March 2013 UILU-ENG-13-2203
CRHC-13-03

ANALYSIS OF SAFETY-CRITICAL COMPUTER FAILURES IN MEDICAL DEVICES

Homa Alemzadeh, Ravishankar K. Iyer, Zbigniew Kalbarczyk, and Jai Raman

Coordinated Science Laboratory
1308 West Main Street, Urbana, IL 61801
University of Illinois at Urbana-Champaign

REPORT DOCUMENTATION PAGE

Form Approved
OMB NO. 0704-0188

1. AGENCY USE ONLY *(Leave blank)*	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	March 2013	

4. TITLE AND SUBTITLE
Analysis of Safety-Critical Computer Failures in Medical Devices

5. FUNDING NUMBERS
2009-DT-2049 (Princeton/MARCO)
27451040-49741-A (Stanford/DTRA)

6. AUTHOR(S)
Homa Alemzadeh, Ravishankar K. Iyer, Zbigniew Kalbarczyk, and Jai Raman

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Coordinated Science Laboratory, University of Illinois at Urbana-Champaign, 1308 W. Main St., Urbana, IL, 61801-2307

8. PERFORMING ORGANIZATION REPORT NUMBER
UILU-ENG-13-2203 (CRHC-13-03)

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
MARCO (via subaward from Princeton): Microelectronics Advanced Research Corporation, Brighton Hall, Suite 120, 1101 Slater Road, Durham, NC 27703
DTRA (via subaward from Stanford): Defense Threat Reduction Agency, 8725 John J. Kingman Rd. Stop 6201, Fort Belvoir, VA 22060-6201

10. SPONSORING/MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES

12a. DISTRIBUTION/AVAILABILITY STATEMENT

Approved for public release; distribution unlimited.

12b. DISTRIBUTION CODE

13. ABSTRACT *(Maximum 200 words)*

Incidents due to malfunctioning medical devices are a major cause of serious injury and death in the United States. During 2006–2011, 5,294 recalls and around 1.2 million adverse events were reported to the U.S. Food and Drug Administration (FDA). Almost 23% of these recalls were due to computer-related failures, of which around 94% presented medium-to-high risk of severe health consequences (such as serious injury or death) to patients. This paper investigates the causes of failures in computer-based medical devices and their impact on patients, by analyzing human-written descriptions of recalls and adverse event reports, obtained from public FDA databases. We characterize computer-related failures by deriving fault classes, failure modes, recovery actions, and number of devices affected by the recalls. This analysis is used as a basis for identifying safety issues in life-critical medical devices and providing insights on the future challenges in the design of safety-critical medical devices.

14. SUBJECT TERMS
Medical devices; Failure analysis; Safety; FDA recalls; FDA adverse events

15. NUMBER OF PAGES
14

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
UNCLASSIFIED	UNCLASSIFIED	UNCLASSIFIED	UL

Analysis of Safety-Critical Computer Failures in Medical Devices

Homa Alemzadeh[1], Ravishankar K. Iyer[1], Zbigniew Kalbarczyk[1], Jai Raman[2]

[1] *University of Illinois at Urbana-Champaign - {alemzad1, rkiyer, kalbarcz}@illinois.edu*
[2] *Rush University Medical Center - jai_raman@rush.edu*

Abstract: Incidents due to malfunctioning medical devices are a major cause of serious injury and death in the United States. During 2006–2011, 5,294 recalls and around 1.2 million adverse events were reported to the U.S. Food and Drug Administration (FDA). Almost 23% of these recalls were due to computer-related failures, of which around 94% presented medium-to-high risk of severe health consequences (such as serious injury or death) to patients. This paper investigates the causes of failures in computer-based medical devices and their impact on patients, by analyzing human-written descriptions of recalls and adverse event reports, obtained from public FDA databases. We characterize computer-related failures by deriving fault classes, failure modes, recovery actions, and number of devices affected by the recalls. This analysis is used as a basis for identifying safety issues in life-critical medical devices and providing insights on the future challenges in the design of safety-critical medical devices.

I. INTRODUCTION

Electronic and computer-based devices are being widely deployed in clinical and personalized settings, facilitated by shrinking technologies, portability, and increasing interconnectedness. With ease of deployment comes significant increase in device complexity and major challenges in reliability, patient safety, and security. Medical devices are often subject to a non-negligible number of failures with potentially catastrophic impact on patients. During 2006-2011 a total of 5,294 recalls and 1,154,451 adverse events were reported to the U.S. Food and Drug Administration (FDA). As shown in Figure 1, since 2006 an overall increase of 69.8% in the number of recalls and 103.3% in the number of adverse events (reaching about 1,190 recalls, 92,600 patient injuries, and 4,590 deaths in 2011) is observed.

In this study, we focus on *Computer-related Recalls* related to failures of computer-based medical devices. During our measurement period, the number of computer-related recalls almost doubled, reaching an overall number of 1,210 (22.9% of all recalls) as shown in Figure 1.a. A previous study conducted during the 6-year period of 1999-2005, attributed 1,261 (33.4%) recalls to software-based medical devices [1]. Our goal is to identify the major causes of *Computer-related Failures* in medical devices that impact patient safety. We define *Computer-related Failure* as any event causing a computer-based medical device to function improperly or present harm to patients/users, due to failures in any of the following components of the device: *software*, *hardware*, *I/O*, or *battery*. Other failures of computer-based devices that could not be easily categorized in these four classes are classified in *other* category.

The data is collected from two public FDA databases: Medical and Radiation Emitting Device Recalls database and Manufacturer and User Facility Device Experience (Adverse Event Reports) database [2]. In-depth study of *recall data* allows us to characterize the computer-related failures based on *Fault Class:* the defective components that led to device failures; *Failure Mode:* the impact of failures on the safe functioning of the devices; *Recovery Action Category:* the types of actions taken by manufacturers to address the recalls; *Number of Recalled Devices:* the quantities of recalled devices distributed in the market; and *Device Category:* the categories and types of recalled devices. We use the overall number of devices that are affected by each recall, as a metric to measure the impact of failures.

