World Academy of Science, Engineering and Technology
International Journal of Computer and Information Engineering
Vol:15, No:8, 2021

A Medical Vulnerability Scoring System Incorporating Health and Data Sensitivity Metrics

Nadir A. Carreón, Christa Sonderer, Aakarsh Rao, Roman Lysecky

***Abstract*—**With the advent of complex software and increased connectivity, security of life-critical medical devices is becoming an increasing concern, particularly with their direct impact to human safety. Security is essential, but it is impossible to develop completely secure and impenetrable systems at design time. Therefore, it is important to assess the potential impact on security and safety of exploiting a vulnerability in such critical medical systems. The common vulnerability scoring system (CVSS) calculates the severity of exploitable vulnerabilities. However, for medical devices, it does not consider the unique challenges of impacts to human health and privacy. Thus, the scoring of a medical device on which a human life depends (e.g., pacemakers, insulin pumps) can score very low, while a system on which a human life does not depend (e.g., hospital archiving systems) might score very high. In this paper, we present a Medical Vulnerability Scoring System (MVSS) that extends CVSS to address the health and privacy concerns of medical devices. We propose incorporating two new parameters, namely health impact and sensitivity impact. Sensitivity refers to the type of information that can be stolen from the device, and health represents the impact to the safety of the patient if the vulnerability is exploited (e.g., potential harm, life threatening). We evaluate 15 different known vulnerabilities in medical devices and compare MVSS against two state-of-the-art medical device-oriented vulnerability scoring system and the foundational CVSS.

***Keywords*—**Common vulnerability system, medical devices, medical device security, vulnerabilities.

I. Introduction

WITH the advent of complex software and network connectivity in medical devices, there is a looming threat of attacks and vulnerabilities. Medical devices are now part of the digital health ecosystem that exacerbates any local security threat throughout the entire health ecosystem. In August 2017, the U.S. Food and Drug Administration (FDA) recalled more than 465,000 implantable cardiac devices after detecting vulnerabilities that could allow an attacker to reprogram the pacemakers [29]. In June 2019, the FDA recalled several models of an insulin pump implanted in over 4000 patients that allowed an unauthorized person to wirelessly access and change settings. And, in November 2019, 1117 insulin pump remote controller units were recalled for similar reasons. These recalls clearly demonstrate the significance and prevalence of security vulnerabilities affecting medical devices, which are likely to continue in the future.

Cybersecurity risk assessment and management must be managed and enhanced throughout the life cycle of medical devices, from design to deployment and long-term maintenance. Regulatory bodies like the FDA have established frameworks, guidelines, best practices, and standards for stakeholders to adopt for efficient cybersecurity risk management [7], [8]. Moreover, the Health Insurance Portability and Accountability Act (HIPAA) specifies a series of security procedures that should be used to ensure the confidentiality of protected electronic health information [9].

One such recommendation is to assess and characterize vulnerabilities that would aid in remediation activities. Found vulnerabilities are analyzed based on the different factors that have an impact on the vulnerability itself, such as remote exploitability, attack complexity, threat privileges, actions required by the user, exploit code maturity, and report confidence. A scoring system provides a consistent framework to quantify the impact, severity and exploitability of these security vulnerabilities. Towards this direction, the open standard CVSS was designed for assessing and quantifying the impact of software vulnerabilities. The CVSS scoring scheme represents the initial step towards comprehensive risk management by scoring vulnerabilities for system risk impact evaluation and assessment. The CVSS currently in its version 3.1 was developed mainly for Information Technology systems and applications [10]. The CVSS rubric is mainly based on three categories of metrics: i) base - represents the properties of the vulnerabilities that do not change over time, such as access complexity, access vector, compromise of confidentiality, integrity, availability of the system, and requirement for authentication to the system, ii) temporal - represents properties that do change over time, such as the existence of an official patch or functional exploit code, and iii) environmental - represents the users' IT environments, such as prevalence of affected systems and potential for loss.

Though the CVSS provides a consistent and standardized way to communicate the severity of a vulnerability, it does not address the additional challenges that are unique to medical devices and the digital health ecosystem. For example, the impact of a particular vulnerability on the health of the patient or sensitivity of patient data that could be potentially exploited should be accounted for. This is a known issue in the community, and there have been efforts to create a scoring system that takes into account the risks of exploiting a vulnerability in medical systems, and the effects it can has on

Nadir A. Carreón, Ph.D. student, and Roman Lysecky are with the Electrical and Computer Department at the University of Arizona, Tucson, AZ, 85721 USA (e-mail: nadir@email.arizona.edu, rlysecky@ece.arizona.edu).

Christa Sonderer is a Ph.D. student at the Biomedical Engineering Department at the University of Arizona, Tucson, AZ, 85721 USA (e-mail: csonderer@email.arizona.edu).

Aakarsh Rao is with Microsoft in Seattle, WA, 98052 USA (e-mail: aakarshrao7@email.arizona.edu).

the patient [5], [6], [27]. However, these approaches do not consider important health factors including the sensitivity of the data stored inside the medical system, and the repercussions of said data being manipulated or stolen. In this paper, we propose the MVSS framework that extends CVSS specifically for medical devices and digital health by carefully considering the impact of a vulnerability on health and data sensitivity. We update the scoring system equations in CVSS to incorporate these impacts in the effective vulnerability score. We evaluate MVSS framework by assessing its scores against field expert rankings for 15 vulnerabilities that impacted medical devices. Our evaluation and experimental results demonstrate that the MVSS more closely aligns with the field expert rankings than the foundational CVSS or other proposed scoring vulnerability systems for medical devices.

