

Therac-25 for Treating Cancer: A computer controlled medical linear accelerator

Adapted from:

http://computingcases.org/case_materials/therac/supporting_docs/therac_case_narr/therac_case_intro.html

Therac-25: Safety is a System Property

Normally, when a patient is scheduled to have radiation therapy for cancer, he or she is scheduled for several sessions over a few weeks and told to expect some minor skin discomfort from the treatment. The discomfort is described as being on the order of a mild sunburn over the treated area. In the case you are about to read, a very abnormal thing happened to several patients: they received severe radiation burns resulting in disability, and, in 3 cases, death.

The Therac-25 was a device that targeted electron or X-ray beams on cancerous tissue to destroy it. Electron beams were used to treat shallow tissue, while photon beams could penetrate with minimal damage to treat deep tissue. Even though operators were told that there were "so many safety mechanisms" that it was "virtually impossible" to overdose a patient, this is exactly what did occur in six documented cases [Leveson].

These massive radiation overdoses were the result of a convergence of many factors including

- 1) simple programming errors
- 2) inadequate safety engineering
- 3) poor human computer interaction design
- 4) a lax culture of safety in the manufacturing organization
- 5) inadequate reporting structure at the company level and as required by the U.S. government

In presenting this case we are not interested in determining who should be blamed for these accidents. All the cases have already gone through the courts and have been settled.

We are interested in helping you learn how to think about the design and use of software in safety-critical applications.

- 1) What are the responsibilities of the organizations and individuals involved?
- 2) What design decisions and organizational structures led to the accidents?
- 3) How might different organizational systems or software design have helped avoid or minimize the harm?
- 4) What is an appropriate level of testing and verification necessary for reasonable safety?

Structure of the Therac-25 Case

Our presentation of the case itself is composed of three parts: introductory materials, a description of the machine, and overviews of the participants in the case. Together, these sections give one a good idea of the information each actor in the case had at the time of the accidents.

We reserve any analysis of this case for the teaching section. However, many of the sections contain broad hints regarding the danger of the machine and the particular ways that inadequate software design might cause harm to patients.

Introductory materials

These provide some background for students to understand the case. There is a general introduction to the case, explanations of how radiation therapy works, and a section on how medical linear accelerators work.

The machine

This section provides an overview of how the Therac-25 machine itself worked. This includes a description of the turntable, the rooms in which the machine is placed, and the role of the operator in setting up the machine.

There is also a section on the design of the software. This is a high-level introduction to the issues involved in the design of the software. The excerpts from Leveson we provide in the resource section provide much more detail, down to two particular coding errors that probably caused some of the accidents.

Finally there is a section specifically on safety. The issues involved in removing the hardware interlocks are explained, as are the issues of sensing the position of the turntable and of reuse of software from older Therac machines.

The Participants

Each of four participants are presented here, along with the accounts of each accident. The perspectives of the designer/manufacture of Therac, of the FDA, of the hospitals, and of the operators of the machines are all presented in some detail. This will allow you to assign individuals to cover the perspectives of each of these groups.

Discussion

The safety of the Therac-25 is not really a property of the machine alone. Accidents that go unreported contribute to (or at least fail to stop) later accidents. When the TV camera in the room is unplugged, the operator cannot see that the patient is in trouble. So safety is really a property of the entire technical and social system (socio-technical system). In a similar manner, an ethical analysis of the issues in this case requires an awareness of the entire socio-technical system.

