DOI: 10.1111/risa.14153

ORIGINAL ARTICLE

# A classification system for characterizing the integrity and quality of evidence in risk studies

**Shital Thekdi**[1] | **Terje Aven**[2]

[1]Robins School of Business, University of Richmond, Richmond, Virginia, USA

[2]University of Stavanger, Stavanger, Norway

**Correspondence**
Shital Thekdi, Robins School of Business, University of Richmond, 102 UR Drive, Richmond, VA 23103, USA.
Email: sthekdi@richmond.edu

**Abstract**

Risk management requires a balance between knowledge and values. Knowledge consists of justified beliefs and evidence, with evidence including data, assumptions, and models. While quality and integrity of evidence are valued in the sciences, risk science involves uncertainty, which suggests that evidence can be incomplete or imperfect. The use of inappropriate evidence can invalidate risk studies and contribute to misinformation and poor risk management decisions. Additionally, the interpretation of quality and integrity of evidence may vary by the risk study mission, decision-maker values, and stakeholder needs. While risk science has developed standards for risk studies, there remains a lack of clarity for how to demonstrate quality and integrity of evidence, recognizing that evidence can be presented in many formats (e.g., data, ideas, and theories), be leveraged at various stages of a risk study (e.g., hypotheses, analyses, and communication), and involve differing expectations across stakeholders. This study develops and presents a classification system to evaluate quality and integrity of evidence that is based on current risk science guidance, best practices from non-risk disciplines, and lessons learned from recent risk events. The classification system is demonstrated on a cyber-security application. This study will be of interest to risk researchers, risk professionals, and data analysts.

**KEYWORDS**
evidence, risk analysis, risk management, validation

## 1 | INTRODUCTION

The strength of risk management decisions is only as strong as the knowledge that forms the basis of those decisions. This knowledge is composed of justified beliefs and evidence. The Cambridge dictionary defines evidence as "One or more reasons for believing that something is or is not true."

Across domains, data and other types of information are being used to form evidence to support stances on various topics. The legal field relies on evidence to determine guilt beyond a reasonable doubt. The statistical field uses evidence for purposes of inferring characteristics about populations using information gathered from sample data. The theoretical math field relies on evidence to prove propositions.

While objectivity is ideal, there is subjectivity in all aspects of analysis and science. For example, students are commonly attuned to "Garbage in Garbage Out," referring to mathematical outputs based on poor assumptions or data. There is also the famous quote, "All models are wrong", (Box, 1976), pointing toward recognizing that no model can be considered entirely accurate and reliable for real systems. Instead, models can be used to serve a purpose, requiring the analyst to understand the implications of model limitations. However, there is no standard for what qualities of objectivity are sufficient for various types of risk problems.

In many basic analysis contexts, the pursuit of objectivity often comes in the form of data, with varying levels of data integrity, that can be used to form a stance. However, evidence can also include qualitative elements such as

opinions and sentiments that could be measured using survey instruments and social media analytics. Risk applications also can rely on aspects of public opinion or expert opinion and other types of knowledge that often cannot be quantified in meaningful mathematical ways.

As a motivating illustration, consider the recent COVID-19 pandemic. The risk event prompted entities to conduct analyses to determine the origin and consider how to address risk due to related and future contagious disease events. Determining the origin and cause of the COVID-19 pandemic involved many hypotheses with varying levels of supporting data and information. Hypotheses included a zoonotic source, thought to spread from a seafood market (Shereen et al., 2020) and artificial origins (Chaturvedi et al., 2020; Rogin, 2020). Some hypotheses were more heavily promoted by various news outlets, with disagreement regarding which hypotheses were credible versus those deemed conspiracy theories (Imhoff & Lambert, 2020; Uscinski et al., 2020). Despite the sentiments from the court of public opinion, evidence continued to be gathered as the pandemic continued. An ongoing investigation led to concerns of credibility for gathered intelligence, inputs from experts, and other stakeholders (Gordon et al., 2021). Thus, the credibility of the various COVID-19 origin hypotheses was heavily dependent on the credibility of the gathered evidence, with some entities putting greater weight on some types of evidence versus others. Without a systematic framework to weigh the integrity of evidence for evaluating hypotheses, there remained a possibility for these types of exercises to be influenced by limitations of human perception and also explicit and implicit biases (Cori et al., 2020; Garcia-Alamino, 2020), and political motivation.

Research and practice in the domain of data analysis has developed momentum in addressing issues related to evidence integrity, such as by exposing methods for misleading audiences using statistics and data visualization (Huff, 1993). Students are often taught about due diligence for investigating evidence and assumptions, but demonstrating transparency related to these issues can often be neglected or forgotten (Barker & Shaw, 2015). This issue is additionally complex as practices for visualizing and communicating with data change over time, particularly as analysts often communicate across multiple channels of rapidly changing platforms, such as print, websites, social media, and video and interact with wide varieties of stakeholders with rapidly changing ideologies and expectations.

The risk field has also conducted significant work in defining guidance and standards for how to conduct and validate risk studies. For example, recent literature defines criteria that can be used to understand and improve the quality of a risk study (Lathrop & Ezell, 2017; SRA, 2021). The topic of validation is a major concern, defining validation as related to the "success at 'measuring' what one sets out to 'measure' in the analysis" (Aven, 2017; Aven & Heide, 2009; Aven & Zio, 2018; Rosqvist, 2010; SRA, 2020). The topic of validation for quantitative risk analysis has been well studied, covering areas of conceptual validity, in relation to various scientific foundations, and pragmatic validity (Goerlandt et al., 2017). This existing literature promotes the need for examining assumptions and limitations and promoting transparency and consistency within risk studies.

Considerable work has also been conducted within the framework of validation and verification (V&V). Validation is about ensuring that one measures or describes what one intends to measure or describe. Validation is also about ensuring study elements meet the expectations of stakeholders, such as employees, customers, and communities. Verification is about ensuring study elements meet relevant specifications; for example, the data used are in fact those referred to or a simulation model is accurately implemented. Various disciplines have developed guidance on V&V (AIAA, 1998; Oberkampf et al., 2004; Pace, 2004; Sargent, 2013). The risk field in particular has witnessed a remarkable culture shift with regard to how to address issues of accuracy and knowledge within V&V activities. The previously held focus on accuracy in risk estimation has been replaced by a more holistic expectation to characterize knowledge and lack of knowledge (Aven & Zio, 2018).

While the existing V&V literature is comprehensive, there are gaps that need to be addressed. Generally, the existing literature calls for guidance and standards to be met with regard to aspects such as data integrity, transparency of assumptions, and statement of biases. However, in practice, it is often impossible to meet the evidence integrity requirements enforced by various stakeholders. For example, in relation to the COVID-19 example presented above, the reality is that evidence can often be very poor but still critical for risk-based decision-making. Additionally, it is often unclear how and when lapses in evidence integrity are large enough to invalidate the risk study. In relation to the COVID-19 example, there was very consistent controversy over the integrity and quality of evidence, as witnessed through academic paper retractions (Ledford & Van Noorden, 2020). While this type of issue exists in many fields, this is a particular problem for risk science, in the sense that it involves the study of uncertainty as it relates to events and dynamics that are often nearly impossible to predict or accurately understand.

Risk researchers and practitioners also have to be careful in recognizing that over-reliance on even the most accurate past data or on the most computationally intensive predictive models can be misleading, and lead to a false confidence in studies of poorly understood phenomena, thereby promoting misleadingly poor risk management decisions. For example, consider data-intensive modeling tasks of predicting future cybersecurity incidents using past data. Due to the fast-paced evolution of technologies and uses of technology, studying frequency of past events might appear to be based on vast quantities of data, but that data could be largely irrelevant in predicting future incidents. As another example, consider epidemiological models used to predict the spread of COVID-19. Data on the presence of COVID-19 are based on evidence emerging from scientific test data, but these data are impacted by test errors, biases in availability of tests, and a largely poor understanding of the virus, causing even the most widely

accepted modeling methods to be questioned. Thus, it is critical to acknowledge that in some situations, insufficient evidence can invalidate the entire risk study, but in other situations, insufficient evidence can be innocuous. There is also a need to acknowledge that this is really a judgment that is dependent on many factors, such as the risk application, stage of the risk study, activity being conducted, values, and mission of decision-makers and stakeholders.

There is need for a framework to evaluate the quality and integrity of evidence, considering the issues described above. This framework should be based on the most important theoretical insights from the literature and also innovate those insights with additional considerations. Those additional considerations include lessons learned from recent risk events, the changing landscape of how evidence can be misrepresented or manipulated, and the practical need for issues of evidence quality to be accessible and understandable to a variety of domains and decision-makers who are learning the importance of high-integrity risk science practices. This is a particularly challenging task because the risk field studies the unknown. Statements about uncertainties may not be proven or disproven. There is a need for the risk science field to investigate the concept of evidence, recognizing that improper use of evidence can be an act of risk malpractice with the potential to cause great harm.

