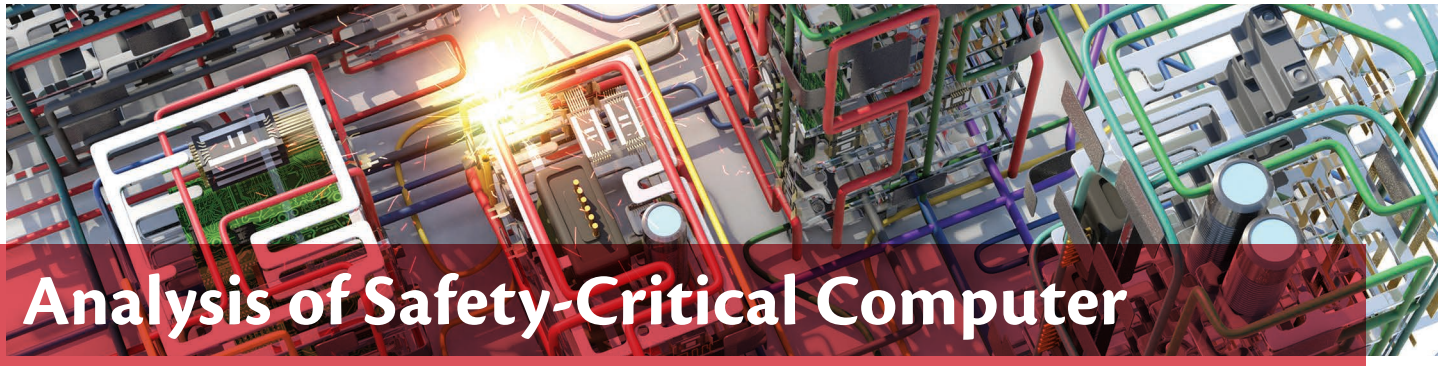


Analysis of Safety-Critical Computer Failures in Medical Devices

Homa Alemzadeh, Ravishankar K. Iyer, and Zbigniew Kalbarczyk |
University of Illinois at Urbana-Champaign
Jai Raman | Rush University Medical Center

Malfunctioning medical devices are one of the leading causes of serious injury and death in the US. Analysis of human-written descriptions of recalls and adverse event reports reveals safety issues in these devices and provides insights on the future challenges in the design of safety-critical devices.

Electronic and computer-based devices are deployed widely in clinical and personalized settings, facilitated by shrinking hardware and increased portability and interconnectedness. But with ease of deployment comes a significant increase in device complexity and major challenges in reliability, patient safety, and security. Medical devices are often subject to a nonnegligible number of failures with potentially catastrophic impacts on patients. Between 2006 and 2011, 5,294 recalls and 1,154,451 adverse events were reported to the US Food and Drug Administration (FDA). As Figure 1 shows, since 2006, there was a 69.8 percent increase in the number of recalls and a 103.3 percent increase in the number of adverse events (reaching approximately 1,190 recalls [see Figure 1a], 92,600 patient injuries, and 4,590 deaths in 2011 [see Figure 1b]).

In this article, 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 percent of all recalls), as Figure 1a shows. A study conducted during the six-year period between 1999 and 2005 attributed 1,261 recalls (33.4 percent) to software-based medical devices.[1] Our goal was 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 or users owing to failures in the device's software, hardware, I/O, or battery.

We collected data from two public FDA databases: the Medical and Radiation Emitting Device Recalls database (referred to as the Recalls database) and the Manufacturer and User Facility Device Experience (MAUDE or Adverse Event Reports) database.[2] Through an in-depth study of recalls data, we characterized the computer-related failures based on

- *fault class*: the defective components that led to device failure,
- *failure mode*: the impact of failures on the device's safe functioning,
- *recovery action category*: the type of actions the manufacturer took to address the recall,
- *number of recalled devices*: the quantity of recalled devices distributed in the market, and