The rest of the paper is organized as follows. Section II presents the related work in security vulnerability scoring systems. We present the background material on CVSS in Section III. Section IV describes the proposed MVSS Model. Section V elaborates on the case studies. Section VI describes our experimental setup and showcases our results, concluding with Section VII.

II. Related Work

A plethora of security risk assessment frameworks exists for cyber-physical systems and critical infrastructure, but they are not the focus in this paper as they represent a separate set of requirements which are distinct from medical devices [1]-[4].

The CVSS provides a generalized scoring system focusing on the confidentiality, integrity, and availability (i.e., C.I.A.) exposed by the vulnerability, along with other relevant parameters including the attack vector, the attack complexity, privileges required, user interaction, and scope of the attack [10]. This widely and commonly used scoring system accurately provides a score for generalized systems, however it does not consider the human component in medical devices, which in turn causes it not to perform as expected in evaluating medical device vulnerabilities.

The MITRE Corporation, under contract to FDA have developed a rubric that provides guidance for utilizing CVSS as part of risk assessment for medical devices [5]. This rubric utilizes decision flowcharts to increase the accuracy of the values selected by a subject matter expert when using CVSS. Although the presented questions and decision flowcharts are specific and well organized, the rubric is used only to guide vulnerability assessment, in contrast with our proposed scoring system developed specifically for medical devices with metrics specifically addressing impacts on health and data sensitivity.

In the context of the potential impact of a software vulnerability to patient safety, the Risk Scoring System for Medical Devices (RSS-MD) was developed [6]. The RSS-MD mainly augments the functional impact and scope of the security vulnerability on patient therapy. However, they do not consider patient data sensitivity in their model, which is an important metric to incorporate for digital health.

Stine et. al. [27] also highlighted the need for having a scoring system that focused on connected medical devices. They developed a cyber risk scoring system for medical devices (CRSS-MD) that has two components: (i) a worst-case assessment of a scenario on which a successful attack was carried out on the device, and (ii) an assessment on the security features of the target system. However, similar to RSS-MD, they do not consider the data sensitivity in their model.

Our proposed MVSS is built upon CVSS to provide a vulnerability impact score for medical devices, which could either be used distinctly or in conjunction with guidance specifications of MITRE. The model incorporates patient data sensitivity along with patient health impact to provide a better understanding of the impact of exploiting a vulnerability in medical devices.

III. CVSS Background

The CVSS is an open framework widely used in many applications. Their scoring system reflects how severe exploiting a vulnerability is by considering its intrinsic characteristic, which are constant over time and always assume the worst-case impact on the system, commonly known as the *base score*. The score is in the range of 0-10, where 10 represents the most severe vulnerability.

CVSS also possess two additional scores, specifically the temporal score, and environmental score. The temporal metrics measure the current state of exploit techniques or code availability, and include the exploit code maturity (E), remediation level (RL) and report confidence (RC). The environmental metrics enable an analysist to customize the base scoring by assigning different values to the base metrics, based on how critical the affected component is to the user/ organization. Although these scores allow for further customization and are encouraged, they are not specifically targeting medical systems, and are not included in our case study.

The base score metrics utilized by CVSS are also used in our proposed model. They reflect the properties of the vulnerable system and are defined as follows:

Attack Vector*:* This metric represents the conduit via which the attacker is able to carry out the attack. As an example, a remote attack via the network would produce a higher base score. This metric has four different possible values:

- Network (N): The vulnerable system is connected to the network, and an attack can be carried out throughout the internet.
- Adjacent Network (A): The vulnerable system is connected to the network but is limited at the protocol level to a logically adjacent topology. Because of this, an attack can be carried out if both the vulnerable system and the attacker are connected to the same network.
- Local (L): The vulnerable system is not connected to a network, and the attacker needs to access the system physically (e.g., keyboard), or remotely (e.g., Secure Shell Protocol).
- Physical (P): The attack requires the attacker to physically manipulate the target system.

Attack Complexity: This metric reflects the amount of effort the attacker needs in order to carry out the attack. As an example, the attacker may need to recollect enough data from the target system, before successfully carrying out the attack. This metric has two possible values:

- Low (L): No special conditions are needed to carry out a successful attack.
- High (H): Successful attacks require time and effort to prepare and/or carry out the attack on the target system.

Privileges Required: This metric refers to the level of privileges on the system the attacker needs in order to carry out the attack. This metric has three possible values:

- None (N): The attacker does not need any privileges to carry out the attack.
- Low (L): The attacker requires a user's basic privileges (i.e., permission to modify only the user's settings).
- High (H): The attacker requires administrative privileges to carry out the attack (i.e., root access).

User Interaction: This metric refers to the interaction of a different user than the attacker with the target system in order to carry out the attack. This metric has two possible values:

- None (N): The attacker does not need any kind of interaction between a user and the target system to carry out an attack.
- Required (R): The attacker needs a user to perform a specific action in order to carry out the attack (i.e., custom settings on the device, installation of a 3rd party application).

