

# UC Irvine

## ICS Technical Reports

### **Title**

An investigation of the Therac-25 accidents

### **Permalink**

<https://escholarship.org/uc/item/5dr206s3>

### **Authors**

Leveson, Nancy G.  
Turner, Clark S.

### **Publication Date**

1992

Peer reviewed

Notice: This Material  
may be protected  
by Copyright Law  
(Title 17 U.S.C.)

ARCHIVES  
Z  
699  
C3  
no. 92-108  
C12

## An Investigation of the Therac-25 Accidents

Nancy G. Leveson  
Clark S. Turner

UCI Technical Report # 92-108  
(University of Washington TR # 92-11-05)

Information and Computer Science Dept.  
University of California, Irvine  
Irvine, CA 92717

Copyright © by the authors, November 1992. All Rights Reserved. Copying without fee is permitted provided that the copies are not made or distributed for direct commercial advantage and credit to the source is given. Abstracting with credit is permitted.

Isihonni eide eide  
hahonni eide eide  
vendi hahonni eide  
(.9.9.9.9.9.9.9)

# Contents

<b>1</b>	<b>Introduction</b>	<b>2</b>
<b>2</b>	<b>The Genesis of the Therac-25</b>	<b>4</b>
<b>3</b>	<b>Accident History</b>	<b>6</b>
3.1	Kennestone Regional Oncology Center (Marietta, Georgia), 1985. . . . .	6
3.2	Ontario Cancer Foundation (Hamilton, Ontario), 1985. . . . .	8
3.3	Yakima Valley Memorial Hospital (Yakima, Washington), 1985 . . . . .	11
3.4	East Texas Cancer Center (Tyler, Texas), March 1986. . . . .	13
3.5	East Texas Cancer Center (Tyler, Texas), April, 1986. . . . .	15
3.6	Yakima Valley Memorial Hospital (Yakima, Washington), January, 1987. .	27
<b>4</b>	<b>Lessons Learned</b>	<b>37</b>
4.1	System Engineering . . . . .	38
4.2	Software Engineering . . . . .	40
4.3	User and Government Oversight and Standards . . . . .	41
<b>5</b>	<b>Final Comments</b>	<b>43</b>
<b>A</b>	<b>Turntable Positioning</b>	<b>45</b>
<b>B</b>	<b>The Operator Interface</b>	<b>46</b>
<b>C</b>	<b>Therac-25 Software Development and Design</b>	<b>47</b>
<b>D</b>	<b>Safety Analysis of the Therac-25</b>	<b>50</b>
<b>E</b>	<b>Major Event Timeline</b>	<b>55</b>
	<b>References</b>	<b>57</b>



# An Investigation of the Therac-25 Accidents

Nancy G. Leveson\*

Clark S. Turner†

## Abstract

Risk in any complex technology is unavoidable. However, important lessons can be learned from accidents which can be used to design procedures for reducing risk in the future. Although descriptions of the Therac-25 medical electron accelerator accidents have been published previously, they are incomplete and often misleading. This paper contains a detailed account of these accidents, along with some lessons that can be learned from them in terms of system engineering, software engineering, and government regulation of safety-critical systems involving software.

## 1 Introduction

Computers are increasingly being introduced into safety-critical systems, and, as a consequence, have been involved in accidents. Some of the most widely cited software-related accidents involved a computerized radiation therapy machine called the Therac-25. Between June 1985 and January 1987, there were six known accidents involving massive overdoses by the Therac-25. Several people died as a result and others were seriously injured. These accidents have been described as the worst series of radiation accidents in the 35-year history of medical accelerators [1].

This paper presents a detailed accident investigation, drawn from publicly available documents, of the factors involved in both the overdoses themselves and the attempts by the users, manufacturers, and governments to deal with them. Our goal is to help others learn from this experience, not to criticize the manufacturer of the equipment or anyone

---

\*Nancy Leveson is a Professor in the Information and Computer Science Department at the University of California, Irvine, and a Professor (on leave for the 1992 academic year) in the Computer Science and Engineering Department, University of Washington.

†Clark Turner is a graduate student at the University of California, Irvine and a practicing attorney.

© Copyright by the authors, November 1992. All Rights Reserved. Copying without fee is permitted provided that the copies are not made or distributed for direct commercial advantage and credit to the source is given. Abstracting with credit is permitted.

else. The mistakes made here are not unique to this manufacturer but are, unfortunately, fairly common in other safety-critical systems. As Frank Houston of the Federal Drug Administration (FDA) has said, "A significant amount of software for life-critical systems comes from small firms, especially in the medical device industry; firms that fit the profile of those resistant to or uninformed of the principles of either system safety or software engineering" [2].

Furthermore, these problems are not limited to the medical industry. It is still a common belief that any good engineer can build software, regardless of whether he or she is trained in state-of-the-art software engineering procedures. Many companies building safety-critical software are not using what are considered proper procedures from a software engineering and safety engineering perspective.

Most accidents are system accidents, that is, they stem from complex interactions between various components and activities. To attribute a single cause to an accident is usually a serious mistake. In this paper, we hope to demonstrate the complex nature of accidents and the need to investigate all aspects of system development and operation in order to understand what has happened and to prevent future accidents.

Despite what can be learned from such investigations, fears of potential liability or loss of business make it difficult to find out the details behind the serious engineering mistakes that are made. When the equipment is regulated by government agencies, some information may be available. Occasionally, major accidents draw the attention of Congress or the President and result in formal accident investigations such as the Rogers Commission investigation of the Challenger accident or the Kemeny Commission investigation of Three Mile Island.

The Therac-25 accidents are the most serious computer-related accidents to date (at least non-military and admitted), and they have even drawn the attention of the popular press<sup>1</sup>. Unfortunately, the previous accounts of the Therac-25 problems have been oversimplified and are misleading by their omissions.

To remedy this, we have obtained information from a wide variety of sources including lawsuits and those government agencies in the U.S. and Canada responsible for regulating such equipment. We have tried to be very careful to present only what we could document from original sources, but there is no guarantee that the documentation itself is correct. When possible, we looked for multiple confirming sources for the more important facts.

We have tried not to bias our description of the accidents, but it is difficult not to filter unintentionally what is described. Also, in this case, we were unable to investigate first hand or get information about some aspects of the accidents that may be very relevant. For example, detailed information about the software development, management, and quality control process of the manufacturer were unavailable to us. Most of this has to be inferred

---

<sup>1</sup>Stories about the Therac-25 have appeared in numerous trade journals and newspapers, People Magazine, 20/20, and the McNeil/Lehrer News Hour.

from statements in correspondence or other means.

As a result, there may be factors left out of our analysis of the accident. But it is still possible to learn a great deal from the facts that we do know. Our account can provide confirmation of some previous hypotheses about the proper development and use of software to control dangerous processes and can suggest hypotheses that need to be further evaluated. Following the account of the accidents and the responses of the manufacturer, government agencies, and users, we present what we believe are the most compelling lessons to be learned from them in the context of software engineering, safety engineering, and government and user standards and oversight.

## 2 The Genesis of the Therac-25

Medical linear accelerators (linacs) accelerate electrons to create high energy beams that can be used to destroy tumors with minimal impact on the surrounding healthy tissue. Relatively shallow tissue is treated with the accelerated electrons; to reach deeper tissue, the electron beam is converted into x-ray photons.

In the early 1970s, Atomic Energy of Canada Limited (AECL)<sup>2</sup> and a French company called CGR went into business together building linear accelerators. The products of this cooperation were the Therac-6, a six million electron volt (MeV) accelerator capable of producing x-rays only and, later, the Therac-20, a 20 MeV dual mode (x-rays or electrons) accelerator. Both were versions of older CGR machines, the Nepture and Sagittaire, respectively, which were augmented with computer control using a DEC PDP-11 mini-computer.

Software functionality was limited in both of these machines: The computer merely added convenience to the existing hardware, which was capable of standing alone. Industry-standard hardware safety features and interlocks in the underlying machines were retained. We know that some of the old Therac-6 software routines were used in the Therac-20 [3] and that CGR developed the initial software.

The business relationship between AECL and CGR faltered after the Therac-20 effort. Citing competitive pressures, the two companies did not renew their cooperative agreement when scheduled in 1981. In the mid-1970's, AECL developed a radical new "double pass" concept for electron acceleration. A double-pass accelerator needs much less space to develop comparable energy levels because it folds the long physical mechanism that is required to accelerate the electrons and is more economic to produce (as it uses a magnetron rather than a klystron as the energy source).

---

<sup>2</sup>AECL is an arms-length entity, called a crown corporation, of the Canadian Government. Since the time of the incidents related in this paper, AECL Medical, a division of AECL, is in the process of being privatized and is now called Theratronics International, Ltd. Currently, the primary business of AECL is the design and installation of nuclear reactors.



Using this double-pass concept, AECL designed the Therac-25, a dual-mode linear accelerator that can deliver either photons at 25 MeV or electrons at various energy levels (see Figure 1). Compared to the Therac-20, the Therac-25 is notably more compact, more versatile and arguably easier to use. The higher energy takes advantage of the phenomenon of "depth dose": as the energy increases, the depth in the body at which maximum dose build-up occurs also increases, sparing the tissue above the target area. Economic advantages also come into play for the customer since only one machine is required for both treatment modalities (electrons and photons).

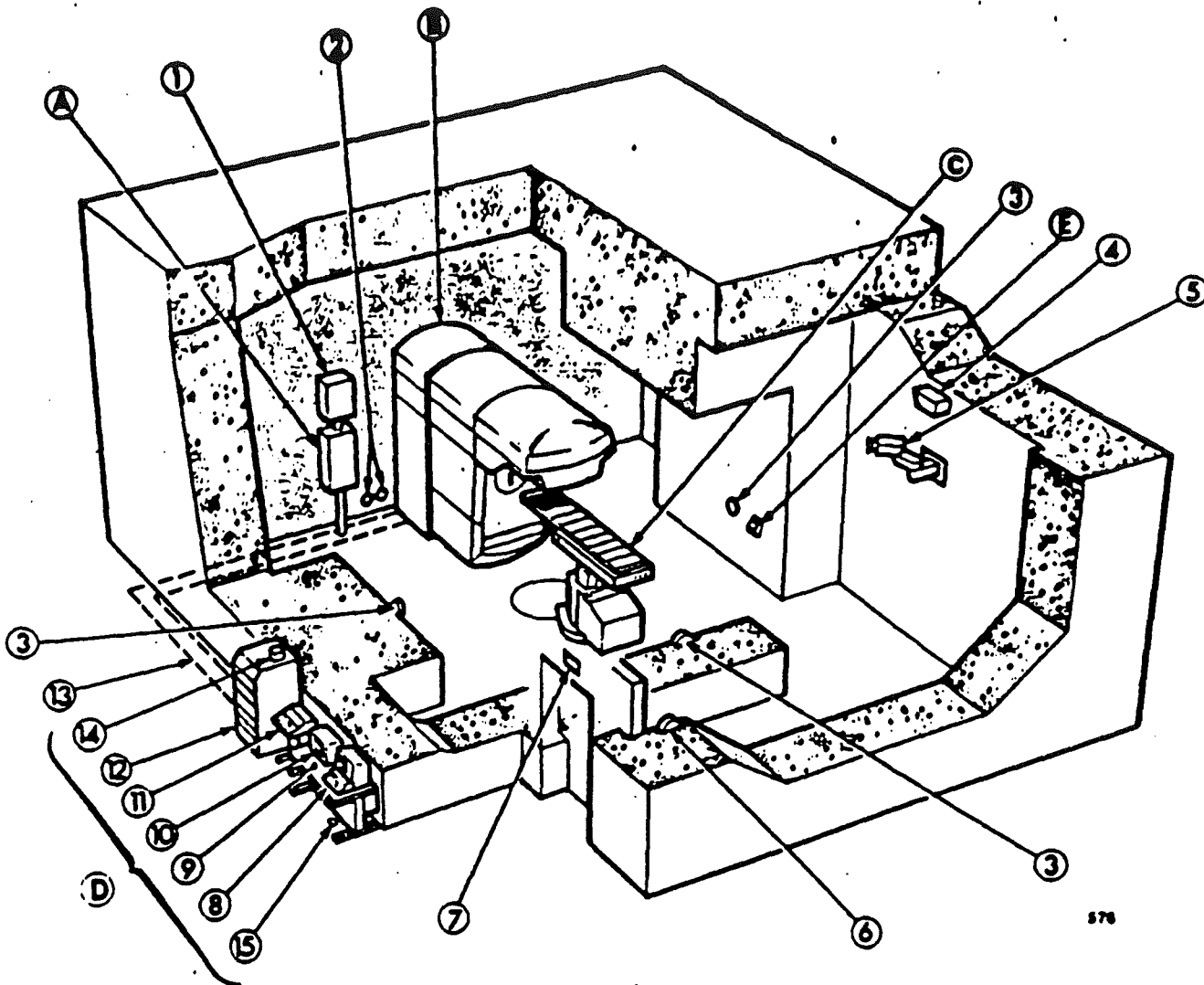
Several features of the Therac-25 are important in understanding the accidents. First, as with the Therac-6 and the Therac-20, the Therac-25 is controlled by a PDP-11 computer. However, AECL designed the Therac-25 to take advantage of computer control from the outset; they did not build on a stand-alone machine. The Therac-6 and Therac-20 had been designed around machines that already had histories of clinical use without computer control.

Also, the Therac-25 software has more responsibility for maintaining safety than the software in the previous machines. The Therac-20 has independent protective circuits for monitoring the electron beam scanning plus mechanical interlocks for policing the machine and ensuring safe operation. The Therac-25 puts more reliance on software for these functions. AECL decided that they would take advantage of the computer's abilities to control and monitor the hardware without having to duplicate all the existing hardware safety mechanisms and interlocks. This approach is becoming more common as companies decide that the hardware interlocks and backups are not worth the expense or they put more faith (perhaps misplaced) on software than on hardware reliability.

Finally, some of the software for the machines was interrelated or reused. The AECL Quality Assurance (QA) Manager, in a letter to a Therac-25 user, stated that "The same Therac-6 package was used by the AECL software people when they started the Therac-25 software. The Therac-20 and Therac-25 software programs were done independently starting from a common base [4]." The reuse of Therac-6 design features or modules may explain some of the problematic aspects of the Therac-25 software (see Appendix C). The QA manager was apparently unaware that some Therac-20 routines were also used in the Therac-25; this was discovered after a bug related to one of the Therac-25 accidents was found in the Therac-20 software.

The first hardwired prototype of the Therac-25 was produced in 1976, and the completely computerized commercial version was available in late 1982. Details about the design of the machine and its controlling software that are important in understanding the accidents are provided in appendices to this paper.

In March 1983, a safety analysis was performed by AECL on the Therac-25. This analysis was in the form of a fault tree and apparently excluded the software. According to the final report, several assumptions formed the basis of the analysis with respect to the computer:



**AECL SUPPLIED EQUIPMENT**

- A. POWER INPUT BOX
- B. THERAC 25 UNIT
- C. TREATMENT TABLE
- D. CONTROL ROOM EQUIPMENT
  - 8. DISPLAY TERMINAL
  - 9. DESK
  - 11. PRINTER
  - 12. CONTROL CONSOLE
  - 14. TURNTABLE POSITION MONITOR
  - 15. MOTION ENABLE SWITCH (FOOTSWITCH)
- E. MOTION POWER SWITCH

**CUSTOMER SUPPLIED EQUIPMENT**

- 1. FUSED DISCONNECT
- 2. WATER SUPPLY AND RETURN
- 3. ROOM EMERGENCY SWITCHES
- 4. THERAPY ROOM INTERCOM
- 5. TV CAMERA
- 6. DOOR INTERLOCK SWITCH
- 7. BEAM ON/OFF LIGHT
- 10. TV MONITOR
- 13. CABLE TROUGH

**Typical Therac 25 Facility**

1. Programming errors have been reduced by extensive testing on a hardware simulator and under field conditions on teletherapy units. Any residual software errors are not included in the analysis.
2. Program software does not degrade due to wear, fatigue, or reproduction process.
3. Computer execution errors are caused by faulty hardware components and by "soft" (random) errors induced by alpha particles and electromagnetic noise.