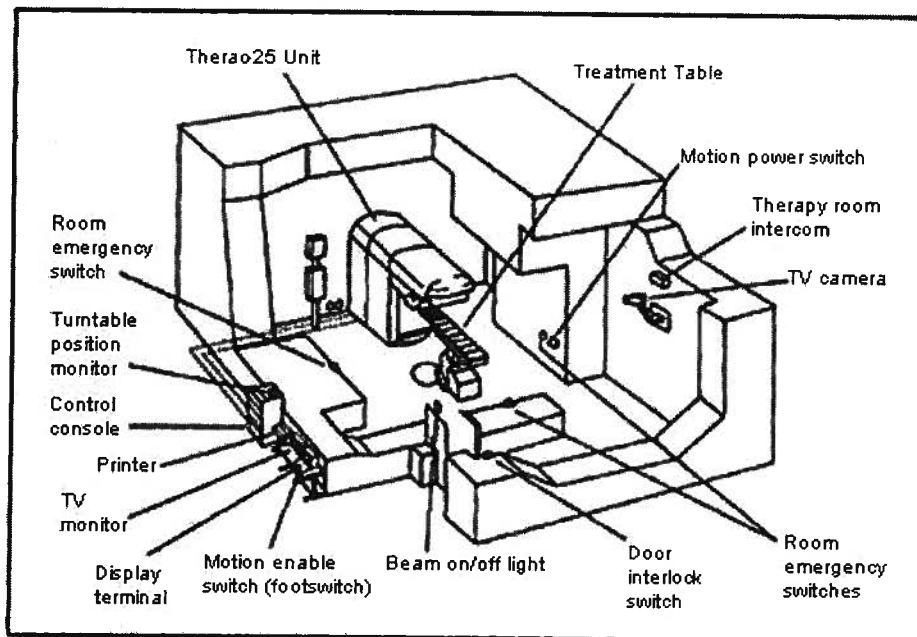


Figure 1. Typical Therac-25 facility

The Therac-25 Medical Linear Accelerator is a large machine that sits in a room designed just for it. We think of the machine itself or the machine-in-the-room as the system. But the larger system, or the Socio-Technical system, that we need to think about includes:

- *Hardware*: The mechanics of the machine itself, including its associated computer
- *Software*: the operating system of the computer and the operating system of the machine
- *Physical surroundings*: the room with its shielding, cameras, locking doors, etc.
- *People*: operators, medical physicists, doctors, engineers, salespeople, managers at AECL, government regulators
- *Institutions*: AECL, FDA, each medical facility, associations of operators, etc.
- *Procedures*
 - *Management models*: AECL's model of how risk is managed
 - *Reporting relationships*: who was required to report accidents to whom
 - *Documentation requirements*: for the software, for the facilities, for the FDA

- *Data flow*: how different parts of AECL shared information, how information was shared among agencies and organizations, how data was used by the Therac software.
- *Rules & norms*: what patients are "normally" told, what operator & physicist responsibilities are, expectations set for the programmer
- *Laws and regulations*: Reporting requirements, FDA enforcement mechanisms, medical liability law
- *Data*: data was collected in FDA approval process, use of data in Therac software,

The following table presents some of these items in a schematic form.

The Socio-Technical System	
<p>The Machine</p> <ul style="list-style-type: none"> • Supporting Systems (video, audio, etc.) • Hardware • Software Systems 	<p>Hospitals and Clinics</p> <ul style="list-style-type: none"> • Doctors, Medical Physicists • Management, User Groups • Operators, Reporting Procedures
<p>Atomic Energy Canada, Limited</p> <ul style="list-style-type: none"> • Management, Reporting Procedures, • Design Teams, Sales Staff, Support and Field Engineers 	<p>Government Medical Device Regulation</p> <ul style="list-style-type: none"> • Food and Drug Administration • Canadian Radiation Protection Bureau • Reporting Procedures

A thorough investigation of the Therac-25 case requires some grasp of most of these items. You will come across most of these items as you read this case. Setting your sights on the entire system will help you avoid the trap of finding a single point of blame. It is easy, for instance, to decide that the programmer made serious mistakes and to end one's analysis there. This is a short-sighted approach. It would miss the problems with maintenance in the cancer therapy facilities; it would miss the incomplete reporting requirements for the FDA; it would miss the inadequate and misleading testing of the Therac-25 system.

Exercises for Therac-25

Some initial considerations in teaching this case

The Therac-25 case is complex and multi-layered enough to require more than a simple once over to understand. There are multiple actors, some of them representing the same entity at different times. There are closely interwoven networks of action and reaction guided by multiple and mixed motives, where the real state of the information available to an actor at any one time is unclear.