 Copublished by the IEEE Computer and Reliability Societies

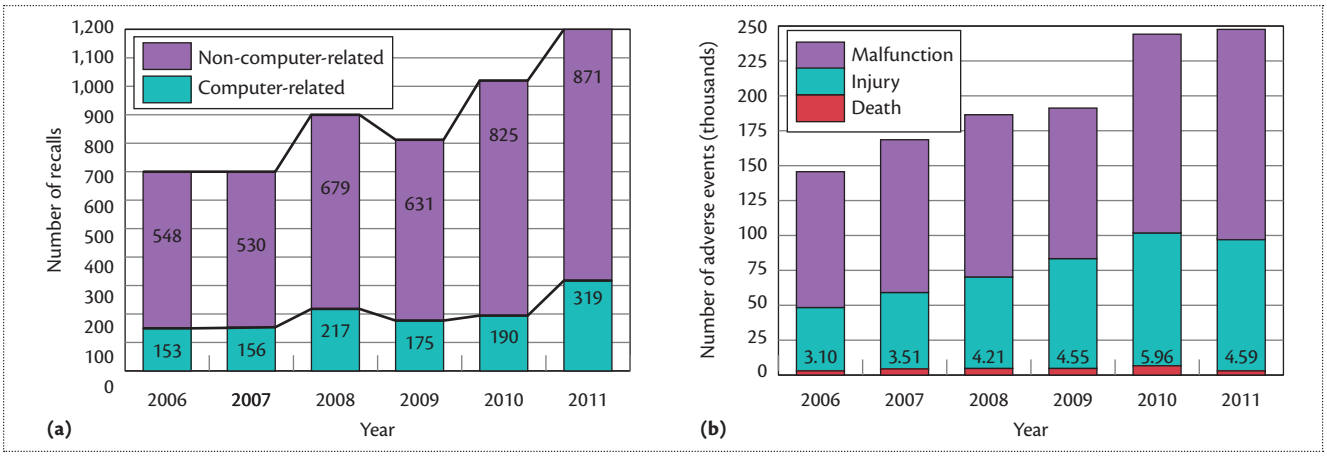


Figure 1. Recalls and adverse events between 2006 and 2011. (a) Total number of computer-related and non-computer-related recalls per year. (b) Total number of adverse events per year, including malfunctions, deaths, and injuries. Numbers on the bars indicate number of deaths in thousands.

- *device category*: the categories and types of recalled devices.

We used the overall number of devices affected by each recall as a metric to measure the impact of failures.

We specifically focused on *safety-critical recalls* and identified them based on the following criteria:

1. recalls that the FDA classified as *class I*, presenting a high risk of severe injury or death to patients;
2. recalls for which the FDA Recalls database's Reason for Recall field specifically indicated a patient safety issue such as injury or death; and
3. recalls for which the FDA Recalls database's Reason for Recall field explicitly indicated potential for exposing patients or users to immediate physical safety hazards such as overdose, overexposure, electrical shock, burning, or fire.

We used these safety-critical recalls as a basis to find categories and types of safety-critical medical devices whose failures will most likely lead to life-critical consequences. Analysis of adverse event reports let us measure the impact of device failures in terms of actual adverse consequences (for example, 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 designing the next generation of medical devices.

Data Sources

The FDA regulates medical devices sold in the US by requiring manufacturers to follow a set of pre- and postmarket regulatory controls. The FDA classifies medical devices into 5,853 distinct types and 19 medical specialties, 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, when necessary, alerts health professionals and the public to ensure proper use of the device and safety of patients.

The FDA's Recalls database contains classified medical device recalls since 1 November 2002. A recall is a voluntary action that a manufacturer, distributor, or other responsible party takes to correct or remove from the market any medical device that violates the laws administrated by the FDA. Recalls are initiated to protect the public health and well-being from devices that are defective or that present health risks 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 might issue a recall order to the manufacturer.

The FDA classifies recalls into three classes based on the relative degree of health hazard the device presents. *Class I* recalls indicate that there's a reasonable chance that use of the device will cause serious adverse health problems or death. *Class II* indicates devices that might cause temporary or medically reversible adverse health consequences or pose a remote chance of serious health problems. *Class III* indicates devices that violate the laws administrated by the FDA but aren't likely to cause adverse health consequences.

The MAUDE database is a collection of adverse events of medical devices that volunteers, user facilities, manufacturers, and distributors reported to the FDA. FDA regulations require firms that receive complaints

