UNDERSTANDING RISK TOLERANCE CRITERIA

by

Paul Baybutt

Primatech Inc., Columbus, Ohio, USA

www.primatech.com

Introduction

Various definitions of risk are used by risk analysts [1]. In process safety, risk is most often defined as the combination of the frequency of a hazard scenario with the severity of its consequence, usually the simple product of them [2]. Consequence severity is the impact on receptors such as people, the environment, and property. Various forms of risk measures can be used [3]. Some forms distill risk estimates into simple measures such as risk indices which are single numbers. Other forms provide more detailed information presented as tables, graphs and plots. Risk tolerance criteria are used to determine if existing risks are tolerable by comparing them with risk etimates to determine if there is a risk gap. If a gap exists, recommendations are developed for risk reduction measures to close the gap. For a valid comparison to be made, the type and form of risk estimates and risk tolerance criteria must match.

For people, both risk to individuals and risk to groups of people are important. *Individual risk* is the frequency at which an individual may experience a given level of harm as a result of exposure to one or more hazards and *group risk* is the relationship between frequency and the number of people in a given population experiencing a specified level of harm from the realization of one or more hazards [2]. Individual and group risk tolerance criteria specify limits to these risks for individuals and groups of individuals.

Both individual and group risk measures are calculated from the same set of frequency and consequence data for the set of hazard scenarios used in a risk analysis [2]. They are related through the individual risk to which receptors are exposed. For group risk, the population at risk must be specified.

In simple terms, for illustrative purposes, if the fatality risk to an individual from a facility is 1×10^{-1} per year, an individual would expect to be killed within a 10 year period. However, if there are 10 individuals working at the facility, the company could expect to see an average of one fatality per year. Furthermore, some hazards may impact more than one person at a time. Clearly, individual and group risk are quite different. The scale of an incident, i.e. the number of people impacted, affects group risk but not individual risk. Group risk provides a measure of the overall risk to a population but only individual risk provides information about the distribution of risk within the population. Individual risk criteria help to prevent any one individual being exposed to an inequitable share of risk. Group risk criteria help to ensure the risk to groups of people is tolerable.

Usually, the risk of a scenario that results in say 10 fatalities is viewed as less tolerable than the risk from 10 scenarios that individually have single fatalities. For example, far more people are killed in automobile accidents than in airplane crashes every year but the latter receive much more publicity owing to the larger number of people impacted. For other types of receptors, such as the environment and property, distinctions are not usually made between impacts on individual and multiple receptors.

Thus, if a process contains multiple pumps, no distinction would be made between the tolerable risk of ten pumps being destroyed in individual scenarios versus ten pumps in a single scenario so long as the risk was the same (total loss is equally likely for both cases). Naturally, in the case of ten pumps being destroyed at the same time, the actual loss may be greater than the replacement cost of the pumps owing to the impacts of such a loss on the process. Of course, such effects can be captured in the estimation of the overall damage costs and the tolerability of such losses are expressed using criteria that define tolerable frequencies as a function of the monetary loss.

Individual Risk Measures for People

Individual risk measures can be presented as single numbers or sets of numbers [3]. They depend on where people are located but not on the number of people present. For example, if two operators work in the same location, they are exposed to the same risk. The individual risk does not increase because two people are present. However, the group risk does increase, as will be discussed later.

The risks of all scenarios that could impact an individual, regardless of the number of people impacted by the scenario, contribute to individual risk. For example, if a hazard scenario can result in the single fatality of Operator A and another hazard scenario can result in 5 fatalities, one of which is Operator A, both scenarios would contribute to the individual risk to Operator A.

The term individual risk can be confusing. Sometimes, it is applied to the risk at a location within a facility regardless of whether an individual is actually present there. This can be viewed as the risk to a hypothetical person at the location, sometimes calculated for all times with no protection from the hazards. This particular risk measure is called *geographic* or *location* individual risk or *hypothetical* individual risk. Individual risk can also be determined for actual individuals and this measure is usually called just individual risk, hence the confusion.

