axis beams is somewhat less than for axial exposures to the macula. Second, the pigmentation and thickness of the neural retina vary from the center of the fovea, through the macula to the paramacular regions. As a general rule, the ED-50 values will be less in the fovea than parafovea or paramacular region. This variation is of the order of two-fold\textsuperscript{4, 17, 23}.

2.7 Intra-ocular Scatter
The state of the ocular media (cornea, aqueous, lens and vitreous) influences the distribution of radiant energy at the retina. Intraocular scatter diffuses the beam, and depending upon the nature and size of the scattering centers in the ocular media, small-angle forward scatter and diffuse scatter will vary. The age of the eye, the care taken by the experimentalist to preserve
corneal clarity by frequent irrigation of the cornea (or use of contact lens), and the potential for subtle lenticular opacities (early stages of cataractogenesis), all play a role in producing scatter.

2.8 Observer Skills and Lesion Detection
The ability of the experimental investigator observing the retina to detect a just-perceptible minimum visible lesion (MVL) also influences the ED-50 value. As investigators have more experience, the reported ED-50 values generally decrease. The visual contrast of the threshold injury is usually low, making the task of correct lesion detection difficult. Furthermore, the chance of scoring the presence of a very small lesion when an injury is actually not present (i.e., false positive) will increase if the retina appears mottled or otherwise exhibits abnormalities that could be mistaken for a laser induced lesion. It may also be difficult to discern the edges of some very large lesions making positive identification difficult. Nevertheless, our review of the published threshold studies that provide dose response slopes show a general trend for a steeper slope (smaller numerical value) for larger images, suggesting that small-lesion experiments are more difficult to perform as shown in Figure 2.

2.9 Beam Quality
The optical quality of the incident laser beam influences the minimal retinal image size. In earlier studies, multi-mode lasers were frequently used, leading to a larger retinal image area and increased values of the ED-50. Today, most lasers used in these experiments have been carefully aligned to achieve a single transverse mode. The presence of apertures in the beam path can also alter the quality of the retinal image.

3. RESULTS
It is evident that aside from false positive identification of a lesion, all of the other aforementioned sources of error will tend to increase the ED-50 values and the spread of the data from which they are derived. This will be discussed in greater detail in the following discussion of the statistical treatment of threshold data.

3.1 Method of Determining a Risk of Injury in Awake Humans
Recent studies typically report slopes of 1.1 to 1.4. Indeed, the choice of the "safety factor" used by committees for derivation of MPEs recognized that experimental difficulties led to a skewed probability distribution and ED-50 values that are too high. The choice of large "safety factor" values, as great as 10 to 20, resulted from this recognition of experimental uncertainties. As experimental techniques improved (often indicated by steeper probit slopes), safety factor values less than 10 have sometimes been applied. It is interesting to note that in deriving exposure limits for skin and cornea, smaller safety factors have been possible because of the reduced uncertainties in experimental determinations of exposure parameters and damage.

An analysis of a number of example data sets reveals that the slope in some experiments could not be explained by biological variation alone. For example, if the reported slope is not very steep due to experimental uncertainties, the laser safety community might assume that a larger-than-necessary safety factor would be required. The consequence of directly applying experimental animal dose-response curves having more shallow slopes (e.g., 2 to 2.5) to estimate risk of injury in humans is that the plot could predict a finite probability for an injury occurring at a dose of as little as 10 % of the ED-50. Yet, from fundamental biophysical principles, this result could be shown clearly to be flawed. If the ED-50 energy corresponds to a retinal temperature elevation of 15°, an energy of 10 % of the ED-50 would correspond to 1.5° (10 % of the ED-50 temperature elevation), which could not produce photoagulation. This is true despite the fact that individual proteins, although denatured at normal body temperatures, are repaired or replaced by cellular maintenance processes that insure proper cellular function and it also fits with one’s own experience, that an elevation of body temperature (as with a mild fever of 1.5° C) although unpleasant, is not fatal. Finney states that “…very extreme probits, say outside the range of 2.5 to 7.5, carry little weight, and may almost be disregarded unless many more subjects were used…” Probit values of 2.5 and 7.5 correspond to probabilities of about 1 % and 99 %, respectively.

The total uncertainty in threshold studies of CO2 laser corneal thermal injury is frequently reported as 10 % or less. For example, Bargeron and colleagues explained that they did not need to apply probit analysis because the final bracket between a lesion and no-lesion was approximately 10 % of the working power level. As a consequence, they required fewer animals to determine their thresholds. Such corneal studies imply a probit slope less than 1.1 because experimental error is minimal and the underlying biological variation for photoagulation is revealed. Clearly, a “safety factor” of 10 in deriving MPE limits cannot be justified.
3.2 Controlling for Refractive Errors

A review of the data from retinal injury studies shows that probit slopes tended to be smaller for the more recent retinal injury studies where laser quality and experimental techniques have improved. It follows that the distributions of experimental data points are more tightly clustered in these more recent studies and should more closely approach the ideal experimental exposure conditions with minimal experimental error. The steepest slopes reported ranged from 1.01 for a large spot 3 µs threshold to about 1.2 for small images for visible wavelengths. This variation in slope with retinal spot size results for two reasons. There are greater experimental difficulties in consistently achieving the smallest possible retinal image due to variations in corneal clarity and intra-ocular light scattering and achieving optimal refraction. This is aptly demonstrated by the studies of Birngruber and colleagues who employed a contact lens delivery system for exposing both rabbits and monkeys with a contact-lens delivery system which eliminated most of the experimental error related to refraction; they consistently obtained steeper slopes of 1.1 to 1.4 in the visible. Also, the small retinal structural and pigmentary inhomogeneities have dimensions approximately the same size at the minimum retinal image and therefore would influence localised energy absorption from exposure to exposure.