This study presents a classification system for understanding the quality and integrity of evidence as it relates to risk management decision-making. Section 2 discusses the topic of qualitative and quantitative evidence that can be used for decision-making, following up the above discussion. Section 3 develops a framework to understand the characteristics of evidence and the applicability of that evidence toward various risk contexts. Section 4 demonstrates the framework on a case study involving cybersecurity. Section 5 provides conclusions as they relate to the applicability of the framework to various risk problems and future directions of research.

## 2 | QUALITATIVE AND QUANTITATIVE EVIDENCE FOR DECISION-MAKING

This section describes evidence as a stand-alone concept and as it relates to the risk discipline. Section 2.1 explains the origins of evidence related to various disciplines, such as law, science, and mathematics. Section 2.2 explains how evidence fits in with fundamental risk concepts.

### 2.1 | Evidence across domains

Evidence is the cornerstone of the scientific approach (Andersen & Hepburn, 2015). Researchers across domains begin by forming a hypothesis containing a proposed idea or assumption. Then, they conduct studies, collect data and analyze, forming evidence. The evidence is then used to support or refute the original hypothesis.

The most common form of evidence is based on measurements. Often, these measurements come in the form of data such as related to surveys, records, sensors, transactions, social media posts, journalism, photos, videos, and satellite imagery. These data can be quantitative, but also qualitative by reflecting thoughts and opinions. These data are assumed to be relevant to the studied problem and the methods used to test the hypotheses. For example, if a statistical study is conducted, the data are assumed to be representative of the total population and free of biases that could invalidate basic assumptions in the selected mathematical method. The analyst's responsibility is to carefully consider the potential for inaccuracy or violation of assumptions, though no systematic standards exist to determine whether the analyst has adequately fulfilled that responsibility.

More specifically, in relation to statistics, the hypothesis test provides an example for the use of evidence. The testing consists of collecting data and performing a mathematical manipulation of that data, resulting in evidence that can either reject or fail to reject the hypothesis. The assumptions of the hypothesis intend to demonstrate the truth of some statements beyond a reasonable doubt, thus carrying parallels with the legal definitions of evidence.

The legal field relies heavily on evidence for making arguments and decisions in a legal justice context . The most commonly discussed use of evidence consists of evidence used as data or information for making judicial decisions. In particular, the legal field makes an important distinction between direct and circumstantial evidence (ABA, 2021). Direct evidence can consist of a testimonial offering first-hand information, potentially from witnesses, experts, or physical objects. Circumstantial evidence could include testimonials or tangible facts, but this evidence is based on inference, suggesting a weaker degree of integrity. There is value in using direct evidence in a legal context, just as there is value of strong evidence integrity in a general analysis context.

However, even with careful attention to the integrity of evidence in a legal process, errors can occur. In the U.S. legal appeals process, there is the use of a substantial evidence standard, defined as: "…more than a mere scintilla; it means such relevant evidence as a reasonable mind might accept as adequate to support a conclusion" (U.S. Courts, 2021). Using these standards that rely on "reasonable mind" judgments, the legal field can utilize various precedents to make decisions, though these standards are relatively more subjective than in a scientific context.

There is the notion of evidence-based practice in healthcare, which promotes consideration of evidence from research and other context-specific factors when making decisions about patient care (Rycroft-Malone et al., 2004). Within this practice is the topic of "quality of evidence" that can be used to assist with decision-making. In this type of setting, one can see a comparison to risk, as physicians are charged with helping patients manage their health risks. Physicians recognize that no two situations are the same, such that the context, relevant data, studies, and supplemental information

can differ drastically with different patients. Factors that can decrease the quality of evidence include study limitations, inconsistent results, indirect evidence (related to circumstantial evidence in a legal field), and biases (Guyatt et al., 2008).

In the natural sciences, one can consider evidence related to various scientific laws, which describe an observed phenomenon. Examples of scientific laws are the law of conservation of mass, Ohm's law, and the commutative law of addition. These laws relate to phenomena that are closely linked to hypotheses that are supported by evidence. These laws do not explain why these phenomena exist, or explain any aspects of causation, only that they have been verified using the evidence. These laws are often the result of scientific experiments that have been repeated and confirmed by a critical mass of researchers. New evidence may refute or redefine these laws, just as evidence is used to support or refute a hypothesis. Thus, evidence has a critical role in this scientific process.

## 2.2 | Evidence as related to a risk context

Knowledge is of different types, and it is common in risk contexts to distinguish between knowledge in a broad sense and a narrow sense (Aven, 2013). According to the latter interpretation, knowledge is ''justified beliefs'' (SRA, 2018). The former interpretation adds data, information, modeling insights, test results, analysis results, and so forth. Thus

$$\begin{aligned}\text{Knowledge (broadsense)} &= \text{Knowledge (narrowsense)}\\ &+ \text{Data, information, assumptions,}\\ &\quad \text{modeling, testing, argumentation, etc.}\end{aligned}$$

Evidence as a concept captures the basis for a belief or statement in the form of data, information, modeling insights, test results, analysis results, and so forth, and consequently, we can schematically write:

$$\text{Knowledge (broadsense)} = \text{Knowledge (narrowsense)} + \text{Evidence.}$$

Knowledge (narrow sense) can be informed by previously encountered evidence. Then, the knowledge (broad sense) can be updated based on newly presented evidence. For the purpose of this study, we assume that a risk study consists of knowledge (narrow sense) that is static at the start of the risk study; then, that knowledge is updated over the course of the risk study.

Consider a risk assessment where the event of interest is $A$. It could represent an accident in a safety context or a violation of legal standards in a court setting. The risk assessment is informed by the probability $P(A|K)$, where $K$ is the knowledge (in a broad sense) supporting the probability assignment. Following (1), this probability can be rewritten as $P(A|B,E)$, where $B$ is a set of justified beliefs, and $E$ is the evidence. The evidence can have varying levels of integrity.

A risk assessment can produce evidence in the form of a risk characterization of relevance for making a judgment about the statement $A$ being true. A court can use a doctor's risk assessment as evidence. Evidence is used both for producing risk characterization and supporting decision-making, see Figure 1.

A more detailed model is presented in Figure 2, using the discussion from Section 2.1, drawing parallels across the various fields in the usage of evidence to form judgments and decisions. One begins by creating a hypothesis. Current evidence and knowledge are collected and described. Then, the analysis is conducted to clarify issues and strengthen the knowledge. Then, judgments are made concerning the correctness of the hypothesis, and final decisions are made.

Figure 2 shows that evidence is not to be treated as a standalone aspect. It is instead assumed to inform every element of the process shown above. For example, in a risk context, one can view the output of Step 3: Analysis as evidence that can inform Step 4: Management review and judgment. Similarly, the risk characterization is informed by the evidence in the form of data, and so forth. The output of Step 5: Decisions, such as risk communications, can also be viewed as evidence that can inform hypotheses of individual stakeholders who are recipients of the communications. Additional discussion about Figure 2 will follow Section 3.

Because evidence used at a particular stage of the process shown in Figure 1 can impact the integrity of work conducted in downstream stages, one can see the critical importance of reviewing evidence integrity at each stage. Deficiencies at each stage should be addressed or acknowledged before conducting later stages. This shows the motivation for a systematic framework to understand evidence integrity, as discussed in Section 3.

# 3 | PROPOSED CLASSIFICATION SYSTEM

This section develops a proposed framework to characterize the overall integrity and quality of a risk study as it relates to four aspects of the process described in Figure 2. In general terms, integrity refers in this study to the reliability and trustworthiness of the evidence, whereas quality is associated with excellence, conformance to requirements, and meeting relevant stakeholders' expectations. For example, integrity could be illustrated by using data from a qualified source, and quality by using analysis methods that have been certified. There is strong overlap between the two terms, and in the study, we commonly refer to the pair integrity and quality. In the following analysis, criteria will be presented and discussed for how to identify and characterize integrity and quality of the evidence.

