*Risk Analysis, Vol. 29, No. 7, 2009* DOI: 10.1111/j.1539-6924.2009.01209.x

*Perspective*

# What's Wrong with Hazard-Ranking Systems? An Expository Note

**Louis Anthony (Tony) Cox, Jr.***

Two commonly recommended principles for allocating risk management resources to remediate uncertain hazards are: (1) select a subset to maximize risk-reduction benefits (e.g., maximize the von Neumann-Morgenstern expected utility of the selected risk-reducing activities), and (2) assign priorities to risk-reducing opportunities and then select activities from the top of the priority list down until no more can be afforded. When different activities create uncertain but correlated risk reductions, as is often the case in practice, then these principles are inconsistent: *priority scoring and ranking fails to maximize risk-reduction benefits.* Real-world risk priority scoring systems used in homeland security and terrorism risk assessment, environmental risk management, information system vulnerability rating, business risk matrices, and many other important applications do not exploit correlations among risk-reducing opportunities or optimally diversify risk-reducing investments. As a result, they generally make suboptimal risk management recommendations. Applying portfolio optimization methods instead of risk prioritization ranking, rating, or scoring methods can achieve greater risk-reduction value for resources spent.

**KEY WORDS:** Homeland security; portfolio optimization; risk management priorities; risk matrices; risk rating; risk scoring; Superfund; vulnerability assessment

## 1. INTRODUCTION

Many organizations currently rate, rank, or score different hazards (sources of risk) or risk-reducing opportunities at least once a year to identify the currently top-ranked opportunities that will be addressed in the current budget cycle. The use of priority scoring and rating systems is widespread and is becoming even more prevalent as they are incorporated into national and international standards and regulations. This note examines some intrinsic limitations in the performance of all possible priority-setting rules and scoring systems, evaluated as guides to rational action. Most of the results are well known in decision analysis and financial risk analysis and/or are mathematically straightforward. However, they are of great practical importance for understanding limitations of risk-scoring methods and developing improved approaches to risk management. In general, *risk-scoring methods are not appropriate for correlated risks.* Indeed, as we will demonstrate, they are not necessarily better than (or even as good as) purely random selection of which risk management activities to fund.

More constructively, when risk-reducing opportunities have correlated consequences, due to uncertainties about common elements (such as potencies of chemicals, effectiveness of countermeasures, etc.), then *methods for optimizing selection of a portfolio (subset) of risk-reducing opportunities can often achieve significantly greater risk reductions for resources spent than can priority-scoring rules.* In general, the best choice of a subset of risk-reducing

Cox Associates and University of Colorado.
* Address correspondence to Louis Anthony (Tony) Cox, Jr., 503 Franklin Street, Denver, CO 80218, USA; tel: 303-388-1778; fax: 303-388-0609; tcoxdenver@aol.om.

activities cannot be expressed by priority scores. Instead, optimization techniques that consider interdependencies among the consequences of different risk-reducing activities are essential. Fortunately, such methods are easy to develop and implement. They can substantially improve the risk-reduction return on investments in risk-reducing activities.

## 2. MOTIVATING EXAMPLES

Examples of currently important applications of priority-scoring systems in risk analysis include the following.

### 2.1. Example: Scoring Information Technology Vulnerabilities

The *common vulnerability scoring system* (CVSS) for rating information technology (IT) system vulnerabilities uses scoring formulas such as the following to help organizations set priorities for investing in security risk reductions:

$$\text{BaseScore} = (0.6 * \text{Impact} + 0.4 * \text{Exploitability} - 1.5) * f(\text{Impact})$$

$$\text{Impact} = 10.41 * (1 - (1 - \text{ConfImpact})(1 - \text{IntegImpact}) * (1 - \text{AvailImpact}))$$

$$\text{Exploitability} = 20 * \text{AccessComplexity} * \text{Authentication} * \text{AccessVector}$$

$f(\text{Impact}) = 0$ if Impact $= 0$; 1.176 otherwise

AccessComplexity = case AccessComplexity of
- high: 0.35
- medium: 0.61
- low: 0.71

Authentication = case Authentication of
- Requires no authentication: 0.704
- Requires single instance of authentication: 0.56
- Requires multiple instances of authentication: 0.45

AccessVector = case AccessVector of
- Requires local access: 0.395
- Local Network accessible: 0.646
- Network accessible: 1

(*Source:* http://nvd.nist.gov/cvsseq2.htm)

Such a rule base, no matter how complex, can be viewed as an algorithm that maps categorized judgments and descriptions (such as that access complexity is "high," and that local access is required) into corresponding numbers on a standard scale. Higher numbers indicate greater vulnerability and the need for remedial action. Proponents envision: "As a part of the U.S. government's SCAP (Security Content Automation Protocol) CVSS v2 will be used in standardizing and automating vulnerability management for many millions of computers, eventually rising to hundreds of millions" (http://www.first.org/cvss/).