Arguably, the preferred measure of individual risk is the actual risk to each individual determined by calculating individual risk at every geographical location where people are present and accounting for time spent at the location. Such a measure allows risk estimates to reflect variations in exposure durations, protections available, individual susceptibility to harm, etc. However, when performed as part of Quantittaive Risk Analysis (QRA), many calculations are needed to address the risk to each exposed individual and the resources required quickly become prohibitive and the analysis of the results becomes complex. Consequently, practitioners use the concept of a hypothetical person to provide a reasonable representation of people actually exposed to the risk, for example, a plant operator with a defined occupancy profile for plant areas during a typical working period. A number of hypothetical persons may need to be constructed to represent different categories of people with different exposures, protections, susceptibilities, etc. Note that this is not the same as the hypothetical person used to give meaning to location individual risk. Location risk is a theoretical construction that does not account for the presence of people and is not a measure of the risk actually experienced. However, it does represent the maximum possible risk that an individual could experience at a particular location.

Location risk criteria are used in the Netherlands for public risk [4]. This usage is understandable from the perspective that no one controls the movements of the public, some members of the public may indeed spend the majority of their time in one location, and a measure of conservatism in public risk criteria is appropriate given that the risk is involuntary. In contrast, the movements and locations of facility workers can be controlled and, usually, no individual is located at a facility 24/7. Consequently, location risk is less applicable to worker risk. The Netherlands does not use risk criteria for workers.

Individual risk calculated for compliance with regulations in the United Kingdom (UK) does account for the presence of people, albeit hypothetical individuals, and risk tolerance criteria are used for both the public and workers [5]. The UK Health and Safety Executive (HSE) views the concept of a hypothetical person as being advantageous because it avoids the assessment of risk to people actually exposed to the risks [5] which avoids the adoption of inappropriate strategies for regulatory compliance such as limiting the time exposure to a high hazard, for example, by rotating personnel. However, these factors are part of the risk analysis model regardless of whether actual or hypothetical people are used and they can be controlled through the assumptions made in the analyses.

Different types of individual risk measures can be used, for example, the maximum geographic individual risk and the maximum individual risk. The maximum geographic individual risk is the maximum value of the geographic individual risk. The maximum individual risk is the individual risk to the person exposed to the highest risk in an exposed population. It can be determined by calculating the individual risk at every location where people are actually present and selecting the maximum value. It is equivalent to the sum of the geographic individual risks at the locations where people are present for the maximum exposed person. Individual risk is usually expressed as a single number, risk contour plot, or risk profile / transect [2] according to the type of individual risk measure used. It can be determined for different levels of harm, e.g. first-aid case, severe injury, and fatality. Often, the latter is used.

Individual risk criteria are often used with the As Low As Reasonably Achievable (ALARP) principle which sets an upper bound below which risk must be reduced (the de manifestus level) and a lower bound that is a target that may not be achieved (the de minimis level) (Figure 1) [6]. The UK HSE has suggested the following values for individual annual fatality risk from all hazards present at a single major industrial activity (i.e. a facility) [5]:

de manifestus		**de minimis**	
Workers	**Public**	**Workers**	**Public**
1×10^{-3}	1×10^{-4}	1×10^{-6}	1×10^{-6}

Note that the word “public’ has been used and not “societal” or “social”. The latter terms are associated with risk to groups of people. The “public” numbers in the table are for individual members of the public located outside a facility. The UK HSE defines a single major industrial activity (i.e. a facility) as:

> “... an industrial activity from which risk is assessed as a whole, such as all chemical manufacturing and storage units within the control of one company in one location or within a site boundary ...”

Historically, individual risk tolerance criteria for facility personnel have been set with reference to the risk levels from accidents that actually have been tolerated in the workplace and individual risk tolerance criteria for the public have been set with reference to the risk levels from non-work-related accidents that actually have been tolerated, although with a reduction factor, typically 1 percent, to account for the involuntary nature of the risk.

Group Risk for People

Historically, group risk has been called *societal risk*, although that is a misnomer as it does not relate to society at large. It was first used as a risk measure for the public, hence its name. It can also be used for employees at a facility.

Group risk is often expressed as f-N or F-N curves [2], although F-N curves are more commonly used. f-N curves display the frequencies of all events that result in N casualties while F-N curves display the cumulative frequencies of all events that lead to N or more casualties (Figure 2). Thus, both curves are functions of N, i.e. f-N = f(N) and F-N = F(N). Usually, both types of curve are plotted on a log-log scale since f, F, and N cover orders of magnitude. f-N and F-N curves are usually displayed for intervals of casualty numbers because risk analysis methods usually organize the consequence severities into bins for computational convenience and uncertainties in the analysis obviate the need for more precision. When f-N or F-N curves are used for group risk tolerance criteria they are called limit lines [7].