We specifically focus on *Safety-critical Recalls* that include (i) recalls classified by FDA as Class I, presenting a high likelihood of severe injury or death to patients; (ii) recalls for which the *Reason for Recall* field specifically indicated a patient *safety* issue such as *injury* or *death*; and (iii) recalls for which the *Reason for Recall* field explicitly indicated potential of exposing patients/users to immediate *"Physical Safety Hazards"* such as *"overdose"*, *"overexposure"*, *"electrical shock"*, *"burning"*, and *"fire"* (classified under *"Physical Safety Hazard"* failure mode).

Safety-critical recalls are used as a basis to find categories and types of *Safety-critical Medical Devices*, whose failures will most likely lead to life-critical consequences for patients. Analysis of *adverse event reports* allows us to measure the impact of device failures in terms of actual adverse consequences (e.g., serious injuries or deaths) reported to the FDA. Finally, based on specific safety issues identified for life-critical medical devices, we discuss the challenges in design of next-generation medical devices.

II. DATA SOURCES

The FDA regulates medical devices sold in the U.S by requiring manufacturers to follow a set of premarket and postmarket regulatory controls. Medical devices are classified by the FDA into 5,853 distinct types and 19 medical specialties (or categories) such as Anesthesiology, Cardiovascular, Clinical Chemistry, General Hospital, General Surgery, and Radiology, indicating their regulatory class and marketing requirements. After a medical device is distributed in the market, the FDA monitors reports of adverse events and other problems with the device and alerts health professionals and the public, when needed, to ensure proper use of the devices and safety of patients.

The Recalls database is a public database of classified medical device recalls since November 1, 2002. A recall is a voluntary action taken by a company (manufacturer, distributor, or other responsible party) to correct or remove from the market any medical devices that violate the laws administrated by the FDA. Recalls are initiated to protect the public health and well-being from devices that are defective or present a risk to health such as disease, injury, or death. In rare cases, if the company fails to voluntarily recall a device that presents a health risk, the FDA may issue a recall order to the manufacturer. The FDA reviews recalls and classifies them into three classes based on the relative degree of health hazard presented by the device. A recall is classified as *Class I* when there is a reasonable chance that use of the device will cause serious adverse health problems or death. *Class II* recalls are related to devices that may cause temporary or medically reversible adverse health consequences, or pose a remote chance of serious health problems. *Class III* recalls are related to devices that violate the laws but are not likely to cause adverse health consequences.

The Adverse Event Reports (MAUDE) database is the collection of adverse events of medical devices reported to the FDA by volunteers, user facilities, manufacturers, and distributors. FDA regulations require firms that receive complaints to notify the FDA of medical device incidents, including device malfunctions, serious injuries, and deaths associated with devices. Not all reported adverse events lead to recalls; manufacturers and the FDA monitor adverse events to detect and correct the problems in a timely manner.

The Total Product Life Cycle (TPLC) database [2] integrates premarket data on medical devices including Device Classifications, Premarket Approvals (PMA), and Premarket Notifications (510(k)) with postmarket data including Adverse Events and Recalls. Each record provides the premarket review information for a device type and a list of adverse events and recalls reported for devices of that type.

III. SAFETY-CRITICAL COMPUTER-RELATED RECALLS

Data Analysis Flow

Step 1) Figure 1.c shows the overall analysis flow, which started with extraction of 13,413 recall records from the database, reported to the FDA from January 1, 2006 to December 31, 2011.