The fault tree resulting from this analysis does appear to include computer failure, although apparently, judging from the basic assumptions above, only hardware failures are considered. For example, in one OR gate leading to the event of getting the wrong energy, there is a box that contains "Computer selects wrong energy" and a probability of  $10^{-11}$  is assigned to this event. For "Computer selects wrong mode," a probability of  $4 \times 10^{-9}$  is given. No justification of either number is provided.

### 3 Accident History

There were eleven Therac-25s installed: five in the U.S. and six in Canada. Six accidents involving massive overdoses to patients occurred between 1985 and 1987, when the machine was finally recalled to make extensive design changes including hardware safeguards against software errors. Related problems were found in the Therac-20 software, although these were not recognized until after the Therac-25 accidents because the Therac-20 included hardware safety interlocks and thus no injuries resulted. In this section, we present a chronological recounting of the accidents as well as the response by the manufacturer, government regulatory agencies, and users.

#### 3.1 Kennestone Regional Oncology Center (Marietta, Georgia), 1985.

Details of this accident are sketchy since it was never carefully investigated, and there was no admission that the injury was caused by the Therac-25 until long after the occurrence, despite (1) claims by the patient that she had been injured during treatment, (2) the suspicions of the radiation physicist involved, and (3) the obvious and severe radiation burns suffered by the patient.

After undergoing a lumpectomy to remove a malignant breast tumor, a 61-year-old woman was receiving followup radiation treatment to nearby lymph nodes on a Therac-25 at the Kennestone Hospital in Marietta, Georgia. The Therac-25 had been operating at Kennestone for about six months; other Therac-25's had been operating, apparently without incident, since 1983.

On June 3, 1985, the patient was set up for a 10 MeV electron treatment to the clavicle area. When the machine turned on, she felt a "tremendous force of heat ... this red-hot sensation [5]." When the technician came in, she said, "you burned me." The technician replied that this was not possible. Though there were no marks on her at the time, the treatment area felt "warm to the touch".

It is unclear exactly when AECL learned about this incident. Dr. Tim Still, the Kennestone physicist, says that he contacted AECL to ask if the Therac-25 could operate in electron mode without scanning to spread the beam. Three days later the engineers at AECL called the physicist back to explain that improper scanning was not possible.

In an August 19, 1986 letter from AECL to the FDA, the AECL QA manager states, "In March of 1986 AECL received a lawsuit from the patient involved ... This incident was never reported to AECL prior to this date, although some rather odd questions had been posed by Tim Still, the hospital physicist." The physicist at a hospital in Tyler, Texas where a later accident occurred reported that: "According to Tim Still, the patient filed suit in October 1985 listing the hospital, manufacturer and service organization responsible for the machine. AECL was notified informally about the suit by the hospital, and AECL received official notification of a law suit in November 1985" [6].

Because of the lawsuit (filed November 13, 1985), some AECL administrators must have known about the Marietta accident — although no investigation occurred at this time. There are further comments by FDA investigators that point to the lack of a mechanism in AECL to follow up with reports of suspected accidents [8]. The lack of followup in this case appears to be evidence of such a problem in the organization.

The patient went home, but shortly afterward she developed a reddening and swelling in the center of the treatment area. Her pain had increased to the point that her shoulder "froze," and she experienced spasms. She was then admitted to West Paces Ferry Hospital in Atlanta, but her oncologists continued to send her to Kennestone for Therac-25 treatments. Clinical explanation was sought for the reddening of the skin, which at first her oncologist attributed to her disease or to normal treatment reaction.

About two weeks later, the physicist at Kennestone noticed that the patient had a matching reddening on her back as though a burn had gone right through her body, and the swollen area had begun to slough off layers of skin. Her shoulder was immobile, and she was apparently in great pain. It was now obvious that she had a radiation burn, but the hospital and her doctors could provide no satisfactory explanation. Shortly afterward, she initiated a lawsuit against the hospital and AECL regarding her injury.

The Kennestone physicist later estimated that she received one or two doses of radiation in the 15,000 to 20,000 rad (radiation absorbed dose) range [7]. He does not believe her injury could have been caused by less than 8,000 rads. Typical single therapeutic doses are in the 200 rad range. Doses of 1,000 rads can be fatal if delivered to the whole body [9] and, in fact, the accepted figure for whole-body radiation, which will cause death in 50% of the cases, is 500 rads [10]. The consequences of an overdose to a smaller part of the body

depends on the radio-sensitivity of the tissue involved. The director of radiation oncology at the Kennestone facility explained their confusion about the accident as due to the fact that they had never seen an overtreatment of that magnitude before [9].

Eventually, the patient's breast had to be removed due to the radiation burns. She completely lost the use of her shoulder and her arm and was in continual pain. She had suffered a serious radiation burn, but the manufacturer and operators of the machine refused to believe that it could have been caused by the Therac-25. The treatment prescription printout feature was disabled at the time of the accident, so there was no hardcopy of the treatment data. The lawsuit was eventually settled out of court.

From what we can determine, the accident was not reported to the FDA until *after* the later Tyler accidents in 1986 (described in Sections 3.4 and 3.5). The medical device incident reporting regulations apply only to equipment manufacturers and importers, not users. The regulations require that manufacturers and importers report deaths, serious injuries, or malfunctions that could result in those consequences. Healthcare professionals and institutions were not required to report incidents to manufacturers at that time<sup>3</sup>. The Comptroller General of the U.S. Government Accounting Office, in testimony before Congress on November 6, 1989, expressed great concern about the viability of the incident reporting regulations in preventing or spotting medical device problems. According to a GAO study, the FDA knows of less than 1% of deaths, serious injuries, or equipment malfunctions that occur in hospitals [11].

At this point, the other Therac-25 users were also unaware that anything untoward had occurred and did not learn about any problems with the machine until after subsequent accidents. Even then, most of their information came through personal communication among themselves.

### 3.2 Ontario Cancer Foundation (Hamilton, Ontario), 1985.

A second accident occurred about seven weeks after the Kennestone patient was overdosed. At that time, the Therac-25 at the Hamilton Clinic had been in use for more than six months. On July 26, 1985, a forty-year-old patient came to the clinic for her twenty-fourth Therac-25 treatment for carcinoma of the cervix. The operator activated the machine, but the Therac shut down after five seconds with an *HTILT* error message. *No dose* was shown by the Therac's dosimetry system display. The display also indicated a *treatment pause*.

Since the machine did not suspend and the control display indicated no dose was delivered to the patient, the operator went ahead with a second attempt at treatment by pressing the **P** key (the Proceed command), expecting the machine to deliver the proper dose this time. This was standard operating procedure, and, as described in Appendix B,

---

<sup>3</sup>The law was amended in 1990 to require healthcare facilities to report incidents to the manufacturer and to the FDA.

Therac-25 operators had become accustomed to frequent malfunctions that had no untoward consequences for the patient [12]. Again, the machine shut down in the same manner. This process was repeated four times after the original attempt — the display showing *NO DOSE* delivered to the patient each time. After the fifth pause, the machine went into treatment suspend, and a hospital service technician was called. The technician found nothing wrong with the machine. This also was not an unusual scenario according to a Therac-25 operator [12].

After the treatment, the patient complained of a burning sensation, described as an “electric tingling shock” to the treatment area in her hip. Six other patients were treated later that day without incident. She came back for further treatment on July 29 and complained of burning, hip pain, and excessive swelling in the region of treatment. The machine was taken out of service, as radiation overexposure was suspected. The patient was hospitalized for the condition on July 30. AECL was informed of the apparent radiation injury and sent a service engineer to investigate. The U.S. FDA, the then Canadian Radiation Protection Bureau (RPB)<sup>4</sup>, and users were informed that there was a problem although the users claim that they were never informed that a patient injury had occurred. Users were told that they should visually confirm the proper alignment of the turntable until further notice (which occurred three months later) [13].

The patient died on November 3, 1985 of an extremely virulent cancer. An autopsy revealed the cause of death as the cancer, but it was noted that had she not died, a total hip replacement would have been necessary as a result of the radiation overexposure [14]. An AECL technician later estimated the patient had received between 13,000 and 17,000 rads [15].

### Manufacturer's Response

AECL could not reproduce the malfunction that had occurred, but suspected a transient failure in the microswitch used to determine the position of the turntable. During the investigation of the accident, AECL hardwired the error conditions they assumed were necessary for the malfunction and, as a result, found some design weaknesses and potential mechanical problems involving the turntable positioning.

The computer senses and controls turntable position by reading a three-bit signal about the status of three microswitches in the turntable switch assembly (see Appendix A for a description of the turntable positioning). Essentially, AECL determined that a one bit error in the microswitch codes (which could be caused by a single open-circuit fault on the switch lines) could produce an ambiguous position message to the computer. The problem was

---

<sup>4</sup>On April 1, 1986, the Radiation Protection Bureau and the Bureau of Medical Devices were merged to form the Bureau of Radiation and Medical Devices (BRMD).

exacerbated by the design of the mechanism that extends a plunger to lock the turntable when it is in one of the three cardinal positions: The plunger could be extended when the turntable was way out of position, thus giving a second false position indication. AECL devised a method to indicate turntable position that tolerated a one bit error, i.e., the code would still unambiguously reveal correct position with any one microswitch failure.

In addition, the software was altered so that the computer checked for "in transit" status of the switches in order to keep further track of the switch operation and position of the turntable and to give additional assurance that the switches were working and the turntable was moving.

As a result of these improvements, AECL claimed in its report and correspondence with hospitals that "analysis of the hazard rate of the new solution indicates an improvement over the old system by at least 5 orders of magnitude" [16]. A claim that safety had been improved by five orders of magnitude seems exaggerated, especially given that in its final incident report to the FDA [17], AECL concluded that they "cannot be firm on the exact cause of the accident but can only suspect ..." which underscored their inability to determine the cause of the accident with any certainty. The AECL QA manager testified that they could not reproduce the switch malfunction and that testing of the microswitch was "inconclusive" [10]. The similarity of the errant behavior and the injury to patients in this accident and a later one in Yakima (attributed to software error) provide good reason to believe that the Hamilton overdose was probably related to software error rather than to a microswitch failure.

### Government and User Response

The Hamilton accident resulted in a voluntary recall by AECL, and the FDA termed it a Class II recall. Class II means "a situation in which the use of, or exposure to, a violative product may cause temporary or medically reversible adverse health consequences or where the probability of serious adverse health consequences is remote." Four users in the U.S. were advised by a letter from AECL on August 1, 1985 to visually check the ionization chamber to make sure it is in its correct position in the collimator opening before any treatment and to discontinue treatment if they get an H-TILT message with an incorrect dose indicated. No mention was made that a patient injury was involved. The FDA audited AECL's subsequent modifications. After the modifications were made, the users were told that they could return to normal operating procedures.

The head of the advanced x-ray systems in the Canadian Radiation Protection Bureau, Gordon Symonds, wrote a report as a result of the Hamilton accident that analyzed the design and performance characteristics of the Therac-25 with respect to radiation safety. Besides citing the flawed microswitch, the report faulted both hardware and software components of the Therac's design [18]. It concluded with a list of four modifications that

were to be made to the Therac-25 as a minimum for compliance with Canada's Radiation Emitting Devices (RED) Act. The RED law, enacted in 1971, gives government officials power to ensure the safety of radiation-emitting devices.

The modifications recommended in the Symonds report included redesigning the microswitch and changing the way the computer handled malfunction conditions. In particular, treatment was to be terminated in the event of a dose-rate malfunction, giving a treatment "suspend". This would have removed the option to proceed simply by pressing the  $\textcircled{P}$  key. The report also made recommendations regarding collimator test procedures and message and command formats. A November 8, 1985 letter signed by Ernest Létourneau, M.D., director of the Canadian Radiation Protection Bureau (PRB), asked that AECL make changes to the Therac-25 based on the Symond's report "to be in compliance with the RED act [19]".

Although, as noted above, AECL did make the microswitch changes, they did not comply with the directive to change the malfunction pause behavior into treatment suspends, instead reducing the maximum number of retries from five to three. According to Mr. Symonds, the deficiencies outlined in the PRB letter of November 8 were still pending when subsequent accidents five months later changed the priorities. If these later accidents had not occurred, AECL would have been required to comply with the requirements in the letter.

Immediately after the Hamilton accident, the Ontario Cancer Foundation hired an independent consultant to investigate. He concluded in a September 1985 report that an independent system (beside the computer) was needed to verify the position of the turntable and suggested the use of a potentiometer. The Canadian Radiation Protection Bureau wrote a letter to AECL in November 1985 requesting that AECL install such an independent upper collimator positioning interlock on the Therac-25. Also, in January 1986, AECL received a letter from the attorney representing the Hamilton Ontario Clinic that put them on notice that there had been continuing problems with the turntable including four occasions at Hamilton and requested the installation of an independent system (potentiometer) to verify the position of the turntable [15]. AECL did not comply: No independent interlock was installed on the Therac-25s at this time.

### **3.3 Yakima Valley Memorial Hospital (Yakima, Washington), 1985**

In this accident, like the Kennestone overdose, machine malfunction was not acknowledged until after later accidents were understood.

The Therac-25 at Yakima had been modified by AECL in September of 1985 in response to the overdose at Hamilton. During December, 1985, a woman came in for treatment with the Therac-25. She developed erythema (excessive reddening of the skin) in a parallel striped pattern at one port site (her right hip) after one of the treatments.. Despite this, she



continued to be treated by the Therac-25 as the cause of her reaction was not determined to be abnormal until January or February of 1986. On January 6, 1986, her treatments were completed.

The skin reaction was monitored closely, and the staff attempted to find possible causes for it. The open slots in the blocking trays in the Therac-25 could have produced such a striped pattern, but by the time that the skin reaction had been determined to be abnormal, the blocking trays had been discarded so the blocking arrangement and tray striping orientation could not be reproduced. A reaction to chemotherapy was ruled out because that should have produced reactions at the other ports and would not have produced stripes. When it was discovered that the woman slept with a heating pad, a possible explanation was offered on the basis of the parallel wires that deliver the heat in such pads. The staff x-rayed the heating pad and discovered that the wire pattern did not correspond to the erythema pattern on the patient's hip.

The hospital staff sent a letter to AECL on January 31, and they also spoke on the phone with the AECL Technical Support Supervisor. On February 24, 1986, the AECL Technical Support Supervisor sent a written response to the Director of Radiation Therapy at Yakima stating that "After careful consideration we are of the opinion that this damage could not have been produced by any malfunction of the Therac-25 or by any operator error." The letter goes on to support this opinion by listing two pages of technical reasons why an overdose by the Therac-25 was impossible along with the additional argument that "There has apparently been no other instances of similar damage to this or other patients." The letter ends with "In closing, I wish to advise that this matter has been brought to the attention of our Hazards Committee as is normal practice."

The hospital staff eventually ascribed the skin/tissue problem to "cause unknown." In a report written on this first Yakima incident after another Yakima overdose a year later (described in Section 3.6), the medical physicist involved wrote:

At that time, we did not believe that [the patient] was overdosed because the manufacturer had installed additional hardware and software safety devices to the accelerator.

In a letter from the manufacturer dated 16-Sep-85, it is stated that "Analysis of the hazard rate resulting from these modifications indicates an improvement of at least five orders of magnitude"! With such an improvement in safety (10,000,000%) we did not believe that there could have been any accelerator malfunction. These modifications to the accelerator were completed on 5,6-Sep-85.

