RISK SCORING SYSTEM FOR MEDICAL DEVICES

RSS-MD

INTRODUCTION

Current methods for analyzing identified cyber vulnerabilities tend to apply traditional information technology focus on impact to system confidentiality, integrity and availability to discern end-user risks. Although sufficient for evaluating traditional information technology systems, this process fails to consider the operational ramifications for complex systems-of-systems.

One of the current industry shortfalls is the lack of a risk scoring system that adequately considers the context of the environment for identified vulnerabilities. Unfortunately, this can lead to organizations improperly prioritizing mitigation efforts. As an example, the current scoring system advocated for by US Government Organizations is the Common Vulnerability Scoring System (CVSS). The CVSS evaluates the severity of an identified vulnerability in the context of system impact. For medical devices, however, it does not take into consideration the impact to patient safety—the true indicator of the severity of the vulnerability.

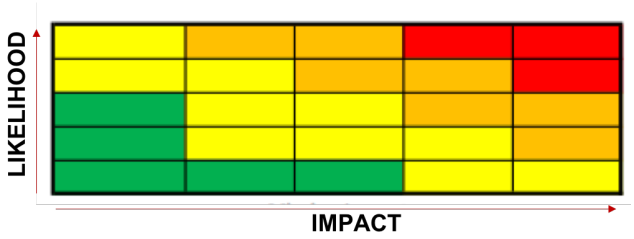
Consider the following demonstration. CVE-2014-5422 identifies a hard-coded password vulnerability in the CareFusion Pyxis SupplyStation. The Pyxis SupplyStations are automated cabinets used for dispensing medical supplies. The hard-coded password vulnerability was assigned a CVSS base score of 9.7. In comparison, CVE-2015-3956 identifies a vulnerability in the Hospira Symbiq Infusion System that directly delivers medication to patients intravenously. The Hospira Symbiq Infusion System vulnerability was assigned a CVSS base score of 7.6. From a comparative standpoint, the CareFusion vulnerability allowed an attacker to compromise the automated supply cabinet, to include removing the contents of the automated supply cabinet. The Hospira vulnerability allowed an attacker to remotely control the infusion system, potentially administering medication directly to a patient. The primary scoring difference from a system

perspective was that the vulnerability on the CareFusion included loss of system confidentiality, whereas the Hospira vulnerability did not. As a result, the CareFusion vulnerability score was more than the Hospira vulnerability, even though the CareFusion vulnerability does not directly impact patient safety and the Hospira vulnerability has direct impact to patient safety via remote access.

In consideration of the potential impact a software vulnerability in a medical device may have on patient safety, the Risk Scoring System for Medical Devices (RSS-MD) was developed that incorporates two primary factors: (i) Functional Impact and (ii) Vulnerability Characterization. The Functional Impact Category considers the impact to patient therapy and the scope of impact. The Vulnerability Characterization Category considers the attributes of the identified vulnerability. Coupled together, these two categories provide an effective means for rating the severity of identified vulnerabilities in the context of patient safety.

TRADITIONAL RISK ASSESSMENT

Risk assessment is used to determine the estimated risk that a threat poses and the magnitude of associated potential loss. Traditional risk assessment techniques evaluate consequence (i.e., impact) and likelihood. The resulting analysis provides a scoring metric that enables organizations to prioritize mitigation efforts. Current risk assessment techniques (as pictured below) for evaluating cyber vulnerabilities, however, are not adequate for vulnerabilities associated with medical devices.



GAPS IN IMPACT

Risk assessment associated with cyber vulnerabilities focus on impact to system confidentiality, integrity and availability. This paradigm evaluates effects on specific system resources, without consideration for the impact to the overall functionality. As

demonstrated in the CareFusion and Hospira example, this can lead to misleading scoring results because the context of the vulnerability is incorrect. Indeed, the context of impact should focus on impact to patient safety and not system-level impacts. This point is critical for assisting organizations in prioritizing mitigation efforts and understanding the magnitude of identified vulnerabilities.