Scope: This metric refers to whether attacking a vulnerable device has an effect on other resources/systems beyond its security scope. This metric has two possible values:

- Unchanged (U): Attacking a vulnerable system does not have an effect on resources/systems outside of its security scope.
- Changed (C): Attacking a vulnerable system has an effect on resources/systems outside of its security scope.

Confidentiality: This metric measures the impact to the confidentiality of the information stored in the target system. Confidentiality refers to the data being revealed only to authorized users. This metric has three possible values:

- High (H): Total loss of confidentiality, such that all data currently stored in the system are completely revealed to the attacker.
- Low (L): Some loss of confidentiality exists, whereby partial access to restricted information is revealed to the attacker.
- None (N): There is no loss of confidentiality within the attacked system.

Integrity: This metric measures the impact to the integrity of the information stored in the target system. Integrity refers to assurance and consistency of the data (e.g., the data have not been tampered). This metric has three possible values:

- High (H): Total loss of integrity, whereby all data currently stored in the system were exposed to the attacker and could be modified.
- Low (L): Partial modification of the data was possible, or the data modification does not have a serious impact on the target system.
- None (N): There is no loss of integrity within the attacked system.

Availability: This metric measures the impact to the availability of the information stored in the target system. Availability refers to being able to access the data anytime the user deems it necessary. This metric has three possible values:

- High (H): Total loss of availability, whereby the attacker can completely deny access to the information stored in the system.
- Low (L): Reduced performance or interruptions exist because of the attack on the target system.
- None (N): No loss of availability within the system.

TABLE I
CVSS 3.1 CALCULATIONS

Impact SubScore (ISS) =	1 - [(1 - Confidentiality) × (1 - Integrity) × (1 - Availability)]
	BaseScore =
If Impact ≤ 0	0, else
If Scope = U	(Min[(Impact + Exp) , 10])
If Scope = C	Min[1.08 × (Impact + Exp) , 10])
	ImpactScore =
If Scope = U	6.42 × ISS
If Scope = C	$7.52 \times (ISS - 0.029) - 3.25 \times (ISS - 0.02)^{15}$
Exploitability (Exp) =	8.22 × AttackVector × Attack Complexity × Privileges Required × User Interaction

CVSS calculates the base score as shown in Table I.

IV. Medical Vulnerability Scoring System

The MVSS vulnerability scoring system incorporates two additional metrics, *Sensitivity* and *Health Impact*.

Sensitivity: measures the impact of the type of data accessible due to a loss of confidentiality. Whereas confidentiality measures how much data on a device is revealed to an attacker, sensitivity indicates the importance or significance of the type of data that can be accessed, which can range from simple device data (e.g., sensor reading) to patients' personal information (e.g., patient name, health history). This metric has three possible values:

- High (H): Sensitive information exposed on device (personal data about one or more patients).
- Low (L): Moderately sensitive information exposed on device (sensor values, no personal data).
- None (N): No sensitive information exposed on device (there is still non-patient data on the device).

Health Impact: measures the potential impact to the physical health of the medical device user due to either a loss of availability or integrity. This metric has three possible values:

- High (H): Critical (life threatening) impact on health.
- Low (L): Non-critical impact on health.
- None (N): No impact on health.

MVSS utilizes both the Impact and Exploitability subscores to calculate a vulnerability score for a medical device. The Impact subscore involves summing the values given from the combinations of Impact variables, where each combination has a different value. The Exploitability subscore involves

multiplying the Exploitability variables. The final vulnerability score (BaseScore) is calculated by multiplying the Impact and Exploitability subscores by two constants, namely Impact weight (wI) and Exploitability weight (wE).

For our proposed vulnerability scoring system, a minimum score of 0 and a maximum score of 10 are utilized. Also, the base score is rounded up to the tenth decimal place and displayed with one decimal place. Similar to CVSS 3.1, in our model, the scope has a direct impact on the final score (i.e., if the scope is changed the final score is multiplied by 1.08).

In MVSS, the values of the data sensitivity and health impact metrics are combined with the confidentiality, integrity and availability values before being used in our scoring system. Data sensitivity is closely related to the confidentiality of the information, since breaking confidentiality would mean an attacker can access the data itself, which can contain different types of data. Because of this, these two are aggregated into the $Confidentiality_{sensitivity}$ value.

TABLE II
MATRICES FOR AGGREGATED SCORE VALUES FOR COMBINATIONS OF SENSITIVITY AND CONFIDENTIALITY, HEALTH IMPACT AND INTEGRITY, AND HEATH IMPACT AND AVAILABILITY

		Confidentiality		
Sensitivity		None	Low	High
	None	0	0.22	0.56
	Low	0	0.65	0.75
	High	0	0.85	0.95
		Integrity		
Health Impact		None	Low	High
	None	0	0.22	0.56
	Low	0.55	0.6	0.75
	High	0.85	0.9	0.95
		Availability		
Health Impact		None	Low	High
	None	0	0.22	0.56
	Low	0.55	0.6	0.65
	High	0.85	0.9	0.95

