

Originally Published [MDDI](#) February 2003

RISK MANAGEMENT

Establishing Overall Risk for Medical Devices

In addition to reducing individual risks, device makers must develop an overall risk index in accordance with ISO/IEC 14971, Clause 7.

Mike W. Schmidt

For medical device makers, identifying and reducing their individual product risks is no longer enough. Clause 7 of ISO/IEC 14971 (published in December of 2000) states that after product developers have reduced all individual risks associated with a medical device as far as is reasonably possible, they must establish an overall risk level. This level of overall risk must reflect the cumulative effects of the individual risks. The clause reads as follows:



After all risk control measures have been implemented and verified, the manufacturer shall decide if the overall residual risk posed by the medical device is acceptable using the criteria defined in the risk management plan. If the overall residual risk is judged unacceptable using the criteria established in the risk management plan, the manufacturer shall gather and review data and literature on the medical benefits of the intended use/intended purpose to determine if they outweigh the overall residual risk. If this evidence does not support the conclusion that the medical benefits outweigh the overall residual risk, then the risk remains unacceptable. The results of the overall residual risk evaluation shall be recorded in the risk management file.

Compliance is checked by inspection of the risk management file.

Currently, the supplementary document ISO/IEC 14971 Amendment 1 is in its final stages of development; no further technical changes will be made. This supplement does not change the normative requirements, but does provide rationales for them.

With regard to Clause 7, "Overall Residual Risk Evaluation," the draft amendment explains it as a requirement for a manufacturer to look not only at the acceptability of individual risks, but to evaluate whether the combination of all individual risks associated with the device exceeds acceptable levels. Amendment 1 states that even when all individual risks are considered acceptable, the cumulative effect of those risks may be unacceptable. It goes on to declare that even in cases where this overall or aggregate risk exceeds acceptable levels, a risk-benefit analysis may still demonstrate that what might otherwise be considered an unacceptable level of risk is acceptable in light of the benefits provided.

Both the original requirement set forth in Clause 7 and the rationale given in Clause H.2.7 exist to ensure that individual (residual) risks associated with a device have been reduced as far as is practicable and that the cumulative effect of those residual risks cannot present an unacceptable risk to patients, users, other parties, or the environment. Although this goal is a worthy one, establishing a method for achieving it that is acceptable to third-party auditors and regulators (that is,

establishing an acceptable level for cumulative risk and evaluating a specific device against that level) is not easy.

What follows is an examination of the issues associated with ISO/IEC 14971:2000 Clause 7 and alternative approaches to compliance.

Establishing Risk Indices for Individual Hazards

Before looking at potential methods for combining individual risks to determine an overall risk level, it is first necessary to review the method defined in ISO/IEC 14971: 2000 for establishing individual risk levels. In essence, the standard requires that all hazards (i.e., potential sources of harm) associated with the device be identified. For each of these hazards, the likelihood that the hazard will occur (i.e., the likelihood that the initiating event will happen and, when it does, that there will be exposure to the hazard) must be estimated. Likelihood can be expressed as a numeric probability, or simply as remote, possible, likely, certain, etc. In addition, the potential severity of the harm, which ranges from minor injury through severe injury to permanent injury or death, must be estimated. Finally, likelihood and severity must be combined to establish a risk index. Acceptable methods for combining likelihood and severity include graphically, as in Figure 1, or mathematically, by the following equation:

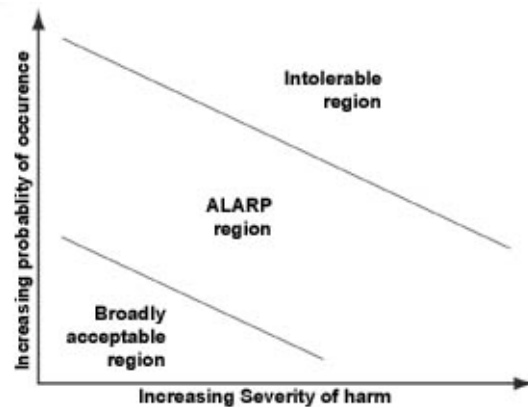


Figure 1. Example of a graph used for determining a risk index.

$$\text{Risk Index} = \text{Severity Index} \times \text{Likelihood Index}$$

(Establishing the likelihood and severity indices typically involves assigning an arbitrary numeric scale—such as 1 through 10—to achieve normalized values.)