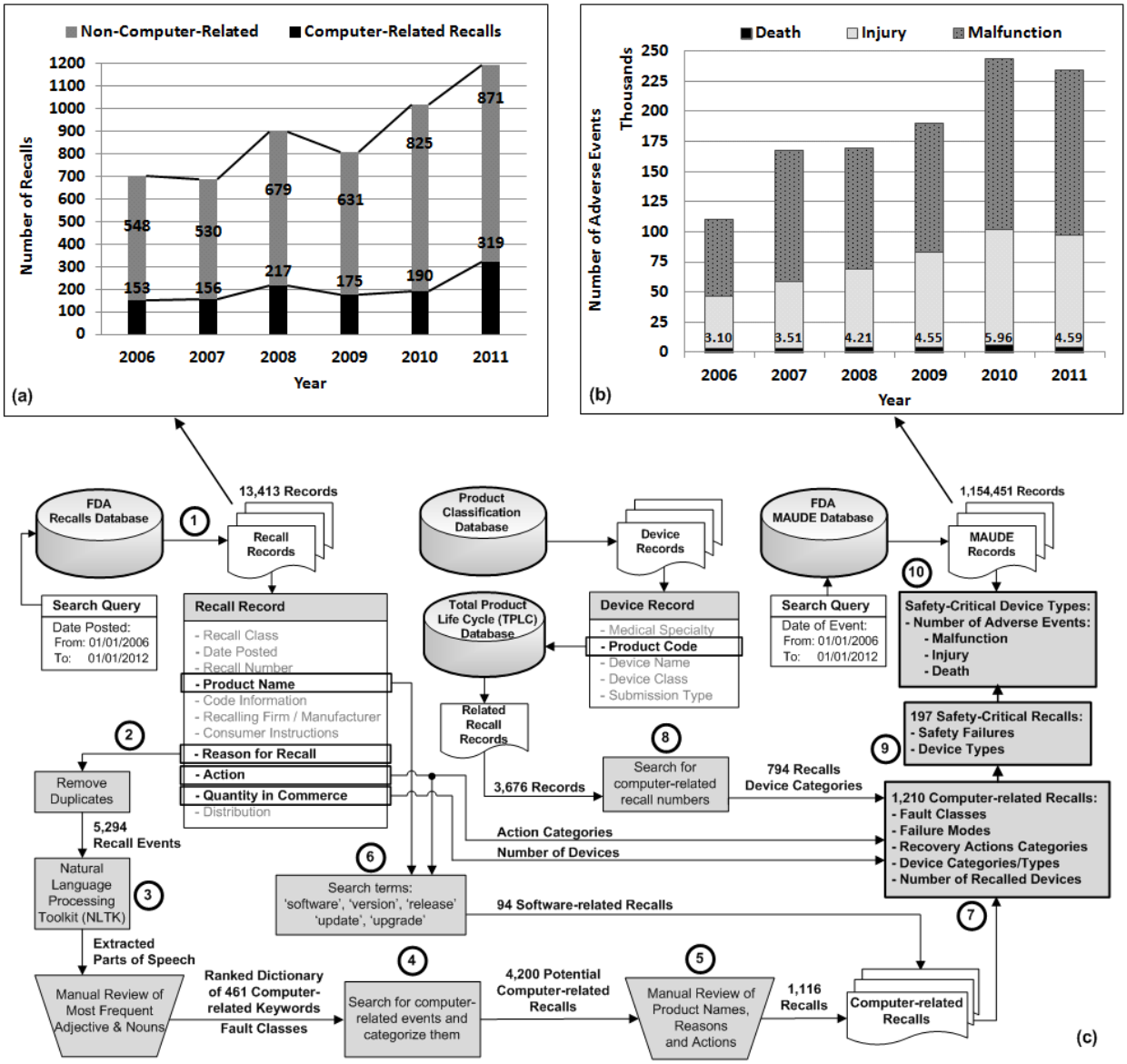


Figure 1 - (a) Total Number of Recalls per Year (2006-2011): Computer-related and Non-Computer-related, (b) Total Number of Adverse Events (2006-2011): Malfunctions, Deaths, and Injuries (Numbers on the bars indicate number of deaths in thousands per year), (c) Methodology for Analysis of Safety-Critical Computer-related Recalls.

Step 2) We identified the computer-related recalls by analysis of the *Reason for Recall* and *Action* fields of records in the FDA recalls database. Those fields contain human-written unstructured text explaining the main reason for the recall and recovery actions taken by the manufacturer to address the recall. Many of the recall records have the same reasons because the same component or part is used in different devices or models manufactured by the same company. After eliminating the duplicate values in the list of *Reasons for Recall*s (using Microsoft Excel feature for removing duplicate rows in a sheet), we came up with 5,294 unique *Recall Events* or what we refer to as *Recalls* in the FDA database.

Step 3) Using Natural Language Toolkit (NLTK), we extracted the most frequently used nouns and adjectives in human-written *Reason for Recall* fields. This list was then reviewed to create a ranked dictionary of 461 common computer-related keywords that could potentially represent failures of computer-based devices. The list of computer-related keywords was further categorized into the classes of Software, Hardware, Battery, Input/Output (I/O), and Other, corresponding to defects in different components of the device.

Step 4) We used the extracted dictionary to identify computer-related recalls by searching for keywords in *Reason for Recall* descriptions. That led us to a reduced list of 4,200 potential computer-related recalls, whose corresponding recall records were manually reviewed for validation and further categorization.

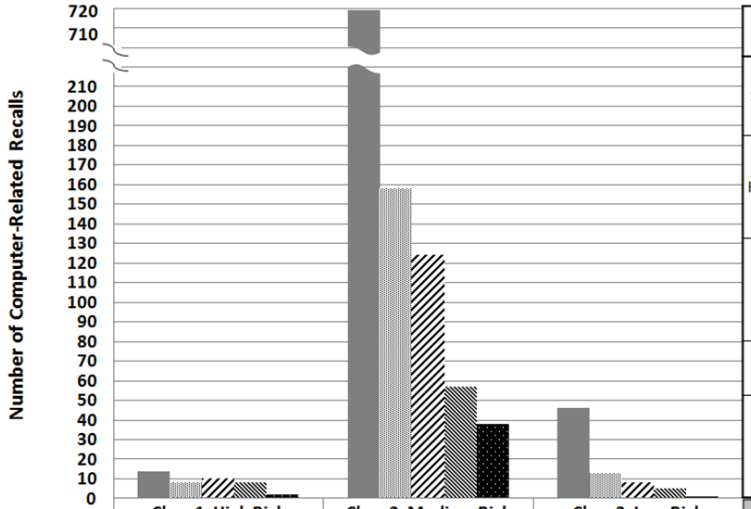
Step 5) In the manual review, we excluded many of the records from the list of computer-related recalls because their *Product Name*, *Reason for Recall* and *Action* fields did not indicate a computer-based device recall. The final list of computer-related recalls included 1,116 unique recall events.

Step 6) We also found 94 additional computer-related recalls reported because of software errors (software-related recalls) that were missed in our reason analysis process because the human-written explanations of reasons did not include any terms from the dictionary related to computers. We extracted these additional recalls by searching for the terms *'software'*, *'version'*, and *'release'* in the *Product Name* field and terms *'software'*, *'update'*, and *'upgrade'* in the *Action* fields.

Step 7) Through manual review of the computer-related recalls, we extracted *Fault Class*, *Failure Mode*, *Recovery Action Category*, and *Number of Recalled Devices* for each recall. The number of recalled devices was calculated by summing up the quantities listed in the recall records related to each recall event. For example, in Figure 2.b the fourth recall event was reported in five records in the recalls database, which together affected a total number of 7,152 devices on the market. In some instances where the total number was entered in all the recall records related to a recall event, we only counted it once.