This is not, however, simply the uniqueness of the Therac-25 case, it is a property of all cases if they are studied closely enough. Finally, it is a property of the real life of technology in use. We provide here some exercises to help students grapple with the complexity of these situations.

But first a comment on simple answers. We recommend you read the section on pitfalls before teaching this case. It outlines ways to approach this case that bring only a shallow level of understanding to the complexities. In the Therac-25 case, one of these pitfalls (single causation) leads to the tendency to fix each error one discovers with a local "patch." This usually increases the complexity of the system, provides false confidence in its safety, and does not address the design issues that led to the existence of the error in the first place. This is clearly the kind of thinking that AECL indulged in during its initial reactions to the early accidents. We recommend you help your students avoid it as they approach this case.

To make this go faster, you might assign groups to construct and present their analysis of the duties and rights of each of the main stakeholders presented in the case: AECL, FDA, hospitals, operators, and patients.

Gather data

1. **List the relevant stakeholders.** Start with some of the groups mentioned in the socio-technical system page. However, do not end there. Notice that our accident victims, the patients, are not included. Other important groups may also be omitted (e.g. "the public"). The ImpactCS framework provides you with a useful guide to different levels of stakeholders that you might overlook.
2. **Outline the duties and rights the stakeholders have toward each other.** This is best done with a drawing of each stakeholder with arrows indicating duties one owes to other and rights one has. Duties always have targets, one has duties to a particular person (even to oneself). Rights may appear to be free floating (e.g. not to be harmed) but they can often be translated into duties that others have toward the individual (avoid harming X). The ImpactCS framework provides a useful guide to outlining these duties and rights. Use the list of ethical issues to remind yourself of rights and duties in the range of likely ethical domains.

Analyze the data

1. **List the relevant opportunities and vulnerabilities that each stakeholder had in the case.** This is the beginning of what Collins and Miller call a utilitarian ethical analysis.

Who is being helped and harmed? What advantages or opportunities does each party receive in this case? What costs or dangers, or vulnerabilities does each party experience?

2. **Determine to what degree each stakeholder's duties were fulfilled or neglected.**
3. **Determine to what degree each stakeholder's rights were violated or protected, and by whom.**

Construct an Alternative Scenario.

1. Construct a promising alternative for some set of actions for a significant actor (e.g. reporting procedures in AECL, FDA procedures, hospital treatment procedures, safety analysis procedures by AECL). For some hints about alternative sets of actions, see the exercises about computer control choices and about reporting procedures.

Judge the Alternative

1. Judge the alternative's effect on each stakeholders' opportunities and vulnerabilities and on each stakeholders' duties and rights.
2. **Imagine each stakeholder in a negotiation** with other stakeholders about whether the alternative should be adopted or not. This certainly helps uncover disagreements about the opportunities and vulnerabilities for each party. One interesting way to stage this negotiation is to have parties that initially represent each stakeholder attempt to don a "veil of ignorance" about which stakeholder they might be when the alternative is adopted. If you might be randomly assigned to any of the stakeholder roles in the case, how would this affect your evaluation of the alternative?
3. **Rank the alternative with other alternatives for that set of actions.** An alternative does not have to be perfect, or even optimal, to be better than the others.

Choosing the Level of Computer Control

In her book *Safeware: System Safety and Computers*, Nancy Leveson lists nine different levels of computer control (taken from Sheridan's analysis):

1. The operator does everything.
2. The computer tells the operator the options available.
3. The computer tells the operator the options available and suggests one.
4. The computer suggests an action and implements it if asked.
5. The computer suggests an action, informs the operator, and implements the action if not stopped in time.
6. The computer selects and implements an action if not stopped in time and then informs the operator.
7. The computer selects and implements an action and tells the operator if asked.
8. The computer selects and implements an action and tells the operator if the designer decides the operator should be notified.
9. The computer selects and implements an action without any human involvement.

In her book *Safeware: System Safety and Computers* (p. 26) Nancy Leveson lists seven myths regarding the safety of software.