TABLE III
MVSS SCORING FOR THE 15 VULNERABILITIES CONSIDERED IN THE CASE STUDY

Vulnerability / Case Number	#1	#2	#3	#4	#5	#6	#7	#8	#9	#10	#11	#12	#13	#14	#15
Attack Vector	A	N	N	P	L	L	A	A	N	A	A	A	P	A	P
Attack Complexity	L	L	L	H	L	L	H	H	L	H	L	L	L	H	L
Privileges Required	N	N	N	N	N	L	N	N	N	N	N	N	N	N	N
User Interaction	N	N	N	N	N	R	N	N	N	N	N	N	N	N	N
Scope	U	C	C	U	U	U	U	U	U	U	U	C	U	U	C
Confidentiality Impact	H	L	H	H	L	H	N	L	H	H	H	L	H	H	L
Integrity Impact	H	H	H	H	L	L	H	H	H	H	H	H	N	H	L
Availability Impact	N	H	H	H	L	L	L	L	L	H	N	H	N	H	N
Health Impact	H	H	H	H	N	N	H	H	N	H	H	H	N	H	N
Sensitivity	L	N	H	L	N	H	N	N	H	L	L	H	H	L	H

A = Adjacent Network, N = Network, P = Physical, L = Local for Attack Vector, U = Unchanged and C = Changed for Scope, and H = High, L = Low, and N = None for every other category.

The health impact is closely related to both the availability and integrity of the systems, since not being able to access a system, and/or the system being compromised, can potentially result in harm to a patient or even loss of life. These values are aggregated into two new values, namely $Integrity_{health}$ and $Availability_{health}$. The calculated aggregated values for $Confidentiality_{sensitivity}$, $Integrity_{health}$, and $Availability_{health}$ are defined by the matrices in Table II. The values have been calculated using the values on CVSS as a starting point and adjusted to avoid saturating the score whenever one impact category is set to high, which is the reason why the High-High combinations does not utilize a value of 1. The first step is to aggregate the CIA values with the Health Impact and Data Sensitivity values, using the previously presented matrix. As an example, a value of Low for sensitivity and High for confidentiality would yield a value of $Confidentiality_{sensitivity}$ = 0.75. Next, the impact subscore is calculated using (1). The impact subscore as described by CVSS defines how significantly certain properties of the vulnerable component will be affected if it is successfully exploited.

$$ISC_{base} = Confidentiality_{sensitivity} + Integrity_{health} + Availability_{health} \quad (1)$$

The exploitability score is based on the Attack Vector (AV), Attack Complexity (AC), Privilege Required (PR) and User Interaction (UI) and is calculated using (2):

$$ESC = AV * AC * PR * UI \quad (2)$$

Finally, we calculate the base score using the previously defined weights, and calculated scores, using (3):

$$Base_{Score} = wI * ISC_{base} + wE * ESC \quad (3)$$

V. CASE STUDIES

We carried out a preliminary study of our scoring system and compared the results against other scoring systems proposed for medical devices [5], [6], [10]. Although we note that a larger study is needed, the purpose of this preliminary study is to provide proof of the added value of our scoring system against other scoring system, and a larger study is left as future work. Furthermore, although no real attack was carried out, the purpose of our proposed scoring system and this case study is to illustrate the advantages of analyzing a vulnerable device using MVSS and highlight the importance of considering the health and sensitivity of a vulnerable

device. The following are well known and documented vulnerabilities we used as test cases in our experiments. The CVSS values used are taken from the official Cybersecurity & Infrastructure Security Agency (CISA) vulnerabilities disclosures [28]. The metrics for each all of them along with the sensitivity and health impact are included in Table III. The following descriptions of the vulnerabilities are the same descriptions provided to the experts in order to rank the vulnerabilities in the case study.

1) Hospira Infusion Pump

Device Function: Inject a drug into a patient's body.

Vulnerability Description summarized from [12]: This vulnerability could allow an attacker to remotely access communications between the insulin pump and a wireless controller. If exploited, a sufficiently skilled attacker could extract sensitive information, such as device serial numbers, or replay captured wireless communications to cause an insulin delivery. This is only possible when non-default options are configured and when the insulin pump is being used with a wireless controller. Additionally, the pump will physically alert the user when insulin is injected, so the user can suspend insulin delivery.

2) Becton Dickinson Syringe Pump

Device Function: Apply precise drug doses to patients.

Vulnerability Description summarized from [13]: This vulnerability could allow an attacker to remotely control the functions of a syringe pump. If exploited, an attacker could start/stop the pump, increase the drug delivery rate up to 1000 times faster, and silence alarms. This vulnerability can be exploited when the syringe pump is connected to the hospital's network via a terminal server bridge, a relatively common practice in many hospitals. The attacker does not require additional authorization but must send proprietary commands to the syringe pump.

3) Qualcomm Datacaptor Terminal Server (DTS)

Device Function: Medical gateway device used by many U.S. hospitals to connect medical devices, such as respirators, bedside monitors, and infusion pumps, to the hospital network.

Vulnerability Description summarized from [14]: This vulnerability could allow an attacker to remotely execute code on the hospital's network to gain administrator-level privileges on the gateway device. This medical gateway device has a web management interface used for remote configuration. An attacker can send a command to the web interface without authentication to cause an arbitrary write to the gateway device's memory to login without credentials or gain administrator-level privileges on the terminal server. If exploited, an attacker could interrupt medical device connectivity to the server by crashing the server. Additionally, an attacker could access sensitive information from connected medical devices. This could cause harm to patients if important life-support or monitoring devices are disconnected.