### 2.2. Example: Scoring Consumer Credit Risks

The practice of rank ordering consumers based on credit scores is ubiquitous in business today. A recent description states that "FICO® risk scores rank-order consumers according to the likelihood that their credit obligations will be paid as expected. The recognized industry standard in consumer credit risk assessment, FICO® risk scores play a pivotal role in billions of business decisions each year. . . . [They] are widely regarded as essential building blocks for devising successful, precisely targeted marketing, origination and customer management strategies by credit grantors, insurance providers and telecommunications companies." Examples include BEACON® at Equifax US and Canada, FICO® Risk Score Classic at TransUnion US, and Experian/Fair Isaac Risk Model at Experian (*Source:* http://www.fairisaac.com/fic/en/product-service/product-index/fico-score/).

### 2.3. Example: Scoring Superfund Sites to Determine Funding Priorities

The State of Connecticut (http://www.ct.gov/dep/lib/dep/regulations/22a/22a-133f-1.pdf) published a *Superfund priority score method* to be used in determining funding priorities for remediation of Superfund sites. Users must score each of the many factors (reflecting exposure potential; groundwater impact; surface water impact; toxicity, persistence, mobility, and quantity of hazardous substances; impact to the environment, including species of special concern; and potential air release and fire hazards) using ordered categories. Each category carries a certain number of points. For example, an area that contains a "rare" species gets a score of 4 on this factor. If it has a "declining or infrequent" species, the score is 3; for a "habitat-limited species," the score is 2. If this factor (species of concern) is not applicable, the score for this factor is 0. The scores for all factors are summed. The resulting total score determines "the priority for funding of remedial action at sites on the SPL" (the State of Connecticut Superfund priority list).

### 2.4. Example: Priority Scoring of Bioterrorism Agents

MacIntyre *et al.* (2006) proposed a risk priority scoring system for bioterrorism agents. They described their approach as follows:

> Disease impact criteria were as follows: infectivity of the agent (person-to-person transmission potential), case fatality rate, stability in the environment and ease of decontamination, incidence of disease per 100,000 exposed persons in the worst-case release scenario, and reports of genetic modification of the agent for increased virulence.
>
> - Probability of attack criteria was [sic] designated as: global availability and ease of procurement of the agent, ease of weaponization, and historical examples of use of the agent for an attack.
> - Prevention/intervention criteria were categorized as: lack of preventability of the disease (such as by vaccination) and lack of treatability of the disease (such as by antibiotics).
> - For each of the scoring categories, a score of 0 to 2 was assigned for each category A agent as follows: 0 = no, 1 = some/low, and 2 = yes/high. The sum of these scores (of a total possible score of 20) was used to rank priority.

This is similar to the Superfund scoring system in that categorical ratings for various factors are assigned numerical scores, and the sum of the scores is used to set priorities. In neither case did the authors verify whether *additive independence* conditions hold, which are required in multiattribute value and utility theory to justify additive representations of preferences (Keeney & Raiffa, 1976). For example, an agent with a score of 2 for "lack of preventability of disease" and 0 for "lack of treatability" would have the same sum for these two factors ($2 + 0 = 2$) as an agent with "lack of preventability of disease" = 0 and "lack of treatability" = 2, or as an agent with "lack of preventability of disease" = 1 and "lack of treatability" = 1. Yet, risk managers who can completely prevent a disease ("lack of preventability of disease" = 0) might not care as much about whether it is treatable as they would if the disease could not be prevented. Likewise, in Superfund site scoring, many decisionmakers might care less about the presence of a declining species near a site that creates no exposure than near a site that creates a large, toxic exposure. Such interactions among factor scores are ignored in purely additive scoring systems.