The characterizations presented in this section are based on principles relating to the mathematical and legal fields,

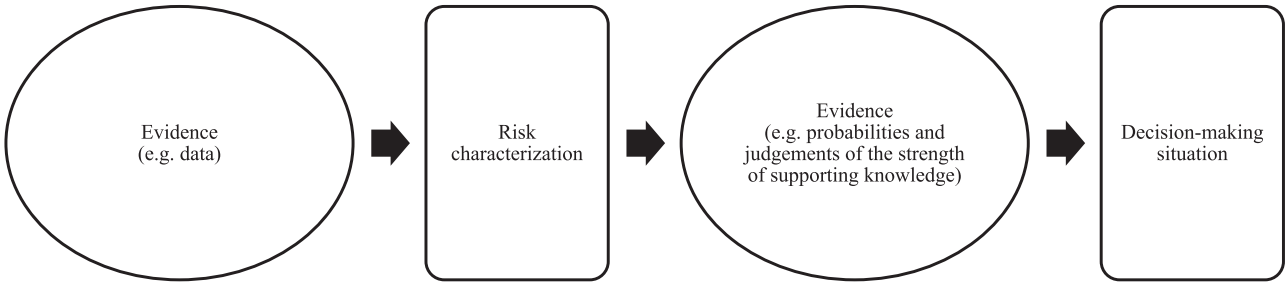


**FIGURE 1** Risk process for using evidence in decision-making activities.

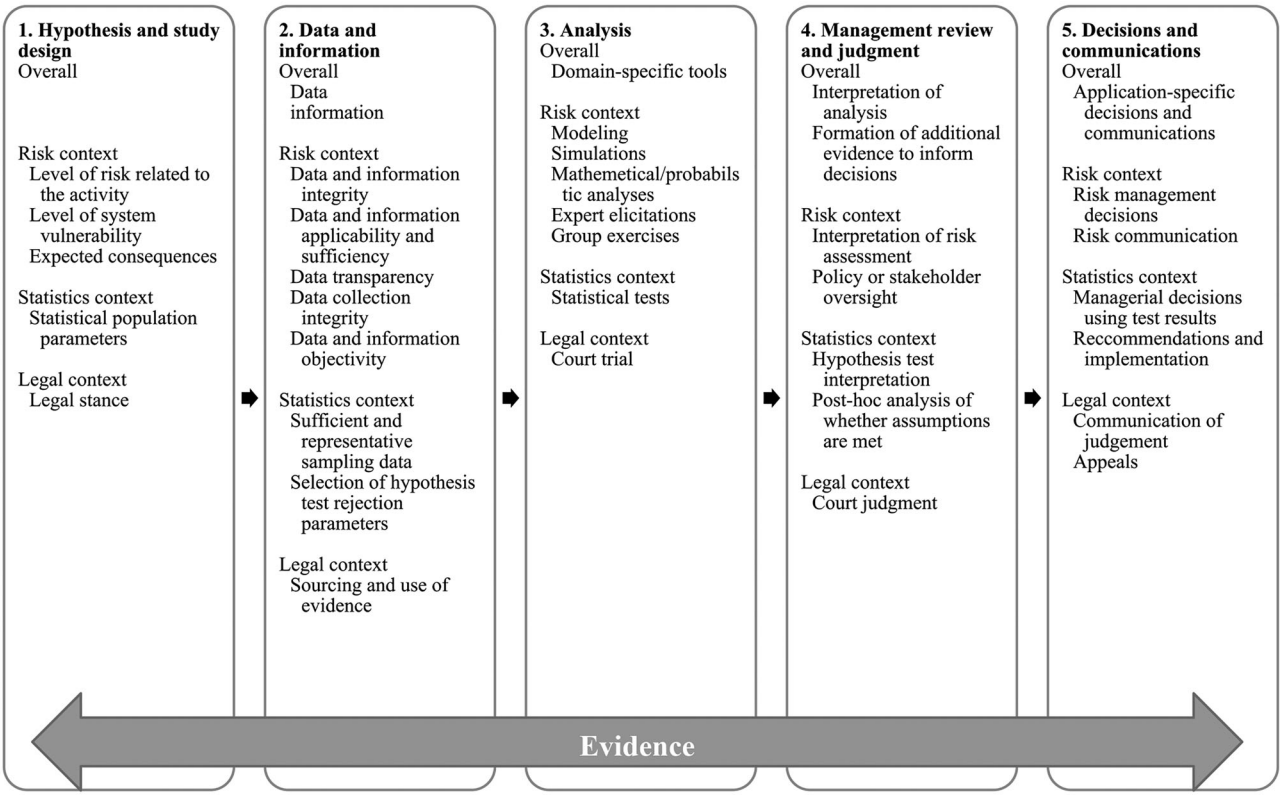


**FIGURE 2** Generic process for using evidence as a basis for judgment.

the most recently defined core principles of risk (SRA, 2020), theoretical insights from existing research, and lessons learned from recent examples of risk practice, as described in earlier sections of this study.

This classification system could be used by a single analyst or in a group setting consisting of risk-teams that include a variety of stakeholders. When this framework is used in a team setting, there are potential issues that could arise. As with any decision-making exercise, there is potential for bias or "gaming" as various parties may be subject to cognitive biases, conflicts of interest, or other types of biases when providing their inputs. It would be necessary to require team members to disclose any conflicts of interest. The lead risk analyst should also be trained on addressing potential biases, as discussed by Kahneman et al. (2011).

While the framework demonstration uses a qualitative high, medium, low rating, scoring systems can be used, such as a 1–100 scale. The analyst may choose to assign their own single scores using information gathered from focus groups, seek quantitative scores that can be evaluated using averages or medians, or adapt consensus-building methods (Landeta, 2006; McMillan et al., 2016). The analyst should also ensure that any information gathering or consensus-building mechanisms are pre-determined and ensure that the pre-determined process is implemented as planned.

Every context and application area may have different concerns when interpreting the results of this framework. As presented, no characteristic is more important than others. There is also no proposed weighting system to aggregate the results. Instead, the risk team should collectively study each

**TABLE 1** Taxonomy of overall integrity and quality for a risk study

| Characteristic | | Integrity and quality (high, medium, low, N/A) |
|---|---|---|
| O.1: Alignment with risk science | • Clear understanding of risk science fundamentals<br>• Clear hypothesis (e.g., to characterize risk and understand vulnerability) | – |
| O.2: Alignment of study design and execution | • Process for monitoring to ensure study aligns with study design<br>• Process for accountability if there are deviations from the study design | – |
| O.3: Expertise | • Qualifications of the assessor<br>• Certifications | – |
| O.4: Relevance and applicability of study design | • Hypothesis is justified (e.g., clear understanding of how hypothesis is formulated; reflects all relevant issues)<br>• Design is informed by understanding of the studied system<br>• Design is informed by previously used and accepted risk studies | – |
| O.5: Biases | • No biases in knowledge or information selected for use in the study<br>• No biases of team members influencing the study design same here | – |
| O.6: Approach | • Transparent and reproducible approach<br>• Approach that can be explained or understood by stakeholders | – |

characteristic and the overall results and then identify the highest-priority needs and actions.

Section 3.1 discusses characteristics of an overall risk study, including the formation of hypotheses and study design (Step 1). Section 3.2 describes characteristics of knowledge and evidence used for the risk study (Step 2). Section 3.3 describes characteristics of analysis (Step 3). Section 3.4 describes characteristics of judgment and decisions (Steps 4 and 5).

## 3.1 | Overall risk study

Table 1 shows characteristics of overall integrity and quality of a risk study related to the formation of hypotheses and the overall design. These characteristics involve many high-level aspects, such as training of the research team, overall consideration of uncertainties, potential biases, and the overall approach. Each characteristic is given an integrity rating of high, medium, or low. This integrity metric measures the confidence in the research team to meet the characteristic described throughout the study. There is no inherent meaning of high, medium, or low beside a relative assessment taken from the perspective of various decision-makers and stakeholders. A summary of the various elements is presented in the following section, with further details discussed in the case example in Section 4.

O.1 describes the need to create a hypothesis and design intending to follow risk science principles. The study should have clear definitions in line with risk science terminologies and ensure documentation demonstrates adherence to risk science concepts. There may be cases in which not all team members are trained in risk science, but this characteristic is in reference to the study, not the individuals involved. There should be a clear hypothesis to be addressed in the risk study.

O.2 refers to ensuring the risk study is executed in accordance with the plan and any applicable standards. While the study has not been implemented at this stage, the team should ensure that processes are in place to monitor the execution of the study. There should be processes for monitoring whether the study execution aligns with the study design, and accountability if there are deviations from the study design.

O.3. refers to the expertise of the assessment team. The assessors should be qualified for conducting a risk study. Qualifications include training in risk science principles and methods, as well as any industry-specific certifications. For example, certifications can include ISO 31000 or Enterprise Risk Management.

O.4 involves the applicability and relevance of the study design. Some studies may be relatable to historically used studies that have generally been acceptable by stakeholders. Some studies may involve systems that are well understood through past studies and research.

O.5 involves the need to minimize biases in the study design. Many types of biases may exist and can influence decision-making (Das & Teng, 1999). For example, suppose an analyst has a pre-determined outcome before initiating the study. This analyst may prefer to show high importance of recent risk events because they are more recent in memory, such as by prioritizing biological weapon risk due to the

**TABLE 2** Taxonomy of overall data and information integrity for a risk study

| Characteristic | | Integrity and quality (high, medium, low, N/A) |
|---|---|---|
| K.1: Data and information integrity | • Data and information are accurate<br>• Information source is reputable | – |
| K.2: Data and information applicability and sufficiency | • Time period of data and information is applicable to the study<br>• Geographic representation of data and information is applicable to study<br>• Data and information are representative of the population studied<br>• Data have a sufficient sample size | – |
| K.3: Data and information transparency | • Documentation and agreement for removal of outliers<br>• Documentation and agreement for handling of missing values<br>• Documentation for data limitations and assumptions related to K.1–K.4<br>• Clear distinction between fact and opinion | – |
| K.4: Data and information collection integrity | • Clear criteria for selecting experts, data, and information<br>• Industry standard norms for data collection<br>• Appropriate phrasing of questions<br>• Appropriate ordering of questions<br>• Appropriate scaling of answers (quantitative, Likert, etc.) | – |
| K.5. Data and information objectivity | • Information sources (e.g., experts, survey respondents, and data sources) have no known conflicts of interest or biases that can influence the risk study<br>• Information sources have not been discredited by third parties | |

recent COVID-19 pandemic. As another example, the analyst may have a pre-determined opinion on the study topic and only select supporting data and information that confirms this opinion.