Even with fairly sophisticated physics support, the hospital as users did not have the ability to investigate the possibility of machine malfunction further. They were not aware of any other incidents, and, in fact, were told that there had been none, so there was no reason for them to pursue the matter. However, it seems that the fact that three

similar incidents had occurred with this equipment should have triggered some suspicion and investigation by the manufacturer and the appropriate government agencies. This assumes, of course, that these incidents were all reported and known by AECL and by the government regulators. If they were not, then it is appropriate to ask why they were not and how this could be remedied in the future.

Approximately a year later (February 1987), after the second Yakima overdose led the hospital staff to suspect that this first injury had been due to a Therac-25 fault, the staff investigated and found that this patient had a chronic skin ulcer, tissue necrosis (death) under the skin, and was in continuous pain. This was surgically repaired, skin grafts were made, and the symptoms relieved. The patient is alive today with minor disability and some scarring related to the overdose. The hospital staff concluded that the dose accidentally delivered to this patient must have been much lower than in the second accident as the reaction was significantly less intense and necrosis did not develop until six or eight months after exposure. Some other factors related to the place on the body where the overdose occurred also kept her from having more significant problems as a result of the exposure.

### 3.4 East Texas Cancer Center (Tyler, Texas), March 1986.

More is known about the Tyler accidents than the others because of the diligent efforts of the Tyler hospital physicist, Fritz Hager, without whom the understanding of the software problems may have been delayed even further.

The Therac-25 was at the East Texas Cancer Center (ETCC) for two years before the first serious accident, and more than 500 patients had been treated. On March 21, 1986, a male patient came into ETCC for his ninth treatment on the Therac-25, one of a series prescribed as followup to the removal of a tumor from his back.

This treatment was to be a 22 MeV electron beam treatment of 180 rads over a 10cm by 17cm field on the upper back and a little to the left of his spine, or a total of 6,000 rads over a 6 1/2 week period. He was taken into the treatment room and was placed face down on the treatment table. The operator then left the treatment room, closed the door, and sat at the control terminal.

The operator had held this job for some time, and her typing efficiency had increased with experience. She could quickly enter prescription data and change it conveniently with the Therac's editing features. She entered the patient's prescription data quickly, then noticed that she had typed "x" (for x-ray) when she had intended "e" (for electron) mode. This was a common mistake as most of the treatments involved x-rays, and she had gotten used to typing this. The mistake was easy to fix; she merely used the  $\uparrow$  key to edit the mode entry.

Since the other parameters she had entered were correct, she hit the return key several times and left their values unchanged. She reached the bottom of the screen where it was indicated that the parameters had been *verified* and the terminal displayed *beam ready*, as

expected. She hit the one key command,  $\textcircled{B}$  for "beam on," to begin the treatment. After a moment, the machine shut down and the console displayed a message saying *Malfunction 54*. The machine also displayed a *treatment pause*, indicating a problem of low priority (see Appendix B). The sheet on the side of the machine explained that this malfunction was a "dose input 2" error. The ETCC did not have any other information available in its instruction manual or other Therac-25 documentation to explain the meaning of Malfunction 54 [20]. An AECL technician later testified that "dose input 2" meant that a dose had been delivered that was either too high or too low.

The machine showed a substantial underdose on its dose monitor display, i.e., 6 monitor units delivered whereas 202 monitor units had been requested by the operator. The operator was accustomed to the quirks of the machine, which would frequently stop or delay treatment; in the past, the only consequences had been inconvenience. She immediately took the normal action when the machine merely paused which was to hit the  $\textcircled{P}$  key to proceed with the treatment. The machine promptly shut down with the same *Malfunction 54* error and the same underdose shown by the dosimetry.

The operator was isolated from the patient, since the machine apparatus was inside a shielded room of its own. The only way that the operator could be alerted to patient difficulty was through audio and video monitors. On this day, the video display was unplugged and the audio monitor was broken [12].

After the first attempt to treat him, the patient said that he felt like he had received an electric shock or that someone had poured hot coffee on his back: He felt a thump and heat and heard a buzzing sound from the equipment. Since this was his ninth treatment, he knew that this was not normal. He began to get up from the treatment table to go for help. It was at this moment that the operator hit the  $\textcircled{P}$  key to proceed with the treatment. The patient stated that he felt like his arm was being shocked by electricity and that his hand was leaving his body. He went and pounded on the door of the treatment room. The operator was shocked and immediately opened the door for him. He appeared visibly shaken and upset.

The patient was immediately examined by a physician who observed intense erythema over the treatment area, but suspected nothing more serious than electric shock [14]. He was discharged and sent home with instructions to return if he suffered any further reactions. The hospital physicist was called in, and he found the machine calibration within specifications. The meaning of the malfunction message was not understood. The machine was then used to treat patients for the rest of the day.

In actuality, but unknown to anyone at that time, the patient had received a massive overdose, concentrated in the center of the treatment area. After the fact simulations of the accident revealed possible doses of 16,500 to 25,000 rads in less than 1 second over an area of about 1 cm [13].

Over the weeks following the accident, the patient continued to have pain in his neck and shoulder. He lost the function of his left arm and had periodic bouts of nausea and

vomiting. He was eventually hospitalized for radiation-induced myelitis of the cervical cord causing paralysis of his left arm and both legs, left vocal cord paralysis (which left him unable to speak), neurogenic bowel and bladder, and paralysis of the left diaphragm. He also had a lesion on his left lung and recurrent herpes simplex skin infections. He died from complications of the overdose five months after the accident.

### User and Manufacturer Response

The Therac-25 was shut down for testing the day after this accident. One local AECL engineer and one from the home office in Canada came to ETCC to investigate. They spent a day running the machine through tests but could not reproduce a Malfunction 54. The AECL engineer from the home office reportedly explained that it was not possible for the Therac-25 to overdose a patient. The ETCC physicist claims that he asked AECL at this time if there were any other reports of radiation overexposure and that the AECL personnel (including the QA manager) told him that no accidents involving radiation overexposure by the Therac-25 were known to AECL [21]. This seems odd since AECL was surely at least aware of the Hamilton accident that had occurred seven months before, the Yakima accident, and, even by their account, learned of the Georgia law suit around this time (which had been filed four months earlier). The AECL engineers then suggested that this accident might have been caused by an electrical problem.

The electric shock theory was checked out thoroughly by an independent engineering firm. The final report indicated that there was no electrical grounding problem in the machine, and it did not appear capable of giving a patient an electrical shock. The physicist at the ETCC checked out the calibration of the Therac-25 and found it satisfactory. They put the machine back into service on April 7, 1986, convinced that it was performing properly.

### **3.5 East Texas Cancer Center (Tyler, Texas), April, 1986.**

Three weeks later, on Friday, April 11, 1986, another male patient was scheduled to receive an electron treatment at ETCC for a skin cancer on the side of his face. The prescription was for 10 MeV to an area of approximately 7 cm by 10 cm. The same technician who had treated the first Tyler accident victim prepared this patient for treatment. Much of what follows is from the deposition of the Tyler Therac-25 operator [12].

As with her former patient, she entered the prescription data and then noticed an error in the mode. Again she used the edit  $\uparrow$  key to change the mode from x-ray to electron. After she finished editing, she pressed the *return* key several times to place the cursor on the bottom of the screen. She saw the *beam ready* message displayed and turned the beam on.

Within a few seconds the machine shut down, making a loud noise audible via the (now working) intercom. The display showed *Malfunction 54* again. The operator rushed into the treatment room, hearing her patient moaning for help. He began to remove the tape that had held his head in position and said something was wrong. She asked him what he felt, and he replied, "fire" on the side of his face. She immediately went to the hospital physicist and told him that another patient appeared to have been burned. The physicist asked the patient about his experience. He explained that something had hit him on the side of the face, he saw a flash of light, and he heard a sizzling sound reminiscent of frying eggs. He was very agitated and asked, "what happened to me, what happened to me?"

This patient died from the overdose on May 1, 1986, three weeks after the accident. He had disorientation which progressed to coma, fever to  $104^{\circ}F$ , and neurological damage. Autopsy showed an acute high-dose radiation injury to the right temporal lobe of the brain and the brain stem [14].

### User and Manufacturer Response

After this latest Tyler accident, the machine was immediately taken out of service by the physicist at the ETCC, and he called AECL to alert them to this second apparent overexposure. The Tyler physicist then began a careful investigation of his own. He worked with the operator, who remembered exactly what she had done on this occasion. After a great deal of effort, they were eventually able to elicit the *Malfunction 54* message. They determined that data entry speed during editing was the key factor in producing the error condition: If the prescription data was edited at a fast pace (as is natural for someone who has repeated the procedure a large number of times), the overdose occurred.

It took some practice before the physicist could repeat the procedure rapidly enough to elicit the *Malfunction 54* message at will. Once he could do this, he set about measuring the actual dose delivered under the error condition. He took a measurement of about 804 rads but realized that the ion chamber had become saturated. After making adjustments to extend his measurement ability, he determined that the dose was somewhere over 4,000 rads.

The next day, an engineer from AECL called and said that he could not reproduce the error. After the Tyler physicist explained that the procedure had to be performed quite rapidly, AECL could finally produce a similar malfunction on its own machine. AECL then set up its own set of measurements to test the dosage delivered. Two days after the accident, AECL said it had measured the dosage (at the center of the field) to be 25,000 rads. They explained that the "frying" sound heard by the patients was the ion chambers being saturated.

In fact, it is not possible to determine the exact dose received by each of the accident victims; the total dose delivered during the malfunction conditions was found to vary

enormously when the faults were simulated at different clinics. The number of pulses delivered in the 0.3 seconds that elapsed before interlock shut-off varied due to the fact that the start-up pulse repetition frequency was adjusted by the software to very different values on different machines. Therefore there is still some uncertainty as to the doses actually received in the accidents [1].

In one of the law suits that resulted from the Tyler accidents, the AECL quality control manager testified that a "cursor up" problem had been found in the service mode at the Bayview and Kennestone clinics in February or March of 1985 and also in the summer of 1985 [10]. Both times it was thought that the software problems had been fixed. There is no way to determine whether there is any relationship between these problems and the Tyler accidents.

### Related Therac-20 Problems

After the Tyler accidents, the users of the Therac-20 (who had heard informally about the Tyler accidents from Therac-25 users) conducted informal investigations to determine if the same problem could occur on their machines. As noted earlier, the software for the Therac-25 and Therac-20 both "evolved" from the Therac-6 software [23]. Additional functions had to be added because the Therac-20 (and Therac-25) operate in both x-ray and electron mode while the Therac-6 has only x-ray mode. The CGR employees modified the software for the Therac-20 to handle the dual modes.

When the Therac-25 development began, AECL engineers adapted the software from the Therac-6, but they also borrowed software routines from the Therac-20 to handle electron mode. The agreements between AECL and CGR gave both companies the right to tap technology employed in joint products for their other products.

After the second Tyler accident, a physicist at the University of Chicago Joint Center for Radiation Therapy heard about the Therac-25 software problem and decided to find out whether the same thing could happen with the Therac-20. At first the physicist was unable to reproduce the error on his machine, but two months later he found the link.

The Therac-20 at the U. of Chicago is used to teach students in a radiation therapy school conducted by the center. The physicist at the center, Frank Borger, noticed that whenever a new class of students started using the Therac-20, fuses and breakers on the machine tripped, shutting down the unit. These failures, which had been occurring ever since they had acquired the machine, might occur three times a week while new students operated the machine and then disappear for months. Borger determined that new students make lots of different types of mistakes and use "creative methods of editing" parameters on the console. Through experimentation he found that certain editing sequences correlated with blown fuses and determined that the same computer bug (as in the Therac-25

software) was responsible [24]. The physicist notified the FDA which notified Therac-20 users [25].

The software error is just a nuisance on the Therac-20 because this machine has independent hardware protective circuits for monitoring the electron beam scanning. The protective circuits do not allow the beam to turn on so there is no danger of radiation exposure to a patient. While the Therac-20 relies on mechanical interlocks for monitoring the machine, the Therac-25 relies largely on software.

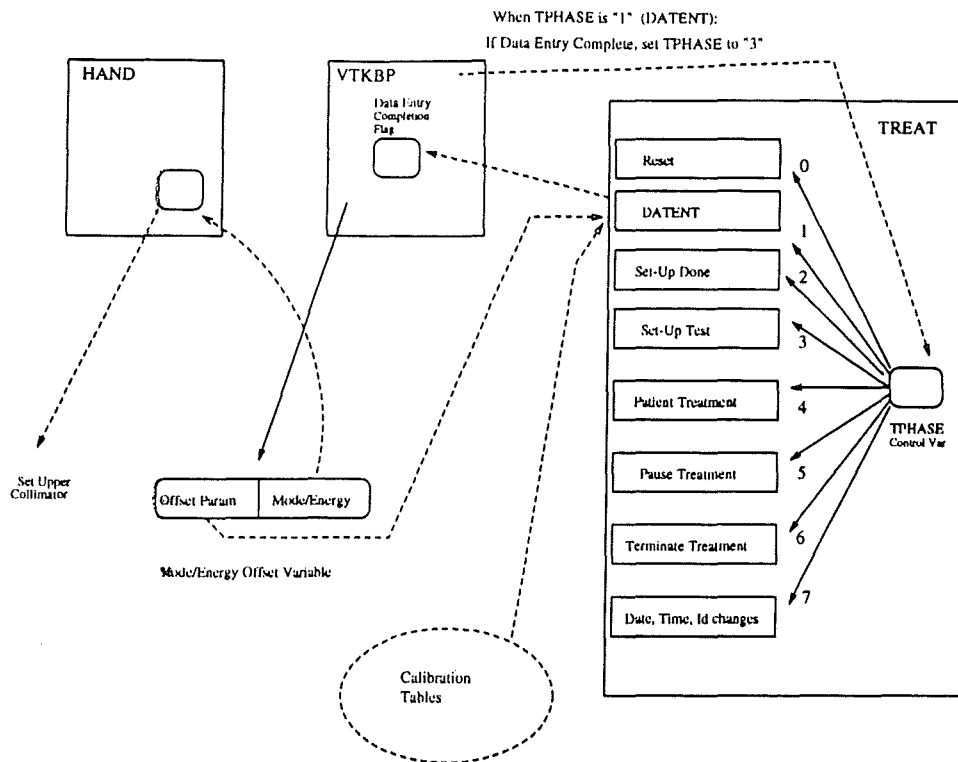
### The Software "Bug"

One of the lessons to be learned from the Therac-25 story is that focusing on particular software "bugs" is not the way to make a safe system. Virtually all complex software can be made to behave in an unexpected fashion under some conditions. The basic mistakes made here involved poor software engineering practices and building a machine that relies on the software for safe operation. Furthermore, the particular coding error is not as important as the general unsafe design of the software overall. Examining the particular part of the code blamed for the Tyler accidents is instructive, however, in demonstrating the overall software design flaws. The following description of the problem is taken from the description provided by AECL for the FDA, although we have tried to clarify it somewhat. The description leaves some unanswered questions, but it is the best we can do with the information that we have.

As described in Appendix C, the treatment monitor task (TREAT) controls the various phases of treatment by executing its eight subroutines. The treatment phase indicator variable (TPHASE) is used to determine which subroutine should be executed. Following the execution of a particular subroutine, TREAT reschedules itself.