Both the graphical and mathematical techniques yield essentially identical results, but when one is evaluating individual hazards, the mathematical approach makes the combination of all individual risk indices into an overall risk index for the device appear more straightforward. However, some who perform risk assessment criticize the use of conventional probabilities in these calculations, asserting that the format of these calculations implies an accuracy that is rarely achieved owing to the extensive effort required to maintain adequately high confidence intervals. These individuals also assert that even if such accuracy were achieved, the meaningfulness of that accuracy becomes moot when the probabilities are combined with the severity values. (The severity values are coarse estimations if they are “less than death” or “greater than none.”)

Whether the mathematical or graphical method is used, the standard suggests risk levels be allocated into three basic categories: unacceptable, as low as reasonably practicable (ALARP), and broadly acceptable. Unacceptable risks are just as the name implies—unacceptable under any terms. ALARP risks may be acceptable if an evaluation shows that the resulting residual risk is justified because there are product benefits that offset it. Broadly acceptable risks are those that are low enough in severity, likelihood, or both to be roughly equivalent to the day-to-day risks encountered in ordinary life. Implicit in all three levels, especially in the latter two, is the recognition that “zero risk” does not exist.

Finally, it should be kept in mind that earlier approaches to assessing the risk of medical devices, which used only risk analysis, have been replaced by the life-cycle model inherent in risk management. Risk management recognizes that risk estimations made during product development are only highly educated guesses. It is imperative that the initial risk analysis be continuously updated based on real-world experience, and that appropriate action be taken to achieve acceptable risk levels based on those updates.

While the recent revision of H.2.7 deleted the directive "add up " to avoid specifying a methodology, circumstances exist in which simply summing the individual risk indices has merit. Similarly, qualitative methods can be valid if they are applied consistently. What follows is an evaluation of some alternative methods for combining individual risks to determine whether the overall risk level is acceptable.

Summing the Severity of Hazards

Summing the severities of all individual hazards to determine their cumulative effect is reasonably valid assuming all hazards occur simultaneously or within a short enough period that their negative effects (i.e., injuries) would, in fact, be cumulative. This principle can be illustrated with the example of a bee sting. Although a single sting may be annoying, it is generally not considered a major injury. However, if an individual receives many stings over a short enough period that the venom from any one sting is not metabolized before subsequent stings occur, the resultant injury could be significant, even fatal.

The bee-sting example demonstrates that for the additive method to reflect accurately the actual result, the effects of the individual hazards must be cumulative; i.e., each injury's severity is magnified by the occurrence of preceding injuries, in addition to the hazards' occurring during a compressed time frame.

Summing the Likelihood That Hazards Will Occur

Similarly, the accuracy of summing the likelihood that individual hazards will occur is affected by the independence of the events. This means that if the likelihood of the occurrence of one event is unaffected by the occurrence of a second, the likelihood or probability that one of the two will occur at any single point in time is represented by adding the likelihood of each occurring. The probability that the two independent events will occur simultaneously, however, is derived by multiplying the probability of each occurring, which results in a reduced probability of occurrence. (For purposes of this discussion, combining the likelihood of individual events is discussed in terms of conventional probability where a value of 1 indicates certainty of occurrence and all other possibilities are expressed as decimal values less than 1.)

Discuss this article on-line!
Share your comments and questions with the authors and other readers in *MD&DI's* [Author Forums](#).