THE LIKELIHOOD TRAP

Assigning ratings for likelihood to cyber vulnerabilities is perhaps the most common mistake when performing cyber risk assessments. Likelihood identifies the probability that a certain consequence will occur. Many organizations attempt to include likelihood as a factor when assessing risk for identified cyber vulnerabilities. The notion of applying traditional likelihood calculations for cyber risk assessment, however, is a trap. Likelihood is a probability based on occurrence of events over time. A cyber event is better aligned to a black swan event (i.e., high impact with low probability). Indeed, the probability of a vulnerability being exploited does not follow a traditional bell curve, and attempts to assign a probabilistic rate of occurrence are not statistically relevant. A more effective means is to evaluate the characteristics of the vulnerability. This slight, but important, difference assigns measurable ratings against a vulnerability as opposed to applying a probabilistic statistic based on historical attributes.

RSS-MD ATTRIBUTES

Through support from the United States Department of Homeland Security Science and Technology (DHS S&T) Directorate, the Risk Scoring System for Medical Devices (RSS-MD) was formalized for assessing cyber vulnerabilities for medical devices. The RSS-MD is derived from actual evaluations and analysis on system effects, working with both the medical community and manufacturers. The framework is founded upon academic rigor and applies a credible and repeatable approach to assigning scoring metrics to identified vulnerabilities. The attributes are consistent with objectives specified in the CVSS and include three primary benefits for the medical community:

- It standardizes vulnerability scoring for risk analysis. Organizations can prioritize risks for implementing mitigation actions because the RSS-MD uses a

standardized scoring algorithm. This allows comparison of severity based on score ratings determined by the RSS-MD.

- It provides an understandable and common framework for evaluating identified vulnerabilities. Factors for determining the RSS-MD scores are well-defined in the Impact Category and Vulnerability Characterization Category. Organizations can readily see why a vulnerability is scored in the manner it is. This prevents confusion that arises from an arbitrary score issued using an unknown method.

- It focuses on impact to patient safety. The RSS-MD is designed to help organizations gain a better understanding of the risk associated with an identified vulnerability as it relates to delivery of patient therapy. The RSS-MD is a useful tool for assessing vulnerabilities in the context of FDA Premarket and Postmarket Guidance, along with addressing other regulatory concerns.

The RSS-MD incorporates functional impact and vulnerability characterization to discern risk scoring. The RSS-MD is designed to be holistic and encompass the complex system-of-systems interactions to provide risk analysis for decision makers to prioritize mitigation efforts.

RSS-MD FUNCTIONAL IMPACT

The functional impact category reflects the impact an exploited vulnerability has on delivery of patient care. Consideration includes direct impact that could cause patient harm or death, impact to patient therapy, impact on diagnosis, or impact to a supporting system. Additionally, the scope of impact is considered to reflect the number of devices effected by an instance of exploiting the vulnerability. Together, these metrics reflect the overall impact to patient safety.

RSS-MD VULNERABILITY CHARACTERIZATION

Vulnerability characterization focuses on the details and attributes of an identified vulnerability and helps provide insight into the exploitability. The vulnerability characteristics include:

- Attack Vector: The context by which vulnerability exploitation is possible. This metric value will be larger the more remote an attacker can be in order to exploit the vulnerable.
- Complexity: The degree of difficulty associated with developing or implementing an exploit for the vulnerability. Factors to consider include amount of publicly available information and maturity of any exploit code.
- Privileges Required: The level of privileges an attacker must possess before successfully exploiting the vulnerability. This metric is greatest if no privileges are required.
- User Interaction: The requirement for a user, other than the attacker, to participate in the successful compromise of the vulnerable component. This metric determines whether the vulnerability can be exploited solely at the will of the attacker, or whether a separate user (or user-initiated process) must participate in some manner. This metric value is greatest when no user interaction is required.
- Duration: This metric captures the ability of an exploit to remain effective against a targeted system. Exploitation of vulnerabilities that remain persistent are generally more concerning than temporal effects.
- Exploit Chain: The reachability of the vulnerability with respect to implemented controls designed to prevent exploitation. Asset owners may implement remediation actions or work-around solutions to mitigate the ability to exploit the vulnerability (e.g., compensating controls). The Controlled chain identifies the scenario when an asset owner has implemented a cybersecurity protection mechanism to prevent exploitation of the vulnerability. The Uncontrolled chain identifies the scenario when cybersecurity protection mechanisms are not available or implemented for the identified vulnerability.
- Information Assurance Principles: The effect on confidentiality, integrity and/or availability specific to the system-level functionality.
    - Confidentiality: The system-level impact to the confidentiality due to a successfully exploited vulnerability. Confidentiality refers to limiting information/data access and disclosure to only authorized assets, as well as preventing access by, or disclosure to, unauthorized ones.