One of TREAT's subroutines, called DATENT (data entry), communicates with the keyboard handler task (a task that runs concurrently with TREAT) via a shared variable (data entry completion flag) to determine if the prescription data has been entered. The keyboard handler recognizes the completion of data entry and changes the data entry completion variable to denote this. Once the data entry completion variable is set, the DATENT subroutine detects the change in status of the variable and changes the value of TPHASE from 1 (Data Entry) to 3 (Set-Up Test). In this case, the DATENT subroutine exits back to the TREAT subroutine, which will reschedule itself and begin execution of the SETUP subroutine. If the data completion variable has not been set, DATENT leaves the value of TPHASE unchanged and exits back to TREAT's mainline. TREAT will then reschedule itself, essentially rescheduling the DATENT subroutine.

The command line at the lower right-hand corner of the screen is the cursor's normal position when the operator has completed all of the necessary changes to the prescription. Editing of the prescription is signified by movement of the cursor off of the command line.



As the program was originally designed, the data entry completion variable by itself is not sufficient since it does not ensure that the cursor is located on the command line and under the right circumstances the data entry phase can be exited before all edit changes are made on the screen.

The keyboard handler parses the mode and energy level specified by the operator and places an encoded result in another shared variable, the two-byte mode/energy offset variable (MEOS). The low order byte of this variable is used by another task (HAND) to set the collimator/turntable to the proper position for the selected mode/energy. The high order byte of the MEOS variable is used by DATENT to set several operating parameters.

Initially, the data entry process forces the operator to enter the mode and energy except when the photon mode is selected, in which case the energy defaults to 25 MeV. The operator can later edit the mode and energy separately. If the data entry completion variable is set by the keyboard handler before the data in MEOS is changed by the operator, the changes in MEOS will not be detected by DATENT since it has already exited and will not be reentered again. The upper collimator, on the other hand, is set to the position dictated by the low order byte of MEOS by another concurrently running task (HAND) and can therefore be inconsistent with the parameters set in accordance with the information in the high-order byte of MEOS. There appear to be no checks included in the software to



detect such an incompatibility.

The first thing that DATENT does when it is entered is to check whether the mode/energy has been set in MEOS. If so, it uses the high-order byte to index into a table of preset operating parameters and places them in the digital-to-analog output table. The contents of this output table are transferred to the D/A converter during the next clock cycle. Once the parameters are all set, DATENT calls the subroutine MAGNET, which sets the bending magnets. The following shows a simplified pseudocode description of relevant parts of the software:

DATENT:

```
    IF mode/energy specified THEN
      BEGIN
        calculate table index
      REPEAT
        fetch parameter
        output parameter
        point to next parameter
      UNTIL all parameters set
      CALL MAGNET
      IF mode/energy changed THEN RETURN
    END
    IF data entry is complete THEN set TPHASE to 3
    IF data entry is not complete THEN
      IF reset command entered THEN set TPHASE to 0
    RETURN
```

MAGNET:

```
    Set bending magnet flag
  REPEAT
    Set next magnet
    Call PTIME
    IF mode/energy has changed, THEN exit
  UNTIL all magnets are set
  RETURN
```

```

PTIME:
  REPEAT
    IF bending magnet flag is set THEN
      IF editing taking place THEN
        IF mode/energy has changed THEN exit
    UNTIL hysteresis delay has expired
  Clear bending magnet flag
  RETURN

```

Setting the bending magnets takes about 8 seconds. MAGNET calls a subroutine called PTIME to introduce a time delay. Since several magnets need to be set, PTIME is entered and exited several times. A flag to indicate that bending magnet setting activity is taking place is set upon entry to the MAGNET subroutine and cleared at the end of PTIME. Furthermore, PTIME checks a shared variable, set by the keyboard handler, that indicates the presence of any editing requests. If there are edits, then PTIME clears the bending magnet variable and exits to MAGNET which then exits to DATENT. But the edit change variable is only checked by PTIME if the bending magnet flag is set. Since PTIME clears it during its first execution, any edits performed during each succeeding pass through PTIME will not be recognized. Thus, an edit change of the mode and/or energy, although reflected on the CRT screen and the mode/energy offset variable, will not be sensed by DATENT so it can index the appropriate calibration tables for the machine parameters.

Recall that the Tyler error occurs when the operator makes an entry indicating the mode/energy, goes to the command line, then moves the cursor up to change the mode/energy and returns to the command line all within 8 seconds. Since the magnet setting takes about 8 seconds and edits were not recognized by MAGNET after the first execution of PTIME, the editing had been completed by the return to DATENT which never detected that it had occurred. Part of the problem was fixed after the accident by clearing the bending magnet variable at the end of MAGNET (after *all* the magnets have been set) instead of at the end of PTIME.

But this is not the only problem. Upon exit from the MAGNET subroutine, the data entry subroutine (DATENT) checks the data entry completion variable. If it indicates that data entry is complete, DATENT sets TPHASE to 3 and DATENT is not again entered. If it is not set, DATENT leaves TPHASE unchanged which means it will eventually be rescheduled. But the data entry completion variable only indicates that the cursor has been down to the command line, not that it is still there. A potential race condition is set up. To fix this, another shared variable was introduced that is controlled by the keyboard handler task and indicates that the cursor is not positioned on the command line. If this variable is set, then prescription entry is still in progress and the value of TPHASE is left unchanged.

## The Government and User Response

The U.S. FDA does not approve each new medical device on the market: All medical devices go through a classification process that determines the level of FDA approval necessary. Medical accelerators follow a procedure called pre-market notification before commercial distribution. In this process, the firm must establish that the product is substantially equivalent in terms of safety and effectiveness as a product already on the market. If that cannot be done to the FDA's satisfaction, a pre-market approval would be required. In the case of the Therac-25, a pre-market notification process only was followed.

The agency is basically reactive to problems and requires manufacturers to report serious ones. Once a problem is identified in a radiation-emitting product, the FDA is responsible to approve the Corrective Action Plan (CAP).

The first reports of the Tyler incidents came to the FDA from the state of Texas health department, and this triggered FDA action. The FDA investigation was well underway when AECL produced a Medical Device Report (MDR) to discuss the details of the radiation overexposures at Tyler. The FDA declared the Therac-25 defective under the Radiation Control for Health and Safety Act and ordered the firm to notify all purchasers, investigate the problem, determine a solution, and submit a corrective action plan for FDA approval.

The final CAP consisted of over twenty changes to the system hardware and software plus modifications to the system documentation and manuals. Some of these changes were unrelated to the specific accidents, but were improvements to the general safety of the machine. The full implementation of the CAP, including an extensive safety analysis, took more than two years after the Tyler accidents to complete.

AECL made their accident report to the FDA on April 15, 1986. On that same date, AECL sent out a letter to each Therac user recommending a temporary "fix" to the machine that would allow continued clinical use. The letter (shown in its complete form) stated:

### SUBJECT: CHANGE IN OPERATING PROCEDURES FOR THE THERAC 25 LINEAR ACCELERATOR

Effective immediately, and until further notice, the key used for moving the cursor back through the prescription sequence (i.e., cursor 'UP' inscribed with an upward pointing arrow) must not be used for editing or any other purpose.

To avoid accidental use of this key, the key cap must be removed and the switch contacts fixed in the open position with electrical tape or other insulating material. For assistance with the latter you should contact your local AECL service representative.

Disabling this key means that if any prescription data entered is incorrect then a 'R' reset command must be used and the whole prescription reentered.

For those users of the Multiport option it also means that editing of dose rate, dose and time will not be possible between ports [27].

On May 2, 1986, the FDA declared the Therac defective, demanded a CAP, and required re-notification of all the Therac customers. In the letter from the FDA to AECL, the Director of Compliance, Center for Devices and Radiological Health wrote:

We have reviewed Mr. Downs' April 15 letter to purchasers and have concluded that it does not satisfy the requirements for notification to purchasers of a defect in an electronic product. Specifically, it does not describe the defect nor the hazards associated with it. The letter does not provide any reason for disabling the cursor key and the tone is not commensurate with the urgency for doing so. In fact, the letter implies the inconvenience to operators outweighs the need to disable the key. We request that you immediately renotify purchasers. . . . [22].

AECL promptly made a new notice to users and also requested an extension to produce a CAP. This request was granted by the FDA.

About this time, the Therac-25 users created a User's Group and held their first meeting at the AAMP (American Association of Physicists in Medicine) annual conference. At the meeting, users discussed the Tyler accident and heard an AECL representative present the company's plans for responding to it. AECL promised to send a letter to all users detailing the CAP.

Several users described additional hardware safety features that they had added to their own machines to provide additional protection. An interlock (that checked gun current values), which the Vancouver clinic had previously added to their Therac-25, was labelled as redundant by AECL; the users disagreed. There were further discussions of poor design and other problems that caused a 10-30% underdosing in both modes.

The meeting notes stated that

“there was a general complaint by all users present about the lack of information propagation. The users were not happy about receiving incomplete information. The AECL representative countered by stating that AECL does not wish to spread rumors and that AECL has no policy to ‘keep things quiet’. The consensus among the users was that an improvement was necessary” [31].

After the first users group meeting, there were two users group newsletters. The first, dated Fall, 1986, contained letters from Tim Still, the physicist at Kennestone in Georgia, who complained about what he considered to be eight major problems he had experienced with the Therac-25. These problems included poor screen refresh subroutines that leave trash and erroneous information on the operator console and some tape loading problems

upon startup that he discovered involved the use of "phantom tables" to trigger the interlock system in the event of a load failure instead of using a checksum. He asked the question, "Is programming safety relying too much on the software interlock routines?" The second user group newsletter, in December, 1986, further discussed the implications of the "phantom table" parameterization.

The first CAP was produced by AECL on June 13, 1986. It contained 6 items:

1. Fix the software to eliminate the specific behavior leading to the Tyler problem,
2. Modify the software "sample and hold" circuits to detect one pulse above a non-adjustable threshold. The software sample and hold circuit is used to monitor the magnitude of each pulse from the ion chambers in the beam. Previously, three consecutive high readings were required to shut off the high voltage circuits, which resulted in a shut down time of 300 msec. The modification to the software results in a reading after each pulse, and a shutdown after a single high reading.
3. Make Malfunctions 1-64 result in treatment *suspend* rather than *pause*,
4. Add a new circuit, which only administrative staff can reset, to shut down the modulator if the "sample and hold" circuits detect a high pulse. This is functionally equivalent to the circuit described above as item two. However, a new circuit board is added that monitors the five sample and hold circuits. The new circuit detects ion chamber signals above a fixed threshold and inhibits the trigger to the modulator after detecting a high pulse. This shuts down the beam independently of the software.
5. Modify the software to limit editing keys to *cursor up*, *backspace*, and *return*.
6. Modify the manuals to reflect the changes.

FDA internal memos describe their immediate concerns regarding the CAP [28]. One memo suggests adding an independent circuit that "detects and shuts down the system when inappropriate outputs are detected," warnings about when ion chambers are saturated, and understandable system error messages. Another memo questions "whether all possible hardware options have been investigated by the manufacturer to prevent any future inadvertent high exposure."

On July 23 the FDA officially responded to AECL's CAP submission. They conceptually agreed with the direction of the plan but complained about the lack of specific information necessary to evaluate the plan, especially with regard to the software. A detailed description of the software development procedures and documentation was requested along with a revised CAP to include revised requirements documents, a detailed description of corrective changes, analysis of the interactions of the modified software with

the system, and detailed descriptions of: (1) the revised edit modes, (2) the changes made to the software setup table and (3) the software interlock interactions. They also made a very detailed request for a documented test plan.

AECL responded on September 26 with several documents describing the software and its modifications but no test plan. They explained how the Therac-25 software evolved from the Therac-6 software and stated that "no single test plan and report exists for the software since both hardware and software were tested and exercised separately and together over many years." AECL concluded that the current CAP improved "machine safety by many orders of magnitude and virtually eliminates the possibility of lethal doses as delivered in the Tyler incident" [29].

An FDA internal memo dated October 20 commented on these AECL submissions, raising several concerns:

Unfortunately, the AECL response also seems to point out an apparent lack of documentation on software specifications and a software test plan.

... concerns include the question of previous knowledge of problems by AECL, the apparent paucity of software QA at the manufacturing facility, and possible warnings and information dissemination to others of the generic type problems.

... As mentioned in my first review, there is some confusion on whether the manufacturer should have been aware of the software problems prior to the ARO's [Accidental Radiation Overdose] in Texas. AECL had received official notification of a law suit in November 1985 from a patient claiming accidental over-exposure from a Therac-25 in Marietta, Georgia. ... If knowledge of these software deficiencies were known beforehand, what would be the FDA's posture in this case?

... The materials submitted by the manufacturer have not been in sufficient detail and clarity to ensure an adequate software QA program currently exists. For example, a response has not been provided with respect to the software part of the CAP to the CDRH's [Center for Devices and Radiological Health of the FDA] request for documentation on the revised requirements and specifications for the new software. In addition, an analysis has not been provided, as requested, on the interaction with other portions of the software to demonstrate the corrected software does not adversely affect other software functions.

The July 23 letter from the CDRH requested a documented test plan including several specific pieces of information identified in the letter. This request has been ignored up to this point by the manufacturer. Considering the ramifications of the current software problem, changes in software QA attitudes are needed at AECL.

On October 30, the FDA responded to AECL's additional submissions, complaining

about the lack of a detailed description of the accident and the lack of sufficient detail in flow diagrams. Many specific questions addressed the vagueness of the AECL response and made it clear that additional CAP work must precede approval.

AECL, in response, created CAP revision 1 on November 12. This CAP contained 12 new items under Software Modifications, all (except for one cosmetic change) designed to eliminate potentially unsafe behavior. Their submission also contained other relevant documents including a test plan.

The FDA responded to CAP Revision 1 on December 11. They explained that the software modifications appeared to correct the specific deficiencies discovered as a result of the Tyler accidents. They agreed that the major items listed in CAP Revision 1 would improve functioning and operation of the Therac. However, the FDA required AECL to attend to several further system problems before CAP approval. AECL had proposed to retain treatment pause for some dose rate and beam tilt malfunctions. Since these are dosimetry system problems, the FDA considered them safety interlocks and believed treatment must be suspended for these malfunctions.

AECL also planned to retain the malfunction codes, but the FDA required better warnings for the operators. Furthermore, AECL had not planned on any quality assurance testing to ensure exact copying of software, but the FDA insisted on it. The FDA further requested assurances that rigorous testing would become a standard part of AECL's software modification procedures.

We also expressed our concern that you did not intend to perform the protocol to future modifications to software. We believe that the rigorous testing must be performed each time a modification is made in order to ensure the modification does not adversely affect the safety of the system [30].

AECL was also requested to draw up an installation test plan to ensure both hardware and software changes perform as designed when installed.

AECL submitted CAP Revision 2 and supporting documentation on December 22, 1986. They changed the CAP to have dose malfunctions suspend treatment, and they included a plan for meaningful error messages and highlighted dose error messages. They also expanded their diagrams of software modifications and expanded their test plan to cover hardware and software.

On January 26, 1987, AECL sent the FDA their Component and Installation Test Plan and explained that their delays were due to the investigation of a new accident on January 17 at Yakima.

### 3.6 Yakima Valley Memorial Hospital (Yakima, Washington), January, 1987.

On Saturday, January 17, 1987, the second patient of the day was to be treated for a carcinoma. This patient was to receive two film verification exposures of 4 and 3 rads plus a 79 rad photon treatment (for a total exposure of 86 rads.)

Film was placed under the patient and 4 rads was administered with the collimator jaws opened to 22 by 18 cm. After the machine paused, the collimator jaws opened to 35 by 35 cm automatically, and the second exposure of 3 rads was administered. The machine paused again.

The operator entered the treatment room to remove the film and verify the patient's precise position. He used the hand control in the treatment room to rotate the turntable to the field light position, allowing him to check the alignment of the machine with respect to the patient's body in order to verify proper beam position. He then either pressed the *set* button on the hand control or he left the room and typed a set command at the console in order to return the turntable to the proper position for treatment; there is some confusion as to exactly what transpired [32]. When he left the room, he forgot to remove the film from underneath the patient. The console displayed "beam ready," and the operator hit the  $\textcircled{B}$  key to turn the beam on.