F-N limit lines are defined by these parameters:

- Base point (the cumulative frequency of 1 or more casualties, also called the exceedance frequency for a single casualty), or another *anchor point* on the line
- Slope
- Cut-offs for cumulative frequency and/or consequence.

A shorthand notation is used to describe F-N limit lines using their anchor point and slope. For example, the triplet (10, 1×10^{-4}, -1) represents a limit line for which N = 10 at F = 1×10^{-4} with a slope -1, i.e. a straight line passing through the anchor point (10, 1×10^{-4}). The slope determines how quickly cumulative frequency falls off with increasing consequence. For slopes of -1, cumulative frequency decreases by an order of magnitude as N increases by an order of magnitude. F-N curves are defined by the equation:

$$F(N) = k / N$$

where k, a constant, is the intercept with the cumulative frequency axis, i.e. F(1).

Group risk measures are a function of the number of casualties and this raises the issue of aversion to higher consequence casualties. Early F-N limit lines used slopes of -1 in the belief that this was risk neutral (the higher the consequence, proportionately, the lower the frequency) and that consequence casualties should drop off at least as fast as the frequency. Risk aversion was introduced using slopes between - 1 and - 2 (the higher the consequence, the greater the reduction in frequency). However, it is now recognized that F-N limit lines with a slope of -1 actually incorporate a degree of risk aversion owing to the relationship between f-N and F-N limit lines [8].

Risk aversion can be shown in either F-N space or f-N space but it is defined in terms of f-N space. Consequently, a f-N limit line with slope -1 is risk neutral but when this is transformed into a F-N limit line, lower risk aversion is displayed for most N values owing to the mathematical properties of F-N curves which are not apparent to the casual observer. F-N limit lines with inflections can be used to incorporate more risk aversion for higher consequence events by changing to a higher gradient line after some large number of casualties has been reached. The equation for a limit line including risk aversion is:

$$F(N) = k / N^{\alpha}$$

where α is a constant ≥ 1. It is the negative slope of the F(N) curve in log space.

In a f-N curve, the frequencies of events with particular numbers of casualties are plotted separately. This can be misleading. For example, if fatality events are being plotted, the frequency of all single fatality events is plotted, followed by the frequencies of multiple fatality events. However, the multiple fatality events are composed of individual fatalities that occur together. F-N curves account for this relationship by showing the frequencies of all events that result in one fatality or more, two fatalities or more, etc. Thus, the data point for a single fatality is a summation of the frequencies of all scenarios in which an individual is killed, regardless of how many more people are killed at the same time. Similarly, the data point for two fatalities is a summation of the frequencies of all scenarios in which two or more individuals are killed, regardless of how many more people are killed at the same time. This display provides a clearer depiction of the risks and helps to account for the preference of F-N over f-N curves.

Furthermore, there are other issues with f-N curves that make them less desirable for displaying risk results but manipulation of F-N curves can be tricky owing to their mathematical properties [7, 9]. Caution must be exercised when manipulating f-N and F-N curves for several reasons:

- Cumulative frequency plots are more difficult to manipulate than frequency plots.
- F-N curves are usually shown in log - log form which complicates their mathematics.
- Often a consequence cut-off is used and sometimes also a frequency cut-off (see Figure 2). These end effects must be addressed in mathematical representations of the curves.
- F-N curves are usually shown as continuous lines, even though casualty numbers are discrete. Thus, the mathematics of the curves can be addressed in continuous or discrete space [7].

Group risk can be expressed in simpler forms similar to individual risk, for example, the frequency of 10 fatalities occurring from a particular facility. Also, group risk can be reduced to a single number, or index. For example, the expectation value for casualties can be calculated by summing f(N) x N for all events at a facility. For fatalities, the expectation value is known as the *Probable Loss of Life (PLL)* or *average rate of death (ROD),* or *average societal (group) risk.* Of course, as simpler group risk measures are used, information is lost about the distribution of group risk. For example, PLL values do not differentiate between low likelihood / high consequence events and high likelihood / low consequence events. However, some of the simpler risk measures are easier to understand and use.