Step 8) We used the FDA TPLC database which integrates the information such as device name, type, category (medical specialty), and regulatory class of recalled devices with a subset (3,676) of recall records. We extracted that information for 794 of computer-related recalls in our study and then used it as a training set for finding the names, types, and categories of the rest of computer-related recalls.

Finally, we ended up with a total of 1,210 computer-related recalls that affected an overall number of 12,024,836 devices distributed in the U.S. and worldwide. The 1,210 recalls were used as the basis for deriving statistics on *Fault Classes*, *Failure Modes*, and *Recovery Actions* of computer-related failures (next section), to identify safety-critical medical devices, their specific safety issues, and patient impacts (Section IV), and to provide insights on the challenges in design of next-generation medical devices (Section V).



Fault Class	Example Dictionary Keywords
Software	software, application, function, code, version, backup, database, program, bug, java, run, upgrade
Hardware	board, chip, hardware, processor, memory, disk, pcb, disk, electronic, electrical, circuit, leak, short-circuit, capacitor, transistor, resistor
Other	error, system, fail, verification, self-test, reboot, web, robotic, calculation, document, performance, workstation
Battery	battery, power, supply, outlet, plug, power-up, discharge, charger
Input/ Output	sensor, alarm, message, screen, signal, interface, monitor, connect, button, scanner, key, speaker, wireless, terminal, communication

	Class 1: High Risk	Class 2: Medium Risk	Class 3: Low Risk	Total Recalls	Num. of Devices
Software	14 (33.3%)	718 (65.6%)	46 (75.3%)	778 (64.3%)	2,303,441 (19.2%)
Hardware	8 (19.0%)	158 (14.4%)	13 (27.4%)	179 (14.8%)	4,228,133 (35.2%)
Other	10 (23.8%)	124 (11.3 %)	8 (12.3%)	142 (11.7%)	2,831,048 (23.5%)
Battery	8 (19.0%)	57 (5.2%)	5 (6.8%)	70 (5.8%)	2,385,613 (19.8%)
Input/Output	2 (4.8%)	38 (3.5%)	1 (2.7%)	41 (3.4%)	276,601 (2.3%)
Total Recalls	**42 (3.5%)**	**1095 (90.5%)**	**73 (6.0%)**	**1210**	**12,024,836**

(a)

Fault Class	Example Computer-related Recall Events					Number of Records	Number of Devices
	Year	Class	Reason for Recall	Summary of Action	Failure Mode		
Software	2008	2	The product has a software interface problem. When the product is connected to the da Vinci IS2000 System it will not allow the System to recognize the Instrument which makes the IS2000 CPG instrument nonfunctional at all sites. Risks associated are loss of operability of the instrument; delay in surgery; and loss of dexterity.	-Urgent Device Recall letter was issued to customers, instructing them to return the product. -Customers were instructed to segregate the product in a secure area for customer service representatives.	Device Operation Failure	1	11
Hardware	2007	1	Failure to Deliver Shock; a defective capacitor may cause the delay or non-delivery of the defibrillating shock which may result in failure to resuscitate the patient	-The firm issued alerts and instructions to customers on how to return their device. -The firm will exchange the recalled defibrillator with a replacement and new five (5) year warranty.	Treatment Interrupt/ Therapy Failure	1	1,794
Other	2010	1	Potential for the device to power off then on by itself; or to power off by itself and requiring the operator to turn it back on; or the device doesn't turn off.	-The firm issued an "Urgent Medical Device Correction" notification and advised customers to keep the affected device in service and to test the units in accordance with operating instructions. -A service visit was scheduled within 60 days.	Device Operation Failure	1	3,609
Battery	2008	1	Unintentional rebooting: Pump products exhibit an intermittent loss of power due intermittent loss of contact between battery cap and battery canister resulting in the device resetting. The failure of the battery cap may result in failure of the device to administer insulin therapy which may result in hyperglycemia.	-The recalling firm issued notification letters to the patients with insulin pumps to inform them of the problem and that they needed to replace the battery.	Treatment Interrupt/ Therapy Failure	5	7,152
Input/ Output	2010	1	Under certain wireless network conditions a communication error can occur; which freezes the PC Unit screen; which may result in a delay of therapy. A delay of therapy may result in serious injury and/or death.	-The firm sent a notification letter to all customers to initiate implementation of the corrective action which will require a hardware update to all affected units. - Customers do not require to return devices, but if they observe this issue, they should remove the device from service and contact the firm immediately.	Display/ Image Error	1	17,081

(b)

Figure 2 - (a) Right: Example Dictionary Keywords, Left: Distribution of Computer-related Recalls in Fault Classes & Risk Levels (b) Example Computer-related Recalls

Data Analysis Results

1) Fault Classes: Figure 2.a (right) lists example keywords from the dictionary used to identify computer-related failures in each fault class. The "Software" class represents failures due to software errors. The "Hardware" category includes both electrical issues and defects of internal circuits, while the "Input/Output (I/O)" category includes failures due to sensors, connections, display, or speakers. The "Battery" category represents defects in batteries, power cords, or power supply units that might cause interruption/failure of computer-based device function or cause harm to patients. "Battery" failures were included as computer-related failures because a typical safety-critical computer system should be able to detect, respond, and manage such failures and prevent harm to patients. The "Other" category includes recalls whose descriptions indicate a computer-related failure, but are not sufficient to be categorized in any of the above categories. Figure 2.b shows example recall records categorized in each fault class.

Figure 2.a (left) illustrates the distribution of recalls across different fault classes and recall classes (risk classes). The last column of the table shows the total number of devices on the market affected by the recalls in each fault class.