The beam came on but there was no dose or dose rate displayed on the console. After 5 or 6 seconds, the unit shut down with a pause and displayed a message. The message "may have disappeared quickly"; the operator was unclear on this point [32]. However, since the machine merely paused, he was able to push the  $\textcircled{P}$  key to proceed with treatment.

The machine paused again, this time displaying "FLATNESS" on the reason line. The operator heard the patient say something over the intercom, but couldn't understand him. He went into the room to speak with the patient who reported "feeling a burning sensation" in the chest. The console displayed only the total dose of the two film exposures (seven rads) and nothing more.

The patient developed a skin burn over the entire treatment area later in the day. Four days later the redness developed a striped pattern matching the slots in the blocking tray. The striped pattern was similar to the burn a year earlier at this same hospital which had been labelled as "cause unknown."

AECL began an investigation, and users were told to confirm the turntable position visually before turning on the beam. All tests run by the AECL engineers indicated that the machine was working perfectly. From the information that had been gathered to that point, it was suspected that the electron beam had come on when the turntable was in the field light position. But they could not reproduce the fault condition that produced the overdose.

On the following Thursday, AECL sent in an engineer from Ottawa to investigate. The hospital physicist had, in the meantime, run some tests with film. He placed a film in the



beam of the Therac and then ran two exposures of x-ray parameters with the turntable in field light position. The film appeared to match the film that was left (by mistake) under the patient during the accident.

After a week of checking the hardware, AECL determined that the "incorrect machine operation was probably not caused by hardware alone" [32]. After checking the software, they discovered a flaw (described in the next section) that could explain the erroneous behavior. The coding problems explaining this accident differ from those associated with the Tyler accidents.

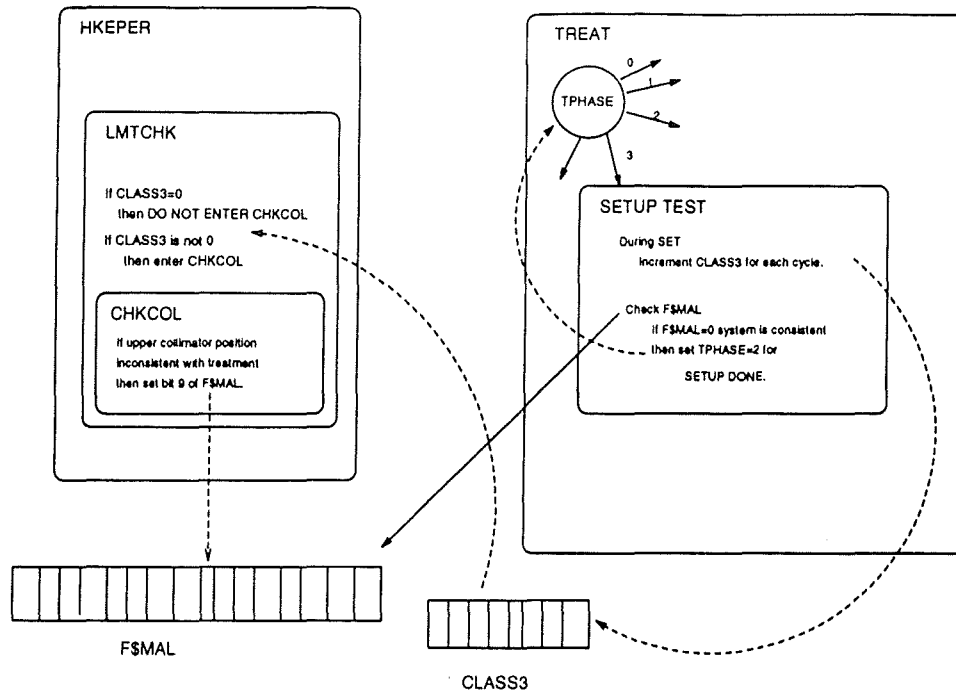
Preliminary dose measurements by AECL indicated that the dose delivered under these conditions, i.e., when the turntable was in the field light position, was on the order of 4,000 to 5,000 rads. After two attempts, the patient could have received 8,000 to 10,000 instead of the 86 rads prescribed. Users were again called on January 26 (nine days after the accident) and given detailed instructions on how to avoid this problem. In an FDA internal report on the accident, an AECL quality assurance manager investigating the problem is quoted as stating that the software and hardware changes to be retrofitted following the Tyler accident nine months earlier (but which had not yet been installed) would have prevented the Yakima accident.

The patient died in April from complications related to the overdose. He had a terminal form of cancer, but law suits were initiated by his survivors alleging that he died sooner than he would have and endured unnecessary pain and suffering due to the radiation overdose. The suit was settled out of court.

### The Yakima Software 'Bug'

The software problem for the second Yakima accident is fairly well-established and different from that implicated in the Tyler accidents. There is no way to determine what particular software design errors were related to the Kennestone, Hamilton, and first Yakima accidents. Given the unsafe programming practices exhibited in the code, it is possible that unknown race conditions or errors could be responsible for them. There is speculation, however, that the Hamilton accident was the same as this second Yakima overdose [33]. In a report of a conference call on January 26, 1987 between the AECL QA manager and Ed Miller of the FDA discussing the Yakima accident, Mr. Miller notes:

This situation probably occurred in the Hamilton, Ontario accident a couple of years ago. It was not discovered at that time and the cause was attributed to intermittent interlock failure. The subsequent recall of the multiple microswitch logic network did not really solve the problem.



Yakima Software Flaw.

The second Yakima accident was again attributed to a type of race condition in the software — this one allowed the device to be activated in an error setting (a “failure” of a software interlock). The Tyler accidents were related to problems in the data entry routines that allowed the code to proceed to SETUP TEST before the full prescription had been entered and acted upon. The Yakima accident involves problems encountered later in the logic after the treatment monitor TREAT reaches SETUP TEST.

The “field light” feature of the Therac-25 allows very precise positioning of the patient for treatment. The Therac-25 may be controlled right at the treatment site using a small hand control offering certain limited functions for patient setup, including setting gantry, collimator and table motions.

Normally, the operator enters all the prescription data at the console (outside the treatment room) before the final set up of all machine parameters is completed in the treatment room. This gives rise to an *unverified* condition at the console. The patient setup is then completed in the treatment room, and all relevant parameters now *verify*. A message is displayed on the console to *Press set button* while the turntable is in the field light position. The operator now presses the *set* button on the hand control or types “set” at the console. That should set the collimator to the proper position for treatment.

In the software, after the prescription is entered and verified by the DATENT routine,

the control variable TPHASE is changed so that the SETUP TEST routine is entered. Every pass through the SETUP TEST routine increments the upper collimator position check, a shared variable called CLASS3. If CLASS3 is nonzero, there is an inconsistency and treatment should not proceed. A zero value for CLASS3 indicates that the relevant parameters are consistent with treatment, and the beam is not inhibited.

After setting the CLASS3 variable, SETUP TEST next checks for any malfunctions in the system by checking another shared variable (set by a routine that actually handles the interlock checking) called F\$MAL to see if it has a non-zero value. A non-zero value in F\$MAL indicates that the machine is not ready for treatment, and the SETUP TEST subroutine is rescheduled. When F\$MAL is zero (indicating that everything is ready for treatment), the SETUP TEST subroutine sets the TPHASE variable equal to two, which results in next scheduling the SETUP DONE subroutine and the treatment is allowed to continue.

The actual interlock checking is performed by a concurrent Housekeeper task (HKEPER). The upper collimator position check is performed by a subroutine of HKEPER called LMTCHK (A/D limit checking). LMTCHK first checks the CLASS3 variable. If CLASS3 contains a non-zero value, LMTCHK calls the Check Collimator (CHKCOL) subroutine. If CLASS3 contains zero, CHKCOL is bypassed and the upper collimator position check is not performed. The CHKCOL subroutine sets or resets bit 9 of the F\$MAL shared variable depending on the position of the upper collimator (which in turn is checked by the SETUP TEST subroutine of DATENT to decide whether to reschedule itself or to proceed to SETUP DONE).

During machine setup, SETUP TEST will be executed several hundred times since it reschedules itself waiting for other events to occur. In the code, the CLASS3 variable is incremented by one in each pass through SETUP TEST. Since the CLASS3 variable is one byte, it can only contain a maximum value of 255 decimal. Thus, on every 256th pass through the SETUP TEST code, the variable will overflow and have a zero value. That means that on every 256th pass through SETUP TEST, the upper collimator will not be checked and an upper collimator fault will not be detected.

The overexposure occurred when the operator hit the "set" button at the precise moment that CLASS3 rolled over to zero. Thus CHKCOL was not executed and F\$MAL was not set to indicate the upper collimator was still in field light position. The software turned on the full 25 MeV without the target in place and without scanning. A highly concentrated electron beam resulted, which was scattered and deflected by the stainless steel mirror that was in the path.

The technical "fix" implemented for this particular software flaw is described by AECL as simple: the program is changed so that the CLASS3 variable is set to some fixed nonzero value each time through SETUP TEST instead of being incremented.

## Manufacturer, Government, and User Response

On February 3, 1987, after interaction with the FDA and others, including the user group, AECL announced to its customers:

- A new software release to correct both the Tyler and Yakima software problems,
- A hardware single pulse shutdown circuit,
- A turntable potentiometer to independently monitor turntable position,
- A hardware turntable interlock circuit.

The second item, a hardware single-pulse shutdown circuit, essentially acts as a hardware interlock to prevent overdosing by detecting an unsafe level of radiation and halting beam output after one pulse of high energy and current. This effectively provides an independent safety mechanism to protect against a wide range of potential hardware failures and software errors. The turntable potentiometer was the safety device that had been recommended by several groups, including the Canadian government RPB, after the Hamilton accident.

After the second Yakima accident, the FDA became concerned that the use of the Therac-25 during the CAP process, even using the interim operating instructions outlined by AECL for the users, involved too much risk to patients. The FDA concluded that the accidents had demonstrated that the software alone cannot be relied upon to assure safe operation of the machine. In a February 18, 1987 internal FDA memorandum, the Director of the Division of Radiological Products wrote:

It is impossible for CDRH to find all potential failure modes and conditions of the software. AECL has indicated the "simple software fix" will correct the turntable position problem displayed at Yakima. We have not yet had the opportunity to evaluate that modification. Even if it does, based upon past history, I am not convinced that there are not other software glitches that could result in serious injury.

For example, we are aware that AECL issued a user's bulletin January 21 reminding users of the proper procedure to follow if editing of prescription parameter is desired after entering the "B" (beam on) code but before the CR [carriage return] is pressed. It seems that the normal edit keys (down arrow, right arrow, or line feed) will be interpreted as a CR and initiate exposure. One must use either the backspace or left arrow key to edit.

We are also aware that if the dose entered into the prescription tables is below some preset value, the system will default to a phantom table value unbeknownst to the operator. This problem is supposedly being addressed in proposed interim revision 7A although we are unaware of the details.

We are in the position of saying that the proposed CAP can reasonably be expected to correct the deficiencies for which they were developed (Tyler). We cannot say that we are reasonable [sic] confident about the safety of the entire system to prevent or minimize exposure from other fault conditions [33].

On February 6, 1987, Ed Miller of the U.S. FDA called Pavel Dvorak of Canada's Health and Welfare to advise them that he would be recommending all Therac-25's be shutdown until permanent modifications can be made. According to Miller's notes on the phone call, Dvorak agreed and indicated that they would coordinate their actions with the FDA.

On February 10, 1987, the FDA gave a Notice of Adverse Findings to AECL declaring the Therac-25 to be defective under U.S. law. The letter to AECL contained in part:

In January 1987, CDRH was advised of another accidental radiation occurrence in Yakima, WA, which was attributed to a second software defect related to the "SET" command. In addition, the CDRH has become aware of at least two other software features that provide potential for unnecessary or inadvertent patient exposure. One of these is related to the method of editing the prescription after the "B" command is entered and the other is the calling of phantom tables when low doses are prescribed.

Further review of the circumstances surrounding the accidental radiation occurrences and the potential for other such incidents has led us to conclude that in addition to the items in your proposed corrective action plan, hardware interlocking of the turntable to insure its proper position prior to beam activation appears to be necessary to enhance system safety and to correct the Therac-25 defect. Therefore, the corrective action plan as currently proposed is insufficient and must be amended to include turntable interlocking and corrections for the three software problems mentioned above.

Without these corrections, CDRH has concluded that the consequences of the defects represents a significant potential risk of serious injury even if the Therac-25 is operated in accordance with your interim operating instructions. CDRH, therefore, requests that AECL immediately notify all purchasers and recommend that use of the device on patients for routine therapy be discontinued until such time that an amended corrective action plan approved by CDRH is fully completed. You may also advise purchasers that if the need for an individual patient treatment outweighs the potential risk, then extreme caution and strict adherence to operating safety procedures must be exercised [34].

At the same time, the Health Protection Branch of the Canadian government instructed AECL to recommend to all users in Canada that they discontinue the operation of the

Therac-25 until "the company can complete an exhaustive analysis of the design and operation of the safety systems employed for patient and operator protection." AECL was told that the letter to the users should include information on how the users can operate the equipment in a safe manner in the event that they must continue with patient treatment. If they could not provide information that would guarantee safe operation of the equipment, AECL was requested to inform the users that they cannot operate the equipment safely. AECL complied by letters dated February 20, 1987 to Therac-25 purchasers. This recommendation to discontinue use of the Therac-25 was to last until August of 1987.

On March 5, 1987, AECL issued CAP Revision 3, which was a CAP for both the Tyler and Yakima accidents. It contained a few additions to the Revision 2 modifications, notably: changes to the software to eliminate the behavior leading to the latest Yakima accident, four additional software functional modifications to improve safety, and a turntable position interlock in the software.

In their response on April 9, the FDA noted that in the Appendix under "turntable position interlock circuit," the descriptions were wrong. AECL had indicated "high" signals where "low" were called for and vice versa. The FDA also questioned the reliability of the turntable potentiometer design and asked whether the backspace key could still act as a carriage return in the edit mode. They requested a detailed description of the software portion of the single pulse shutdown and a block diagram to demonstrate the PRF generator, modulator and associated interlocks.

AECL responded on April 13 with an update on the Therac CAP status and a schedule of the nine action items pressed by the users at a Users Group meeting in March. This unique and highly productive meeting provided an unusual opportunity to involve the users in the CAP evaluation process; it brought together all concerned parties in one place and at one time so that a course of action could be decided upon and approved as quickly as possible. The attendees included representatives from:

- the manufacturer (AECL),
- all users, including their technical and legal staffs,
- the FDA from the U.S. and the Bureau of Radiation and Medical Devices from Canada,
- the Canadian Atomic Energy Control Board
- the Province of Ontario, and
- the Radiation Regulations Committee of the Canadian Association of Physicists.

According to Gordon Symonds, from the Canadian BRMD, this meeting was very important to the resolution of the problems since the regulators, users, and the manufacturer arrived at a consensus in one day.

At this second users meeting, the participants carefully reviewed all the six known major Therac-25 accidents to that date and discussed the elements of the CAP along with

possible additional modifications. They came up with a prioritized list of modifications that they wanted included in the CAP and expressed concerns about the lack of independent evaluation of the software and the lack of a hard copy audit trail to assist in diagnosing faults.

The AECL representative, who was the QA manager, responded that tests had been done on the CAP changes, but that the tests were not documented and that independent evaluation of the software "might not be possible." He claimed that two outside experts had reviewed the software, but he could not provide their names. In response to user requests for a hard copy audit trail and access to source code, he explained that memory limitations would not permit including such options and that source code would not be made available to users.