### 2.5. Example: Threat-Vulnerability-Consequence (TVC) Risk Scores and Risk Matrices

Many organizations use numerical priority-scoring formulas such as *Risk* = *Threat* × *Vulnerability* × *Consequence*, or *Risk* = *Threat* × *Vulnerability* × *Criticality*, or *Risk* = *Threat* × *Vulnerability* × *Impact*. The Department of Homeland Security, the Department of Defense, and the armed services all use this approach to prioritize antiterrorism risk-reduction efforts (Jones & Edmonds, 2008; Mitchell & Decker, 2004; http://www.ncjrs.gov/pdffiles1/bja/210680.pdf.) The formula *Risk* = *Threat* × *Vulnerability* × *Consequence* also provides the conceptual and mathematical basis for the RAMCAP™ (risk analysis and management for critical asset protection) standard and related compliance training and software (http://www.ramcapplus.com/). Law enforcement officers have been trained to use *Risk* = *Threat* × *Vulnerability* × *Impact* scoring systems to set priorities for managing security risks at major special events (http://www.cops.usdoj.gov/files/ric/CDROMs/Planning Security / modules / 3 / module %203 %20ppt. ppt). Unfortunately, when the components on the right-hand side (e.g., *Threat*, *Vulnerability*, and *Consequence*) are correlated random variables—for example, because attackers are more likely to attack facilities with high *Vulnerability* and *Consequence*, or because larger storage facilities have higher *Vulnerability* and *Consequence* than small ones—then the product of their means differs from the mean of their product, and it is not clear what either one has to do with risk. Correct expressions require additional terms to adjust for nonzero covariances (Cox, 2008b). Similar comments apply to widely used "risk matrices" based on formulas such as *Risk* = *Frequency* × *Severity*, with the right-hand side variables assessed using ordered categories (such as high, medium, and low), and *Risk* ratings or priorities then being determined from these component ratings. In general, such risk matrices order some pairs of risks incorrectly and, in some cases, can perform even worse than setting priorities randomly (Cox, 2008a).

## 3. PRIORITIES FOR KNOWN RISK REDUCTIONS

To enable formal analysis of priority-scoring systems in a reasonably general framework, we define a *priority-setting process* as consisting of the following elements:

(1) *A set of items to be ranked or scored.* The items may be hazards, threats, customers, interventions, assets, frequency-severity pairs,

threat-vulnerability-consequence triples, threat-vulnerability-consequence remediation cost quadruples, Superfund sites, construction projects, or other objects. We will refer to them generically as "items," "hazards," "prospects," or "opportunities."

(2) *An ordered set of priority scores* that are used to compare hazards. These may be ordered categorical grades, such as "high," "medium," and "low"; nonnegative integers indicating relative priority or ranking, or nonnegative real numbers representing values of a quantitative priority index such as *risk = threat × vulnerability × consequence*, or *priority index = expected benefit of remediation/expected cost of remediation*, where the italicized variables are nonnegative numbers.

(3) *A priority-scoring rule.* A scoring rule is a mathematical function (or a procedure or algorithm implementing it) that assigns to each hazard a unique corresponding priority score. (This implies that any two hazards having identical attribute values, or identical joint distributions of attribute values, must have the same priority score.)

The priority-scoring rule determines a *priority order* in which hazards are to be addressed (possibly with some ties). Addressing a hazard is assumed to reduce risk and hence assumed to be valuable to the decisionmaker: it increases expected utility. For example, it may stochastically reduce the flow of illnesses, injuries, or fatalities resulting from a hazardous process, activity, or environment.

Although items might have multiple attributes, and value tradeoffs might make preferences among them difficult to define clearly in practice, we shall assume that the decisionmaker has perfectly clear, consistent preferences for the consequences of addressing different hazards. For example, suppose that addressing hazard $j$ reduces loss, measured on a scale such as dollars (for financial risks) or quality-adjusted life years (QALYs) (Doctor *et al*., 2004), for health risks, by an amount, $x_j$, defined as the difference between the loss if hazard $j$ is left unaddressed and the loss if hazard $j$ is addressed. Suppose that all value units (e.g., dollars or QALYs) are considered equally *intrinsically valuable*, with twice as many being worth twice as much to the decisionmaker. More generally, we assume that addressing hazards creates gains on a *measurable value scale*, satisfying standard axioms (Dyer & Sarin, 1979) that allow preferences for changes in or differences between situations, from before a hazard is addressed to after it is addressed, to be coherently ranked and compared. Let $x_j$ be the measurable value from addressing hazard $j$. We assume that the value of addressing a hazard, expressed on such a measurable value scale, depends only on its attributes and works directly with the measurable values, rather than with the underlying attributes. (The value scale need not be measured in QALYs, but thinking of such a concrete example may aid intuition.) If it costs the same amount to address any hazard, and if the resulting increases in value are known with certainty, then, *for any budget, total benefits are maximized by addressing the hazards in order of their decreasing values,* $x_j$. This provides one useful model for priority-based risk management decision making.