- Note that all Class I recalls are classified as safety-critical according to criterion (i) in Section I. Our analysis shows among Class I recalls, 42 were due to computer-related failures (Figure 2.a, column 2). Software failures accounted for 33.3% (14) of Class I recalls, while Hardware (8), Other (10), Battery (8), and I/O (2) combined were the reason for 66.7%. Clearly, a non-negligible fraction of computer-related recalls are due to non software-related failures.

- Majority (90.5%) of computer-related recalls were classified by FDA as Class II, with a medium risk of health consequences. Of these, we classified 66 to be safety-critical based on criterion (ii) for safety-critical recalls in Section I. In each case the manufacturer's description explicitly indicated that the device failure resulted in or had the potential to result in a patient *safety* issue, *injury*, or *death*.

- When we simply look at the overall number of recalls, similar to what other studies (e.g., [1][3][4]) reported, software is a major cause (14.7%) of the total recalls. Additionally, 64.3% of computer-related recalls are due to software failures. However, we get a very different perspective by considering the total number of devices on the market that were impacted by specific recall types (software, hardware, other, battery, and I/O). This analysis can be derived from the last column of Figure 2.a. If we look at the total number of devices, hardware-related recalls had a larger impact (almost 84% more) compared to software. Of all the recalled devices on the market, 57.3% were recalled because of hardware, battery, or I/O failures, and only 19.2% because of software faults.

2) Failure Modes: To show the breadth of failures that might impact the safe functioning of a computer-based medical device, we group the failures under six different categories shown in Table I (upper part). Number of recalls in different FDA recall classes categorized in each failure mode along with example failures of each category is shown in the table. For example, 84 of 1,210 computer-related recalls were due to failures affecting the alarm functionality of the device and were grouped under "Alarm/Message Error" category.

Table I - Computer-related Failure Modes and Recovery Action Categories in Medical Devices

Failure Mode	Recalls Count by Recall Class I	II	III	Example Failures	Example Safety-Critical Recalls: Recall Class	Recall Record Number: Reason for Recall Summary	Criteria
Alarm/ Message Error	4	76	4	- Alarm reset - Lack of audible alarms - Missed alarms - Unexpected/false alarms	**I**	Z-0051-2012: Pumps stop infusing, backup alarm sounds; but the "Run" LEDs advance as if the pumps were infusing.	(i)
Physical Safety Hazard	2	89	0	- Electrical shock - Smoke/fire/explode - Unintended movement - Overdose/over exposure	II	Z-0119-2009: A short-circuit (e.g. in a cable or the control units) can result in **uncontrolled and unstoppable movement** of the Video Fluoroscopy table. This failure might lead to serious deterioration in state of patient health.	(iii)
Display/ Image Error	1	156	11	- Blank image - Display freeze - Image distortion/corruption - Loss of image data	**I**	Z-0006-2011: Under certain wireless network conditions a communication error can occur that freezes the PC Unit screen. This failure may result in delay of therapy and serious **injury or death.**	(i) (ii)
Treatment Interruption/ Therapy Failure	18	129	3	- Delayed/failed shock delivery - Infusion/ventilation failure - Signal analysis failure - Loss of monitoring	II	Z-0689-2007: Defective integrated circuit board could result in loss of the system pump and **patient injury (hot fluid 90 degree C into uterus)**.	(ii)
Device Operation Failure	12	234	23	- Device inoperable - Failure at startup - Failure to stop exposure - Hangup/Freeze	II	Z-1474-2009: Unusual occurrence of system lockups of cardiovascular X-ray imaging systems causes image acquisition failure and user has to reset the system. **One patient death has been reported related to this issue**.	(ii)
Calculation/ Output Error	4	311	20	- Corrupted patient files - Inconsistent output - Incorrect calculation/display - Miscalculation	**I**	Z-0263-2012: Drug dosage calculation may indicate incorrect values; misalignment of ECG-ART waveforms was observed on the central station.	(i)

Recovery Action Category	Example Recovery Actions	Recalls Count
Safety Notification	"Consignees were notified by letter on/about December 1, 2005."	223
Safety Instructions	"In the notice letter, Agfa HealthCare is also providing customers with the recommended workaround. The workaround is to only print from the Viewer screen or to print the ECG once confirmed. The Viewer screen, however, does not allow the user to print batches of reports as does the Index screen."	
Software Update	"The letters stated that the recall was to the user level and requested that the user perform the software upgrade, which will eliminate the possibility of shock and burn."	632
Repair	"The notice asks that the customers inspect their units for signs of discoloration indicative of a faulty connector. The customers were instructed to return the product to CSZ for repair by contacting their Customer Service division and obtaining a Return Authorization number and specific instructions concerning packaging and returning of the unit(s) for repair."	95
Replace/ Remove	"The letter includes a response form, the firm's contact information, and indicates that the firm will exchange the recalled defibrillator with a replacement and new five (5) year warranty."	139

We draw the attention to safety-critical recalls identified based on criterion (iii) of Section I. For these 91 recalls, devices had potential to expose patients/user to immediate safety hazards (e.g., overdose, electrical shock, and fire), and are grouped under the "Physical Safety Hazards" failure mode. It is interesting that nearly all physical safety hazards were in Class II, but it is important to consider them as safety-critical because the manufacturer's description explicitly indicated a possibility of immediate harm to patients/users.

The last three columns of Table I (upper part) show example recalls in each failure mode category that are classified as safety-critical according to the criteria defined in Section I. In each example, information used for identification of the safety-critical recall is highlighted in bold in the reason description and the corresponding criteria are shown in the last column. For example, the third recall is related to a hardware defect that might lead to loss of the system pump and injection of hot fluid into patient's uterus. Although this recall was classified in Class II, it was selected as a safety-critical recall by criterion (ii).

For 113 of the 1,210 recalls there was not enough details on failure symptoms, or the event could not be classified in any of the defined failure modes.