On May 1, AECL issued CAP Revision 4 as a result of the FDA comments and users meeting input. The FDA response on May 26, approved the CAP subject to submission of the final test plan results and an independent safety analysis, distribution of the draft revised manual to customers, and completion of the CAP by June 30, 1987. The FDA concluded by rating this a Class I recall. The basis for such a recall is a situation in which there is a reasonable probability that the use of, or exposure to, a violative product will cause serious adverse health consequences or death [35].

AECL sent more supporting documentation to the FDA on June 5, 1987, including the CAP test plan, a draft operators manual, and the draft of the new safety analysis (described in Appendix D). Their safety analysis revealed four potentially hazardous subsystems that were not covered by CAP Revision 4:

1. Electron beam scanning,
2. Electron energy selection,
3. Beam shut-off,
4. Calibration and/or steering.

AECL planned a fifth revision of the CAP to include the testing and safety analysis results.

Referring to the test plan at this, the final stage of the CAP process, an FDA reviewer said:

Amazingly, the test data presented to show that the software changes to handle the edit problems in the Therac-25 are appropriate prove the exact opposite result. A review of the data table in the test results indicates that the final beam type and energy (edit change) has no effect on the initial beam type and energy. I can only assume that either the fix is not right or the data was entered incorrectly. The manufacturer should be admonished for this error. Where is the QC [Quality Control] review for the test program? AECL must: (1) clarify this situation, (2) change the test protocol to prevent this type of error from occurring, and (3) set up appropriate QC control on data review [36].

A further FDA memo indicates:

[The AECL QA Manager] could not give an explanation and will check into the circumstances. He subsequently called back and verified that the technician completed the form incorrectly. Correct operation was witnessed by himself and others. They will repeat and send us the correct data sheet [37].

At the AAPM meeting in July of 1987, a third Users meeting was held. The AECL representative gave the status of CAP Revision 5 and explained that the FDA had given verbal approval and that he expected full implementation by the end of August of 1987. He went on to review and comment on the prioritized concerns of the last meeting. Three of the user-requested hardware changes had been included in the CAP. Changes to tape load error messages and checksums on the load data would wait until after the CAP was done.

Two user-requested hardware modifications had not been included in the CAP. One of these, a pushbutton energy and selection mode switch, would be worked on after the CAP was completed, according to the AECL QA Manager. The other, a fixed ion chamber with dose/pulse monitoring, was being installed at Yakima, had already been installed by Halifax on their own, and would be an option for other clinics. Software documentation was described as a lower priority task that needed definition and would not be available to the FDA in any form for over a year.

On July 6, 1987, AECL sent a letter to all users to update them on the FDA's verbal approval of the CAP and to delineate how AECL would proceed. Finally, on July 21, 1987, AECL issued the final and fifth CAP revision. The major features of the final CAP are:

- All interruptions related to the dosimetry system will go to a treatment suspend not a treatment pause. Operators will not be allowed to restart the machine without reentering all parameters.
- A software single pulse shutdown will be added.
- An independent hardware single pulse shutdown will be added.
- Monitoring logic for turntable position will be improved to ensure that the turntable is in one of the three legal positions.
- A potentiometer will be added to the turntable. The output is used to monitor exact turntable location and provide a visible signal of position to the operator.
- Interlocking with the 270-degree bending magnet will be added to ensure that the target and beam flattener are in position if the x-ray mode is selected.
- Beam on will be prevented if the turntable is in the light field or an intermediate position.
- Cryptic malfunction messages will be replaced with meaningful messages and highlighted dose-rate messages.
- Editing keys will be limited to the *cursor up*, *backspace*, and *return*. All other keys will be inoperative.



- A motion enable footswitch will be added which the operator is required to hold closed during movement of certain parts of the machine to prevent unwanted motions when the operator is not in control (i.e., a type of "dead man's switch.")
- Twenty three other changes will be made to the software to improve its operation and reliability including disabling of unused keys, changing the operation of the *set* and *reset* commands, preventing copying of the control program on-site, changing the way various detected hardware faults are handled, eliminating errors in the software that were detected during the review process, adding several additional software interlocks, disallowing changing to the service mode while a treatment is in progress, and adding meaningful error messages.
- The known software problems associated with the Tyler and Yakima accidents will be fixed.
- The manuals will be fixed to reflect the changes.

In a paper written in 1987, Ed Miller, the director of the Division of Standards Enforcement, Center for Devices and Radiological Health at the FDA wrote about the lessons learned from the Therac-25 experiences [37]. The first was the importance of safe versus "user friendly" operator interfaces, i.e., making the machine as easy as possible to use may conflict with safety goals. The second is the importance of providing fail-safe designs:

The second lesson is that for complex interrupt driven software, timing is of critical importance. In both of these situations, operation action within very narrow time frame windows was necessary for the accidents to occur. It is unlikely that software testing will discover all possible errors that involve operator intervention at precise time frames during software operation. These machines, for example, have been exercised for thousands of hours in the factory and in the hospitals without accident. Therefore, one must provide for prevention of catastrophic results of failures when they do occur.

I, for one, will not be surprised if other software errors appear with this or other equipment in the future.

Miller concluded the paper with:

FDA has performed extensive review of the Therac-25 software and hardware safety systems. We cannot say with absolute certainty that all software problems that might result in improper dose have been found and eliminated. However, we are confident that the hardware and software safety features recently added will prevent future catastrophic consequences of failure.

## 4 Lessons Learned

Often it takes an accident to alert people to the dangers involved in technology. A medical physicist wrote about the Therac-25 accidents:

In the past decade or two, the medical accelerator "industry" has become perhaps a little complacent about safety. We have assumed that the manufacturers have all kinds of safety design experience since they've been in the business a long time. We know that there are many safety codes, guides, and regulations to guide them and we have been reassured by the hitherto excellent record of these machines. Except for a few incidents in the 1960's (e.g., at Hammersmith, Hamburg) the use of medical accelerators has been remarkably free of serious radiation accidents until now. Perhaps, though we have been spoiled by this success [1].

Accidents are seldom simple — they usually involve a complex web of interacting events with multiple contributing technical, human, and organizational factors. One of the serious mistakes that led to the multiple Therac-25 accidents was the tendency to believe that the cause of an accident had been determined (e.g., a microswitch failure in the case of Hamilton) without adequate evidence to come to this conclusion and without looking at all possible contributing factors. Another mistake was the assumption that they could just patch one factor (e.g., eliminate the current software bug) and future accidents would be eliminated. There is always another software bug.

Accidents are often blamed on a single cause like human error. But virtually all factors involved in accidents can be labelled human error except perhaps for hardware wear-out failures. Even such wear-out failures could be attributed to human error, for example, a failure of the designer to provide adequate redundancy or a failure of operational personnel to properly maintain or replace parts: Concluding that an accident was the result of human error is not very helpful or meaningful.

It is nearly as useless to ascribe the cause of an accident to a "computer error" or a "software error." Certainly software was involved in the Therac-25 accidents, but it was only one contributing factor. By assigning "software error" as *the* cause of the Therac-25 accidents, one is forced to conclude that the only way to prevent such accidents in the future is to build perfect software that will never behave in an unexpected or undesired way under any circumstances (which is clearly impossible) or not to use software at all in these types of systems. Both of these conclusions are overly pessimistic.

The problem of accidents in complex systems must be approached from a system engineering point of view and all possible contributing factors considered and handled. In the case of the Therac-25 accidents, contributing factors included such things as management inadequacies and lack of procedures for following through on all reported incidents, overconfidence in the software and removal of hardware interlocks (i.e., making the software into

a single point of failure that could lead to an accident), presumably less than acceptable software engineering practices, and unrealistic risk assessments along with overconfidence in the results of these assessments. The exact same accident may not happen a second time, but if we examine and try to ameliorate the contributing factors to the accidents we have had, we may be able to prevent different accidents in the future. In the following sections, we present some lessons learned from the Therac-25 that we feel are important. You may draw different or additional conclusions to ours.

## 4.1 System Engineering

A common mistake in engineering, in this case and in many others, is to put too much confidence in software. There seems to be a feeling among non-software professionals that software will not or cannot fail, which leads to complacency and overreliance on computerized functions. Although it is true that software is not subject to random wear-out failures like hardware, software design errors are much harder to find and eliminate. Furthermore, hardware failure modes are generally much more limited and therefore usually easier to handle in terms of building in protection against them. A lesson to be learned from the Therac-25 accidents is not to remove standard hardware interlocks when adding computer control.

Hardware backups, interlocks, and other safety devices are currently being replaced by software in many different types of systems including commercial aircraft, nuclear power plants, and weapon systems. Where the hardware interlocks are still used, they are often controlled by software. Designing any dangerous system such that one failure can lead to an accident violates basic system engineering principles. In this respect, software needs to be treated as a single component. Software should not be assigned sole responsibility for safety, and systems should not be designed such that a single software error or software engineering error can be catastrophic.

A related tendency among engineers is to ignore software. The first safety analysis on the Therac-25 did not include software (although nearly full responsibility for the safety rested on the software). When problems started occurring, it was assumed that hardware had caused them, and the investigation focused only on the hardware. Investigation of the possible contribution of software to an accident should not be the last avenue explored after all other possible explanations are eliminated. In fact, a software error can always be attributed to a transient hardware failure since software (in these types of process control systems) reads and issues commands to actuators. Without a thorough investigation (and without on-line monitoring or audit trails that save internal state information), it is not possible to determine whether the sensor provided the wrong information, the software provided an incorrect command, or the actuator had a transient failure and did the wrong thing on its own. In the case of the Hamilton accident, a transient microswitch failure was assumed to be the cause even though the engineers were unable to reproduce the failure

or find anything wrong with the microswitch.

Patient reactions were the only real indications of the seriousness of the problems with the Therac-25. There were no independent checks that the software was operating correctly (including software checks). Such verification cannot be assigned to operators without providing them with some means of detecting errors; the Therac-25 software "lied" to the operators, and the machine itself was not capable of detecting that a massive overdose had occurred. The ion chambers on the Therac-25 could not handle the high density of ionization from the unscanned electron beam at high beam current; they thus became saturated and gave an indication of a low dosage. Engineers need to design for the worst case.

Every company building safety-critical systems should have audit trails and incident analysis procedures that are applied whenever any hint of a problem is found that might lead to an accident. The first phone call by Tim Still should have led to an extensive investigation of the events at Kennestone. Certainly learning about the first law suit should have triggered an immediate response. Although hazard logging and tracking is required in the standards for safety-critical military projects, it is less common in non-military projects. Every company building hazardous equipment should have hazard logging and tracking and incident reporting and analysis as part of their quality control procedures. Such followup and tracking will not only help prevent accidents, but will easily pay for itself in reduced insurance rates and reasonable settlement of lawsuits when they do occur.

Finally, overreliance on the numerical output of safety analyses is unwise. The arguments over whether very low probabilities are meaningful with respect to safety are too extensive to summarize here. But, at the least, a healthy scepticism is in order. The claim that safety had been increased five orders of magnitude as a result of the microswitch fix after the Hamilton accident seems hard to justify. Perhaps it was based on the probability of failure of the microswitch (typically  $10^{-5}$ ) ANDed with the other interlocks. The problem with all such analyses is that they exclude aspects of the problem (in this case, software) that are difficult to quantify but which may have a larger impact on safety than the quantifiable factors that are included.

Although there is a great desire to obtain such numbers and engineers are often pressed by management and regulatory agencies to provide them, they should insist that any risk assessment numbers that are used are meaningful and that statistics of this sort are treated with caution. In our enthusiasm to provide measurements, we should not be attempting to measure the unmeasurable. William Ruckelshaus, two-time head of the U.S. Environmental Protection Agency has cautioned that "risk assessment data can be like the captured spy; if you torture it long enough, it will tell you anything you want to know" [38, pp. 157-158]. E.A. Ryder of the British Health and Safety Executive has written that the numbers game in risk assessment "should only be played in private between consenting adults as it is too easy to be misinterpreted [39, p. 12]."

## 4.2 Software Engineering

The Therac-25 accidents were fairly unique in having software coding errors involved — the majority of computer-related accidents have not involved coding errors but rather errors in the software requirements such as omissions and mishandled environmental conditions and system states. Although using good basic software engineering practices will not prevent all software errors, it is certainly required as a minimum. Some companies that are introducing software into their systems for the first time do not take software engineering as seriously as they should. Some basic software engineering principles that apparently were violated in the case of the Therac-25 include:

- Documentation should not be an afterthought.
- Software quality assurance practices and standards should be established.
- Designs should be kept simple.
- Ways to get information about errors, i.e, software audit trails, should be designed into the software from the beginning.
- The software should be subjected to extensive testing and formal analysis at the module and software level; system testing alone is not adequate.

In addition, special safety analysis and design procedures must be incorporated into safety-critical software projects. Safety must be built into software, and, in addition, safety must be assured at the system level despite software errors [40, 41]. The Therac-20 contained the same software error as has been implicated in the Tyler deaths, but the machine included hardware interlocks that mitigated the consequences of the error. Protection against software errors can also be built into the software itself.

Furthermore, important lessons about software reuse can be found here. A naive assumption is often made that reusing software or using commercial off-the-shelf software will increase safety since the software will have been exercised extensively. Reusing software modules does not guarantee safety in the new system to which it is transferred and sometimes leads to awkward and dangerous designs. Safety is a quality of the system in which the software is used; it is not a quality of the software itself. Rewriting the entire software in order to get a clean and simple design may be safer in many cases.

Taking a couple of programming courses or programming a home computer does not qualify anyone to produce safety-critical software. Although certification of software engineers is not yet required, more events like those associated with the Therac-25 will make such certification inevitable. There is activity currently in Britain to specify required courses for those working on critical software. Any engineer is not automatically qualified to be a software engineer — an extensive program of study and experience is required. Safety-critical software requires training and experience in addition to that required for non-critical software.

Although the user interface of the Therac-25 has attracted a lot of attention, it was really a side issue in the accidents. Certainly, it could have been improved, just as many other aspects of this software could have been improved. Either software engineers need better training in interface design or more input should be obtained from human factors engineers. There also needs to be more recognition of potential conflicts between "user friendly" interfaces and safety. One of the goals of interface design is to make it as easy as possible for the operator to use. But in the Therac-25 example, some design features (e.g., not requiring the operator to reenter patient prescriptions when mistakes are made) and later changes (e.g., allowing a carriage return to indicate that information has been entered correctly) enhanced usability at the expense of safety.

Finally, not only must safety be considered in the initial design of the software and its operator interface, but the reasons for design decisions should be recorded so that they are not inadvertently undone in future modifications.

### 4.3 User and Government Oversight and Standards

Once the FDA got involved in the Therac-25, their response was impressive, especially considering how little previous experience they had with similar problems in computerized medical devices. Since the Therac-25 events, the FDA has moved to improve the reporting system and to augment their procedures and guidelines to include software. The problem of deciding when to forbid the use of medical devices that are also being used to save lives has no simple answer and involves ethical and political issues that cannot be answered by science or engineering alone. However, at the least, better procedures are certainly required for reporting problems to the FDA and to users.

The issues involved in regulation of risky technology are complex. Overly strict standards can inhibit progress, require techniques that are behind the state of the art, and transfer responsibility from the manufacturer to the government. The fixing of responsibility requires a delicate balance. Someone needs to represent the needs of the public, which may be subsumed by a company's desire for profits. On the other hand, standards can have the undesirable effect of limiting the safety efforts and investment by companies who feel that their legal and moral responsibility is fulfilled as long as they follow the standards.