- Integrity: The system-level impact to integrity of a successfully exploited vulnerability. Integrity refers to the trustworthiness of data and information.
- Availability: The system-level impact to the availability of the impacted component resulting from a successfully exploited vulnerability.

Assigning ratings to the vulnerability characterization attributes provides a means of characterizing vulnerabilities without attempting to assign traditional likelihood probabilities.

RSS-MD SCORING GUIDE

The RSS-MD scoring provides three ratings: Functional Impact Score; Vulnerability Score; and Total Score. The scores are evaluated on a 0 to 10 scale in increments of .1, with higher numbers indicating more severe ratings. The scoring algorithm was developed using academic research principles in coordination with medical professionals and healthcare providers. The scoring algorithm uses a weighted scale to focus on impact to patient safety, while incorporating the attributes of vulnerability characterization. The results provide a scoring mechanism that can be used to help prioritize mitigation efforts in a consistent and measurable manner.

Users select one rating for each category. The ratings available for selection are listed in increasing severity, with the rating associated with the lowest score listed first and the rating associated with the highest score listed last. When *No Rating Selected* is chosen, a score of '0' is assigned for that associated category, and the category is not considered in the overall scoring. Note that the initial default value for each category is set to No Rating Selected. For completeness, a rating should be assigned for each category. If an attribute for a category is unknown, then further analysis is recommended so that an appropriate rating can be provided. Indeed, the categories associated with Functional Impact and Vulnerability Characterization have been identified according to extensive research in order to provide a sufficient means for classifying and scoring risks.

FUNCTIONAL IMPACT

Impact Category

This metric reflects the impact that an exploit of the identified vulnerability would have on delivery of patient therapy.

Impact Category Options	Rating Description
Potential to impact supporting systems	The targeted system supports patient care. Impact to the system does not have an immediate impact on delivery of patient therapy or diagnosis.
Potential to impact diagnosis	The targeted system supports medical diagnosis in support of patient care. Impact to the system may alters the physician's ability to adequately diagnose medical conditions.
Potential to impact patient therapy	The targeted system is important to delivery of patient care. Impact to the system may result in negative consequences to delivery of patient therapy.
Direct potential to cause patient safety event	The targeted system is vital to delivery of patient care. Impact to the system may result in a patient safety event that could cause harm or death to the patient.

Scope of Impact

This metric reflects the number of assets effected by an instance of exploiting the vulnerability.

Scope of Impact Options	Rating Description
Single	Triggering an exploit for the vulnerability affects a single susceptible system.
Multi	Triggering an exploit for the vulnerability affects multiple susceptible systems.
All	Triggering an exploit for the vulnerability affects all susceptible systems.

VULNERABILITY CHARACTERIZATION

Attack Vector

This metric reflects the context by which vulnerability exploitation is possible.

Attack Vector Options	Rating Description
Local	A vulnerability exploitable with direct access to the target system that may require the attacker to physically touch or manipulate the vulnerable component.
Adjacent	A vulnerability exploitable from an authorized system or a system that has authorized/direct access to the target system.
Remote	A vulnerability exploitable through an external access point.

Complexity

This metric describes the degree of difficulty associated with developing or implementing an exploit for the vulnerability. Factors to consider include amount of publicly available information, maturity of any exploit code, and vendor remediation level.

Complexity Options	Rating Description
High	Limited information is available to the public and there is no known automated or demonstration of exploit code.
Medium	General information is available to the public. A proof of concept exploit is available, or the effect has been demonstrated.
Low	Information is openly available to the public and working exploit code exists.