4) Medtronic MyCare Patient Monitor

Device Function: Wirelessly communicates with implantable cardiac devices to transmit data such as heart rhythm directly to the patient's clinician.

Vulnerability Description summarized from [15]: This vulnerability could allow an attacker to access the monitor's operating system and read and write arbitrary memory values to an implanted cardiac device. The patient monitor has a hard-coded password, which an attacker could exploit by physically accessing the monitor. After removing the monitor's case, an attacker could connect to the debug port and use the hard-coded password to gain access to the operating system. From there, an attacker could use the debug function to read and write arbitrary memory values to an implanted cardiac device, provided that a patient is in close proximity to monitor at the time of the attack. Once the patient monitor device is compromised, an attacker could potentially write an invalid configuration to the implanted cardiac device, resulting in incorrect pacing and possibly fibrillation.

5) Phillips Cardiograph

Device Function: Acquires ECG signals (heart signals) from the surface of the body and records, displays, analyzes, and stores these signals for review by the clinician.

Vulnerability Description summarized from [16]: This vulnerability could allow an attacker to modify settings on the device. The cardiograph device has improper input validation (no sanitization of data entered by user) and uses hard-coded credentials. If exploited, which requires physical access to the device, an attacker could input an administrative-level password to access and modify all settings on the device.

6) Phillips CT Scanner

Device Function: Acquires medical images.

Vulnerability Description from [17]: "This vulnerability could allow an attacker to attain elevated privileges and access unauthorized system resources, including access to execute software or to view/update files including patient health information (PHI), directories, or system configuration".

7) Medtronic Insulin Pump

Device Function: Delivers medications directly into the bloodstream of a patient while in the hospital.

Vulnerability Description from [18]: "Successful exploitation of these vulnerabilities may allow an attacker to replay captured wireless communications and cause an insulin (bolus) delivery. This is only possible when non-default options are configured. Additionally, the pump will annunciate this by providing a physical alert, and the user has the capability to suspend the bolus delivery".

8) Johnson & Johnson Insulin Pump

Device Function: Injects insulin into a patient's body.

Vulnerability Description summarized from [19]: This vulnerability could allow an attacker to cause the insulin pump to administer too little or too much insulin. This insulin pump is not connected to the internet or any external network, rather it can be accessed manually or via a wireless remote, which patients use to regulate the amount of insulin administered. An attacker could exploit the wireless communication from up to

25 feet and could order the pump to inject a specified dose of insulin. An overdose of insulin could cause hypoglycemia (low blood sugar), which is very dangerous for diabetics and in extreme cases.

9) Phillips Medical Imaging Archiving Communications Systems

Device Function: Supports medical image management (via software packages): acquires, stores, distributes, processes, and displays medical images and data in a clinical environment.

Vulnerability Description from [20]: "This vulnerability could allow a low-skill attacker to compromise patient data and system availability. An attacker could use the security flaw to execute arbitrary code, alter the intended control flow of the system, access sensitive information, or cause a system crash".

10) Abbott Pacemaker

Device Function: Provides external electrical stimulation to the heart to correct improper pacing (arrhythmias)

Vulnerability Description summarized from [21]: This vulnerability could allow an attacker to access and change critical pacemaker settings. There are 465,000 affected pacemakers, all of which are RF-enabled, allowing an attacker to remotely access the pacemaker while nearby. The vulnerability itself includes improper authentication (which can be bypassed or compromised) and the lack of encryption while transmitting sensitive patient information. If exploited, an attacker could drain pacemaker battery life, change programmed settings, or change the pacing parameters of the device, all of which pose life-threatening consequences.

11) Smith's Infusion Pump

Device Function: Delivers medications directly into the bloodstream of a patient while in the hospital.

Vulnerability Description from [22]: "This vulnerability could allow an attacker to remotely alter the pump's operations, causing the incorrect drug dosage to be administered to a patient (over or under dosage). This vulnerability includes several software issues including 3rd party components, which could cause crashes or allow remote code to be executed on the devices, and issues with the device's wireless and wired network configuration and credentials. A third-party component used in the pump does not verify input buffer size prior to copying, leading to a buffer overflow, allowing remote code execution on the target device. The pump receives the potentially malicious input infrequently and under certain conditions, increasing the difficulty of exploitation".

12) Medtronic Conexus Radio Frequency Telemetry Protocol (Used in MyCareLink Monitor, CareLink Monitor, CareLink 2090 Programmer, etc.).

Monitors enable users to remotely monitor their heart condition, the implanted heart device and obtain information from the implanted heart device on an as-needed basis. Programmers allow the user to make any necessary programming changes in the cardiac device to ensure that patient is receiving the right therapy.

Device Function: Radio Frequency Communications

Vulnerability Description from [23]: "The Conexus telemetry protocol utilized within this ecosystem does not implement authentication or authorization. An attacker with adjacent short-range access to an affected product, in situations where the product's radio is turned on, can inject, replay, modify, and/or intercept data within the telemetry communication. This communication protocol provides the ability to read and write memory values to affected implanted cardiac devices; therefore, an attacker could exploit this communication protocol to change memory in the implanted cardiac device".