Some of the most effective standards and efforts for safety come from users. Manufacturers have more incentive to satisfy customers than to satisfy government agencies. The American Association of Physicists in Medicine had established a task group to work on problems associated with computers in radiation therapy in 1979, long before the Therac-25 problems began. The accidents intensified these efforts and user-written standards are under development. A report written by J.A. Rawlinson of the Ontario Cancer Institute attempted to define the physicist's role in assuring adequate safety in medical accelerators.

We could continue our traditional role which has been to provide input to the

manufacturer on safety issues but to leave the major safety design decisions to the manufacturer. We can provide this input through a number of mechanisms . . . . These include participation in standards organizations such as the IEC, in professional association groups such as the new AAPM task group #35 on accelerator safety and a similar ACR committee and in accelerator user groups such as the Therac-25 user group. It includes also making use of the Problem Reporting Program for Radiation Therapy Devices operated by USP/CDRH and it includes consultation in the drafting of the government safety regulations. Each of these if pursued vigorously will go a long way to improving safety. It is debatable however whether these actions would be sufficient to prevent a future series of accidents.

Perhaps what is needed in addition is a mechanism by which the safety of any new model of accelerator is assessed independently of the manufacturer. This task could be done by the individual physicist at the time of acceptance of a new machine. Indeed many users already test at least the operation of safety interlocks during commissioning. Few however have the time or resources to conduct a comprehensive assessment of safety design.

A more effective approach might be to require that prior to the use of a new type of accelerator in a particular jurisdiction, an independent safety analysis is made by a panel (including but not limited to medical physicists). Such a panel could be established within or without a regulatory framework [1].

It is clear that users need to be involved. It was users who found the problems with the Therac-25 and forced AECL to respond. The process involved in fixing the Therac-25 was user-driven — the manufacturer was slow to respond. The Therac-25 user meetings were, according to participants, important to the final resolution of the problems. But if users are to be involved, then they must be provided with information and the ability to perform this function. Manufacturers need to understand that the adversarial approach and the attempt to keep government agencies and users in the dark about problems will not be to their benefit in the long run.

The U.S. Air Force has one of the most extensive programs to inform users. Contractors who build space systems for the Air Force must provide an Accident Risk Assessment Report (AFAR) to system users and operators that includes a comprehensive description of the hazardous subsystems and operations associated with that system and its interfaces. The AFAR also provides a comprehensive identification and evaluation of the accident risks of the system; provides a means of substantiating compliance with safety requirements; summarizes all system safety analyses and testing performed on each system and subsystem; and identifies design and operating limits to be imposed upon system components to preclude or minimize accidents that could cause injury or damage.

One of the interesting aspects of the Air Force AFAR is the requirement to include

a section containing a record of all safety-related failures or accidents related to system acceptance, test, and checkout along with an assessment of the impact on flight and ground safety and action taken to prevent recurrence. It also must address failures, accidents, or incidents from previous missions of this system or from other systems using similar hardware. All corrective action taken to prevent recurrence must be documented. The accident and correction history must be updated throughout the life of the system and if any design or operating parameters change after government approval, the AFAR must be updated to include all changes affecting safety.

Unfortunately, the Air Force program is not practical for commercial systems. There are some informational requirements, however, that government agencies might require manufacturers to provide to users. If these were required for everyone, competitive pressures to withhold information might be lessened. Manufacturers might find that providing such information actually increases customer loyalty and confidence. An emphasis on safety can be turned into a competitive advantage.

## 5 Final Comments

Most previous accounts of the Therac-25 accidents have blamed them on a software error and stopped there. This is not very useful and, in fact, can be misleading and dangerous: If we are to prevent such accidents in the future, we must dig deeper. Most accidents involving complex technology are caused by a combination of organizational, managerial, technical and, sometimes, sociological or political factors; preventing accidents requires paying attention to *all* the root causes, not just the precipitating event in a particular circumstance.

Accidents are unlikely to occur in exactly the same way again. If we patch only the symptoms and ignore the deeper underlying causes or we fix only the specific cause of one accident, we are unlikely to have much affect on future accidents. The series of accidents involving the Therac-25 is a good example of exactly this problem: Fixing each individual software flaw as it was found did not solve the safety problems of the device. Virtually all complex software will behave in an unexpected or undesired fashion under some conditions — there will always be another software “bug.” Instead, accidents need to be understood with respect to the complex factors involved and changes made to eliminate or reduce the underlying root causes and contributing factors that increase the likelihood or resulting loss associated with accidents.

Although these particular accidents occurred in software controlling medical devices, the lessons to be learned apply to all types of systems where computers are controlling dangerous devices. In our experience, the same types of mistakes are being made in non-medical systems. We must learn from our mistakes so that they are not repeated.



## Acknowledgements

Ed Miller of the U.S. FDA was especially helpful in providing both information to be included in this report and in reviewing and commenting on the final version. Gordon Symonds of the Canadian Government Health Protection Branch also reviewed and commented on a draft of the report. Finally, the referees, several of whom were apparently intimately involved in some of the accidents, were also very helpful in providing additional information about the accidents..

## A Turntable Positioning

The Therac-25 turntable design is important in understanding the accidents. The upper turntable (see Figure A) is a rotating table, as the name implies. The turntable rotates accessory equipment into the beam path to produce two therapeutic modes: electron mode and photon mode. A third position (called the field light position) involves no beam at all, but rather is used to facilitate correct positioning of the patient.

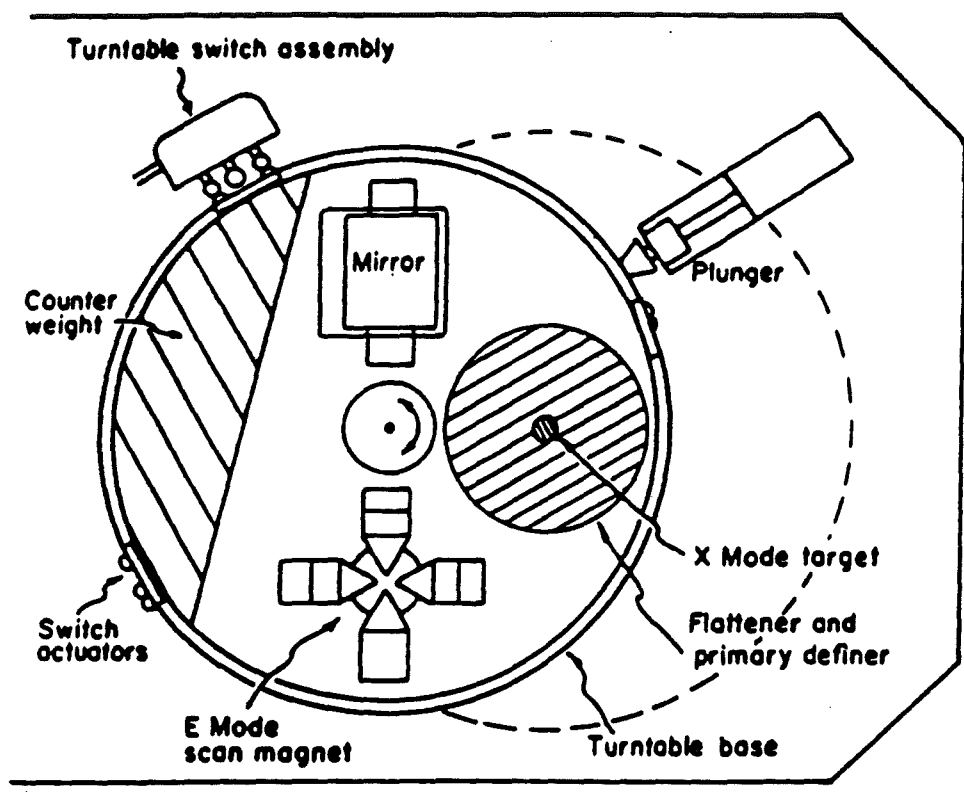
Proper operation of the Therac-25 is heavily dependent on the turntable position; the accessories appropriate to each mode are physically attached to the turntable. The turntable position is monitored by three sets of microswitches, corresponding to the three cardinal turntable positions: electron beam, x-ray, and field light. These microswitches are attached to the turntable and are engaged by hardware stops at the appropriate positions. The position of the turntable, sent to the computer as a three-bit binary signal, is based upon which of the three microswitches are depressed by the hardware stops.

The raw, highly concentrated accelerator beam is dangerous to living tissue. In electron therapy, the computer controls the beam energy (from 5 to 25 MeV) and current while scanning magnets are used to spread the beam to a safe, therapeutic concentration. These scanning magnets are mounted on the turntable and moved into proper position by the computer. Similarly, an ion chamber to measure electrons is mounted on the turntable and also moved into position by the computer. In addition, operator-mounted electron trimmers can be used to shape the beam if necessary.

For x-ray therapy, only one energy level is available: 25 MeV. Much greater electron-beam current is required for photon mode (some 100 times greater than that for electron therapy) [1] to produce comparable output. Such a high dose rate capability is required because a "beam flattener" is used to produce a uniform treatment field. This flattener, which resembles an inverted ice cream cone, is a very efficient attenuator; to get a reasonable treatment dose rate out, a very high input dose rate is required. If the machine should produce a photon beam with the beam flattener not in position, a huge output dose rate results. This is the basic hazard of dual-mode machines: if the turntable is in the wrong position the beam flattener will not be in place. In the Therac-25, the computer is responsible for positioning the turntable (and for checking the position of the turntable) so that a target, flattening filter, and x-ray ion chamber are directly in the beam path. With the target in the beam path, electron bombardment produces x-rays. The x-ray beam is shaped by the flattening filter and measured by the x-ray ion chamber.

No accelerator beam is expected in the field light position. A stainless steel mirror is placed in the beam path and a light simulates the beam. This allows the operator to see precisely where the beam will strike the patient so that necessary adjustments can be made before treatment starts. There is no ion chamber in place at this turntable position, since no beam is expected.

Traditionally, electromechanical interlocks have been used on these types of equipment



Upper Turntable Assembly

Figure A

to ensure safety — in this case, to ensure that the turntable and attached equipment are in the correct position when treatment is started. In the Therac-25, software checks were substituted for many of the traditional hardware interlocks.

## B The Operator Interface

The description of the operator interface here applies to the version of the software used during the accidents. Changes made as a result of an FDA recall are described in the text.

The operator of the Therac-25 controls the machine by use of a DEC VT-100 terminal. In the general case, the operator positions the patient on the treatment table, manually sets the treatment field sizes and gantry rotation, and attaches accessories to the machine. Leaving the treatment room, the operator returns to the VT-100 console to enter the patient identification, treatment prescription (including mode, energy level, dose, dose rate, and time), field sizing, gantry rotation, and accessory data. The system then compares the manually set values with those entered at the console. If they match, a *verified* message is displayed and treatment is permitted. If they do not match, treatment is not allowed to proceed until the mismatch is corrected. The screen layout is shown in figure B.

When the system was first built, operators complained that it took too long to enter the treatment plan. In response, the software was modified before the first unit was installed so that instead of reentering the data at the keyboard, a carriage return could be used to merely copy the treatment site data [42]. A quick series of carriage returns could thus be used to complete the entry of the data. This was to figure in several of the accidents.

The Therac-25 can shut down in two ways after an error condition is detected. One is a *treatment suspend*, which requires a complete machine reset to restart. The other, not so serious, is a *treatment pause*, which only requires a single key command to restart the machine. If a *treatment pause* occurs, the operator can press the  $\textcircled{P}$  key to “proceed” and resume treatment quickly and conveniently. The previous treatment parameters remain in effect, and no reset is required. Though convenient and simple, this feature could be invoked a maximum of five times before the machine automatically suspends treatment and requires the operator to perform a system reset.

Error messages provided to the operator were cryptic and some merely consisted of the words *MALFUNCTION XX* where *XX* was a number from 1 to 64 denoting an analog/digital channel number. According to an FDA memorandum written after one of the accidents [49]:

The operator's manual supplied with the machine does not explain nor even address the malfunction codes. The Maintenance [sic] Manual lists the various malfunction numbers but gives no explanation. The materials provided give no indication that these malfunctions could place a patient at risk.

PATIENT NAME : TEST  
TREATMENT MODE: FIX

BEAM TYPE: X ENERGY (MeV): 25

	ACTUAL	PRESCRIBED
UNIT RATE/MINUTE	0	200
MONITOR UNITS	50 50	200
TIME (MIN)	0.27	1.00

GANTRY ROTATION (DEG)	0.0	0	VERIFIED
COLLIMATOR ROTATION (DEG)	359.2	359	VERIFIED
COLLIMATOR X (CM)	14.2	14.3	VERIFIED
COLLIMATOR Y (CM) .....	27.2	27.3	VERIFIED
WEDGE NUMBER	1	1	VERIFIED
ACCESSORY NUMBER	0	0	VERIFIED

DATE : 84-OCT-26  
TIME : 12:55: 8  
OPR ID: T25V02-R03

SYSTEM: BEAM READY  
TREAT : TREAT PAUSE  
REASON: OPERATOR

OP.MODE: TREAT AUTO  
X-RAY 173777  
COMMAND:

Figure B

The program does not advise the operator if a situation exists wherein the ion chambers used to monitor the patient are saturated, thus are beyond the measurement limits of the instrument. This software package does not appear to contain a safety system to prevent parameters being entered and intermixed that would result in excessive radiation being delivered to the patient under treatment.

One of the operators involved in an overdose accident testified that she had become insensitive to machine malfunctions [12]. Malfunction messages were commonplace — most of them do not involve patient safety. Service technicians would fix the problems or the hospital physicist would realign the machine and make it operable again. She states that “...it was not out of the ordinary for something to stop the machine [...]. It would often give a low dose rate in which you would turn the machine back on. [...] They would give messages of low dose rate, V-tilt, H-tilt, and other things, I can’t remember all the reasons it would stop, but there was a lot of them” [12, pp. 41 and 42]. A radiation therapist at another clinic reported an average of 40 dose-rate malfunctions, attributed to underdoses, occurred on some days [43].

The operator further testified that during instruction she had been taught that there were “so many safety mechanisms” that she understood it was virtually impossible to overdose a patient.

## C Therac-25 Software Development and Design

We know that the software for the Therac-25 was developed by a single person, using PDP-11 Assembly language, over a period of several years. The software “evolved” from the Therac-6 software, which was started in 1972: As stated in a letter from AECL to the FDA, the “program structure and certain subroutines were carried over to the Therac 25 around 1976” [44].

Apparently very little software documentation was produced during development. In a 1986 internal FDA memo, a reviewer lamented, “[u]nfortunately, the AECL response also seems to point out an apparent lack of documentation on software specifications and a software test plan” [45].

The manufacturer says that the hardware and software were “tested and exercised separately or together over many years” [44]. In his deposition for one of the law suits, the Quality Assurance (QA) Manager explained that testing was done in two parts. A “small amount” of software testing was done on a simulator, but most of the testing was done as a system. It appears that unit and software testing was minimal with most of the effort directed at the integrated system test. At a Therac-25 users meeting, the same QA manager stated that the Therac-25 software was tested for 2700 hours. Under questioning by the users, he clarified this as meaning “2700 hours of use.” [46].

The programmer left AECL in 1986. In a lawsuit connected with one of the accidents, the lawyers were unable to obtain information about the programmer from AECL. In the depositions connected with that case, none of the AECL employees questioned could provide any information about his educational background or experience. Although an attempt was made to depose the programmer, the law suit was settled before this was accomplished. We have been unable to find out anything about his background.

AECL claims proprietary rights to its software design. However, from voluminous documentation regarding the accidents, the repairs, and the eventual design changes, we can build a rough picture of it.

The software is responsible for monitoring the machine status, accepting input about the treatment desired, and setting the machine up for this treatment. It turns the beam on in response to an operator command (assuming that certain operational checks on the status of the physical machine are satisfied) and also turns the beam off when treatment is completed, when an operator commands it, or when a malfunction is detected. The operator can print out hardcopy versions of the CRT display or of machine set-up parameters.