3) Recovery Actions: We classified the recovery actions taken by manufacturers to five categories shown in the lower part of Table I. Also Figure 2.b (column 5) shows actions taken for the example recalls. In the following, the denominators refer to the number of recalls/devices in the specified fault classes:

- For 18.4% (223/1,210) of recalls the recovery action was limited to sending notifications to customers about the device problem, or providing instructions on how to avoid or workaround the problem.

- The majority of computer-related recalls due to software faults (80% = 623/778) were addressed by releasing a new software version or software patch to fix the problem. Sending notifications or instructions was the next most common action to address software-related recalls (16.7% =130/778).

- For hardware-related recalls, in most cases, customers were required to completely remove the device and/or return it to the company for replacement (36.3% = 65/179), or the device or part of it had to be corrected/repaired by the company (38.5% = 69/179). Interestingly, 4.5% (8/179) of hardware-related recalls were addressed by a software update.

- Of all the devices affected by the recalls, approximately 17.8% (2,145,087/12,024,836) required replacement of parts or a complete removal. Additionally, the majority of these replacements were because of battery (52.9% = 1,135,478/2,145,087) or hardware (37.6% = 805,868/2,145,087) failures.

These results show the importance of non software-related (e.g., hardware and battery) failures in terms of higher cost for manufacturers, caregivers, and patients. For example, for Implantable Cardioverter-Defibrillators recalled during 1990-2000, a total cost of $870 million including device checks/analyses ($83 million) and replacements ($787 million) was estimated [5]. These costs could be considerably reduced by the use of fault-tolerance techniques to enable recovery from such failures without requiring complete removal of the devices.

For 10% (121/1,210) of the records, the *Action* field information was not available or sufficient to categorize them.

IV. SAFETY-CRITICAL MEDICAL DEVICES

In the final stage of analysis we focused on safety-critical devices whose failures present the highest likelihood of severe life-critical consequences to patients **(Figure 1.c, Step 9)**. A total of 197 (16.3%) computer-related recalls were identified as safety-critical, including (i) 42 Class I recalls, (ii) 66 Class II recalls whose *Reason for Recall* field specifically indicated a patient *"safety"* issue such as *"injury"* or *"death"*, and (iii) 89 Class II recalls with a *"Physical Safety Hazards"* failure mode. Those 197 recalls together affected 2,447,894 devices on the market.

We found that the majority (80.7% = 159/197) of safety-critical recalls were for devices used in Radiology (e.g., linear accelerators), Cardiovascular (e.g., automated external defibrillators), General Hospital (e.g., infusion pumps), Anesthesiology (e.g., ventilators), and General Surgery (e.g. electrosurgical accessories). More importantly, 73.8% (31/42) of Class I recalls were for Cardiovascular and General Hospital devices such as defibrillators, patient monitors, and infusion pumps. Almost all those devices were approved by the FDA under a medium level of regulatory controls (510(k) clearance).

Table II shows example types of safety-critical medical devices that were recalled because of potential harm to patients. The total number of computer-related recalls, safety-critical computer-related recalls, and example fault classes and failures for each device type were extracted from the recalls database. The last three columns present the number of adverse events reported for these devices in the MAUDE database. We obtained these numbers by searching for the devices based on their *Names* and *Product Codes* in the MAUDE database **(Figure 1.c, Step 10)**. For extraction of computer-related adverse events, we used the *Product Problems* of the reports.

Out of 75,267 identified computer-related adverse events, around 50% (representing 398 deaths, 18,241 injuries, and 18,937 malfunctions) were related to the devices shown in Table II. However, our observation is similar to other studies ([4]) that there are inaccuracies and under-reporting in the MAUDE database and inconsistencies between MAUDE and Recalls databases. As an example of the inconsistency, we see that although safety-critical computer-related recalls affected a significant number of devices used in Radiology, very few severe adverse events due to computer problems could be identified for these devices in the MAUDE database. Nonetheless, implantable pacemakers, defibrillators, and infusion pumps dominate the computer-related failures (35 recalls) and fatalities (392 deaths). This observation can be explained by the large number of these devices in use, for treatment of critical conditions such as sudden cardiac arrest.

Table II - Safety-Critical Medical Devices:
Computer-related Recalls (Numbers and Example Safety Issues) and Adverse Event Reports

Device Category	Device Type (Product Codes)	Safety Critical Computer-related Recalls				Number of Computer-related Adverse Events		
		Num of Recalls	Num of Devices	Example Faults Classes	Example Failures	Death	Injury	Mal-function
Radiology	Linear Accelerator (IYE)	13	4,415	Software	- No interlock of beam delivery - Unexpected gantry rotation - Incorrect treatment plan - Dose delivered to wrong location - Overdose	0	0	5
				Other	- Unexpected flat panel movement - Unexpected couch movement			
	Image Processing System (LLZ)	15	15,069	Software	- Mismatched/wrong image orientation - Inaccurate annotation/data printed - Unintended images displayed - Incorrect/incomplete data displayed - Overestimated image scales	1	0	4
	Image-Intensified Fluoroscopic X-Ray System (JAA)	7	3,468	Software	- Unexpected system lockup - Inaccurate detection - Incorrect dose exposure - Unstoppable X-ray exposure - Image storage failure	0	1	2,186
Cardiovascular	Defibrillator: External (Non-Wearable) (MKJ)	17	415,537	Hardware Battery	- Delayed/failed shock delivery - Energy discharge failure	16	1	281
				Software	- Premature shutdown - Incorrect energy/shock delivery			
				Other	- Unexpected power on/off			
	Defibrillator: Implantable (NIK/LWS/ MRM)	2	170,542	Software	- Loss of rate response - Premature battery depletion - Loss of telemetry - Aborted therapy	293	14,281	11,028
	Implantable Pacemaker/ Pulse Generator (DXY/LWP/NVZ)	1	40,164	Hardware	- Loss of rate response - Premature battery depletion - Loss of telemetry	60	3,301	2,742
	Physiological Patient Monitor/ Arrhythmia Detector/Alarm (MHX/DSI)	10	38,394	Software	- Incorrect dosage - Misaligned waveforms displayed - Delayed audible alarms - Failure to restart - Burn or electrical shock hazard	4	79	276
	Pulse Oximeter (DQA)	1	14,964	Other	- Incorrect messages/alarms - Overheating - Burn, electrical shock, or fire hazard	1	0	1
General Hospital	Infusion Pump (FRN/LZH/ LKK/MEA)	15	945,300	Software	- Incorrect safety alarms	23	574	2,399
				Hardware Battery	- Delayed/over/under infusion - Infusion failure without alarms - Electrical shock, burn, or fire hazard			
	Insulin Infusion Pump (LZG)	2	13,756	Battery	- Insulin delivery failure - Unexpected shutdown w/o warning	0	4	15