Integrals of the areas under F-N curves can also be used as measures of group risk. For example, the UK HSE has developed a Scaled Risk Integral (SRI) as a measure of group risk [10]. It is used as a screening tool for land use planning in the vicinity of hazardous facilities.

A group risk measure can be calculated for individual consequence severities, e.g. single fatality events. However, this is not the same as individual risk for a single fatality. The difference is that the group risk measure represents the total frequency of all single fatalities in an exposed population, regardless of the individuals involved, while the individual risk of a fatality is for specific individuals. For example, if there are 100 people who work in a plant, the group risk of a single fatality is the sum of the frequencies of all hazard scenarios that can result in a single fatality regardless of the individual. However, the individual risk of a fatality is the sum of the frequencies of only those hazard scenarios that can result in a fatality for each individual. Individual risk will vary according to the individual and their job, location, etc.

The UK HSE is considering the development of group risk tolerance criteria [11]. However, few F-N curves are available showing actual societal risks that are tolerated. The UK HSE has proposed a group risk tolerance criterion: the risk of an accident causing the death of 50 or more people in a single event should be regarded as intolerable if the frequency is estimated to be more than one in 5,000 per year [5].

F-N curves can be calculated using techniques such as QRA and LOPA. They can also be constructed from historical data where they are available. Not only can F-N curves be used for comparison with limit lines but they can also be used to compare the risks of facilities or to benchmark a new facility with an existing one.

Risk Measures for Other Receptors

Risk measures for hazard scenarios that impact receptors such as the environment and property are usually expressed by converting impacts such as property losses and environmental remediation into monetary terms and specifying risk as the frequency of the monetary loss. Difficulties may be encountered if the impacts cannot be converted into monetary terms, for example, in the case of permanent environmental impacts and the loss of irreplaceable property. There are no commonly-used risk tolerance criteria for these receptors.

Allocation of Risk Tolerance Criteria

Practitioners often use risk tolerance criteria for hazard scenarios or hazardous events in the belief that it is easier to calculate their risk rather than the overall risk of a facility. However, risk tolerance criteria for individual hazard scenarios or hazardous events have no meaning by themselves. They must be derived by allocating overall facility risk to the scenarios or events by dividing facility risk tolerance criteria by the number of scenarios or events which requires the number of scenarios or events to be estimated. These numbers are uncertain unless all scenarios and events are identified in which case their overall risk can be summed to produce an estimate of facility risk that can be compared with faclity risk tolerance criteria.

Of course, reliance solely on meeting overall facility risk tolerance criteria may result in the inequitable distribution of risk across a facility. While the overall facility risk criterion may be met, there may be processes, areas, units, process modes, people, etc. that bear the brunt of the risk resulting from the disproportionate allocation of risk across the facility. Consequently, overall risk determination should be accompanied by the allocation of the overall risk tolerance across a facility, particularly to all receptors as ultimately that is what matters.

Scaling of Risk Tolerance Criteria

Some companies adjust, or scale, their group risk tolerance criteria to reflect the number of operations and their size. Scaling does not apply to individual risk criteria as it does not depend on these factors. Scaling addresses inconsistencies that would otherwise arise. For example, a company may operate 10 facilities and another company a single similar facility. If both companies used the same overall tolerable risk criterion, the company with 10 facilities would have to operate them all at one tenth the risk of the company with the single facility to meet the criterion. More specifically, if a small company operates one facility in which the individual fatality risk to each employee is 1×10^{-3} per year and ten people work in the facility at any time, the company can expect to experience one fatality every 100 years or a group risk of 1×10^{-2} single fatalities per year. However, if a large company operates one hundred such facilities, that company faces an enterprise group risk of one fatality per year. Similarly, it is not logical to assign the same group risk tolerance criterion to a large facility as for a small facility. Either the small facility would be allowed to operate with undue risk or more stringent requirements than necessary would be imposed on the large facility. If a company operated a facility and determined that its risk was twice the tolerable value, theoretically, the company could split the facility into two and meet the facility risk tolerance criterion without actually having reduced risk at all.