## 4. PRIORITIES FOR INDEPENDENT, NORMALLY DISTRIBUTED RISK REDUCTIONS

Next, suppose that the value achieved by addressing hazard $j$ is uncertain. This might happen, for example, if the quantities or potencies of hazardous chemicals stored at different waste sites are uncertain, or if the sizes of exposed populations and their susceptibilities to exposure are not known, or if the effectiveness of interventions in reducing risks is in doubt. To model priority-based risk management decisions with uncertainty about the sizes of risk-reduction opportunities, we assume that their values are random variables, and that the decisionmaker is risk-averse. For a risk-averse decisionmaker with a smooth (twice-differentiable) increasing von Neumann-Morgenstern utility function for the value attribute, the conditions in Table I are all mutually equivalent and all imply that the utility function is exponential. If one or more of these conditions is considered normatively compelling, then an exponential utility function should be used to choose among prospects with uncertain values.

The expected value of an exponential utility function for any random variable corresponds to its moment-generating function. For example, let $X_j$ represent the uncertain measurable value of addressing hazard $j$, modeled as a random variable on the value axis. Let $\mathrm{CE}(X_j)$ denote the certainty equivalent of $X_j$, i.e., the value (such as QALYs saved) received with certainty that would have the same expected utility as (or be indifferent to) random variable $X_j$. Then, if $X_j$ is normally distributed with

Let $X$ and $Y$ be any two risky prospects (random variables) measured on the intrinsic value scale. They represent the uncertain values (e.g., QALYs saved) by addressing two different hazards.

- *Strong Risk Independence:* Adding the same constant to both $X$ and $Y$ leaves their preference ordering unchanged. Thus, if $X + w$ is preferred to $Y + w$ for *some* value of the constant $w$, then $X$ is preferred to $Y$ for *all* values of $w$.
- *Risk Premium Independence:* The decisionmaker's risk premium (amount she is willing to pay to replace a prospect with its expected value) for any risky prospect depends only on the prospect. (Thus, it is independent of background levels of the value attribute.)
- *Certainty Equivalent Independence:* If a constant, $w$, is added to every possible outcome of a prospect $X$, then the certainty equivalent of the new prospect thus formed is $\mathrm{CE}(X) + w$, where $\mathrm{CE}(X)$ denotes the certainty equivalent (or "selling price" on the intrinsic value scale) of prospect $X$. (This is sometimes called the "delta property," due to Pfanzagl, 1959.) Thus, for any constant, $w$, $\mathrm{CE}(w + X) = \mathrm{CE}(X) + w$.
- *Equal Buying and Selling Prices:* For any prospect $X$ and any constant $w$, the decisionmaker is indifferent between $w + \mathrm{CE}(X) - X$ and $w + X - \mathrm{CE}(X)$.
- *No Buying Price/Selling Price Reversals:* The ranking of prospects based on their certainty equivalents (i.e., "selling prices," e.g., how many QALYs would have to be saved with certainty to offset the loss from abandoning the opportunity to save $X$ QALYs) never disagrees with their ranking based on "buying prices" (e.g., how many QALYs a decisionmaker would give up with certainty to save $X$ QALYs). (This assumes the decisionmaker is risk-averse; otherwise, the linear risk-neutral utility function $u(x) = x$ would also work.)
- *Exponential Utility:* $u(x) = 1 - e^{-kx}$.

**Table I.** Equivalent Characterizations of Exponential Utility Functions

*Source:* Dyer and Jia (1998) and Hazen and Sounderpandian (1999).

mean $E(X_j)$ and variance $\mathrm{Var}(X_j)$, it follows (from inspection of the moment-generating function for normal distributions) that its certainty equivalent is:

$$\mathrm{CE}(X_j) = E(X_j) - (k/2)\mathrm{Var}(X_j),$$

where $k$ is the coefficient of risk aversion in the exponential utility function (Infanger, 2006, p. 208).