1. The cost of computers is lower than that of analog or electromechanical devices.
2. Software is easy to change.
3. Computers provide greater reliability than the devices they replace.
4. Increasing software reliability will increase safety.
5. Testing software and formal verification of software can remove all the errors.
6. Reusing software increases safety.
7. Computer reduce risk over mechanical systems.

CHBE 410 Handout - A History of the Introduction and Shut Down of Therac-25

Normally, when a patient is scheduled to have radiation therapy for cancer, he or she is scheduled for several sessions over a few weeks and told to expect some minor skin discomfort from the treatment. The discomfort is described as being on the order of a mild sunburn over the treated area. In the case you are about to read, a very abnormal thing happened to several patients: they received severe radiation burns resulting in disability, and, in 3 cases, death.

The Therac-25 was a device that targeted electron or X-ray beams on cancerous tissue to destroy it. Electron beams were used to treat shallow tissue, while photon beams could penetrate with minimal damage to treat deep tissue. Even though operators were told that there were "so many safety mechanisms" that it was "virtually impossible" to overdose a patient, this is exactly what did occur in six documented cases.

Therac-25 was released on the market in 1983. In 1987, all treatment with the eleven machines in operation was suspended. Those machines were refitted with the safety devices required by the FDA and remained in service. No more accidents were reported from these machines. At about that time, the division of AECL that designed and manufactured Therac-25 became an independent company.

The major innovations of Therac-25 were the double pass accelerator (allowing a more powerful accelerator to be fitted into a small space, at less cost) and the move to more complete computer control. The move to computer control allowed operators to set up the machine more quickly, giving them more time to speak with patients and making it possible to treat more patients in a day. Along with the move to computer control, most of the safety checks for the operation of the machine were moved to software and the hardware safety interlocks removed.

AECL's FDA Testing and Safety Analysis

Before release of Therac-25 on the US market, AECL obtained approval to market it from the FDA. This approval was obtained by declaring what FDA called pre-market equivalence. Since the software was based on software already in use, and the linear accelerator was a minor modification of existing technology, designation of Therac-25 as equivalent to this earlier technology meant that Therac-25 bypassed the rigorous FDA testing procedures. In 1984, 94% of medical devices entered the market in this manner. This declaration of pre-market equivalence seems optimistic, since most of the safety mechanisms were moved into the software, a major change from previous version of the machine.

In 1983, just after AECL made the Therac-25 commercially available, AECL performed a safety analysis of the machine using Fault Tree Analysis. This involves calculating the probabilities of the occurrence of varying hazards (e.g. an overdose) by specifying which causes of the hazard must jointly occur in order to produce the hazard.

In order for this analysis to work as a Safety Analysis, one must first specify the hazards (not always easy), and then be able to specify the all possible causal sequences in the system that could produce them. It is certainly a useful exercise, since it allows easy identification of single-

point-of-failure items and the identification of items whose failure can produce the hazard in multiple ways. Concentrating on items like these is a good way to begin reducing the probabilities of a hazard occurring.

In addition, if one knows the specific probabilities of all the contributing events, one can produce a reasonable estimate of the probability of the hazard occurring. This quantitative use of Fault Tree Analysis is fraught with difficulties and temptations, as AECL's approach shows.

In order to be useful, a Fault Tree Analysis needs to specify all the likely events that could contribute to producing a hazard. Unfortunately, AECL's analysis left out consideration of the software in the system almost entirely. Since much of the software had been taken from the Therac-6 and Therac-20 systems, and since these software systems had been running many years without detectable errors, the analysts assumed there were no design problems in the software. The analysts considered software failures like "computer selects wrong mode" but assigned them probabilities like 4×10^{-9} .

These sorts of probabilities are likely assigned based on the remote possibility of random errors produced by things like electromagnetic noise. They do not at all take into account the possibility of design flaws in the software. This shows a major difficulty with Fault Tree Analysis as it is often practiced. If the only items considered are "failure" items (e.g. wear, fatigue, etc.) a Fault Tree Analysis really only gives one a reliability for the system.