Privileges Required

This metric describes the level of privileges an attacker must possess before successfully exploiting the vulnerability. This metric is greatest if no privileges are required.

Privileges Required Options	Scoring Description
High	The attacker requires privileges that provide significant control over the vulnerable component.
Low	The attacker requires standard privileges that provide general authorization to the vulnerable component.
None	The attacker is unauthorized prior to attack, and therefore does not require any access to carry out an attack.

User Interaction

This metric captures the requirement for a user, other than the attacker, to participate in the successful compromise of the vulnerable component. This metric determines whether the vulnerability can be exploited solely at the will of the attacker, or whether a separate user (or user-initiated process) must participate in some manner. This metric value is greatest when no user interaction is required.

User Interaction Options	Rating Description
Required	Successful exploitation of this vulnerability requires a user to take some action before the vulnerability can be exploited.
None	The vulnerable system can be exploited without interaction from any user.

Duration

This metric captures the ability of an exploit to remain effective against a targeted system. Exploitation of vulnerabilities that remain persistent are generally more concerning than temporal effects.

Duration Options	Scoring Description
Volatile	The attack has a one-time effect or minimal ability to remain persistent.
Persistent	A single instantiation of an attack has a sustained ability to create an effect (e.g., installed malware that remains effective even after a power cycle).

Exploit Chain

This metric identifies if any compensating controls or work-arounds are implemented to prevent an attacker from fully executing an attack that exploits the vulnerability.

Exploit Chain Options	Rating Description
Controlled	Cybersecurity protection mechanisms are in place to prevent the realization of a full exploit chain against identified vulnerabilities.
Uncontrolled	Cybersecurity protection mechanisms are not available to prevent the full exploit chain against an identified vulnerability.

Confidentiality

This metric measures the system-level impact to the confidentiality due to a successfully exploited vulnerability. Confidentiality refers to limiting information/data access and disclosure to only authorized assets, as well as preventing access by, or disclosure to, unauthorized ones.

Confidentiality Options	Rating Description
None	There is no loss of confidentiality within the impacted component.
Low	There is some loss of confidentiality. Access to some restricted information is obtained, but the attacker does not have control over what information is obtained, or the amount or kind of loss is constrained.
High	There is total loss of confidentiality, resulting in all resources within the impacted component being divulged to the attacker.

Integrity

This metric measures the system-level impact to integrity of a successfully exploited vulnerability. Integrity refers to the trustworthiness of data and information.

Integrity Options	Rating Description
None	There is no loss of integrity within the impacted component.
Low	Modification of data is possible, but the attacker does not have control over the consequence of a modification, or the amount of modification is constrained.
High	There is a total loss of integrity.

Availability

This metric measures the system-level impact to the availability of the impacted component resulting from a successfully exploited vulnerability.

Availability Options	Rating Description
None	There is no impact to availability within the impacted component.
Low	There is reduced performance or interruptions in resource availability.
High	There is total loss of availability, resulting in the attacker being able to fully deny access to resources in the impacted component

SCORING ALGORITHM

The scores are evaluated on a 0 to 10 scale in increments of .1, with higher numbers indicating more severe ratings. The scoring algorithm was developed using academic research principles in coordination with the medical community. The scoring algorithm uses a weighted scale to focus on impact to mission effectiveness, while incorporating the attributes of vulnerability characterization.

When No Rating Selected is chosen, a score of 0 is assigned for that associated category, and the category is not considered in the overall scoring. For completeness, a rating should be assigned for each category. If an attribute for a category is unknown, then further analysis is recommended so that an appropriate rating can be provided. The categories associated with Functional Impact and Vulnerability Characterization have been identified according to extensive research in order to provide a sufficient means for classifying and scoring risks.

The formula for computing the Functional Impact score is:

$$Functional\_Impact = IC + SI$$

where *IC* is Impact Category and *SI* is Scope of Impact. Rating assignments for each variable are associated with the values from the table below.