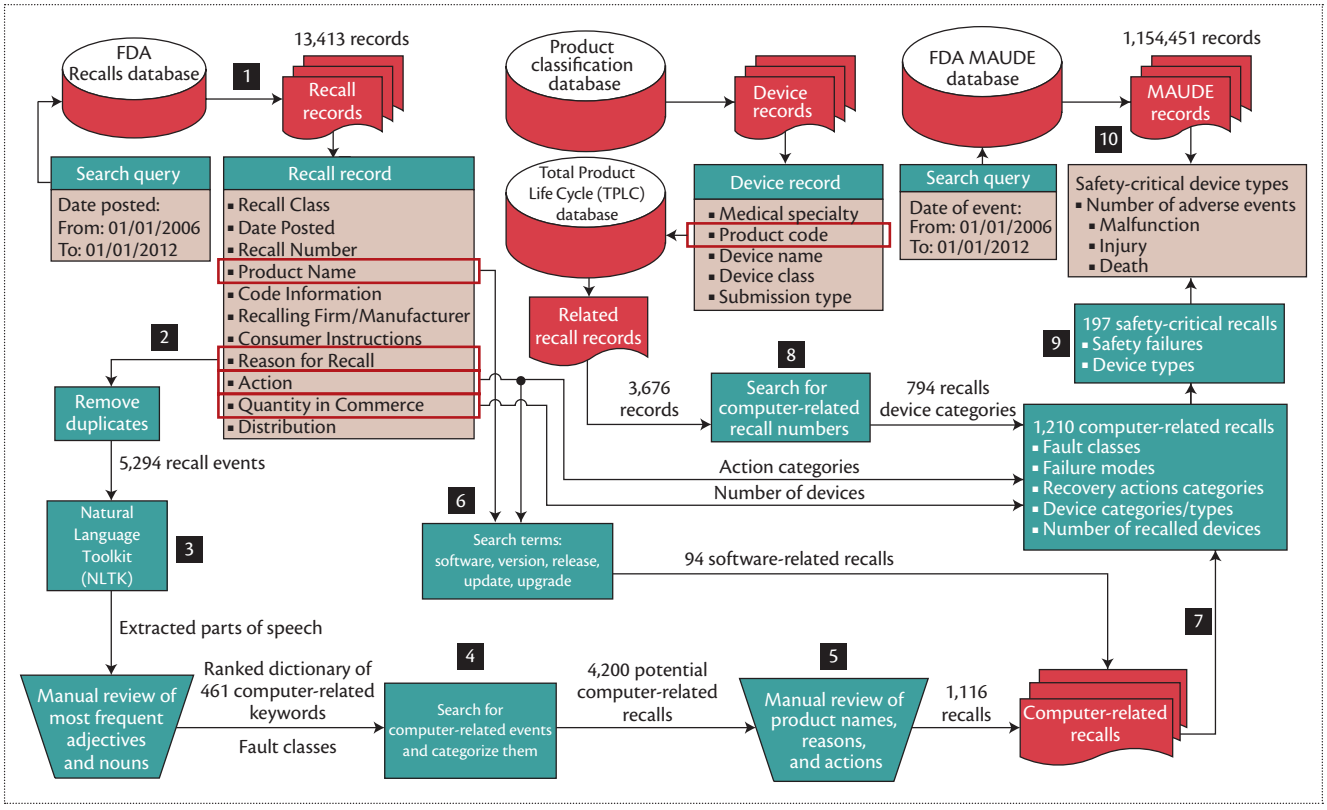


Figure 2. Methodology for analyzing safety-critical computer-related recalls.

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 problems in a timely manner.

The Total Product Life Cycle (TPLC) database 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.[2] Each record provides the premarket review information for a device type and a list of adverse events and recalls reported.

Data Analysis Flow

Figure 2 shows our analysis flow. We extracted 13,413 records from the Recalls database, reported to the FDA between 1 January 2006 and 31 December 2011 (see Figure 2, step 1).

We then identified the computer-related recalls by analyzing the Reason for Recall and Action fields in the FDA Recalls database records (see Figure 2, step 2). Those fields contain human-written, unstructured text explaining the main reason for the recall and recovery actions the manufacturer took 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 Reason for Recalls list (using Microsoft Excel to remove duplicate entries), we came up with 5,294 unique *recall events* (or what we call *recalls*) in the FDA database.

Using the Natural Language Toolkit (NLTK; a suite of Python libraries for natural language processing), we extracted the most frequently used nouns and adjectives in the human-written Reason for Recall fields (see Figure 2, step 3). We reviewed this list to create a ranked dictionary of 461 common computer-related keywords that could potentially represent failures of computer-based devices. We further categorized the list of computer-related keywords into five classes—Software, Hardware, Battery, I/O, and Other—corresponding to defects in the device's different components (see Table 1).

We then used the extracted dictionary to identify computer-related recalls by searching for keywords in the Reason for Recall descriptions (see Figure 2, step 4). This led to a reduced list of 4,200 potential computer-related recalls, whose corresponding recall records we manually reviewed for validation and further categorization.

In the manual review, we excluded many of the records from the list of computer-related recalls because

Table 1. Example dictionary keywords.

Fault class	Keywords
Software	Software, application, function, code, version, backup, database, program, bug, Java, run, upgrade
Hardware	Board, chip, hardware, processor, memory, disk, PCB, 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
I/O	Sensor, alarm, message, screen, signal, interface, monitor, connect, button, scanner, key, speaker, wireless, terminal, communication

their Product Name, Reason for Recall, and Action fields didn't indicate a computer-based device recall (see Figure 2, step 5). The final list of computer-related recalls included 1,116 unique recall events.

We found 94 additional computer-related recalls because of software errors (software-related recalls) that we missed in our reason analysis process because the human-written explanations of reasons didn't include any computer-related terms from our dictionary. We extracted these additional recalls by searching for the terms "software," "version," and "release" in the Product Name fields and the terms "software," "update," and "upgrade" in the Action fields (see Figure 2, step 6).

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 (see Figure 2, step 7). We calculated the number of recalled devices by summing up the quantities listed in the recall records related to each recall event. For example, in Table 2, the fourth recall event was reported in five records in the Recalls database, which together affected a total of 7,152 devices on the market. In some instances in which the total number was entered in all the recall records related to a recall event, we counted it only once.

We used the FDA's TPLC database, which integrates the information such as device name, type, category (medical specialty), and the regulatory class of recalled devices with a subset (3,676) of recall records. We extracted that information for 794 computer-related recalls in our study and then used it as a training set to find the names, types, and categories of the remaining computer-related recalls (see Figure 2, step 8).

Finally, we ended up with a total of 1,210 computer-related recalls that affected 12,024,836 devices distributed in the US and worldwide.

Data Analysis Results

We used the 1,210 identified computer-related recalls as the basis for deriving statistics on fault classes, failure modes, and recovery actions of computer-related failures to identify safety-critical medical devices, their specific safety issues, and patient impact.