AECL's Response to the Accidents

In July of 1985, AECL was notified that a patient in Hamilton, Georgia had been overdosed. AECL sent a service engineer to the site to investigate. AECL also informed the United States Food and Drug Administration (FDA), and the Canadian Radiation Protection Board (CRPB) of the problem. In addition they notified all users of the problem and issued instructions that operators should visually confirm hardware settings before each treatment. AECL could not reproduce the malfunction, but its engineers suspected that a hardware failure in a microswitch was at fault. They redesigned the hardware and claimed that this redesign improved the safety of the machine by five orders of magnitude. After modifications were made in the installed machines, AECL notified sites that they did not need to manually check the hardware settings anymore.

In November of 1985, AECL heard of another incident in Georgia. The patient in that incident (Linda Knight) filed suit that month based on an overdose that occurred in June. There is no evidence that AECL followed up this case with the Georgia hospital. Though this information was clearly received by AECL, there is no evidence that this information, was communicated internally to engineers or others who responded to later accidents.

In January of 1986, AECL heard from a hospital in Yakima, Washington that a patient had been overdosed. The AECL technical support supervisor spoke with the Yakima hospital staff on the phone, and contacted them by letter indicating that he did not think the damage they reported was caused by the Therac-25 machine. He also notified them that there have "apparently been no other instances of similar damage to this or other patients."

In March of 1986, AECL was notified that the Therac-25 unit in Tyler, Texas had overdosed a patient. They sent both a local Texas engineer and an engineer from their Canada home office to investigate the incident the day after it occurred. They spent a day running tests on the machine but could not reproduce the specific error. The AECL engineer suggested that perhaps an electrical problem had caused the accident. He also said that AECL knew of no accidents involving radiation overexposure with the Therac-25. An independent engineering firm checked out the electric shock theory and found that the machine did not seem capable of delivering an electric shock to a patient.

On April 11th of 1986, AECL was alerted to another overdose that had occurred in Tyler. After communication with the medical physicist at Tyler, AECL engineers were able to reproduce the overdose and the sequences leading up to it.

AECL filed a medical device report with the FDA on April 15, 1986 to notify them of the circumstances that produced the two Tyler accidents.

At this point, the FDA, having been notified of the first Tyler accident by the hospital, declared Therac-25 defective and ordered the firm to contact all sites that used the machine, investigate the problem, and submit a report called a corrective action plan. AECL contacted all sites and recommended a temporary fix involving removing some keys from the keyboard at the computer console.

The FDA was not satisfied with the notification that AECL gave sites, and in May 1986 required AECL to re-notify all sites with more specific information about the defect in the product and the hazards associated with it. AECL was also at this time involved in meetings with a "user's group" of Therac-25 sites to help formulate its corrective action plan. After several exchanges of information among AECL and the FDA (in July, September, October, November, and December of 1986), AECL submitted a revised corrective action plan to FDA.

In January 1987, AECL was notified of another overdose occurring again at the Yakima, Washington hospital. After sending an engineer to investigate this incident, AECL concluded that there was a different software problem that allowed the electron beam to be turned on without the device that spread it to a safe concentration being placed in the beam.

Therac-25 is Shut Down

In February, 1987, the FDA and its Canadian counterpart cooperated to require all units of Therac-25 to be shut down until effective and permanent modifications were made. After another 6 months of negotiation with the FDA, AECL received approval for its final corrective action plan. This plan included numerous software fixes, the installation of independent, mechanical safety interlocks, and a variety of other safety related changes.

Several of the surviving victims or the deceased victim's families filed suit in US courts against AECL and the medical facilities using Therac-25. All of these suits were settled out of court.

AECL Medical Goes Independent

The division of AECL that designed and manufactured Therac-25 has become an independent private Canadian company. They still make radiation therapy machines.

Government and FDA response to the Accidents

The Therac-25 case pointed to significant weak links in communication between FDA, medical device manufacturers, and their customers or users. Users were not required to report injuries to any government office, or to the manufacturers of the devices that had caused injury.

A 1986 GAO study found 99% of injuries caused by medical devices were not reported to the FDA. At that time, hospitals reported only about 51% of problems to the manufacturer. The hospitals mostly reported dealing with problems themselves. Problems were mainly the result of wear and tear on machines and design flaws.