A set of equally costly risk-reducing measures with independent, normally distributed values can be prioritized in order of decreasing $\mathrm{CE}(X_j)$ values. *For any budget, total expected utility is maximized by funding risk-reduction opportunities in order of decreasing priority* until no more can be purchased. Moreover, even if the risk-reducing measures do not have identical costs, an optimal (expected utility maximizing, given the budget) policy maximizes the sum of certainty equivalents, subject to the budget constraint. (This follows from the additivity of means and variances for independent risks. Finding an optimal subset in this case is a well-studied combinatorial optimization problem, the knapsack problem.) Thus, for any two feasible portfolios of risk-reducing measures, the one with the greater sum of certainty equivalents is preferred. Certainty equivalents therefore serve as satisfactory priority indices for identifying optimal risk-reducing investments in this case.

## 5. PRIORITY RATINGS YIELD POOR RISK MANAGEMENT STRATEGIES FOR CORRELATED RISKS

Priority-based risk management successfully maximizes the risk-reduction value (expected utility or certainty equivalent value of risk-reducing activities) of defensive investments in the special cases discussed in the preceding two sections. However, it fails to do so more generally. Selecting a best portfolio of hazards to address (or of risk-reducing measures to implement) cannot, in general, be accomplished by priority setting if uncertainties about the sizes of risks (or of risk-reduction opportunities) are correlated. Unfortunately, this is the case in many applications of practical interest. No priority rule can recommend the best portfolio (subset) of risk-reducing opportunities when the optimal strategy requires diversifying risk-reducing investments across two or more types of opportunities or when it requires coordinating correlated risk reductions from opportunities of different types (having different priority scores).

### 5.1. Example: Priority Rules Overlook Opportunities for Risk-Free Gains

A priority-setting rule that rates each uncertain hazard based on its own attributes only, as all the real priority-scoring systems in Section 1 do, will, in general, be unable to recommend an optimal subset of correlated risk-reducing opportunities. For example, any risk-averse decisionmaker prefers a single random draw from a normal distribution with mean 1 and variance 1, denoted $N(1, 1)$, to a single draw from normal distribution, $N(1, 2)$, having mean 1 but variance 2. Therefore, a scoring rule would assign a higher priority to draws from $N(1, 1)$ than to draws from $N(1, 2)$. But suppose that $X$ and $Y$ are two $N(1, 2)$ random variables that are perfectly negatively correlated with $Y = 2 - X$. (This might happen, for example, if effects depend only on the sum of $X$ and $Y$, which has a known value of 2, but the relative contributions of $X$ and $Y$ to their sum are uncertain.) Then, drawing once from $X$ and once from $Y$ (each of which is $N(1, 2)$) would yield a sure gain of 2. Any risk-averse decisionmaker prefers this sure gain to two draws from $N(1, 1)$. Unfortunately, any priority rule that ignores correlations among opportunities would miss this possibility of constructing a risk-free gain by putting $X$ and $Y$ in the same portfolio, as it would always assign draws from $N(1, 1)$ a higher priority than draws from $N(1, 2)$.

This example shows that priority-setting rules can recommend dominated portfolios, such as allocating all resources to risk reductions drawn from $N(1, 1)$ instead of pairing negatively correlated $N(1, 2)$ risk reductions, because *they cannot describe optimal portfolios that depend on correlations among risk-reducing opportunities*, rather than on the attributes of the individual opportunities. The next example shows that priority rules can, in principle, not only recommend a dominated decision but, in some cases, also recommend the worst possible decision.

### 5.2. Example: Priority Setting Can Recommend the Worst Possible Resource Allocation

#### *5.2.1. Setting*

Suppose that an environmental risk manager must decide how to allocate scarce resources to remediate a large number of potentially hazardous sites. There are two main types of sites. Hazards at type A sites arise primarily from relatively long, thin chrysotile asbestos fibers. Hazards at type B sites arise from somewhat shorter and thicker amphibole asbestos fibers. The risk manager is uncertain about their relative potencies but knows that removing *mixtures* of approximately equal parts of the chrysotile and amphibole fibers significantly reduces risks of lung cancer and mesothelioma in surrounding populations. She believes that the following two hypotheses are plausible, but is uncertain about their respective probabilities. (This is intended for purposes of a simple illustration only, not as a realistic risk model.)

- H1: Relative risk from a type A site is 0; relative risk from a type B site is 2 (compared with the risk from a hypothetical site with equal mixtures of chrysotile and amphibole fibers, which we define as 1). This hypothesis implies that all risk is from the amphibole fibers.
- H2: Relative risk from a type A site is 2; relative risk from a type B site is 0. This hypothesis implies that all risk is from the chrysotile fibers.