Fault Classes

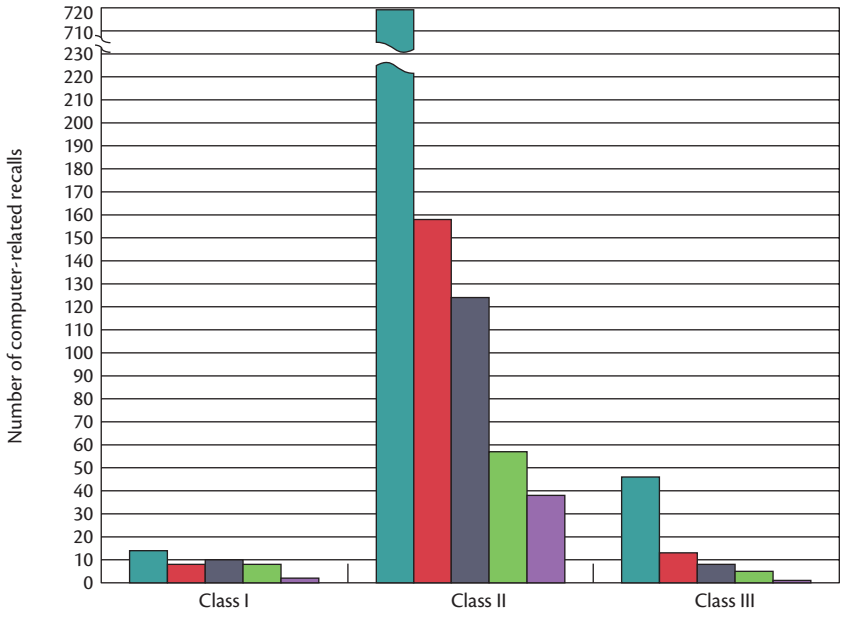
Table 1 lists example keywords from the dictionary we used to identify computer-related failures in each fault class. The Software class represents failures due to software errors. The Hardware category includes both electrical issues and defects of internal circuits, whereas the I/O category includes 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 or failure of computer-based device function or cause harm to patients. We included battery failures as computer-related failures because a typical safety-critical computer system should be able to detect, respond to, and manage such failures and prevent harm to patients. The Other category includes recalls whose descriptions indicate a computer-related failure but aren't sufficient to be classified in any of the other categories. Table 2 shows example recall records categorized in each fault class.

Figure 3 illustrates the distribution of recalls across different fault classes and recall classes (risk classes). The following are our observations based on these results:

- Note that we classify all FDA class I recalls as safety critical according to criterion 1 above. Our analysis shows that among class I recalls, 42 were due to computer-related failures (see Figure 3). Software failures accounted for 33.3 percent (14) of class I recalls, while Hardware (8), Other (10), Battery (8), and I/O (2) combined caused 66.7 percent of failures. Clearly, a nonnegligible fraction of computer-related recalls are due to non-software-related failures.
- The majority (90.5 percent) of computer-related recalls were classified as class II, with a medium risk of health consequences. Of these, we classified 66 as safety critical based on criterion 2 for safety-critical

Table 2. Example computer-related recalls.

Fault class	Example computer-related recall event					Number of records	Number of devices
	Year	Class	Reason for recall	Action summary	Failure mode		
Software	2008	II	This product has a software interface problem. When connected to the main system, it won't allow the system to recognize the instrument that makes the 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 separate the product in a secure area for customer representatives.	Device Operation Failure	1	11
Hardware	2007	I	A defective capacitor may cause the delay or nondelivery of the defibrillating shock, which might result in failure to resuscitate the patient.	■ The firm issued alerts and instructions to customers on how to return the device. ■ The firm will exchange the recalled defibrillator with a replacement and new five-year warranty. ■ A service visit was scheduled within 60 days.	Treatment Interruption or Therapy Failure	1	1,794
Other	2010	I	The device potentially powers off then on by itself; powers off, then requires the operator to turn it back on; 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.	Device Operation Failure	1	3,609
Battery	2008	I	Pump products exhibit an intermittent loss of power due to intermittent loss of contact between battery cap and battery canister, resulting in the device resetting. The failure of the battery cap might result in failure of the device to administer insulin therapy, which might 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 Interruption or Therapy Failure	5	7,152
I/O	2010	II	Speakers on the patient monitors may fail, causing absence of an audible alarm and delaying patient treatment.	■ The firm sent notification letters and instructions to customers on actions to take while awaiting their replacement speaker assemblies. ■ Affected products may continue to be used provided that the user routinely checks for the display of the Speaker Malfunction warning message at power-up. If this warning is experienced or there is no sound from speaker, device should be removed from use and the service representative should be contacted.	Alarm or Message Error	2	21,654



	Class I: high risk	Class II: medium risk	Class III: low risk	Total recalls	Number 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%)
I/O	2 (4.8%)	38 (3.5%)	1 (2.7%)	41 (3.4%)	276,601 (2.3%)
Total recalls	42 (3.5%)	1,095 (90.5%)	73 (6.0%)	1,210	12,024,836

Figure 3. Distribution of computer-related recalls in fault classes and risk levels. The last column of the table shows the total number of devices on the market affected by the recalls in each fault class.

recalls. 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 reported,[1,3,4] software is a major cause (14.7 percent) of the total recalls. In addition, 64.3 percent 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 (see the last column of Figure 3). If we look at the total number of devices, hardware-related recalls had a larger impact (almost 84 percent more) than software. Of all the recalled devices on the market, 57.3 percent were recalled because of hardware, battery, or I/O failures, and only 19.2 percent because of software faults.