Because the ISO/IEC 14971 standard requires that hazards (events resulting in specific injuries) be identified before evaluating the likelihood that the initiating event will occur, a robust risk evaluation process will tend to incorporate all dependent events that contribute to the realization of a specific hazard. Therefore, for purposes of evaluation of risks associated with a medical device according to ISO/ IEC 14971, it is reasonable to assume that the risks are independent.

Summing probabilities ignores the issue of a compressed time frame when one evaluates the summing of severity levels, however. In effect, this approach assumes that all injuries resulting from the use of the device over its lifetime will

compound. Obviously, this assumption ignores whether the device will be used on many patients over several years.

Summing Individual Risk Levels

Based on the critical factors associated with summing the severity of individual hazards as well as summing the individual probabilities of their occurrence, it becomes clear that to achieve an accurate result, the level of interdependence of the hazards in terms of severity and likelihood is crucial. While the likelihood of the occurrence of individual injuries can reasonably be considered independent within the context of risk analysis according to ISO/IEC 14971, the dependence of those injuries in terms of cumulative severity will vary dramatically depending on the device, the number of patients on which it is intended to be used, and the specific hazards involved.

However, it is certain that the additive technique will always yield the absolute worst-case cumulative risk levels, since it effectively takes as given that all severities are cumulative (that is, that each injury increases the effective severity of the others). This technique also takes for granted that individual events will occur simultaneously or within a short enough period to compound each other's severity. (Any such case will be described as simultaneous.) From the perspective of probabilities, it must be kept in mind that the likelihood that any one event will occur addresses all cases, including those where the other events occur simultaneously.

Consider the example of events A and B, each of which can occur with or without the other. To understand better how the additive method can artificially inflate the aggregate probability, assume that either event will occur without the other 50% of the time, and therefore will occur simultaneously 50% of the time. If A and B were each to occur 10 times in 1000 opportunities, then five of those 10 occurrences would be simultaneous. This means that the actual probability that the two events would occur simultaneously is 5 times in 1000 opportunities. However, the additive method would indicate that the probability of simultaneous occurrence is 20 times in 1000 opportunities (i.e., $10 + 10$), thereby artificially inflating the likelihood four-fold.

Furthermore, it is clear that performing an in-depth analysis of the severities and probabilities of individual risks to accurately indicate how these risks will combine is almost certainly impractical in light of the urgency to get beneficial products to market. The primary shortcoming of this approach is that if the same scale or threshold to determine acceptability is used for cumulative risk as is used for individual risks, the cumulative risk value will nearly always exceed acceptable limits, since hundreds of individual risks are associated with most medical devices.

This issue can of course be overcome by simply creating a separate scale for cumulative risk with higher thresholds for broadly acceptable, ALARP, and unacceptable risks. However, manufacturers who decide to take such an approach should invest significant resources into documenting how their alternate scale was developed and justifying the thresholds selected. Failing to do so could easily leave the impression that these cumulative levels were developed to ensure that all products will fall within acceptable levels. Leaving this impression on regulators or a civil court during liability litigation could be quite costly.

An alternative to creating a separate cumulative scale for acceptability is to simply exclude all individual risks that are within the broadly acceptable limits, summing only those risks that fall within the ALARP category.

This alternative method sums the risk indices for all individual risks that are greater than the broadly acceptable threshold, and so ensures that the result will tend to be conservative (i.e., erring on the side of safety because they are based on the assumptions that risks are dependent and compounding). But results will not be so conservative as to require a higher level of safety than is achieved in ordinary day-to-day life, because risks that are broadly acceptable were disregarded. This method permits maximum flexibility, by allowing multiple ALARP risks at the lower end of the range or only one that falls within the upper limits.

The main drawbacks of this approach, however, are that by dealing with calculated values, the accuracy of the system can be perceived as far greater than it actually is, and that in an attempt to achieve such high accuracy, the system can become inherently cumbersome and unworkable.