For purposes of illustration only, we assume that only these two hypotheses are considered plausible, although clearly others (especially, that the two types of fibers are equally potent) would be considered in reality.

#### *5.2.2. Problem*

If the risk manager can afford to clean $N = 10$ sites, then how should she allocate them between type A and type B sites? Assume that she is risk-averse, and that more than 10 sites of each type are available.

#### *5.2.3. Solution*

If the risk manager cleans $x$ type A sites and $(N - x)$ type B sites, then the total expected utility from cleaned sites is: $pu(N - x) + (1 - p)u(x)$. Here, $p$ denotes the probability that hypothesis H1 is correct, $1 - p$ is the probability that H2 is correct, $N = 10$ is the total number of sites that can be cleaned, and $u(x)$ is the utility of cleaning $x$ sites with relative risk of 2 per site cleaned. For *any* risk-averse (concave) utility function $u(x)$, and for *any* value of $p$ between 0 and 1, Jensen's inequality implies that expected utility is maximized for some $x$ strictly between 0 and $N$. For example, if $u(x) = x^{0.5}$ and $p = 0.5$, then $x = 5$ maximizes expected utility. The worst possible decision (minimizing expected utility) is to allocate all

resources to only one type of site (either type A or type B). Yet, this is precisely what a priority system that assigns one type a higher priority than the other must recommend. Hence, in this case, any possible priority order (either giving type A sites precedence over type B sites or *vice versa*, perhaps depending on whether $p < 0.5$) will recommend a subset of sites that has lower expected utility than even a randomly selected subset of sites. The best subset (e.g., 5 type A sites and 5 type B sites, if $p = 0.5$) can easily be constructed by optimization if $p$ is known. But even if both $p$ and $u(x)$ are unknown, it is clear that *a priority order is the worst possible decision rule*.

### 5.3. Example: Priority Setting Ignores Opportunities for Coordinated Defenses

#### *5.3.1. Setting*

Suppose that an information security risk manager can purchase either of two types of security upgrades for each of 100 web servers. Type A prevents undetected unauthorized access to a web server, and type B prevents unauthorized execution of arbitrary code with the privileges of the web server, even if the web server is accessed. (For examples of real-world historical vulnerabilities in an Apache web server, see http://www.first.org/cvss/cvss-guide.html#i1.2.) For simplicity, suppose that installing a type A upgrade reduces the annual incidence of successful attacks via web servers from 0.03 to 0.02 per web server year, and that installing a type B upgrade reduces it from 0.03 to 0.025. Installing both reduces the average annual rate of successful attacks via these machines from 0.03 to 0.

#### *5.3.2. Problem*

If the security risk manager can afford 100 security upgrades (of either type), what investment strategy for reducing the average annual frequency of successful attacks would be recommended based on: (1) priority ranking of options A and B, and (2) minimization of remaining risk? (Assume that the frequency of attempted attacks remains constant, because hackers only discover the defenses of a web server when they attempt to compromise it.)

#### *5.3.3. Solution*

(1) A vulnerability-scoring system could assign top priority to installing a type A upgrade on each of the 100 web servers because a type A upgrade achieves a larger reduction in the vulnerability score of each server than a type B upgrade. Following this recommendation would leave a residual risk of $0.02 \times 100 = 2$ expected successful attack per year. (2) In contrast, a risk-minimizing budget allocation installs both A and B upgrades on each of 50 machines, leaving 50 machines unprotected. The residual risk is then $0.03 \times 50 = 1.5$ expected successful attack per year, less than that from giving A priority over B.

#### *5.3.4. Comment*

In this example, a scoring system that considers interactions among vulnerability-reducing activities could give "install A and B" a higher priority for each server than either "install A" or "install B." But most deployed scoring systems do not encourage consideration of interactions among vulnerabilities or among vulnerability-reducing countermeasures. In many applications, doing so could lead to combinatorial explosion. (For example, the guidance for CVSS 2.0 offers this advice: "SCORING TIP #1: Vulnerability scoring should not take into account any interaction with other vulnerabilities. That is, each vulnerability should be scored independently." http://www.first.org/cvss/cvss-guide.html#i1.2.)