V. DISCUSSION

With growing costs of healthcare, technological advances in data gathering and computing are being employed to reduce the costs while offering high quality services. Medical devices are becoming smaller, more portable, and increasingly networked via both wired and wireless networks, to provide rapid access and remote patient monitoring. By relying on complex software, sophisticated hardware, batteries, sensors, and network communications, future medical devices face several challenges in terms of reliability, safety, and security: Increased complexity raises the possibility of component interaction accidents [6]; portability makes the devices more vulnerable to power outages; and interconnectedness increases the chance of error propagation and facing failure storms that devices will not be able to handle in a fail-safe manner. Additionally, medical devices are prone to major security and privacy vulnerabilities such as unauthorized control of sensing and communication functions of devices and access to private patient data [7].

Our study found:

- While software failures remain the major cause (64.3%) for recalls of computer-based medical devices, hardware, battery and I/O are also significant contributors to failures that can lead to potential life-critical hazards.
- Hardware, battery, and I/O failures had a larger impact (57.3%) in terms of the number of devices affected by the recalls (almost 3 times) and the cost of device removals/repairs.
- By looking at example safety-critical failures studied here (e.g., hardware defect that might lead to injection of hot fluid into patient's body shown in Table I) we see that many of the recalled devices were either designed without identifying and handling the safety issues or the safety mechanisms were not designed/implemented correctly.

This emphasizes the importance of designs with well-defined safety requirements and implementations that employ robust error detection techniques and fail-safe mechanisms that are rigorously validated. In what follows we discuss major challenges in design of next-generation safety-critical medical devices.

A. Hazard and Requirements Analysis

Current hazard analysis techniques usually start with identifying possible hazards (See [8] for example hazard categories identified by FDA for infusion pumps) and safety requirements of system (See [9] for example safety requirements for a generic infusion pump model) based on an existing design. Most of the safety efforts are focused on checking and proving that design satisfies the safety requirements rather than performing a safety-driven design. Moreover, traditional safety analysis (such as Fault-tree Analysis (FTA) and Event Tree Analysis) and reliability techniques (such as Failure Mode and Effect Analysis (FMEA)) only focus on the *reliability* of individual components in the system and have limited capability in identifying other contributing factors to the system *safety* such as complex software errors, component interaction failures, human errors, complex decision-making, and flawed management in the design [6].

[6] proposes a new hazard analysis method (STPA) for safety-driven design of future complex systems. In this method instead of checking the safety requirements of a completed design, the design process is treated as an optimization problem in which the identified hazards are translated into safety constraints on system states. These constraints are enforced by design of safety features to detect, control, and mitigate the hazards in the system while making trade-off between safety, performance and cost.

B. Error Detection and Validation

Some residual faults/errors may escape even the most rigorous design and testing processes performed by manufacturers and manifest during the device's operation. Catastrophic failures caused by hard to test corner cases could be prevented by robust techniques such as formal modeling and validation. Fault injection-based validation (especially symbolic fault-injection) of key failure modes of system is an example technique that has been applied to safety-critical systems such as air traffic control software and is proven to be successful in detecting corner cases that might evade the error detection [10]. Unintentional rebooting of system due to battery failures or integrated circuit errors (such as examples in Table II) can be detected before presenting harm to patients using runtime detection techniques such as runtime assertions, watchdog timers, self-test mechanisms and periodic system checks (e.g., for critical parts such as batteries, sensors, processor, and memory). The study reported in [11] demonstrates use of static program analysis for design of application-specific runtime assertions that can detect data errors leading to failures with high coverage and very low cost.

C. Safety-based Design

Despite measures to build highly robust devices, failures that impact patient safety will inevitably happen and full device removal may not be an acceptable option because of the cost for both patients and manufacturers. There are many well-understood fail-safe mechanisms and failure-recovery techniques used in modern computing systems that can be brought into medical devices in order to manage the failures at lower cost. For example battery or hardware failures leading to power loss could be managed by turning off power to unused components of the system to maintain power for the critical parts of device and avoid sudden power outages (e.g., sleep modes are used in modern embedded systems such as cell phones). Also techniques such as fault containment (used in aerospace and commercial systems) can be used for isolating the faulty units or components (e.g. damaged batteries) and moving the system into a fail-safe mode without presenting harm to patients/users. Online detection and reconfiguration strategies for switching to backup batteries or redundant hardware units in case of failure can be employed. For example, in the second recall in Table I, the uncontrolled movement of table was not stoppable even by disconnecting the power, since a backup battery was still providing power to the device. In this case, detection of a failure in the system before deciding on recovery to backup battery could prevent the potential patient injury.