O.6 refers to the approach. As with any scientific study, studies that are reproducible by third-party analysts show stronger integrity. Thus, the study team should have methods to document all aspects of the process to support reproducibility and transparency. Additionally, stakeholders using the information from the process should be able to understand the process. For example, stakeholders should clearly understand the hypothesis, the study design, and how to use the study outcomes. This type of explainability is also essential to promote buy-in or funding for the study.

## 3.2 | Data and information evidence

Table 2 shows characteristics of overall integrity and quality related to data and information in a risk study. These characteristics involve qualities of information sources, data, and also qualitative information that can be used to inform any step within the risk study.

K.1. refers to data and information integrity. The risk team should ensure that data and information represent properties intended to be studied. An accuracy check for the information could include focus groups or vetting with experts, stakeholders, and industry groups with experience with the risk study scope. More granular accuracy checks could include reviewing the data collection methodologies provided by the source. If the data are collected by the risk team, an accuracy check could consist of manually testing for errors in a subset of the data. High-level accuracy checks could include investigating whether the data characteristics (summary statistics, trends, etc.) agree with comparable datasets and industry knowledge; if they do not agree, take steps to understand why the disagreement exists.

K.2 refers to the applicability and sufficiency of the data and information. These data and information should ideally represent the time period applicable to the risk study. This is particularly important because risk studies involve future projections, requiring analysts to combine past data and assumptions to project the future. Additionally, the data and information should reflect an appropriate geography and population that is applicable to the risk study. In many cases, discrepancies will exist, but those should be documented as

part of K.3.

The issue of sufficient quantities of data should also be addressed. In cases where the risk problem represents a rare event or some new situation, there may be little or no usable data. Instead, the analyst may depend on qualitative approaches, modeling, or simulation. Questions surrounding the simulations or other methods for data generation would be addressed in the Analysis (A1–A4) characteristics described in Section 3.3. The analyst may interpret a minimal bias and properly conducted qualitative approach as high integrity since this is the best data and information available, but this is dependent on the context of the risk application. In cases of statistical analysis, sample sizes should be sufficient for the study, recognizing clear sampling standards (Ross, 2020). In cases of big data resources, the analyst may interpret large datasets as high integrity. However, regardless of the size of the dataset, biases in the data and information are dangerous as they can mislead or promote overconfidence in results based on incorrect information. Thus, any data and assumptions that are used have a critically high influence on the risk study.

K.3. refers to transparency related to the use of data. All relevant details and assumptions about the data should be disclosed to stakeholders. This disclosure should include a description about how outliers and missing values were addressed. There should also be a clear distinction between fact and opinion. Facts could involve data that represent measured quantities (temperatures, amounts, etc.). Those facts can involve some subjectivity. For example, if an expert provides an estimate of the temperature or a probability distribution of the temperature, the estimate and probability distributions can be viewed as facts reflecting their understanding, but is a judgment or opinion about the temperature. As another example, surveys using Likert scales are quantitative but measure subjective opinions.

K.4. refers to the integrity of data collection. Standard norms should be used for collecting data in particular contexts. For example, sampling for water quality involves particular procedures that are commonly used within scientific research and industry. When using surveys, the way in which questions are phrased or ordered could influence how respondents answer those questions (Pew Research, 2021). For example, asking "Were the work conditions safe?" versus "Did the work conditions meet all applicable workplace safety standards?" have very different wording. The question of safety is a management judgment, while the question of meeting standards is measurable. Thus, the statements could evoke very different responses. Additionally, the scaling of survey responses should also be considered (Revilla et al., 2014; Weijters et al., 2010).

K.5. refers to objectivity for data and information sources. Examples of sources include the owners of the data, experts who are providing technical knowledge, published studies, and any other expertise that is used for the risk study. These information sources should have no conflicts of interest or biases that can influence the risk study. Additionally, care should be taken to ensure that these information sources have not been discredited by third parties.

## 3.3 | Analysis

Table 3 shows characteristics of the overall integrity and quality related to analysis in a risk study. The risk study can be composed of approaches, such that each approach consists of one or more methods. Together, these approaches and methods can be used to analyze and describe risk. The characteristics relate to the selection, implementation, interpretation, and interaction among approaches/methods. When evaluating these characteristics, it is recommended to consider each method individually, then take a holistic perspective to understand if the broader approach is suitable for the risk application.

The use of various risk methods involves making assumptions about variables, constants, and relationships among those aspects. Consider examples of broad categories of methods that are often used for risk applications:

- Conceptual based on opinions and general statements (e.g., system diagramming, group exercises, expert elicitation)
- Decision analysis
- Optimization (e.g., linear programming, integer programming, network flow, and goal programming)
- Queuing
- Bayesian networks
- Probabilistic risk assessment
- Empirical (e.g., statistics, correlation, regression, experiments)
- Simulation (e.g., deterministic or stochastic Monte Carlo)
- Game theory
- Markov decision processes
- Artificial intelligence, machine learning, and deep learning

While the method list shown above is not exhaustive, it demonstrates that the general approach studied in this section should apply to this wide array of methods used for risk applications.

All of the characteristics described below relate to the validation of the analysis related to scientific quality and risk. Meeting the characteristics implies that the analysis complies with rules, assumptions, and constraints; the analysis is explainable, transparent, and subject to feedback from experts and stakeholders; and the analysis is relevant and useful, such that it addresses the scope and intent of the risk study (Aven & Zio, 2018). In a risk study Stage 4 Judgement and Stage 5 Decisions & Communications (see Table 2), one will rely on these characteristics to promote appropriate judgment and decisions. See Rosqvist and Tuominen (2004), Aven and Heide (2009), and Aven and Zio (2018) for a comprehensive discussion of quality for risk assessment, and Oberkampf et al. (2007) for discussion of quality in modeling and simulation initiatives.

**TABLE 3** Taxonomy of analysis for a risk study

| Characteristic | | Integrity and quality (high, medium, low, N/A) |
|---|---|---|
| A.1: Analysis approaches/methods appropriate for the application | • Risk is adequately characterized<br>• Overall modeling strategy is appropriate<br>• Approaches/methods and related assumptions are credible<br>• Mathematical properties and assumptions for approaches/methods are appropriate | – |
| A.2: Procedures appropriately performed | • Analysis standards/protocols met<br>• No use of misleading procedures | – |
| A.3: Analysis properly interpreted | • Correct reading of analysis<br>• Overall findings justified by analysis<br>• Causative relationships correctly interpreted | – |
| A.4: Analysis approaches/methods adequately interact with each other | • Assumptions, metrics, information, etc. can be shared across analysis methods | – |

The characteristics described below are also not intended to solely represent scientific accuracy. Instead, they recognize that risk is the study of the future and involves the confluence of scientific concepts, interacting systems, values, decisions, and uncertainties. Thus, the characteristics encompass a broader range of criteria, many of which address recognizing scientific knowledge and lack of scientific knowledge. These characteristics prompt the risk team to explore appropriateness in both the selection and use of approaches/methods, including assessment of conditions and assumptions related to those approaches/methods. The characteristics are as follows:

A.1. refers to the selection of the approaches/methods. This characteristic relates to both the selection of the overall approaches/methods and the related assumptions. Several criteria can be used to ensure the selected approaches/methods and related assumptions are appropriate for the risk study, as follows:

First, there should be an adequate characterization of risk. Following the recommendations by Society for Risk Analysis (SRA, 2018), risk in the most general form is described by the $(C',Q,K)$ and $(A',C',Q,K)$, where $A'$ is a set of specified events, $C'$ some specified consequences, $Q$ a measurement or description of uncertainties, and $K$ is the basis for $Q$ and $(A',C')$. To form this characterization, there is a need to seek input from sources, such as measurements, experts, models, and testing.

Second, the analysts should consider the appropriateness and usefulness of the overall modeling strategy, ensuring that the strategy is aligned with foundational risk concepts (Aven, 2020; Kaplan & Garrick, 1981; Singpurwalla, 1988). There is a wide array of existing guidance for modeling related to quantitative and probabilistic methods (Apostolakis, 1990, 2004) and decision-making (Borgonovo et al., 2018; Paté-Cornell & Dillon, 2006). Mission-driven risk study modeling strategies show usefulness by clearly addressing the risk characterization mentioned above. Because risk problems study the future, models are used to study these future conditions under uncertainty. Risk models and analyses are useful for developing thought-constructed systems and activities. Choosing to develop a model implies that the system is understood to the level of being able to define inputs, outputs, and functions that demonstrate the system. Often, this is appropriate for systems that have been studied in the past. In those cases, there is a balance to be made between using the model to accurately represent the system and creating a simple and effective model. However, consider a new risk problem that is poorly understood, such as related to emerging risks. For example, consider the new phenomena of cyber-attacks involving ransomware. The technologies, financial systems, actors, and intrusion methods may be changing so rapidly that any current model can be quickly deemed obsolete. In these cases, because the uncertainties are too large, there may be a preference for avoiding mathematical methods that require many assumptions, as these assumptions would be without basis.