The basis for scaling should be a facility metric that addresses the importance of the facility to the people it serves. However, the people it serves may not be the people it affects. Societal factors enter into the picture. However, a possible simple importance measure for a facility is the value of the products it produces as measured by their sale price. Consequently, overall company group risk tolerance criteria could allocated to individual facilities by the ratio of the value of the products from the facility to the total value of products produced by all facilities.

More complex scaling metrics may be a function of additional facility metrics such as the number of employees, production volume, value of the facility to the community, and/or national strategic importance. Failure to scale group risk criteria likely will result in higher risk being tolerated for small facilities and large amounts of risk reduction.

References

1. T. Aven, Misconceptions of Risk, Wiley, 2010.
2. Guidelines for Chemical Process Quantitative Risk Analysis, CCPS/AIChE, 2nd Edition, 2000.
3. Guidelines for Developing Quantitative Safety Risk Criteria, CCPS/AIChE, 2009.
4. Netherlands risk criteria.
5. Reducing Risks, Protecting People, United Kingdom Health and Safety Executive Books, 2001.

6. ALARP Suite of Guidance, United Kingdom Health and Safety Executive, http://www.hse.gov.uk/risk/theory/alarp.htm.

7. D. C. Cox and P. Baybutt, Limit Lines for Risk, Nuclear Technology, Vol. 57, pps. 320 - 330, June 1982.

8. D. J Ball and P. J. Floyd, Societal Risks, United Kingdom Health and Safety Executive report, 1998.

9. R. F. Griffiths, Problems in the Use of Risk Criteria, in: Dealing with Risk, ...R.F. Griffiths (Editor), Wiley, New York, 1982.

10. I. L. Hirst and D. A. Carter, A "worst-case methodology for obtaining a rough but rapid indication of the societal risk from a major hazard installation", J. of Hazardous Materials, A92, pps. 223 – 237, 2002.

11. CD212, Consultative document on proposals for revised policies to address societal risk around onshore non-nuclear major hazard installations, United Kingdom Health and Safety Executive, 2007.

Figure 1. The ALARP Principle.

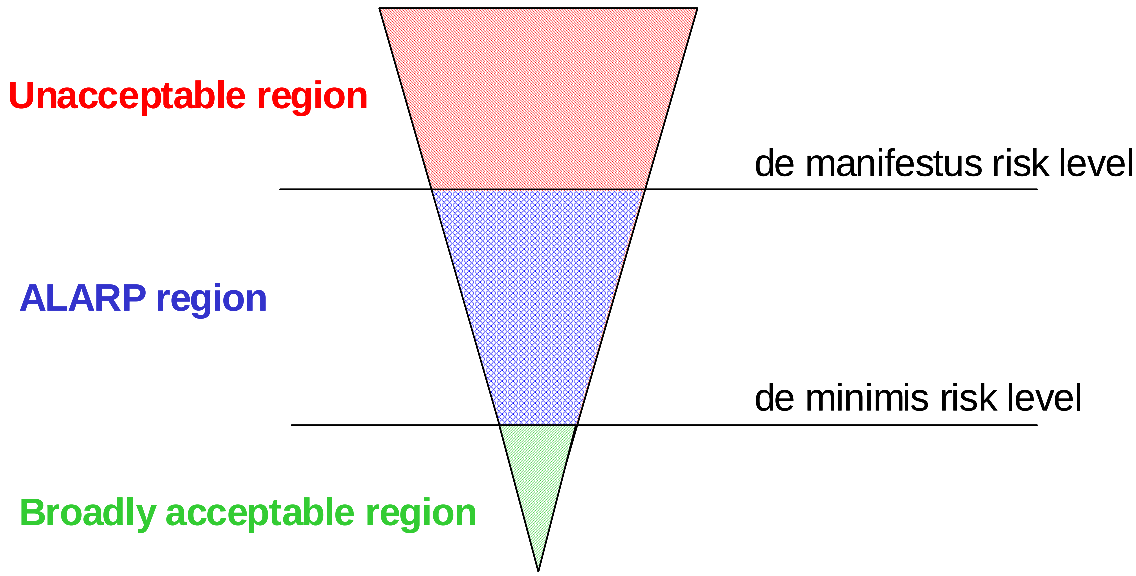

Figure 2. F-N Limit Line.