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 categories: Alarm or Message Error, Physical Safety Hazard, Display or Image Error, Treatment Interruption or Therapy Failure, Device Operation Failure, and Calculation or Output Error. Table 3 shows these failure modes, along with the number of recalls in different FDA recall classes in each failure mode and example failures in each category. For example, 84 of 1,210 computer-related recalls were due to failures affecting the device's alarm functionality, so we grouped these under Alarm or Message Error. The last three columns of Table 3 show example recalls in each failure mode category that are classified as safety critical according our criteria. For example, the fourth recall relates to a hardware defect that might lead to loss of the system pump and injection of hot fluid into a patient's uterus. Although this recall was classified as class II by the FDA, our criteria classified it as a safety-critical recall by criterion 2.

For the 91 safety-critical recalls identified based on criterion 3, the devices had the potential to expose patients or users to immediate safety hazards (for example, overdose, electrical shock, and fire) and are grouped

Table 3. Computer-related failure modes

Failure mode	Recall class and count			Example failures	Example safety-critical recalls		
	I	II	III		Recall class	Recall record number: reason for recall summary	Criteria
Alarm or Message Error	4	76	4	■ Alarm reset ■ Lack of audible alarms ■ Missed alarms ■ Unexpected/false alarms	I	Z-0051-2012: pumps stop infusing and backup alarm sounds, but the "Run" LEDs advance as if the pumps were infusing.	1
Physical Safety Hazard	2	89	0	■ Electrical shock ■ Smoke, fire, or explosion ■ Unintended movement ■ Overdose or overexposure	II	Z-0119-2009: a short circuit (for example, 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 of patient health.	3
Display or 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's screen. This failure might result in delay of therapy and serious injury or death.	1 2
Treatment Interruption or Therapy Failure	18	129	3	■ Delayed/failed shock delivery ■ Infusion or ventilation failure ■ Signal analysis failure ■ Loss of monitoring	II	Z-0689-2007: defective integrated circuit board could result in the loss of the system pump and patient injury (sending 90-degree Celsius fluid into uterus).	2
Device Operation Failure	12	234	23	■ Device inoperable ■ Failure at startup ■ Failure to stop exposure ■ Hang-up or 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.	2
Calculation or Output Error	4	311	20	■ Corrupted patient files ■ Inconsistent output ■ Incorrect calculation or display ■ Miscalculation	I	Z-0263-2012: drug dosage calculation might indicate incorrect values; misalignment of electrocardiogram (ECG) waveforms was observed on the central station.	1

under the Physical Safety Hazard failure mode. It's interesting that nearly all physical safety hazards were in FDA class II, but it's important to consider them as safety critical because the manufacturer's description explicitly indicated a possibility of immediate harm to patients or users.

For 113 of the 1,210 recalls, there weren't enough details on failure symptoms, or we couldn't classify the event in any of the defined failure modes.

Recovery Actions

We classified the manufacturers' recovery actions in five categories: safety notification, safety instructions, software update, repair, and replace or remove (see Table 4). In Table 2, we showed examples of recovery actions in association with specific recalls. The following are our observations based on the recovery action results (the denominators of fractions indicate the total number of recalls or total number of devices in the specified fault classes; see Figure 3):

- For 18.4 percent (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 work around the problem.
- Manufacturers addressed 80 percent (623/778) of computer-related recalls due to software faults by releasing a new software version or patch to fix the problem. Sending notifications or instructions was the next most common action (16.7 percent; 130/778).
- For most hardware-related recalls, customers were required to completely remove the device or return it to the company for replacement (36.3 percent; 65/179), or the device or part of it had to be corrected or repaired by the company (38.5 percent; 69/179). Interestingly, 4.5 percent (8/179) of hardware-related recalls were addressed by a software update.
- Of all the devices affected by the recalls, approximately 17.8 percent (2,145,087/12,024,836) required replacement of parts or complete removal. In addi-

Table 4. Recovery action categories and examples.

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, [the manufacturer] 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 or 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

tion, the majority of these replacements were because of battery (52.9 percent; 1,135,478/2,145,087) or hardware (37.6 percent; 805,868/2,145,087) failures.

These results show the importance of non-software-related (for example, hardware and battery) failures in terms of higher cost for manufacturers, caregivers, and patients. For example, implantable cardioverter-defibrillators recalled between 1990 and 2000 cost an estimated US$870 million, including device checks and analyses ($83 million) and replacements ($787 million).[5] These costs could be considerably reduced by using fault-tolerance techniques to enable recovery from such failures without requiring complete removal of the devices.

For 10 percent (121/1,210) of the records, the Action field information wasn't available or sufficient for categorization.

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 (see Figure 2, step 9). We identified a total of 197 (16.3 percent) computer-related recalls as safety critical, including

- *criterion 1*: 42 class I recalls;
- *criterion 2*: 66 class II recalls whose Reason for Recall field specifically indicated a patient safety issue such as injury or death; and
- *criterion 3*: 89 class II recalls with a Physical Safety Hazard failure mode.

Together, those 197 recalls affected 2,447,894 devices on the market.

We found that 80.7 percent (159/197) of safety-critical recalls were for devices used in radiology (for instance, linear accelerators), cardiovascular (for instance, automated external defibrillators), general hospital (for instance, infusion pumps), anesthesiology (for instance, ventilators), and general surgery (for instance, electrosurgical accessories). More important, 73.8 percent (31/42) of class I recalls were for cardiovascular and general hospital devices, such as defibrillators, patient monitors, and infusion pumps. The FDA approved almost all those devices under a medium level of regulatory controls (510(k) clearance).