Third, there is also a need to evaluate whether the approaches/methods and related assumptions are credible. This question of credibility can refer to mathematical methods, conceptual designs, or assumptions related to inputs, outputs, and functions representing the features of the system. Credibility can be demonstrated by ensuring industry standards support the use of each approach/method. However, this is the bare minimum, as industry standards are not always founded on risk science, but rather on ad-hoc consensus in industry. The approaches/methods should also be acceptable to stakeholders who may not be experts in the system but rely on trust of the system for their own risk and professional purposes. Ideally, the approaches/methods should be acceptable to experts within the domain of the studied system and also the risk field to allow for consensus among those disciplines. The analyst should also consider whether the approaches/methods reflect all relevant dimensions, such as scientific, social, or environmental factors. For example, consider climate change models that reflect the impact on business and society but neglect to reflect the impact on

biological diversity.

Once the model is tested, there is also a need to show credibility by validating the analysis results. This is particularly important because if there are problems with analysis outputs, these should be addressed before progressing to subsequent stages of the risk study. Validation exercises may include evaluating sensitivity and conducting external checks with industry and risk professionals. While validation is commonly included in field-specific uses of approaches/methods, these validation exercises are often limited in scope and may not be entirely adequate for the risk study. For example, industry standards may not be appropriately informed by the risk science approach. Therefore, validation can benefit from input from analysts trained in risk science. If this task is a particular challenge, it could be helpful to hire consultants who can use their own experience to validate the approaches/methods and train the analysts.

Finally, the analyst should ensure that the properties and assumptions associated with the chosen approaches/methods are appropriate. For example, consider the use of a stochastic simulation model, which uses probability distributions to represent the behavior of variables. These types of models are highly dependent on assumptions related to probabilities, but these probabilities can be based on little or no knowledge. Thus, overreliance on probabilities with limited supporting knowledge can be highly problematic. Also consider the use of probabilistic models, such as Bayesian inference, which allows probabilities to be updated using new evidence. This updating using new data can help models and related predictions to become increasingly accurate, particularly when this new data are relevant for the risk study. But this use of new data can also be problematic when models are over-fit with recent data, potentially reducing accuracy in new or surprise risk-related situations. More complex examples include relationships found in artificial intelligence approaches. The selection of these methods should consider the relevance and accuracy of these methods in new or surprise situations, as related to risk events.

As a very simple example, consider the decision regarding approaches to manage risk for water quality. The topic of water quality risk has become very visible following the Flint Water Crisis (Masten et al., 2016) and broader issues with ransomware impacting critical infrastructures (CISA, 2021). Because water quality has been highly studied, the analyst can look to expert and industry-standard guidance for testing water quality (ASTM, 2021; EPA, 2021a), though the broader assessment of risk is still in its infancy with few guidelines that can adequately apply to a wide variety of communities (EPA, 2021b). The analyst may refer to the literature and adopt models and associated assumptions that are based on strong scientific evidence and precedence. Then, the analyst can decide on the most appropriate risk metric related to water quality. The analyst may consider using the central tendency of Total Dissolved Solids by choosing between using a mean or a median. While these metrics are both very commonly used, the mean is more sensitive to outliers when compared to the median. Thus, the analyst should consider whether this type of sensitivity can interfere with the intent of the risk study. Additionally, the analyst should consider whether addressing only central tendency is sufficient and consider how to measure the spread by using range or standard deviation.

The analyst should also consider how these types of decisions about metrics fit into the broader approaches/methods. For example, with little guidance from industry, the analyst must decide how to model the system using various inputs, outputs, and functions. The analyst also decides what stakeholders to consider, such as residents, agriculture, industry, and other stakeholders who influence and use the system. Then, the analyst can consider the scope of the problem, for example, considering the types of pipes involved in the water transport, the vulnerabilities related to the cyber-infrastructure, or bad actors who may intentionally try to harm the system. Suppose the analyst then chooses to use a simulation model that addresses the aspects listed above.

A.2. refers to appropriate procedures for implementation of the selected approaches/methods. Complete understanding of the procedures and how to implement the procedures is not a trivial task. This often involves training, mentorship, and experience with these tasks. Analysis standards and protocols should be met, such as testing conditions, software use, and technological instruments. This also calls for the need to avoid human error, such as by training, incorporating redundancies, auditing, reporting, or using other types of managerial oversight that ensure repeatability and reliability of the approach. There may also be cases of intentional misuse of methods. For example, p-hacking, which results from the analyst manipulating the data or analysis process to achieve statistical significance, is a common problem in academia (Head et al., 2015). In addition, consider biases and broader conflicts of interest related to the risk study. While all of the criteria described here are important for demonstrating credibility in correctly implementing the approaches/methods, there is also a need to recognize that different stakeholders may have different criteria when assessing the credibility of the approaches/methods and may have formed opinions about the approaches/methods without sufficient understanding of the problem, methods, or related issues (Aven, 2017).

In cases of expert elicitation or group exercises, these tasks should focus on using these methods for analysis only, not judgment (judgment to follow in Section 4). There could be confusion in this regard as it could be a simple mistake to unintentionally use these methods to collect expert judgments about risk due to unclear phrasing of questions and prompts (Aven, 2017). The data collection should be conducted in settings that promote honest and unbiased feedback, allowing participants to fully understand questions and context and communicate freely in response.

Consider the example of water quality again. The analyst can choose to conduct a statistical analysis of sample data for analyzing water quality risk. The analyst first needs to fully understand factors such as the importance of obtaining

a random sample, where to physically locate the sample, the appropriate timing of the sample, how to physically collect the sample, and how to test the sample. Then, the analysis team must also understand basic principles related to the broader modeling approach, such as the chosen simulation model. Then, the analysts must know how to operate the software for analysis. There needs to be some type of managerial oversight, such as required reporting, to ensure that the analyst appropriately follows the procedures. Any deficiencies in these processes can have major implications on the quality of analysis results.

A.3. refers to the proper interpretation of the analysis. While the next section of this study will discuss judgment, this characteristic refers to an accurate reading of results (non-judgment). The analyst should be able to understand how to interpret the output of the analysis (e.g., software and expert elicitation) and ensure that the analysis justifies the overall findings. It can be effective to ensure that others can confirm the interpretation of the analysis output with familiarity with the risk problem and application.

Consider the water quality example. Suppose the analyst uses the software output to view the 95% confidence interval, using an output:

$$300 \pm 20\,\text{ppmTotalDissolvedSolids}.$$

The analyst must correctly interpret this confidence interval with an understanding of aspects of significance, sampling assumptions, and statistical thinking. Then, the analyst should be able to formulate study findings that are supported by this software output. For example, suppose an analyst with limited statistical experience reported the following:

95% of the samples had between 280 and 320 ppm Total Dissolved Solids.

This is an incorrect interpretation. A more experienced analyst may instead report the following:

We are 95% confident that the true population average Total Dissolved Solids is between 280 and 320 ppm.

This is a better-informed interpretation, but also could be improved by further explaining concepts such as confidence and significance.

Analysts should also always be able to distinguish causation and correlation in results correctly.

As seen with A.2, there is a need to consider the potential for biases and conflicts of interest. In cases of expert elicitation or group exercises, analysts need to ensure that input from respondents is appropriately interpreted. There may be a need to confirm responses before progressing with the risk study.

A.4. recognizes that the risk study may involve the use of many methods that interact with one another. Poor model selection, data, or assumptions for one method can lead to poor implementation of related or interacting methods. Care should be taken to ensure that integrity is high for all methods used. For example, consider the earlier example of an analyst choosing an appropriate risk metric. If the analyst chooses to use a mean, they can combine the assessment with a confidence interval, including consideration of confidence intervals. Choosing the median would preclude any further statistical analysis in a later stage of the risk study.

## 3.4 | Management review and judgment, decisions and communications

Table 4 shows characteristics of overall integrity and quality related to management review and judgment, decisions and communications (MRJDC) in a risk study.

We define management review and judgment as the process of interpreting and deliberating the results of the risk assessment, as well as of other contextual issues not included in the assessment, in order to make a decision (Aven & Thekdi, 2021). There is recognition that limitations exist, and they should be documented and shared with relevant parties. Then, the decision-making process considers all of these factors (assessments, context, limitations, etc.) when making decisions. This process explicitly considers assumptions and beliefs and also places importance on stakeholders' values, goals, criteria, and preferences. The process of decisions and communications consists of decision-makers (policymakers, stakeholders, etc.) deciding on the most appropriate course of action to address risk. Those decisions are translated into communication with stakeholders, such as executives, policy-makers, and communities.