13) Medtronic 9790, 2090 CareLink, and 29901 Encore Programmers

Device Function from [24]: "As part of the intended functionality of this device, it may store protected health information (PHI) or personally identifiable information (PII)".

Vulnerability Description from [24]: "Exploitation of the vulnerability may allow an attacker with physical access to an affected programmer to access PHI or PII stored on the device. The affected products do not encrypt or do not sufficiently encrypt PHI or PII that may allow the identification of an individual".

14) Abbott Defibrillator

Device Function: Implantable Cardioverter Defibrillator (ICD) designed to correct a heart's irregular beating.

Vulnerability Description summarized from [25]: "Successful exploitation of these vulnerabilities may allow a nearby attacker to gain unauthorized access to an ICD to issue commands, change settings, or otherwise interfere with the intended function of the ICD. The device's authentication algorithm, which involves an authentication key and timestamp, can be compromised or bypassed, which may allow a nearby attacker to issue unauthorized commands to the ICD via RF communications. Additionally, the ICD does not restrict or limit the number of correctly formatted "RF wake-up" commands that can be received, which may allow a nearby attacker to repeatedly send commands to reduce device battery life."

15) HealthSuite Health Android App

Device Function: Health monitoring app for smartphones.

Vulnerability Description summarized from [26]: The software uses simple encryption that is not strong enough for the level of protection required. Successful exploitation of this vulnerability may allow an attacker with physical access to impact the product. These data include body measurements from different connected health devices like heart rate, activity, sleep patterns, blood pressure, body weight, calories burnt, etc. The following devices can be connected to the app: Health watch, health band, upper arm blood pressure monitor, wrist blood pressure monitor, body analysis scale, and ear thermometer.

TABLE IV
SCORES GIVEN TO ALL THE STUDIED VULNERABILITIES BY MVSS, RSS-MD, CRSS-MD, AND CVSS

Scoring System / Vulnerability	#1	#2	#3	#4	#5	#6	#7	#8	#9	#10	#11	#12	#13	#14	#15
MVSS	9.03	7.39	10	8.89	2.54	4.81	6.38	7.11	6.28	9.04	9.03	10	3.29	9.04	3.98
RSS-MD	8.67	9.1	10	8.8	5.3	5.1	8.3	8.5	4.1	8.8	8.67	9.4	3.2	9	3.2
CRSS-MD	5.89	7.88	6.05	7.13	2.14	2.14	3.37	6.52	1.5	7.31	6.6	4.8	0	6.74	1.39
CVSS	8.1	10	10	6.4	5.9	6.1	5.9	6.4	9.4	7.5	8.1	9.6	4.6	7.5	4

The gray shaded columns highlight how different the scores of a generalized vulnerability scoring system might be, versus a medical oriented scoring system.

TABLE V
RANKING FOR ALL FIFTEEN VULNERABILITIES CONSIDERED IN THE CASE STUDY

Scoring System / Vulnerability	#1	#2	#3	#4	#5	#6	#7	#8	#9	#10	#11	#12	#13	#14	#15
Expert Ranking	8	4	1	10	14	13	9	7	11	5	3	2	12	6	15
MVSS Ranking	6	8	1	7	15	12	10	9	11	3	5	2	14	4	13
RSS-MD Ranking	7	2	1	6	11	12	10	9	13	5	8	3	15	4	14
CRSS - MD Ranking	8	1	7	3	12	11	10	6	13	2	5	9	15	4	14
CVSS Ranking	6	2	1	9	13	11	12	10	4	8	5	3	14	7	15

VI. EXPERIMENTAL SETUP AND RESULTS

The weights in our system were calculated based on a survey filled out by field experts, which consisted of ranking five different known vulnerable systems targeting medical devices from the most vulnerable to the least vulnerable. The group of experts included a total of eight researchers at public universities in the U.S. and Europe whose research focus medical device security. The weights were tuned to match the average ranking yielded by the survey, and a second set containing the 15 known vulnerable systems described in Section V that also included the original five was used as the test case. For each one of the vulnerabilities, their individual score was calculated using the four different systems: CVSS 3.1, RSS-MD, CRSS-MD, and MVSS. The results are presented in Table IV. The metrics utilized to evaluate the previously mentioned vulnerabilities were taken directly from the ICS-CERT Advisories website from the U.S. Department of Homeland Security. The metrics remained mostly untouched, only changing whenever the advisory used an old version of CVSS (2.0), or discrepancies existed between the values given by their metrics and the experts' opinions.