Table 5 shows example types of safety-critical medical devices that were recalled because of potential harm to patients. We extracted the total number of safety-critical computer-related recalls, number of affected devices, and example fault classes and failures for each device type from the Recalls database. Product Code fields in FDA databases uniquely identify each device type; for example, LLZ is the FDA product code for "image processing system" device type. 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 in the MAUDE database using their Product Name and Product Code information (see Figure 2, step 10). To extract computer-related adverse events, we used the reports' Product Problem fields.

Of the 75,267 identified computer-related adverse events, approximately 50 percent (representing 397 deaths, 18,237 injuries, and 18,916 malfunctions) were related to the devices in Table 5. However, our observation is similar to other studies that found inaccuracies and underreporting in the MAUDE database and inconsistencies between the FDA MAUDE and

Table 5. Safety-critical medical devices: computer-related recalls and adverse event reports.

Device category	Device type (product codes)	Safety-critical computer-related recalls			Number of computer-related adverse events		
		Number of recalls (number of devices)	Example faults classes	Example failures	Death	Injury	Malfunction
Radiology	Image processing system (LLZ)	15 (15,069)	Software	■ Mismatched or wrong image orientation ■ Inaccurate annotation or data printed ■ Unintended images displayed ■ Incorrect or 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	External defibrillator (nonwearable) (MKJ)	17 (415,537)	Hardware Battery	■ Delayed or failed shock delivery ■ Energy discharge failure	16	1	281
			Software	■ Premature shutdown ■ Incorrect energy or shock delivery			
			Other	■ Unexpected power on/off			
	Implantable defibrillator (NIK/LWS MRM)	2 (170,542)	Software	■ Loss of rate response or telemetry ■ Premature battery depletion ■ Aborted therapy	293	14,281	11,028
	Implantable pacemaker or 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 or 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
General hospital	Infusion pump (FRN/LZH/LKK/ MEA)	15 (945,300)	Software	■ Incorrect safety alarms	23	574	2,399
			Hardware Battery	■ Delayed or over/under infusion ■ Infusion failure without alarms ■ Electrical shock, burn, or fire hazard			

Recalls databases.[4] For example, we see that although safety-critical computer-related recalls affected a significant number of radiology devices (approximately 18,537 devices), the MAUDE database included very few severe adverse event reports due to computer problems for these devices (only one death and one injury). 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.

Discussion

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 (for example, the first recall in Table 2), portability makes the devices more vulnerable to power outages (for example, the fourth recall in Table 2), and interconnectedness increases the chance of error propagation (for example, the third recall in Table 3) and failure storms that devices won't be able to handle in a fail-safe manner.

In addition, medical devices are prone to major security and privacy vulnerabilities, such as unauthorized control of devices' sensing and communication functions and access to private patient data.[7] Despite the significance of challenges related to security and privacy, these issues are severely underreported in the FDA databases. A previous study indicated that 142 instances of malware infections affecting medical devices occurred between 2009 and 2011, but none of them were reported in the MAUDE database.[7] Our analysis of FDA data found three adverse events related to computer malware and virus infections in a defibrillator, a radiology workstation, and an imaging system reported by the manufacturers and user facilities, and one voluntary report of unauthorized access to a glucose monitor. We found only one FDA recall related to computer malware affecting an imaging system and categorized it under the Other fault class.

Our study found that

- although software failures remain the major cause (64.3 percent) 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 percent) in terms of the number of devices affected by the recalls (almost three times) and the cost of device removal or repair; and
- by looking at example safety-critical failures here (such as the hardware defect that might lead to injection of hot fluid into patient's body shown in Table 3), we see that many of the recalled devices were either designed without identifying and handling the safety issues or their safety mechanisms weren't designed or implemented correctly.

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

Hazard and Requirements Analysis

The international standard for risk management of medical devices (ISO 14971) and the FDA require manufacturers to maintain a process for identifying foreseeable hazardous situations of a device, estimating the risks associated with each hazard and controlling the risks by defining safety requirements and implementing effective risk control measures. The FDA guidance document for premarket notification (510(k)) submission of infusion pumps provides example hazard categories the FDA identified for infusion pumps.[8] "Generic Infusion Pump Hazard Analysis and Safety Requirements Version 1.0" shows example safety requirements derived based on these hazard categories for a generic infusion pump model.[9]

However, traditional safety analysis techniques (such as hazard and operability study [HAZOP], fault-tree analysis [FTA], and event-tree analysis) as well as reliability techniques (such as failure mode and effect analysis [FMEA]) used for probabilistic risk analysis focus only on the *reliability* of a system's individual components and have limited capability in identifying other contributing factors to safety, such as complex software errors, component interaction accidents, human errors, complex decision making, and flawed management in the design.[6]

System theoretic process analysis (STPA) is a new hazard analysis method that treats the design process as a *control optimization* problem rather than a *component failure* problem and is able to capture new causal factors (such as social, organizational, and management) contributing to the accidents that traditional hazard analysis techniques often miss.[6] STPA is a top-down process developed based on the STAMP (system-theoretic accident model and process) causality model, which focuses on identifying unsafe control actions and flawed process models of human operators, automated controllers, and other causal factors (such as contextual and environmental) to hazardous system states leading to accidents. The identified inadequate controls are translated into safety constraints on system states and are enforced in the design to detect, control, and mitigate hazards.