The characteristics for this section primarily ensure that MRJDC addresses context-specific concerns, which may not be fully understood using scientific process alone, ensuring there is a clear link or logic among the stages of the risk study, and ensuring that the decision-making and communication process adhere to best practices.

D.1. refers to ensuring MRJDC is informed by context-specific concerns. There is danger in over-reliance on the analysis and scientific process involved with the risk study. These elements should be supplemented by consideration of stakeholder needs and concerns. For example, issues of equity, inclusion, and fairness are those that are not easily studied using scientific principles, but instead require input from stakeholders and consideration of values and ideals for the risk study.

The structure and key features of the decision-making process should also be pre-defined, such that the logic and procedures are known and agreed upon in advance. This process should be documented if questions arise. There should also be accountability to ensure the process is carried out as planned.

Additionally, the decision-making process must address context-specific concerns, such as community values and ethical issues. There may not be a science-based process for addressing these issues. However, the decision-making process should address these concerns in a transparent manner. For example, transparency can be demonstrated by

**TABLE 4** Taxonomy of judgment and decisions for a risk study

| Characteristic | | Integrity and quality (high, medium, low, N/A) |
|---|---|---|
| D.1: Judgement, decisions, and communications informed by context-specific concerns | • MRJDC informed by stakeholder needs and concerns<br>• Decision-making process addresses context-specific concerns in a transparent manner<br>• Limitations are addressed | – |
| D.2: Clear links among risk study mission, analysis output, judgement, decisions, and communications | • MRJDC is informed by analysis results<br>• MRJDC addresses mission of the risk study<br>• Decision-making approach is appropriate for the risk study<br>• Communication approach is appropriate for the risk study | – |
| D.3: Decision-making and communications follow best practice for risk science knowledge and practice | • Pre-defined process for decision-making, with accountability for ensuring the decision-making process is as-planned<br>• Decision-making methodologies and communication efforts are founded in literature and/or improve upon existing methodologies | |

Abbreviation: MRJDC, management review and judgment, decisions and communications.

documenting the guiding values used to make decisions. If stakeholders agree with the guiding values, they can better understand and possibly even agree with the risk study outcome. Similarly, the risk communication efforts should address context-specific concerns by strategizing efforts based on community needs. For example, different communities may more effectively respond to particular platforms or communication characteristics.

Finally, there is the acknowledgment that all risk studies have limitations. These limitations can relate to a variety of factors, such as assumptions made, not addressing all relevant risk contributors, representing risk poorly, or basing results on assumptions and beliefs that could be wrong or are weakly supported.

Transparency of those limitations can help demonstrate credibility in decision-making activities in particular.

D.2 refers to ensuring there is a clear link between analysis output, judgment, decisions, and communications. The risk team should consider questions, such as:

- Does the MRJDC address the question originally posed in the risk study?
- Does the analysis output provide sufficient information to inform the risk characterization, judgment, decisions, and communication?
- Is the decision-making approach appropriate for the risk study? For example, are consensus-based approaches appropriate for politically charged topic areas? Should the values of particular stakeholders matter more than others for this particular topic area?
- Is the communication approach appropriate for the risk study? For example, is the mode of communication accessible by the intended audience? Or is the language of the message understandable by the intended audience?
- D.3. refers to ensuring the decision-making and communication activities follow best practices for risk science knowledge and guidance.

This includes having a pre-defined structure for the decision-making process, demonstrating a clear logic related to what information is used, who are the decision-makers, how to address disagreement among decision-makers, and how final decisions are made. This pre-defined process should be documented and have some accountability mechanism to ensure the process is followed. Additionally, the decision-making methodologies and communication efforts should be either founded in the literature or improve upon existing methodologies. Refer to Goodwin and Wright (2014) and Lundgren and McMakin (2018) for resources on best practices.

## 4 | DEMONSTRATION OF FRAMEWORK

This section demonstrates the framework using a hypothetical case study involving cybersecurity. Consider the case of a non-technology firm that is concerned about direct cybersecurity attacks. Suppose the firm has historically viewed cybersecurity as a nonessential task, but now understands the urgency, considering extreme consequences related to system functionality, data privacy, and financial viability. For example, consider the Kaseya ransomware attack, with attackers demanding a $70 million ransom (Bobrowski, 2021), and consequently disrupting the business functionality for the impacted firms.

While the firm has not been directly attacked in the past, they have invested in cybersecurity software, such as antivirus software. They have also trained their employees to be cautious about phishing attacks. However, leaders in the firm also recognize that they could do more because these types of attacks are becoming increasingly commonplace and more difficult to detect.

**TABLE 5** Taxonomy of overall integrity and quality for a risk study: Application to cybersecurity risk

| Characteristic | | Integrity and quality (high, medium, low, N/A) |
|---|---|---|
| O.1: Alignment with risk science | • Clear understanding of risk science fundamentals<br>• Clear hypothesis (e.g., to characterize risk and understand vulnerability) | H |
| O.2: Alignment of study design and execution | • Process for monitoring to ensure study aligns with study design<br>• Process for accountability if there are deviations from the study design | H |
| O.3: Expertise | • Qualifications of the assessor<br>• Certifications | M |
| O.4: Relevance and applicability of study design | • Hypothesis is justified (e.g., clear understanding of how hypothesis is formulated; reflects all relevant issues)<br>• Design is informed by understanding of the studied system<br>• Design is informed by previously used and accepted risk studies | M |
| O.5: Biases | • No biases in knowledge or information selected for use in the study<br>• No biases of team members influencing the study design same here | H |
| O.6: Approach | • Transparent and reproducible approach<br>• Approach that can be explained or understood by stakeholders | H |

The firm has formed an internal task force to address cybersecurity risk, led by a manager who has been trained in risk science. The other members of the task force are not trained but have the knowledge of the technology systems and functionality of the firm. Because this is the firm's first attempt to manage cybersecurity risk proactively, there is the expectation to keep the risk study very general but be informed by input from those with detailed system expertise.

The first step involves understanding and planning the overall integrity and quality of the risk study. The task force uses a consensus approach to complete the information in Table 5. The task force has a clear mission: To properly manage risk of cyber-attacks for the firm. The task force has developed reporting mechanisms to ensure the risk study is conducted in a clear and transparent manner. The team is somewhat concerned because only the team leader is qualified to conduct the risk study. However, the team has confidence in the manager who is leading the risk study and affirm that they have a high rating for the study's alignment with risk science. The team is also concerned because they have very little knowledge about how these types of risk studies are conducted in similar firms, as these types of studies are kept confidential for security reasons. The team discusses potential biases and finds no major issues to exist. Senior executives are informed about the risk study plan and are energetic about the proactivity and mission of the study.

The team then compiles a set of data sources that can be used to gauge cybersecurity risk. They find information on overall cybersecurity attacks across the world. They also gather capabilities of third-party providers who sell cybersecurity management services. They have concerns over the data regarding past cybersecurity attacks because the capabilities of attackers and technologies may have changed significantly in only the last few months. These past attacks are also generalized for an array of industry types, not specific to the firm. They are also doubtful of the data because it only represents attacks that have been reported, recognizing that the majority of ransomware attacks go unreported. They also recognize that they have very little data about the security of their internal software and hardware. Their IT department is overtasked and is not able to handle these types of requests.

The team then begins the process of analyzing risk. Because they have limited information, they decide to rely on a simple conceptual model of their system, designating inputs, outputs, and internal mechanisms. They are careful not to make any assumptions outside of their domain knowledge. While being cautious is appropriate, they also feel they are oversimplifying the situation because they are not including detailed information about their system design and specific strategies being used by attackers.

The risk characterization effort results in disagreement and hesitancy in defining key components. The team is unsure if they have been comprehensive in defining *A'*, the set of specified events. There is immense disagreement in defining *C'*, the specified consequences. The knowledge, *K*, is also very weak.

While the risk team would like to create a more detailed conceptual model of their cyber-system and conduct a scenario analysis to see how the system would react to various types of cyber-attacks, they do not have the IT resources to make appropriately informed generalizations.

**TABLE 6** Taxonomy of overall data and information integrity for a risk study: Application to cybersecurity risk

| Characteristic | | Integrity and quality (high, medium, low, N/A) |
|---|---|---|
| K.1: Data and information integrity | • Data and information are accurate<br>• Information source is reputable | L |
| K.2: Data and information applicability and sufficiency | • Time period of data and information is applicable to the study<br>• Geographic representation of data and information is applicable to study<br>• Data and information is representative of the population studied<br>• Data have a sufficient sample size | L |
| K.3: Data transparency | • Documentation and agreement for removal of outliers<br>• Documentation and agreement for handling of missing values<br>• Documentation for data limitations and assumptions related to K.1–K.4<br>• Clear distinction between fact and opinion | L |
| K.4: Data collection integrity | • Clear criteria for selecting experts, data, and information<br>• Industry standard norms for data collection<br>• Appropriate phrasing of questions<br>• Appropriate ordering of questions<br>• Appropriate scaling of answers (quantitative, Likert, etc.) | H |
| K.5. Data and information objectivity | • Information sources (e.g., experts, survey respondents, data sources) have no known conflicts of interest or biases that can influence the risk study<br>• Information sources have not been discredited by third parties | H |

The team does find that the approach used was credible, albeit simplistic. They show the conceptual models to various experts in the organization, including IT managers, who find the model to be acceptable. The team does not have any other model results to share for purposes of validation. All agree that the properties and assumptions are mostly appropriate and address requests for minor changes.