![Figure 4. Schematic effect of error in focus for only 0.25 D error. Actual focal error can be greater.](image)

To estimate the magnitude of effect upon the probit slope of refractive errors in an experimental study of retinal thresholds for small images, we mathematically simulated a range of likely refractive errors. It is generally accepted that the best refractive correction in a laboratory setting is approximately 0.25 diopter (0.25 D). In the model, which is schematically depicted in Figure 4, a series of hypothetical exposures were delivered in one primate eye having an effective focal length in air of 15.00 mm and the minimal, nearly diffraction-limited image of 10 µm was achieved for every exposure and the ED-50 was set at 1 µJ with a probit slope of 1.1. Next, a similar series of exposures were made under the assumption that all were delivered with varying degrees of refractive error, but not exceeding 0.25 D. A maximal refractive error of 0.25 D resulted in a retinal image blur circle of 34 µm for a 7-mm pupil. When it was assumed that the thermal injury threshold varied linearly with spot size, then the slope increased to 1.6 and the ED-50 increased to 2.5 µJ (for a beam diameter of 7 mm). When it was assumed that the threshold was constant with retinal irradiance (as postulated for some sub-microsecond durations and shown by Lin et al.), then the slope increased to values greater than 2.0 and the ED-50 increased to approximately 6 µJ. Because the rabbit eye has a shorter focal length (about 10 mm) than the monkey eye (about 15 mm), these calculated results would be even more dramatic for a rabbit model. The results are illustrated in Figure 5. Note that there is an inflection in the slope because there is a minimum spot size of 10 µm, which applies to all probabilities below the inflection point. This model actually fits quite well with published experimental laboratory experience. If the beam diameter at the cornea were 3 mm rather than 7 mm, the slope above the inflection point would become steeper than as shown in Figure 5 because of the increased geometrical depth of focus for the 3 mm beam. In addition to this geometrical factor, the greater corneal spherical aberrations over a 7 mm corneal area will further spread the retinal focal diameter of the larger beam; however, we did not add this effect to our simulation.
Figure 5. The change in the dose response curve modelled by a slight refractive error of only 0.25 D. Note the increase in the ED-50 value, which can become much greater for a slightly greater refractive error.

4. CONCLUSIONS

By adding the effects of misfocus and the change in the impact of pigmentation variation, we show a reduction in the spot-size dependence. In setting retinal laser safety limits, committees have always agreed that human exposures at levels less than 10% of the apparent ED-50 assured safe exposure. In other words a “safety factor” of 10 was quite adequate even though the “real safety factor”, i.e. the ratio of the “true” (“ideal” experiment) ED-50 and the EL, would be less. In the future, committees deriving human exposure limits must take full account of the impact of improved experimental techniques. As techniques improve and uncertainties reduce, probit slopes will become steeper, and it must be recognised that the former “safety factor” of 10 should be reduced accordingly without sacrificing safety. When all contributing physical factors are considered, the true dose-response function cannot be considered to follow a log-normal distribution at extremes. The slope of the dose-response curve is very steep at low probabilities and becomes shallower for higher probabilities where the physical variables of pigment variation, focusability, etc. strongly influence the slope.

Collective plots of spectral dependence, temporal dependence or spot-size dependence, will be skewed if only ED-50 values are plotted without consideration of the variation in their corresponding slopes. The most obvious impact would relate to plots of the variation of ED-50 values with retinal spot-size, since the slope (and “true ED-50”) vary greatly with spot size as was shown in Figure 3 (where the reported ED-50 for the smallest image could be a factor of at least four too large). By adding the effect of misfocus and the variability of pigment motting, the spot-size dependence of some recent studies of Zuclich et al (2000), can be nearly predicted, as shown in Figure 6. This problem has a direct impact upon setting human exposure limits for extended sources. While both theory and experiment predict a spot-size dependence for CW laser exposures, this is not the case for short-pulse laser exposures, and this review suggests that the value for alpha-max needs to be reduced for short pulses. Finally, we can conclude the “safety factor” applied to IR-B and IR-C (corneal hazard) can be two-fold and not ten-fold.
Figure 6. The retinal injury thresholds expressed as radiant exposure calculated at the retina in J/cm² (■) reported by Zuclich et al. (2000) are the uppermost data points. The solid sloped line is the current MPE value for visible microsecond pulsed exposures. The horizontal solid line is the theoretically expected spot-size dependence for such a short pulse. The effect of uncertainties and data spread upon the ED-50 due to mottling is shown by the (●) points. Finally, the combined effect of both pigment-mottling and focal errors to increase the ED-50 values for the smaller images is shown as triangular data points (▲). This suggests that past efforts to plot ED-50 values may have led committees to set limits with a spot-size dependence that was largely artifactual.

REFERENCES