Error Detection and Validation

Our study on example safety-critical recalls indicates that some residual faults or errors might escape manufacturers' most rigorous design and testing processes and manifest as failures and unforeseeable hazards during the device's operation. For example, consider a defect in the integrated circuit board that might lead to

injection of hot fluid into a patient's body. Such catastrophic failures probably caused by hard-to-test corner cases could be identified by fault injection–based validation and formal modeling of key failure modes of the system. For instance, symbolic fault injection has been successful in detecting corner cases that might evade error detection in safety-critical systems, such as air traffic control software.[10]

Alternatively, techniques such as runtime assertions, watchdog timers, self-test mechanisms, and periodic system checks (for critical parts of the system such as batteries, sensors, processor, and memory) can detect the failures at runtime before leading to hazardous situations and harming patients. Techniques such as static program analysis have been demonstrated to be effective for designing application-specific runtime assertions that can detect data errors leading to application failures with high coverage and very low overhead.[11]

Fail-Safe Mechanisms

Despite measures to build highly robust devices, failures that impact patient safety will inevitably happen, and full device removal might not be an acceptable option because of the high cost for manufacturers, user facilities, and patients. Many well-understood fail-safe mechanisms and failure-recovery techniques used in modern computing systems can be brought into medical devices to manage failures at a lower cost. For example, battery or hardware failures leading to power loss and unexpected shutdowns (such as the examples in Table 5) could be managed by turning off power to unused system components to maintain power for the critical parts of the device and to avoid sudden power outages. For instance, standby or idle modes are used in low-power microcontrollers and modern embedded systems such as cell phones.

Also, techniques such as fault containment (used in aerospace and commercial systems) can be used to isolate the faulty units or components (for example, damaged batteries) and move the system into a fail-safe mode without presenting harm to patients or users. Online detection and smart reconfiguration strategies can be employed for switching to backup batteries or redundant hardware units in case of failure. For example, in the second recall in Table 3, disconnecting the power didn't stop the uncontrolled table movement because the device automatically switched to a backup battery. In this case, identifying the type of failure and the reason for power loss (whether intentional or accidental) before deciding to switch to a backup battery could stop the unintentional movement and potential patient injury.

Recalls and Adverse Events Reporting

FDA mechanisms for reporting recalls and adverse events can assist in preventing future adverse events. However, current FDA databases for reporting recalls and adverse events suffer from underreporting, inaccuracies, and inconsistencies that often make it difficult to identify the root causes of failures and their impact on patients, which in turn makes it difficult to determine how to improve design of future devices.

To improve reporting mechanisms, we recommend providing robust and systematic interfaces for reporting recalls and adverse events so that

- more accurate and complete information (for instance, Device Name, Product Code, and Product Problem) is entered in the recall reports; and
- the MAUDE database's list of keywords representing different product problems more precisely reflects causes of device failures, especially computer-related failures.

We also recommend creating integrated databases of recalls and adverse events so that

- the recall records' Product Name and Reason for Recall fields correspond to standard device names, product codes, and FDA-defined device categories; and
- recall records can be cross-referenced with related adverse event reports in the MAUDE database.

The FDA's Role in Device Regulation and Approval

FDA guidelines and safety recommendations (for instance, the 2010 FDA initiative for external defibrillator improvement,[12] 2010 industry guidance for infusion pumps,[8] and 2013 guidance for pulse oximeters[13]) emphasize the 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 medical devices' resiliency.

FDA initiatives to harden life-critical devices, such as infusion pumps and external defibrillators, recommend formal mechanisms to improve premarket review and approval of devices. One FDA study introduced the idea of developing use models for different device classes to provide generic safety features and test cases that manufacturers can use.[14] A more recent idea has involved the use of assurance cases for formal communication of claims about device functionality based on arguments supported by evidence from companies to the FDA. In the FDA's guidance document for infusion pumps, manufacturers are specifically recommended to submit assurance case reports for device approval.[8] A case study considering cardiac pacemaker software

presented an approach for constructing assurance cases for model-driven development of real-time software in safety-critical systems.[15]

Significant technological advances in data gathering and computing provide the potential to reduce healthcare costs by 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. However, we still face major challenges related to the reliability, safety, and security of medical devices. The following are our insights on some of the future challenges in design of these devices:

- Develop hazard analysis tools and safety-driven design procedures that consider flawed requirements, unsafe organizational and management decisions in design, complex software errors, accidents due to dysfunctional interactions among components, and human errors.
- Employ advanced techniques such as model checking, symbolic fault injection for comprehensive validation of the system, and application-specific techniques for runtime detection of hard-to-test anomalies.
- Introduce application- and situation-aware monitoring techniques to enable precise diagnosis of errors, smart recovery from safety-critical failures, and mitigation of undesired and harmful events.
- Improve premarket approval and postmarket surveillance of devices by introducing formal mechanisms in reviewing device applications and robust interfaces for reporting the problems.