Table 8 describes the resulting taxonomy of judgment and decisions for the risk study. Because the team leader is very well versed in risk science, the MRJDC process has been well designed, with transparent decision-making processes, transparency, and accountability.

Finally, the team uses the output to share with decision-makers, which consists of executive management. The decision-makers are very concerned by Tables 6–8. They are surprised by the low level of integrity in several categories and are most concerned about the low integrity levels associated with data, information, and analysis. The decision-makers feel the in-house expertise in the area of cyber-security is insufficient.

The team discusses each characteristic separately and then discusses their overall concerns in the findings. Those concerns are shared with the decision-makers who consider the following choices: (1) to do nothing about this risk issue, (2) mitigate the risk, such as by investing in a cybersecurity department or contracting with a third-party cybersecurity firm, (3) transfer the risk by purchasing cybersecurity insurance, (4) transfer the risk by outsourcing the most vulnerable business functions, or (5) allocate appropriate IT resources in order to more appropriately conduct the risk study. The decision-makers are not considering the option to discontinue any IT-related practices. Suppose the decision-makers decide to outsource their cybersecurity initiatives. This makes sense, as these third parties have the knowledge and technical expertise to address a very rapidly changing cybersecurity landscape. They also have detailed information about how other technology firms are addressing cybersecurity.

In this example, the risk team carried out a well-planned and documented risk study. The team was qualified to perform this task. However, the assessment of integrity in the risk study brought attention to severe deficiencies, which is the intent of the presented framework. Even with a trained risk expert leading the risk study, this does not compensate for poor integrity and evidence for the risk study. Gauging the integrity of each facet of the risk study proved to be very important for the decision-makers, as they saw the barriers to obtaining a high level of integrity and found it to be more effective to outsource the entire cybersecurity risk function.

**TABLE 7** Taxonomy of analysis for a risk study: Application to cybersecurity risk

| Characteristic | | Integrity and quality (high, medium, low, N/A) |
|---|---|---|
| A.1: Analysis approaches/methods appropriate for the application | • Risk is adequately characterized<br>• Overall modeling strategy is appropriate<br>• Approaches/methods and related assumptions are credible<br>• Mathematical properties and assumptions for approaches/methods are appropriate | L |
| A.2: Procedures appropriately performed | • Analysis standards/protocols met<br>• No use of misleading procedures | L |
| A.3: Analysis properly interpreted | • Correct reading of analysis<br>• Overall findings justified by analysis<br>• Causative relationships correctly interpreted | H |
| A.4: Analysis approaches/methods adequately interact with each other | • Assumptions, metrics, information, etc. can be shared across analysis methods | H |

**TABLE 8** Taxonomy of judgment and decisions for a risk study: Application to cybersecurity risk

| Characteristic | | Integrity and quality (high, medium, low, N/A) |
|---|---|---|
| D.1: Judgement, decisions, and communications informed by context-specific concerns | • MRJDC informed by stakeholder needs and concerns<br>• Decision-making process addresses context-specific concerns in a transparent manner<br>• Limitations are addressed | H |
| D.2: Clear links among risk study mission, analysis output, judgement, decisions, and communications | • MRJDC is informed by analysis results<br>• MRJDC addresses mission of the risk study<br>• Decision-making approach is appropriate for the risk study<br>• Communication approach is appropriate for the risk study | H |
| D.3: Decision-making and communications follow best practice for risk science knowledge and practice | • Pre-defined process for decision-making, with accountability for ensuring the decision-making process is as-planned<br>• Decision-making methodologies and communication efforts are founded in literature and/or improve upon existing methodologies | H |

Abbreviation: MRJDC, management review and judgment, decisions and communications.

## 5 | CONCLUSIONS

This study has presented a framework for understanding how to classify evidence for risk problems. The classification system acknowledges that evidence can be created and evaluated across multiple stages of a risk study. Any limitations or insufficiencies of evidence at any stage of the risk study have the potential to undermine the overall mission of risk analysis and management.

The case study involving cybersecurity risk demonstrates that risk studies conducted by highly qualified and trained individuals can have severe limitations due to poor integrity of evidence. In the cybersecurity industry, risk events are particularly difficult to predict and detect, thereby exacerbating consequences. That issue combined with poor evidence used for modeling and understanding the cyber infrastructure system can be overwhelmingly concerning for decision-makers and stakeholders.

Without the use of a framework as presented in this study, there is high potential for researchers and practitioners to be easily misled about the quality of evidence. All risk studies can have limitations, despite the time, resources, data, and information used to inform the study. For example, risk studies using large datasets can have limitations due to assumptions and issues with compatibility with the risk study topic area. Thus, the value of this framework is to provide a systematic tool that can be used to carefully consider each individual characteristic that can gauge the quality of evidence, such that no individual characteristic is more or less important than another. Decision-makers can then investigate each characteristic and decide how to address issues in the risk study.

This study builds theoretical insights from literature and innovates using lessons learned from recent risk events, issues of misinformation, and the need to build transparency in topics of evidence quality and integrity. The present study addresses gaps in existing literature by acknowledging that various stakeholders may have different expectations for quality and integrity of evidence. Additionally, the approach acknowledges that in some situations, but not all, lapses in the quality and integrity of evidence can invalidate the risk study. Additionally, this study provides a critical lens for other aspects of quality and integrity of evidence, such as emerging from data applicability, biases, selection of metrics, and selection of analytical approaches. The present study also recognizes that there is not necessarily one standard that is appropriate for all domains and applications. Understanding where criteria are met, and where not met, is critical for involved decision-makers and stakeholders.

The use of this framework will be important for a wide array of risk researchers and practitioners. Potential users of this framework include data analysts, managers, and executives. This framework can also include input from other types of stakeholders, such as operational partners, oversight committees, third-party evaluators, and regulators. With a systematic checklist, researchers can include this step as part of a more extensive risk science toolkit. This framework should also supplement the chosen framework for risk, such as ISO 31000 (ISO, 2021) or Enterprise Risk Management (COSO, 2021) processes. The framework of this study also can serve as a supplement to overall quality tests for risk studies (SRA, 2021).

Consistency in gauging the integrity of evidence is very important for the risk field for several reasons. First, having standards for evidence integrity helps elevate risk science as its own scientific discipline. Second, this framework allows the topic area of risk science to be more accessible, making risk science more transparent and easier to use. With the framework, conducting a risk study can be less intimidating for individuals and organizations. Third, this framework promotes consistency in evidence integrity. As there is concern over disciplines developing their own risk sub-disciplines without consistency or without cross-learning among fields, this framework can help promote a consistent evaluation across disciplines. In other words, because this framework is not discipline specific, it can be adapted and applied across disciplines.

The work of this study is intended to be a starting point for discussions about integrity of evidence. Increased discussion on this topic can also leak out to broader disciplines and society. It promotes all disciplines, whether in the realm of risk or not, to pause and consider how valid the presented evidence is prior to forming conclusions. At a high level, this framework is a first step in battling issues of misinformation. At a pragmatic level, this framework encourages risk researchers to understand how to effectively present their work and also to ask the right questions when information is being presented to themselves. Thus, issues with understanding the integrity of evidence will remain widespread, and this study presents fundamental steps to address these issues.

## ACKNOWLEDGMENTS

The authors thank the anonymous reviewers for their helpful feedback.

## ORCID

*Shital Thekdi* https://orcid.org/0000-0003-4145-508X

## REFERENCES

AIAA (1998). *AIAA guide for the verification and validation of computational fluid dynamics simulations*. American Institute of Aeronautics and Astronautics.

Andersen, H., & Hepburn, B. (2015). Scientific method. https://plato.stanford.edu/entries/scientific-method/?source=post_page

Apostolakis, G. E. (1990). The concept of probability in safety assessments of technological systems. *Science*, *250*(4986), 1359–1364.

Apostolakis, G. E. (2004). How useful is quantitative risk assessment? *Risk Analysis*, *24*(3), 515–520.

ASTM (2021). Water testing standards. https://www.astm.org/Standards/water-testing-standards.html

Aven, T. (2013). A conceptual framework for linking risk and the elements of the data-information-knowledge-wisdom (DIKW) hierarchy. *Reliability Engineering & System Safety*, *111*, 30–36.

Aven, T. (2017). Risk analysis validation and trust in risk management: A postscript. *Safety Science*, *99*(B), 255–256.

Aven, T. (2020). Three influential risk foundation papers from the 80s and 90s: Are they stillstate-of-the-art? *Reliability Engineering & System Safety*, *193*, 106680.

Aven, T., & Heide, B. R. (2009). Reliability and validity of risk analysis. *Reliability Engineering & System Safety*, *94*(11), 1862–1868.