Combining Individual Risks Qualitatively

Taking a qualitative approach typically involves bringing together a review panel of individuals who have an adequate level of distance from the specific device's development to review the risk analysis objectively and determine whether the overall, cumulative level of residual risk for the device is acceptable. The panel must consider the benefits provided by the device compared to those of the available clinical alternatives, such as pharmaceuticals, and the level of residual risk provided by those alternatives, as well as by similar devices.

Clause 3.3 of ISO/IEC 14971 also requires the manufacturer to define a policy or procedure for determining the acceptability of all risks, including the overall residual risk. Such policies, like all other aspects of risk management, are, of course, subject to ongoing review and revision based on field experience. In essence, this requirement is intended to ensure consistency from device to device.

Thus, when developing a qualitative system for assessing risk, it is critical that a company take the following actions:

1. Establish a set of rules that will be applied when evaluating individual and cumulative risks for all products.
2. Document the specific qualifications (training, education, etc.) of those who will apply the rules.
3. Confirm that personnel assigned to apply the rules are not directly involved in the design or development of the product being evaluated.

Because the second and third items are typical for documented quality systems in general, the following discussion is limited to the principles used to generate the rules of acceptability.

Probably the simplest approach to setting such rules is to place a limit on the number of risks falling in the broadly acceptable and ALARP regions of the graphical analysis. Although this technique is simple in theory, it can cause difficulties when all risks are categorized as broadly acceptable and none fall within the ALARP range, yet the limit for broadly acceptable risks is exceeded.

Of course, this same principle can be applied to actually overcome this difficulty. Since broadly acceptable risks are just that, a reasonable argument can be made that such risks should not be considered in establishing the acceptability of cumulative risks. Applying this argument eliminates the problem of deeming a

product with only broadly acceptable risks unacceptable.

With this issue resolved, one final potential problem remains: If the quantity of ALARP risks serves as the final determinant of the acceptability of cumulative risk, how should the acceptable number of ALARPs for the device be determined? If the number is set assuming risks are at the high end of ALARP range (so that the maximum acceptable number of such risks is low), a device could have a quantity of individual risks that are barely within that range—realistically presenting a reasonably low overall risk level—but be rejected as unacceptable. If the acceptance level for the quantity of ALARP risks is set based on the assumption that they are all at the low end of the range (which then allows a larger number), a device that should be rejected for overall risk level could be considered acceptable.

While these extreme examples can be addressed by assuming that the individual risks are at the median value of the ALARP range, the resulting limits will still be somewhat coarse and will carry the potential for inappropriate decisions. Furthermore, this method could be viewed as driving determinations of individual risk levels. For example, if a group making evaluations is readily aware that one more ALARP will put them over the acceptable-risk limit, the risk level selected for the next hazard identified might be perceived as suspect.

An alternative qualitative approach is to establish a group of experts with backgrounds in device regulation, liability, engineering (both electrical and mechanical, as appropriate), medical practice, and other areas. After this group reviews the risk assessments and all relevant safety-related information, each member of the group makes an acceptability determination from his or her own perspective. The group then attempts to reach consensus on the acceptability of the overall residual risk. Although this technique is subjective, it ensures that an appropriate evaluation is performed using predetermined limits for overall risk, and that various perspectives are considered.

Conclusion

As the ISO/IEC 14971 standard recognizes, the risk evaluation and estimation process will always be, at best, only partially scientific. What constitutes a broadly acceptable level of risk is driven by societal values, and perceptions are rarely based on quantifiable and predictable values.

Mike W. Schmidt is a senior standards compliance associate at Ethicon Endo-Surgery Inc. in Cincinnati, OH.

Discuss this article on-line! The authors will be responding to questions and comments about this article in *MD&DI's* Author Forums through March 14, 2003. Visit the [Author Forums](#) now to join the discussion of this and other articles.

Copyright ©2003 **Medical Device & Diagnostic Industry**

Comments about this article?
Post them in our [Members' Discussion Forums](#).