### 5.4. Example: Priority Rules Ignore Aversion to Large-Scale Uncertainties

#### *5.4.1. Setting*

A bioterrorism risk manager must choose which of two defensive programs to implement this year: (A) a prevention program (e.g., vaccination) that, *if it works*, will reduce the risk of fatal infection from 10% to 0% for each affected person in the event of a bioterrorism attack with a certain agent; or (B) a treatment program (e.g., stockpiling an antibiotic) that will reduce the risk of mortality from 10% to 5% for each affected individual in the event of such an attack. For simplicity, suppose that program A will prevent either $N$ expected deaths (if it works) or none (if it does not) following an attack, and that its success probability is $p$. Program B prevents $0.5N$ expected deaths with certainty, leaving $0.5N$ remaining expected deaths in the event of an attack.

#### *5.4.2. Problem*

(1) For a risk-averse decisionmaker with utility function $u(x) = 1 - e^{-kx}$, where $x$ is the number

of expected deaths prevented, which risk-reduction measure, A or B, is preferable? (Express the answer as a function of $p$, $k$, and $N$.) (2) How does this compare with the results of a priority-ranking system, for $p = 0.8$ and $k = 1$?

### *5.4.3. Solution*

(1) The expected utility of risk reduction is $pu(N) = p(1 - e^{-kN})$ for program A and $u(0.5N) = 1 - e^{-0.5kN}$ for program B. Program A is preferable to program B if and only if $p(1 - e^{-kN}) > 1 - e^{-0.5kN}$ or, equivalently, if $p > (1 - e^{-0.5kN})/(1 - e^{-kN})$. For example, if $kN = 1$, then $p$ must be at least 62.2% to make A preferable to B. If $kN = 10$, then $p$ must be at least 99.3% to make A preferable to B. (2) If the probability that program A will work is $p = 0.8$ and the coefficient of absolute risk aversion is $k = 1$, then A is preferred to B for $N = 1$ or 2, and B is preferred to A for $N \geq 3$. In this case, diversification is not an issue (i.e., either A or B is definitely preferable, depending on the value of $N$.) However, *no priority ranking of interventions A and B is best for both $N = 2$ and $N = 3$.* The reason is that a risk-averse decisionmaker who prefers A to B for small $N$ prefers B to A for larger $N$. Any priority-scoring system that ranks either one of A or B above the other, and that is not sensitive to $N$, will recommend the less valuable decision for some values of $N$. In practice, most scoring systems use qualitative or ordered categorical descriptions that are not sensitive to quantitative details such as $N$. (For example, the CVSS rates "collateral damage potential," which scores "potential for loss of life, physical assets, productivity or revenue," as high if "[a] successful exploit of this vulnerability may result in catastrophic physical or property damage and loss. Or, there may be a catastrophic loss of revenue or productivity" (http://www.first.org/cvss/cvss-guide.html#i1.2). Such a qualitative description does not discriminate between $N = 2$ and $N = 3$.)

### *5.4.4. Discussion*

Precisely analogous examples hold for consumer credit risk-reducing interventions, information security, homeland security, and other applications in which the success of some proposed interventions is uncertain. Suppose that intervention A reduces the average rate of successful attacks per target (e.g., secure facility or web server) per year from 10% to 0% *if it works*, whereas intervention B reduces the rate from 10% to 5% with certainty. The probability that A will work (i.e., that an attacker cannot circumvent it) is $p$. If the choice between A and B affects $N$ similar targets, then, by analogy to the above example, a risk-averse risk manager should prefer A to B for sufficiently small $N$ and B to A for larger values of $N$. Any priority system that is applied to a small number of targets at a time (possibly only 1, by the target's owner, operator, or security manager) will then consistently recommend A, even though B should be preferred when the complete set of $N$ targets is considered. That scoring systems are blind to the total number of similar targets that they are applied to (i.e., to the scale of application) can lead to excessively high-risk exposures arising from large-scale application of priorities that hold for small numbers of targets, but that should be reversed for larger numbers of targets.

## 6. DISCUSSION AND CONCLUSIONS

Applied risk analysis is in a curious state today. Highly effective optimization methods for selecting subsets of risk-reducing investments to maximize the value of risk reductions achieved for a given budget are readily available. They can draw on a rich and deep set of technical methods developed in financial risk analysis and operations research over the past half-century. Yet, these methods are having little or no impact on the management of some of the world's most critical risks. Instead, extremely simplistic priority-setting rules and scoring systems are being widely used to set priorities and allocate resources in important practical risk management applications. Scoring systems are being used in important real-world applications as diverse as Superfund site cleanups, computer and IT security vulnerability assessment, counterterrorism, military asset protection, and risk matrix systems (used in everything from designing and defending federal buildings and facilities, to managing construction project and infrastructure risks, to regulating risks of financial and business enterprises) (Cox, 2008a). Yet, these risk-scoring systems achieve less value of risk reduction than could easily be obtained if resources were allocated by other methods (including randomized decision making, in extreme cases.)