Aven, T., & Thekdi, S. (2021). *Risk science: An introduction*. Routledge.

Aven, T., & Zio, E. (2018). Quality of risk assessment: Definition and verification. In T. Aven & E. Zio (Eds.), *Knowledge in risk assessment and management* (pp. 143–164). John Wiley & Sons.

ABA (2021). How Courts Work. https://www.americanbar.org/groups/public_education/resources/law_related_education_network/how_courts_work/evidence/

Box, G. E. (1976). Science and statistics. *Journal of the American Statistical Association*, *71*(356), 791–799.

Barker, L. E., & Shaw, K. M. (2015). Best (but oft-forgotten) practices: Checking assumptions concerning regression residuals. *The American Journal of Clinical Nutrition*, *102*(3), 533–539.

Bobrowski, M. (2021). Kaseya Ransomware Attack: What We Know as REvil Hackers Demand $70 Million. https://www.wsj.com/articles/kaseya-ransomware-attack-11625593654

Borgonovo, E., Cappelli, V., Maccheroni, F., & Marinacci, M. (2018). Risk analysis and decision theory: A bridge. *European Journal of Operational Research*, *264*(1), 280–293.

Chaturvedi, P., Ramalingam, N., & Singh, A. (2020). Is COVID-19 man-made? *Cancer Research, Statistics, and Treatment*, *3*(2), 284.

CISA (2021). Ransomware guide. https://www.cisa.gov/publication/ransomware-guide

Cori, L., Bianchi, F., Cadum, E., & Anthonj, C. (2020). Risk perception and COVID-19. *International Journal of Environmental Research and Public Health*, *17*(9), 3114.

COSO (2021). Enterprise Risk Management—Integrated Framework. https://www.coso.org/Pages/erm-integratedframework.aspx

Das, T. K., & Teng, B. S. (1999). Cognitive biases and strategic decision processes: An integrative perspective. *Journal of Management Studies*, *36*(6), 757–778.

EPA (2021a). How are water quality standards developed? https://www.epa.gov/standards-water-body-health/how-are-water-quality-standards-developed

EPA (2021b). New risk assessment and emergency response plan requirements for community water systems. https://www.epa.gov/sites/production/files/2019-03/documents/awia_risk_assessments_and_emergency_response_plan_frn.pdf

Garcia-Alamino, J. M. (2020). Human biases and the SARS-CoV-2 pandemic. *Intensive & Critical Care Nursing*, *58*, 102861.

Goerlandt, F., Khakzad, N., & Reniers, G. (2017). Validity and validation of safety-related quantitative risk analysis: A review. *Safety Science*, *99*, 127–139.

Goodwin, P., & Wright, G. (2014). *Decision analysis for management judgment*. John Wiley & Sons.

Guyatt, G. H., Oxman, A. D., Kunz, R., Vist, G. E., Falck-Ytter, Y., & Schünemann, H. J. (2008). What is "quality of evidence" and why is it important to clinicians? *Bmj*, *336*(7651), 995–998.

Gordon, M., Stroebel, S., & Hinshaw, D. (2021). Intelligence on sick staff at Wuhan lab fuels debate on Covid-19 origin. https://www.wsj.com/articles/intelligence-on-sick-staff-at-wuhan-lab-fuels-debate-on-covid-19-origin-11621796228</bib>

Head, M. L., Holman, L., Lanfear, R., Kahn, A. T., & Jennions, M. D. (2015). The extent and consequences of p-hacking in science. *PLoS Biology*, *13*(3), e1002106.

Huff, D. (1993). *How to lie with statistics*. WW Norton & Company.

Imhoff, R., & Lamberty, P. (2020). A bioweapon or a hoax? The link between distinct conspiracy beliefs about the Coronavirus disease (COVID-19) outbreak and pandemic behavior. *Social Psychological and Personality Science*, *11*(8), 1110–1118.

ISO (2021). ISO 31000 risk management. https://www.iso.org/iso-31000-risk-management.html

Kahneman, D., Lovallo, D., & Sibony, O. (2011). Before you make that big decision. https://hbr.org/2011/06/the-big-idea-before-you-make-that-big-decision

Kaplan, S., & Garrick, B. J. (1981). On the quantitative definition of risk. *Risk Analysis*, *I*(I), 11–27.

Landeta, J. (2006). Current validity of the Delphi method in social sciences. *Technological Forecasting and Social Change*, *73*(5), 467–482.

Lathrop, J., & Ezell, B. (2017). A systems approach to risk analysis validation for risk management. *Safety Science*, *99*, 187–195.

Ledford, H., & Van Noorden, R. (2020). High-profile coronavirus retractions raise concerns about data oversight. *Nature*, *582*(7811), 160–161.

Lundgren, R. E., & McMakin, A. H. (2018). *Risk communication: A handbook for communicating environmental, safety, and health risks*. John Wiley & Sons.

Masten, S. J., Davies, S. H., & Mcelmurry, S. P. (2016). Flint water crisis: What happened and why? *Journal-American Water Works Association*, *108*(12), 22–34.

McMillan, S. S., King, M., & Tully, M. P. (2016). How to use the nominal group and Delphi techniques. *International Journal of Clinical Pharmacy*, *38*(3), 655–662.

Oberkampf, W. L., Pilch, M., & Trucano, T. G. (2007). Predictive capability maturity model for computational modeling and simulation. http://prod.sandia.gov/techlib/access-control.cgi/2007/075948.pdf

Oberkampf, W. L., Trucano, T. G., & Hirsch, C. (2004). Verification, validation, and predictive capability in computational engineering and physics. *Applied Mechanics Reviews*, *57*(5), 345–384.

Pace, D. K. (2004). Modeling and simulation verification and validation challenges. *Johns Hopkins APL Technical Digest*, *25*(2), 163–172.

Paté-Cornell, M. E., & Dillon, R. L. (2006). The respective roles of risk and decision analyses in decision support. *Decision Analysis*, *3*(4), 220–232.

Pew Research (2021). Writing survey questions. https://www.pewresearch.org/our-methods/u-s-surveys/writing-survey-questions/

Revilla, M. A., Saris, W. E., & Krosnick, J. A. (2014). Choosing the number of categories in agree–disagree scales. *Sociological Methods & Research*, *43*(1), 73–97.

Rogin, J. (2020, April 14). State Department cables warned of safety issues at Wuhan lab studying bat coronaviruses. *Washington Post*. https://www.washingtonpost.com/opinions/2020/04/14/state-department-cables-warned-safety-issues-wuhan-lab-studying-bat-coronaviruses/

Rosqvist, T. (2010). On the validation of risk analysis: A commentary. *Reliability Engineering & System Safety*, *95*(11), 1261–1265.

Rosqvist, T., & Tuominen, R. (2004). Qualification of formal safety assessment: An exploratory study. *Safety Science*, *42*(2), 99–120.

Ross, S. M. (2020). *Introduction to probability and statistics for engineers and scientists*. Elsevier Academic Press.

Rycroft-Malone, J., Seers, K., Titchen, A., Harvey, G., Kitson, A., & McCormack, B. (2004). What counts as evidence in evidence-based practice? *Journal of Advanced Nursing*, *47*(1), 81–90.

Sargent, R. G. (2013). Verification and validation of simulation models. *Journal of Simulation*, *7*(1), 12–24.

Shereen, M. A., Khan, S., Kazmi, A., Bashir, N., & Siddique, R. (2020). COVID-19 infection: Origin, transmission, and characteristics of human coronaviruses. *Journal of Advanced Research*, *24*, 91–98.

Singpurwalla, N. D. (1988). Foundational issues in reliability and risk analysis. *SIAM Review*, *30*(2), 264–282.

SRA (2020). Risk analysis: Fundamental principles. https://www.sra.org/wp-content/uploads/2020/04/SRA-Fundamental-Principles-R2.pdf

SRA (2018). SRA Glossary. https://www.sra.org/wp-content/uploads/2020/04/SRA-Glossary-FINAL.pdf

SRA (2021). SRA risk analysis quality test. https://www.sra.org/wp-content/uploads/2020/08/SRA-Risk-Analysis-Quality-Test-R6.pdf

Uscinski, J. E., Enders, A. M., Klofstad, C., Seelig, M., Funchion, J., Everett, C., Stefan, W., Kamal, P., & Murthi, M. (2020). Why do people believe COVID-19 conspiracy theories? Harvard Kennedy School Misinformation Review, *1*(3).

U.S. Courts (2021). I.DEFINITIONS. https://cdn.ca9.uscourts.gov/datastore/uploads/guides/stand_of_review/I_Definitions.html

Weijters, B., Cabooter, E., & Schillewaert, N. (2010). The effect of rating scale format on response styles: The number of response categories and response category labels. *International Journal of Research in Marketing*, *27*(3), 236–247.

**How to cite this article:** Thekdi, S., & Aven, T. (2023). A classification system for characterizing the integrity and quality of evidence in risk studies. *Risk Analysis*, 1–17. https://doi.org/10.1111/risa.14153