Impact Category	**Value**
Potential to impact supporting systems	1.0
Potential to degrade general flight operations	3.5
Potential to impact safety of flight operations	6.0
Direct potential to cause a catastrophic event	8.5
Scope of Impact Value	**Value**
Single	0.5
Multi	1.0
All	1.5

The formula for computing the Vulnerability Characterization score is:

$$Vulnerability\_Characterization = AV + CM + PR + UI + D + EC + C + I + A$$

where AV is the Attack Vector, CM is the Complexity, PR is the Privileges Required, UI is the User Interaction, D is the Duration, EC is the Exploit Chain, C is the Confidentiality, I is the Integrity, and A is the Availability. Rating assignments for each variable are associated with the values from the table on the next page.

The Total Score is calculated using a weighted scale to emphasize mission impact. The formula for computing the Total Score is:

$$Total\_Score = \frac{(Mission\_Impact \times 2) + Vulnerability\_Characterization}{3}$$

The algorithm ensures that vulnerabilities are rated with respect to severity in the context of mission impact.

Attack Vector	**Value**
Local	0.4
Adjacent	0.7
Remote	1.1
Complexity	**Value**
High	0.4
Medium	0.7
Low	1.1
Privileges Required	**Value**
High	0.4
Medium	0.7
Low	1.1
User Interaction	**Value**
Required	0.6
None	1.1
Duration	**Value**
Volatile	0.6
Persistent	1.1
Exploit Chain	**Value**
Controlled	0.6
Uncontrolled	1.1
Confidentiality	**Value**
None	0.0
Low	0.7
High	1.1
Integrity	**Value**
None	0.0
Low	0.7
High	1.1
Availability	**Value**
None	0.0
Low	0.7
High	1.1

SCORING INTERPREATIONS

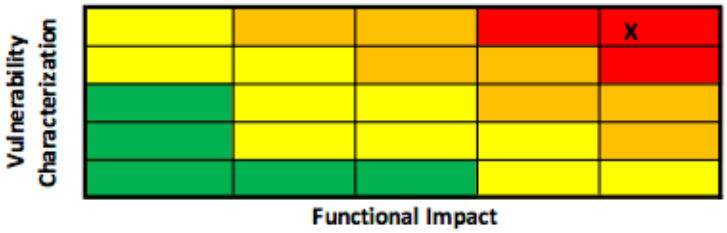
The RSS-MD provides scoring on a 0 to 10 scale, with higher numbers indicating higher severity. The example below shows the RSS-MD scoring for the previously mentioned Hospira Symbiq Infusion System.

Functional Impact	**Rating**
Impact Category	Direct potential to cause patient safety event
Scope of Impact	Single
Vulnerability Characterization	
Attack Vector	Remote
Complexity	Medium
Privileges required	None
User Interaction	None
Duration	Persistent
Exploit Chain	Uncontrolled
Confidentiality	High
Integrity	High
Availability	High
Scoring	
Functional Impact Score	9.0
Vulnerability Score	9.6
Total Score	**9.2**

Using RSS-MD, the identified vulnerability is a 9.2 Total Score. The factors include direct potential to cause a patient safety event and a single scope of impact (an exploit of the vulnerability would be a one-to-one attack). The vulnerability can be exploited remotely and a proof of concept exists. There are no privileges required on the system, no user interaction is necessary, and an exploit can remain persistent on the device. There are no protection mechanisms to prevent the full exploit chain against the vulnerability.

Finally, the vulnerability results in complete loss of system-level confidentiality, integrity and availability.

The RSS-MD also maps the scoring to a 5x5 matrix to aid in visualization of the total score. The Functional Impact Score is identified along the horizontal axis, and the Vulnerability Characterization Score is identified along the vertical axis. The Total Score is annotated via an 'X' in the matrix with implications of the finding increase in severity moving from the lower left corner of the matrix to the highest consequence associated with the top right corner of the matrix.



CONCLUSIONS

The RSS-MD provides a means to characterize identified vulnerabilities and numerically score the potential severity. The two main components include functional impact and vulnerability characterization. The RSS-MD was developed in response to gaps for evaluating associated risks of identified vulnerabilities in medical devices. The RSS-MD has been validated by myriad healthcare professionals and is currently being leveraged by major medical device manufacturers to assist in their risk assessment frameworks.