The requirements that scoring systems must meet before being adopted and recommended in standards are not very stringent. In the applications examined in this article, there appears to be no requirement that risk-scoring systems should

produce effective risk management decisions (or even that they should not produce the lowest-value decision possible) before they are standardized for widespread use. In all of the applications mentioned, the common elements found in multiple risky systems create correlated vulnerabilities, criticalities, consequences, or threats. Priority lists do not generally produce effective risk management decisions in such settings. Applying investment portfolio optimization principles (such as optimal diversification, consideration of risk aversion, and exploitation of correlations among risk reductions from different activities) can create better portfolios of risk-reducing activities in these situations than any that can be expressed by priority scores.

In summary, risk priority scoring systems, although widely used (and even required in many current regulations and standards), ignore essential information about correlations among risks. This information typically consists of noting common elements across multiple targets (e.g., common vulnerabilities). These common features induce common, or strongly positively correlated, uncertainties about the effectiveness of different risk-reducing measures. It is easy to use this information, in conjunction with well-known decision analysis and optimization techniques, to develop more valuable risk-reduction strategies, for any given risk management budget, than can be expressed by a priority list. Thus, there appears to be abundant opportunity to improve the productivity of current risk-reducing efforts in many important applications using already well-understood optimization methods.

Nothing in this note is intended to be new or surprising to experts in decision and risk analysis. Techniques for optimizing investments in risk-reducing (and/or benefit-producing) interventions have been extensively developed in operations research and management science for decades. What is perhaps startling is that these methods are so little exploited in current risk assessment and risk management systems. Risk priority scores can never do better (and often do much worse) than optimization methods in identifying valuable risk-reducing strategies. Perhaps it is time to stop using risk priority scores to manage correlated risks, recognizing that they often produce simple but wrong answers. Optimization techniques that consider dependencies among risk-reducing interventions for multiple targets should be used instead.

## REFERENCES


Cox LA Jr. What's wrong with risk matrices? Risk Analysis 2008a; 28(2):497–512.

Cox LA Jr. Some limitations of "*Risk = Threat × Vulnerability × Consequence*" for risk analysis of terrorist attacks. Risk Analysis, 2008b; 28(6):1749–1762.

Doctor JN, Bleichrodt H, Miyamoto J, Temkin NR, Dikmen S. A new and more robust test of QALYs. Journal of Health Economics, 2004; 23(2):353–367.

Dyer JS, Jia J. Preference conditions for utility models: A risk-value perspective. Annals of Operations Research, 1998; 80(1):167–182. Available at: http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.39.5480, Accessed March 14, 2009.

Dyer JS, Sarin RK. Measurable multiattribute value functions. Operations Research 1979; 27(4):810–822.

Hazen, G., Sounderpandian J. Lottery acquisition versus information acquisition: Price and preference reversals. Journal of Risk and Uncertainty, 1999;18(2):125–136.

Infanger G. Dynamic asset allocation strategies using a stochastic dynamic programming approach. Pp. 200–205 in Zenios SA, Ziemba WT (eds). Handbook of Assets and Liability Management, Vol. 1. New York: North Holland, 2006.

Jones P, Edmonds Y. Risk-based strategies for allocating resources in a constrained environment. Journal of Homeland Security, 2008. www.homelandsecurity.org/newjournal/Articles/displayArticle2.asp?article=171, Accessed March 14, 2009.

Keeney RL, Raiffa H. Decisions with Multiple Objectives: Preferences and Value Trade-Offs. New York: Wiley, 1976.

MacIntyre CR, Seccull A, Lane JM, Plant A. Development of a risk-priority score for category A bioterrorism agents as an aid for public health policy. Military Medicine, 2006; 171(7):589–594.

Mitchell C, Decker C. Applying risk-based decision-making methods and tools to U.S. navy antiterrorism capabilities. Journal of Homeland Security, 2004. www.homelandsecurity.org/journal/Articles/Mitchell_Decker.html