The breakdown in communication with hospitals and clinics using medical devices prevented FDA from knowing about the isolated and recurring problems with the Therac-25 until after two deaths occurred in Tyler, TX.

Even when the FDA became aware of the problem, they did not have the power to recall Therac-25, only to recommend a recall. After the Therac-25 deaths occurred, the FDA issued an article in the Radiological Health Bulletin (Dec. 1986) explaining the mechanical failures of Therac-25 and explaining that "FDA had now declared the Therac-25 defective, and must approve the company's corrective action program."

After another Therac-25 overdose occurred in Washington state, the FDA took stronger action by "recommending that routine use of the system on patients be discontinued until a corrective plan had been approved and implemented" (Radiological Health Bulletin, March 1987). AECL was expected to notify Therac-25 users of the problem, and of FDA's recommendations.

After the Therac-25 deaths, the FDA made a number of adjustments to its policies in an attempt to address the breakdowns in communication and product approval. In 1990, health-care facilities were required by law to report incidents to both the manufacturer and FDA.

How to Produce a Malfunction 54 on a [AECL] Therac-25 Linear Accelerator

This statement was written by the East Texas Cancer Center physicist after he discovered how to reproduce the "Malfunction 54 error"

Enter the room and set up the machine for an electron beam treatment by selecting a field size and installing the trimmers. Press the set button. Leave the room and close the door. At the control console proceed to the patient set-up display. For Mode enter "X". The machine will default to 25 MeV and go to dose rate of 250 rads/min. Use return key to go to dose. Enter 200. Use return key to go to time. Enter 0.8 min. Use the return key to rapidly advance to the bottom of the display. Immediately use the up arrow to move from the bottom of the display. You are now in the edit mode. Use the up arrow to go to the top of the display and change the mode "X" to "E" for electrons. Change the energy from 25 to 10. Use the return key to go back down to the bottom of the display. Wait for the "beam ready" message then type "B" return. The unit will have no indications on dose rate or dose 1 or dose 2 for about 3 to 4 seconds. Then the dose rate will flash 550 to 575 for one cycle and return to zero. Dose 1 and Dose 2 will count to -6. A malfunction 54 message will appear at the bottom of the display. You have just delivered a dose of approximately 25,000 rads of 25 MeV electrons in less than two seconds.

Impacts of Malfunction 54 - Malfunction 54, produced in this way would deliver a dose 25,000 rads of 25 MeV electrons in less than two seconds. The standard therapeutic dose is about 200 rads at any one time. A dose of 500 rads over the entire body is considered lethal to 50% of individuals who receive it. Two persons were killed from the malfunction 54 overdose. One died in 5 months, the other within one month.

Operator Interview - Susan operated a Therac-4 linear accelerator machine in the mid 1980's. She enjoyed operating AECL's new Therac machine because it was the first computerized linear accelerator. She remembered that while operating the machines, she did not think about whether there could be computer software "bugs" in the system. The technology was new, and she trusted the machine's components and designers. Susan reported being able to move more patients through during the day. She also remembered feeling good about the extra time she had to talk with patients when she was working with a computerized machine.

Susan learned about the Therac-25 incidents at a national radiation conference. A radiation therapist spoke about how many times therapists involved in the accidents attempted to resume treatment *in spite of* the computer error messages. How many attempts to resume treatment is too many? What is the possibility of establishing institutional policies and limits on the number of times an operator could resume treatment after having received an error message, such as the cryptic "malfunction 54" messages that the operator received during the two fatal accidents in Texas. Even today, there are no industry-wide standards for these situations. Susan felt that she was lucky to have worked where there was a physicist available to help with error messages operators received. She also felt that in other clinics, where this assistance is not available, there was, and still is, a great deal more pressure on therapists to just keep going despite the error messages. An operator might attempt, for example, to deliver the prescribed dose in 12 increments instead of 1 by continually clearing the faults generated by the computer. Susan stated that this type of activity happens all the time in medical radiation therapy, particularly in clinics where there is more pressure from the administration to keep patients moving through quickly.