Studying recalls and adverse event reports for computer-based medical devices allows understanding causes of safety-critical failures that can lead to life-threatening incidents. By learning from past accidents, we can identify the potential hazards, safety requirements, and risk mitigation techniques and strategies to design the next generation of devices and prevent reoccurrence of similar adverse events in the future. ■

Acknowledgments

A grant from the National Science Foundation (CNS 10-18503 CISE), an unrestricted grant from Infosys, and a faculty award from IBM partially supported this work. We thank Carol Bosley and Jenny Applequist for their editing of early drafts of the manuscript and our colleagues in the DEPEND group for many useful comments and suggestions.

References

1. B. Zhivko, G. Mitalas, and N. Pallikarakis, "Analysis and Classification of Medical Device Recalls," *Proc. Int'l Federation Medical and Biological Engineering*, Springer, 2006, pp. 3782–3785.
2. "Medical Device Databases," US Food and Drug Administration, 5 Mar. 2013; 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," *Int'l J. Reliability Quality and Safety Eng.*, vol. 8, no. 4, 2001, pp. 351–372.
4. K. Fu, "Trustworthy Medical Device Software," *Public Health Effectiveness of the FDA 510(k) Clearance Process*, Nat'l Academies Press, 2011, p. 102.
5. W.H. Maisel et al., "Recalls and Safety Alerts Involving Pacemakers and Implantable Cardioverter-Defibrillator Generators," *J. American Medical Assoc.*, vol. 286, no. 7, 2001, pp. 793–799.
6. N. Leveson, *Engineering a Safer World: Systems Thinking Applied to Safety*, MIT Press, 2011.
7. D.B. Kramer et al., "Security and Privacy Qualities of Medical Devices: An Analysis of FDA Postmarket Surveillance," *PLoS ONE*, vol. 7, no. 7, 2012, pp. 1–7.
8. Guidance for Industry and FDA Staff—Total Product Life Cycle: Infusion Pump—Premarket Notification [510(k)] Submissions," US Food and Drug Administration, Apr. 2010; 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*, tech. report MSCIS-08-31, Dept. Computer and Information Science, University of Pennsylvania, Feb. 2009.
10. K. Pattabiraman et al., "SymPLFIED: Symbolic Program-Level Fault Injection and Error Detection Framework," *Proc. IEEE Int'l Conf. Dependable Systems and Networks* (DSN), IEEE, 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 Trans. Dependable and Secure Computing*, vol. 8, no. 1, 2011, pp. 44–57.
12. "External Defibrillator Improvement Initiative," US Food and Drug Administration, Nov. 2010; www.fda.gov/downloads/MedicalDevices/ProductsandMedicalProcedures/CardiovascularDevices/ExternalDefibrillators/UCM233824.pdf.
13. "Pulse Oximeters—Premarket Notification Submissions [510(k)s] Guidance for Industry and FDA Staff," US Food and Drug Administration, Mar. 2013; www.fda.gov/MedicalDevices/DeviceRegulationandGuidance/GuidanceDocuments/ucm341718.htm.
14. R. Jetley, S.P. Iyer, and P.L. Jones, "A Formal Methods Approach to Medical Device Review," *Computer*, vol. 39, no. 4, 2006, pp. 61–67.

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*, LNCS 6416, 2010, pp. 343–356.

Homa Alemzadeh is a PhD candidate in electrical and computer engineering at the University of Illinois at Urbana-Champaign. Her research interests include measurement-based dependability evaluation, hardware-based techniques for safety and reliability, and design of medical monitoring systems. Alemzadeh received an MS in computer engineering from the University of Tehran, Iran. She's a student member of IEEE, the IEEE Computer Society, and the IEEE Engineering in Medicine and Biology Society. Contact her at alemzad1@illinois.edu.

Ravishankar K. Iyer is the George and Ann Fisher Distinguished Professor of Engineering at the University of Illinois at Urbana-Champaign. His research contributions have led to major advances in the design and validation of dependable computing systems. Iyer has a PhD in electrical engineering from the University of Queensland, Australia. He's a Fellow of the American Association for the Advancement of Science, IEEE, and the ACM. Contact him at rkiyer@illinois.edu.

Zbigniew Kalbarczyk is a research professor at the University of Illinois at Urbana-Champaign. His research interests include design, implementation, and evaluation of dependable and secure computing systems. Kalbarczyk received a PhD in computer science from the Bulgarian Academy of Sciences. He's a member of IEEE, the IEEE Computer Society, and the IFIP Technical Committee on Fault-Tolerant Computing (WG 10.4). Contact him at kalbarcz@illinois.edu.

Jai Raman is director of the Section of Adult Cardiac Surgery and surgical director, Advanced Heart Failure, Heart Transplant and Mechanical Circulatory Support Program at Rush University Medical Center in Chicago, Illinois. He's also a professor in the Department of Surgery. Raman's research interests include surgery for heart failure, diastolic function and dysfunction, extracardiac approaches to functional mitral valve disease, and topical therapy of the heart. He received an MD from St. John's Medical College in Bangalore, India. He's also a member of Society of Thoracic Surgeons, American Heart Association, American Medical Association, and the American Association for Thoracic Surgery and a Fellow of International College of Surgeons. Contact him at jai_raman@rush.edu.

Selected CS articles and columns are also available for free at http://ComputingNow.computer.org.