D. Recalls and Adverse Events Reporting

FDA mechanisms for reporting recalls and adverse events by manufacturers can assist in preventing future adverse events through lessons learned from the earlier problems. However, current FDA databases for reporting recalls and adverse events suffer from underreporting, inaccuracies, and inconsistencies that make it difficult in many cases to identify the causes of failures and their impact on patients to determine how the design of future devices could be improved. The following are recommendations based on our study on how the reporting mechanisms could be improved:

- Providing robust and systematic interfaces for reporting recalls and adverse events such that:
  - More accurate and complete information (e.g., *Device Name, Product Code,* and *Product Problem*) is entered in the reports.
  - List of keywords representing different *Product Problems* in the MAUDE database [2] more precisely reflects causes of device failures, especially computer-related failures.
- Creating integrated databases of recalls and adverse events such that:
  - *Product Name* and *Reason for Recall* fields of each recall record correspond to standard device names, product codes, and device categories defined by the FDA
  - Recall records can be cross-referenced with related adverse event reports in MAUDE database

E. FDA's Role in Device Regulation and Approval

FDA guidelines and safety recommendations (e.g., 2010 FDA initiative for external defibrillator improvement [12], 2010 industry guidance for infusion pumps [8], and 2013 guidance for pulse oximeters [13]) emphasize use of safety design and manufacturing practices, proper correction and communication of device problems by manufacturers, and better reporting and monitoring of adverse events to prevent the reoccurrence of failures, and enhance the resiliency of medical devices.

FDA initiatives to harden life-critical devices such as infusion pumps and external defibrillators recommend introduction of formal mechanisms for improving premarket review and approval of devices. One FDA study [14] introduced the idea of developing usage models for different classes of devices to provide generic safety features and test cases that can be used by manufactures. A more recent idea has involved the use of assurance cases for formal communication of claims, arguments, and evidences about devices from companies to the FDA. In the FDA guidance document for infusion pumps [8], manufacturers are specifically recommended to submit assurance case reports for the approval of devices. Adelard's ASCE software tool for creation and management of safety cases is currently used for the development of assurance cases for infusion pumps. In [15] a model-driven approach to derive assurance cases for cardiac pacemakers is proposed.

REFERENCES

[1] B. Zhivko, G. Mitalas, and N. Pallikarakis. “Analysis and Classification of Medical Device Recalls,” *World Congress on Medical Physics and Biomedical Engineering 2006*, pp. 3782-3785, 2006.

[2] U.S. Food and Drug Administration, “Medical Device Databases [Recalls, MAUDE, and TPLC],” http://www.fda.gov/MedicalDevices/DeviceRegulationandGuidance/Databases/default.htm.

[3] D. Wallace and D. Kuhn, “Failure Modes in Medical Device Software: An Analysis of 15 Years of Recall Data,” *International Journal of Reliability Quality and Safety Engineering*, vol. 8, pp. 351–372, 2001.

[4] K. Fu. “Trustworthy Medical Device Software. In Public Health Effectiveness of the FDA 510(k) Clearance Process: Measuring Postmarket Performance and Other Select Topics: Workshop Report,” Washington, DC, 2011. Institute of Medicine (IOM), National Academies Press.

[5] W. H. Maisel, et al., “Recalls and Safety Alerts Involving Pacemakers and Implantable Cardioverter-Defibrillator Generators,” *JAMA: The Journal of the American Medical Association* 286.7, pp.793-799, 2001.

[6] N. Leveson, *Engineering a Safer World: Systems Thinking Applied to Safety*. MIT Press, 2011.

[7] W. Burleson et al., “Design Challenges for Secure Implantable Medical Devices,” Proc. Design Automation Conf. (DAC 12), 2012, pp. 12-17.

[8] U.S. Food and Drug Administration, “Guidance for Industry and FDA Staff - Total Product Life Cycle: Infusion Pump - Premarket Notification [510(k)] Submissions,” Apr. 2010; http://www.fda.gov/medicalDevices/DeviceRegulationandGuidance/GuidanceDocuments/ucm206153.htm.

[9] D. E. Arney, *et al.*, “Generic Infusion Pump Hazard Analysis and Safety Requirements Version 1.0,” Department of Computer and Information Science, University of Pennsylvania, Technical Report No. MS-CIS-08-31, Feb. 2009.

[10] K. Pattabiraman, et al., “SymPLFIED: Symbolic program-level fault Injection and error detection framework,” *Proc. IEEE International Conference on Dependable Systems and Networks (DSN)*, 2008, pp. 472-481.

[11] K. Pattabiraman, Z. Kalbarczyk, and R. K. Iyer. “Automated derivation of application-aware error detectors using static analysis: The trusted illiac approach.” *IEEE Transactions on Dependable and Secure Computing*, vol. 8, no. 1, pp. 44-57, 2011.

[12] U.S. Food and Drug Administration, "External Defibrillator Improvement Initiative," Nov. 2010; http://www.fda.gov/downloads/MedicalDevices/ProductsandMedicalProcedures/CardiovascularDevices/ExternalDefibrillators/UCM233824.pdf.

[13] U.S. Food and Drug Administration, "Pulse Oximeters - Premarket Notification Submissions [510(k)s] - Guidance for Industry and FDA Staff," Mar. 2013; http://www.fda.gov/MedicalDevices/DeviceRegulationandGuidance/GuidanceDocuments/ucm341718.htm.

[14] R. Jetley, S. P. Iyer, and P. Jones, "A Formal Methods Approach to Medical Device Review," IEEE Computer, vol. 39, no. 4, pp. 61–67, Apr. 2006.

[15] E. Jee, I. Lee, and O. Sokolsky, "Assurance Cases in Model Driven Development of the Pacemaker Software," *Leveraging Applications of Formal Methods, Verification, and Validation*, volume 6416 of LNCS, pages 343–356. 2010.