Once calculated, the risk scores of every vulnerability for each ranking system were compared against each other. As expected, discrepancies between the vulnerability values output by the generalized CVSS and the evaluated medical devices-oriented systems were found. A few examples of these discrepancies are the following cases: Case #1 Hospira Infusion pump was given a vulnerability score of 8.1 by CVSS, 9.03 by MVSS, 8.67 by RSS-MD, and 5.89 by CRSS-MD. Exploiting the vulnerability on an infusion pump puts at risk the health of a human being. However, due to the CVSS scoring model not considering the health impact of exploiting the vulnerability, it produces a lower score for this vulnerability. Another example is Case #4 Medtronic's MyCare Patient Monitor. This vulnerability was assigned a score of only 6.4 by CVSS, while MVSS, RSS-MD, and CRSS-MD assigned scores of 8.89, 8.8, and 7.13, respectively. On the other hand, Case #9 Phillips Medical imaging archiving communication systems was given a score of 9.4 by CVSS, while MVSS, RSS-MD, and CRSS-MD assigned scores of 6.28, 4.1, and 1.5, respectively. This is due to CVSS only considering the damage exploiting the vulnerability can have on the systems, but MVSS analyzes the potential harm on human lives if the vulnerability is exploited. In this scenario, although an attack would disrupt the archiving system, it does not necessarily pose a risk to human lives, and thus is given a lower score. Similarly, Case #5 Phillips Cardiograph was given a score of 5.3 and 5.9 by RSS-MD and CVSS respectively. This device can be attacked and is highly vulnerable to modifying the settings of the device, but an attacker would not be able to steal personal information from the device, nor directly harming the patient. Although we note that it might give an improper assessment of the patient, the attack requires physical access to the cardiograph and hence, both MVSS and CRSS-MD reflects this and assign scores of 2.54 and 2.14, respectively.

Although several of the studied cases were assigned similar scores by several of the scoring systems (e.g., Cases #3, #6, #10, #12, #15), the advantages of having a medical devices-oriented scoring system can be observed by looking specifically at the scores given by RSS-MD and MVSS. Cases #1, #4, #6, #8, #10, #11, #13, and #14 were assigned similar scores, while remaining distant to the score assigned by CVSS.

As previously mentioned, a second study containing all 15 vulnerabilities was conducted. Like the first set, the second set was also sent to the experts to be ranked from the most dangerous to the least. Similarly, the results from the four different systems were ranked from most dangerous to least dangerous, the rankings can be seen on Table V. The similarity of each system's ranking against the expert ranking was calculated utilizing the average step distance, maximum step distance, and the Jaro-Winkler distance [11]. The step distance consists of measuring the distance between the expected ranking and the obtained ranking, (e.g., an obtained rank of 8, against an expected rank of 3, would yield a distance of 5). Analyzing the rankings using the step distance allows us to see how far away the vulnerability was ranked

from the expected ranking. Similarly, the average step distance allows us to see on average, how different the rankings were.

The second method was the Jaro-Winkler distance, which consists of measuring the edit distance that exists between two string sequences, and it is normalized to a value between 0 and 1. A value of 1 means no similarity between the strings, and a value of 0 would mean an exact match. The Jaro-Winkler distance is particularly useful in this scenario due to utilizing a prefix scale p, which is used to increase the rating between two strings that matches from the beginning to a specified length *l*, thus producing a higher similarity value if the highest-ranking vulnerabilities match the expert ranking in the correct order. This method also considers the number of matching characters in the string, and the distance between non-matching characters, which is crucial in making the comparison against the expert ranking. The resulting rankings can be seen in Table VI.

By looking at the Jaro-Winkler distance, our proposed system does better at classifying the vulnerabilities with a similarity of 0.11, compared against 0.12 for RSS-MD and CVSS, and 0.16 for CRSS-MD. Our system considers not only the exploitability score, but also the health and sensitivity impact exploiting the vulnerability can incur, which is a critical property that needs to be considered. Otherwise, like with the original CVSS, a vulnerability that can potentially harm humans could be assigned a low-risk score.

Finally, by looking at the step distance we can see that MVSS has the lowest average step distance, and the lowest maximum distance found across all evaluated vulnerabilities. This means that the maximum difference our system ranked a vulnerability compared against the expert's ranking was lower, compared to the others. The difference against CVSS is due to CVSS not incorporating the health impact and the sensitivity of the information. Although RSS-MD and CRSS-MD both yield better results than CVSS due to the systems focusing on medical devices, MVSS yields an improvement over both of these approaches, with a maximum step distance of only 4, compared to 5 for RSS-MD and 7 for both CRSS-MD and CVSS, along with an average step distance of 1.6 compared to 1.87, 2.8 and 2 for RSS-MD, CRSS-MD, and CVSS, respectively.

TABLE VI
JARO-WINKLER AND STEP DISTANCE OF THE ANALYZED SCORING SYSTEMS COMPARED TO THE EXPERT'S RANKING

Jaro-Winkler	Jaro-Winkler Similarity	Jaro-Winkler Distance
MVSS	0.89	0.11
RSS-MD	0.88	0.12
CRSS-MD	0.84	0.16
CVSS	0.88	0.12
Step Distance	Avg. Step Distance	Max Step Distance
MVSS	1.6	4
RSS-MD	1.87	5
CRSS-MD	2.8	7
CVSS	2	7

VII. CONCLUSIONS AND FUTURE WORK

Our proposed scoring system, MVSS, provides a framework to calculate the risk of vulnerabilities on medical devices by considering the impact exploiting the vulnerabilities can have on the patient health and data sensitivity. We demonstrate an improvement in accuracy when scoring vulnerabilities for medical devices over a general vulnerability ranking system, and two ranking systems targeting medical devices. We considered two different edit distance metrics for the comparison and verification, namely Jaro-Winkler distance, and step distance. We also evaluated the MVSS scoring system by comparing it against experts' assessment and proved that the proposed system yields a ranking closer to the experts.

As future work, a bigger case study would further add validity to our proposed system. This would not only include additional vulnerabilities, but also a larger group of experts. By doing this, the weights of our scoring system could be more carefully tuned, and our model would yield more accurate scores. Lastly, after polishing our model with a bigger case study, our goal is to develop an online version to be used by the public, in the same fashion CVSS did with their model.

ACKNOWLEDGMENTS

This research was partially supported by the National Science Foundation under Grant CNS-1615890.

REFERENCES

[1] Ralston, P.A., Graham, J.H. and Hieb, J.L., "Cyber security risk assessment for SCADA and DCS networks", ISA transactions, 46(4), pp.583-594, 2007.

[2] Vellaithurai, C., Srivastava, A., Zonouz, S. and Berthier, R., "CPIndex: Cyber-physical vulnerability assessment for power-grid infrastructures", IEEE Transactions on Smart Grid, 6(2), pp.566-575, 2014.

[3] Collier, Z.A., DiMase, D., Walters, S., Tehranipoor, M.M., Lambert, J.H. and Linkov, I., "Cybersecurity standards: Managing risk and creating resilience", Computer, 47(9), pp.70-76, 2014.

[4] DeSmit, Z., Elhabashy, A.E., Wells, L.J. and Camelio, J.A., "An approach to cyber-physical vulnerability assessment for intelligent manufacturing systems", Journal of Manufacturing Systems, 43, pp.339-351, 2017.

[5] The MITRE Corporation, "Rubric for applying CVSS to Medical Devices", October 27 2020, (Online), Available: https://www.mitre.org/sites/default/files/publications/pr-18-2208-rubric-for-applying-cvss-to-medical-devices.pdf, Accessed: January 2021.

[6] QED Secure Solutions, "Risk Scoring System for Medical Devices", 2019, (Online), Available: https://www.riskscoringsystem.com/medical/, Accessed: January 2020.

[7] FDA - U.S. Food and Drug Administration, "Guidance for the Content of Premarket Submissions for Software Contained in Medical Devices", 2005, (Online), Available: http://www.fda.gov/downloads/MedicalDevices/DeviceRegulationandGuidance/GuidanceDocuments/ucm089593.pdf, Accessed: January 2020.

[8] FDA - U.S. Food and Drug Administration, "Postmarket Management of Cybersecurity in Medical Devices", 2016, (Online), Available http://www.fda.gov/downloads/MedicalDevices/DeviceRegulationandGuidance/GuidanceDocuments/UCM482022.pdf, Accessed: January 2020.

[9] Department of Health and Human Services. Health insurance reform: security standards; final rule. Fed Regist 2003;68(34): 8334-81.

[10] FIRST.org Inc., "Common Vulnerability Scoring System v3.1", 2019, (Online), Available: https://www.first.org/cvss/v3-1/cvss-v31-specification_r1.pdf, Accessed: January 2020.

[11] Winkler, W. E., "String Comparator Metrics and Enhanced Decision Rules in the Fellegi-Sunter Model of Record Linkage", 1990, Proceedings of the Section on Survey Research Methods. American Statistical Association: 354–359.
[12] Hospira Infusion pump https://www.us-cert.gov/ics/advisories/ICSA-15-174-01
[13] Becton Dickinson Syringe pump https://www.us-cert.gov/ics/advisories/ICSMA-18-235-01
[14] Qualcomm Datacaptor Terminal Server (DTS) https://www.us-cert.gov/ics/advisories/ICSMA-18-240-01
[15] Medtronic MyCare Patient Monitor https://www.us-cert.gov/ics/advisories/ICSMA-18-179-01
[16] Phillips Cardiograph https://www.us-cert.gov/ics/advisories/ICSMA-18-228-01
[17] Phillips CT Scanner https://www.us-cert.gov/ics/advisories/ICSMA-18-123-01
[18] Medtronic insulin pump https://www.us-cert.gov/ics/advisories/ICSMA-18-219-02
[19] Johnson & Johnson Insulin pump https://www.us-cert.gov/ics/advisories/ICSMA-16-279-01
[20] Phillips Medical imaging archiving communications systems https://ics-cert.us-cert.gov/advisories/ICSMA-18-088-01
[21] Abott Pacemaker https://www.us-cert.gov/ics/advisories/ICSMA-17-241-01
[22] Smiths Infusion pump https://www.us-cert.gov/ics/advisories/ICSMA-17-250-02A
[23] Medtronic Conexus Radio Frequency Telemetry Protocol https://www.us-cert.gov/ics/advisories/ICSMA-19-080-01
[24] Medtronic 9790, 2090 CareLink, and 29901 Encore Programmers https://www.us-cert.gov/ics/advisories/ICSMA-18-347-01
[25] Abbot Defibrillator https://www.us-cert.gov/ics/advisories/ICSMA-18-107-01
[26] HealthSuite Health Android App https://ics-cert.us-cert.gov/advisories/ICSMA-18-340-01
[27] Ian Stine, M. Rice, S. Dunlap, and J. Pecarina. “A cyber risk scoring system for medical devices”. Int. J. Crit. Infrastruct. Prot. 19, C (December 2017), 32–4.
[28] ICS-CERT Advisories (Online), https://us-cert.cisa.gov/ics/advisories, Accessed: January 2021.
[29] FDA Abott Pacemaker Recall https://www.fda.gov/medical-devices/medical-device-recalls/abbott-formally-known-st-jude-medical-recalls-assuritytm-and-enduritytm-pacemakers-potential