The treatment unit has an interlock system designed to remove power to the unit when there is a hardware malfunction. The computer monitors this interlock system and provides diagnostic messages. Depending on the fault, the computer either prevents a treatment from being started or, if the treatment is in progress, creates a pause or a suspension of the treatment.

There are two basic operational modes: Treatment Mode and Service Mode. Treatment Mode controls the normal treatment process. In Service Mode, the unit can be operated with some of the operational and treatment interlocks bypassed and additional operational commands and characteristics may be selected. Service mode is entered only through the use of a password at the service keyboard.

The Therac-25 software is described by the manufacturer as having a stand-alone, real-time treatment operating system. The system is not built using a standard operating system or executive. Rather, the real-time executive was written especially for the Therac-25 and runs on a 32K PDP-11/23. Cycles are allocated to the critical and non-critical tasks using a preemptive scheduler.

The software, written in PDP-11 assembly language, has four major components: stored data, a scheduler, a set of critical and non-critical tasks, and interrupt services. The stored data includes calibration parameters for the accelerator setup as well as patient treatment data. The interrupt routines include:

- clock interrupt service routine,
- scanning interrupt service routine,
- traps (for software overflow and computer hardware generated interrupts,
- power up (initiated at power up to initialize the system and pass control to the scheduler),

- treatment console screen interrupt handler,
- treatment console keyboard interrupt handler,
- service printer interrupt handler,
- service keyboard interrupt handler

The scheduler is responsible for controlling the sequences of all non-interrupt events and coordinating all concurrent processes. Tasks are initiated every 0.1 second with the critical tasks executed first and the non-critical tasks executed in any remaining cycle time. Critical tasks include:

- Treatment monitor (TREAT): this task directs and monitors patient set-up and treatment via eight operating phases. These are called as subroutines depending on the value of the TPHASE control variable. The subroutines are: Reset, Data Entry, Set-up Done, Set-up Test, Patient Treatment, Pause Treatment, Terminate Treatment, and Date, Time and ID Changes. Following the execution of a particular subroutine, TREAT reschedules itself. TREAT interacts with the keyboard processing task, which handles operator console communication. The prescription data is cross-checked and verified by other tasks (e.g., keyboard processor and parameter set-up sensor, etc.) that inform the treatment task of the verification status via shared variables.
- Servo task: This is used to control gun emission, dose rate (pulse repetition frequency), symmetry (beam steering), and machine motions. The servo task also sets up the machine parameters and monitors the beam tilt error and the flatness error interlocks.
- Housekeeper Task: This task takes care of system status interlocks and limit checks and puts appropriate messages on the CRT display. It decodes some information and checks the set-up verification.

Non-critical tasks include:

- Checksum processor (scheduled to run periodically)
- Treatment console keyboard processor (scheduled to run only if it is called by other tasks or by keyboard interrupts): This task acts as the interface between the other software and the operator communication.
- Treatment console screen processor (run periodically): This task lays out appropriate record formats for either CRT displays or hard copies.
- Service keyboard processor (run on demand): This task arbitrates non-treatment-related communication between the therapy system and the operator.
- Snapshot (run periodically by scheduler): Snapshot captures pre-selected parameter values and is called by the Treatment task at the end of a treatment.
- Hand control processor (run periodically)



- Calibration processor: This task is responsible for a package of tasks that allow the operator to examine and change system set-up parameters and interlock limits.

It is clear from the AECL documentation on the modifications that the software allows concurrent access to shared memory, that there is no real synchronization aside from data that are stored in shared variables, and that the "test" and "set" for such variables are not indivisible operations. Race conditions resulting from this implementation of multitasking played an important part in the accidents.

## D Safety Analysis of the Therac-25

As part of the CAP following the Tyler and Yakima accidents, AECL performed a safety analysis on the redesigned system, including software. An interim safety report was completed on January 29, 1988 and the final report was completed on November 3, 1988.

The stated purpose of the safety analysis was to identify any means whereby the Therac-25 might cause an FDA Class I Incident, which it defined as "death or serious injury by any means." Because the phrase "by any means" is too vague to be operationally useful, they interpreted this for the Therac-25 to mean "by delivery of improper radiation, by unwanted motion, by mechanical part failure, by electrical shock, by suffocation, or by scalding."

In order to determine whether a particular condition should be considered an FDA Class I Incident, the conditions or hazards that can cause death or serious injury must be defined. Since there were no established hazard and probability criteria for this type of equipment, they defined and used the following:

Failure type A, target probability  $10^{-6}$  per patient.

- 1) Faulty beam which delivers a dose of
  - a) 200% or more of prescribed dose per fraction to any part of field when there is no indication of malfunction "clear enough to generate immediate corrective action."
  - b) 1000 cGy (centigray) or more to any part of field even if clear indication of malfunction would generate immediate correction action.
- 2) Primary radiation delivered is 2.0 cm or more outside the intended field
- 3) Beam incorrect by 10% or more is delivered to a patient.
- 4) Patient-machine collision due to machine motion or mechanical failure.
- 5) Electric shock, suffocation or scalding to a patient.

Failure type B, target probability  $10^{-4}$  per patient

- 1) 125% or more of prescribed dose per fraction to any part of the field where there is no indication of malfunction clear enough to generate immediate corrective action.

## FMEA SUMMARY

ITEM	TOTAL NUMBER OF FAILURE MODES ANALYZED	NUMBER OF SINGLE FAILURES MODES THAT DEFEAT PURPOSE OF ITEM
		NUMBER
TURNTABLE POSITIONING SYSTEM	24	0
TURNTABLE POSITION INTERLOCK CIRCUIT	93	11
DOSE MONITORING SYSTEM	26	1
OVERDOSE INTERLOCK CIRCUIT	33	5
GANTRY AND COLLIMATOR MOTION ENABLE INTERLOCK	6	1

5. CONCLUSIONS5.1 Interlock Assessment

The FMEA performed indicates that the hardware interlocks perform their intended purpose:

- 5.1.1 Beam will not come on if the turntable is in the wrong position. No failure mode was found which disabled both the hardware and software turntable interlocks. Eleven failures disable the Turntable Position Interlock Circuit Board. All eleven failure modes are detected by the daily test, as specified by the Corrective Action Plan.
- 5.1.2 If a high dose-per-pulse occurs with the turntable in the x-ray or electron position, the Overdose Interlock Circuit will shut down the beam even if the software dose-per-pulse interlocks fail. Five failure modes were identified which disabled the Overdose Interlock Circuit. Three of the five are detected by the daily test as specified by the Corrective Action Plan. The other two failures are immediately detected by the operator.
- 5.1.3 During the FIX mode of operation, the gantry and collimator cannot be moved via the computer unless the footswitch is pressed.

During an ARC treatment, however, the gantry can fail to stop after the treatment, due to a software fault. This failure could cause a Class 1 Hazard if the operator does not respond in time (taking foot off footswitch).

Table 1

- 2) 400 cGy or more to any part of the field even if there is a malfunction message clear enough to generate immediate corrective action.

The report notes that some conclusions will not be strictly derivable from the given criteria, i.e., those involving safety-related functions under computer control, since they cannot be quantified in the same way as hardware malfunctions.

The methodology used included three types of analysis: Failure Mode and Effects Analysis (FMEA), Fault Tree Analysis (FTA), and software examination.

**Failure Modes and Effects Analysis.** An FMEA describes the associated system response to all failure modes of the individual components of the system, considered one by one. In this particular FMEA, the response was recorded as a 'Pass' or 'Fail,' where 'Fail' means a Class I hazard was present. When the system response was not obvious, a test was performed. When software was involved, no assessment was made of the "how and why" of software faults and any combination of software faults were taken as a single event. The latter means that if the software is the initiating event, then no credit was taken for the software to mitigate the effects. This seems like a reasonable and conservative approach to handling software faults.

Three Therac-25 system functions were addressed in the FMEA: turntable positioning, dose-rate monitoring, and gantry and collimator motion control. Table 1 (taken from the Therac-25 FMEA report) shows a summary of the results, including the total number of failure modes analyzed and the number of single failure modes that defeat the purpose of the item.

During the FMEA, four software-related potential Class I hazards were identified:

1. Electron Beam Scanning: Scanning is both controlled and interlocked by software. Localized large overdoses can occur with collapsed electron beams, particularly at high energies.
2. Electron Beam Selection: There is no simple means whereby the operator can ascertain that the computer has correctly set up the electron beam energy-dependent parameters. Such information can be obtained with the maintenance mode printout, but there is no direct feedback to the operator video terminal.
3. Beam Shut Off: The computer has complete control to turn off the beam after the prescribed dose is achieved. If it failed to do so, the beam would continue to run until the operator intervened by pressing the *Emergency Off* switch. The amount of excess radiation delivered then depends on the operator's response time, which "could be considerable." There is no hardware backup timer.

4. Calibration And/Or Steering: All parameters for proper beam flatness and calibration are computer controlled. An out of calibration unit can generate a beam that produces an overdose condition, operating below hardware interlock threshold values.

The authors of the FMEA recommended that a further safety study of these systems be conducted.

**Fault Tree Analysis.** An FMEA identifies single failures leading to Class I hazards. To identify multiple failures and quantify the results, AECL used Fault Tree Analysis. An FTA starts with a postulated hazard, e.g., two of the top events for the Therac-25 are 'High Dose Per Pulse' and 'Illegal Gantry Motion.' The immediate causes for the event are then generated in an AND/OR tree format, using a basic understanding of the machine operation to determine the causes. The tree generation continues until all branches end in "basic events." Operationally, a basic event is sometimes defined as an event that can be quantified (e.g., resistor fails open).

For the Therac-25, the safety analysis included fault tree analysis to investigate the following hazards: Calibration or Tilt Error (two trees), Beam Fails to Stop (one tree), High Dose per Pulse (one tree), Improper Scanning (two trees), Wrong Beam Energy (two trees), Illegal Motion (three trees), and Beam On at Wrong Time (one tree). We have been unable to examine any of the fault trees directly.

A "generic failure rate" of  $10^{-4}$  per hour was used for software events. This number was justified as being based on the historical performance of the Therac-25 software. The final report on the safety analysis says that many of the fault trees for the Therac-25 have a computer malfunction as a causative event, and the outcome of quantification is therefore dependent on the failure rate chosen for software.

Leaving aside the general question of whether such failure rates are meaningful or measurable for software in general, it seems rather difficult to justify a single figure of this sort for every type of software error or software behavior. It would be equivalent to assigning the same failure rate to every type of failure of a car, no matter what the particular failure being considered. The authors of the safety study did note that despite the uncertainty that software introduces into quantification, fault tree analysis provides valuable information in showing single and multiple failure paths and the relative importance of different failure mechanisms. This is certainly true.

**Software Examination.** Because of the difficulty of quantifying software behavior, a detailed inspection of the code was conducted in order to "obtain more information on which to base decisions" [47]. The particular software functions selected for examination were those related to the Class I software hazards identified in the FMEA: Electron Beam Scanning, Energy Selection, Dose Calibration, and Beam Shut-off.

The inspection was performed under contract by an outside consultant and included a detailed examination of the implementation of each function, a search for coding errors, and a qualitative assessment of its reliability. Program changes were recommended to correct any shortcomings, improve reliability, or improve the software package in a general sense. No information is given in the final safety report as to whether any particular methodology or tools were used in the software inspection or whether someone just read the code looking for errors.

**Supplemental Analyses.** In order to analyze structural failure (which is not handled well by the above techniques), stresses in all relevant hardware were calculated and compared to part strength and standard design safety margins. Relevant hardware included all parts that might fall on the patient or allow the patient to fall suddenly.

**Conclusions of the Safety Analysis.** The conclusions of the safety analysis are summarized in the final report:

The conclusions of the analysis call for 10 changes to Therac-25 hardware; the most significant of these are interlocks to back up software control of both electron scanning and beam energy selection.

Although it is not considered necessary or advisable to rewrite the entire Therac-25 software package, considerable effort is being expended to update it. The changes recommended have several distinct objectives: improve the protection it provides against hardware failures; provide additional reliability via cross-checking; and provide a more maintainable source package. Two or three software releases are anticipated before these changes are completed.

The implementation of these improvements including design and testing for both hardware and software is well under way. All hardware modifications should be completed and installed by mid 1989, with final software updates extending into late 1989 or early 1990.

The recommended hardware changes appear either to add protection against software errors, to add extra protection against hardware failures, or to increase safety margins. The software conclusions included:

The software code for Beam Shut-Off, Symmetry Control, and Dose Calibration was found to be straight-forward and no execution path could be found which would cause them to perform incorrectly. A few improvements are being incorporated, but no additional hardware interlocks are required.

Inspection of the Scanning and Energy Selection functions, which are under software control, showed no improper execution paths; however, software

inspection was unable to provide a high level of confidence in their reliability. This was due to the complex nature of the code, the extensive use of variables, and the time limitations of the inspection process. Due to these factors and the possible clinical consequences of a malfunction, computer-independent interlocks are being retrofitted for these two cases.

Given the complex nature of this software design and the basic multitasking design, it is difficult to understand how any part of the code could be labeled "straightforward" or any confidence be achieved that "no execution paths" exist for particular types of software behavior. However, it does appear that a conservative approach, i.e., including computer-independent interlocks, was taken in most cases. Furthermore, few examples of such safety analyses being done on software exist in the literature. One such software analysis was performed in 1989 on the shutdown software of a nuclear power plant that was written by a different division of AECL [48]. Much still needs to be learned about how to accomplish these goals.

## E Major Event Timeline

### 1985

- June 3 Marietta, Georgia overdose
- June Tim Still calls AECL and asks is overdose by Therac-25 is possible
- July 26 Hamilton, Ontario overdose, AECL notified and determines microswitch failure was the cause.
- Sept. AECL makes changes to microswitch and notifies users of increased safety
- Sept. Independent consultant (for Hamilton Clinic) recommends potentiometer on turntable
- Oct. Georgia patient files suit against AECL and hospital
- Nov. 8 Letter from RPB to AECL asking for additional hardware interlocks and software changes.
- Dec. Yakima clinic overdose.

### 1986

- Jan. Attorney for Hamilton clinic requests that potentiometer be installed on turntable.
- Jan. 31 Letter to AECL from Yakima reporting overdose possibility
- Feb. 24 Letter from AECL to Yakima saying overdose was impossible and no other incidents had occurred.
- Mar. 21 Tyler, Texas overdose. AECL notified; claims overdose impossible and no other accidents had occurred previously. AECL suggests hospital might have an electrical problem
- Apr. 7 Tyler machine put back in service after no electrical problem could be found.
- Apr. 11 Second Tyler overdose. AECL again notified. Software problem found.
- Apr. 15 AECL files accident report with FDA
- May 2 FDA declares Therac-25 defective. Asks for CAP, and proper renotification of Therac-25 users
- June 13 First version of CAP sent to FDA
- July 23 FDA responds and asks for more information.
- Sept. 26 AECL sends FDA additional information
- Aug. First users group meeting
- Oct. 30 FDA requests more information
- Nov 12 AECL submits revision of CAP
- Dec. 11 FDA requests further changes to CAP
- Dec. AECL submits second revision of CAP
- Dec. Therac-20 users notified of a software bug.

**1987**

- Jan. 17 Second overdose at Yakima
- Jan. 26 AECL sends FDA their revised test plan
- Feb. Hamilton clinic investigates first accident and concludes there was an overdose.
- Feb. 3 AECL announces changes to Therac-25
- Feb. 10 FDA sends notice of Adverse Findings to AECL declaring Therac-25 defective under U.S. law and asking AECL to notify customers that machine should not be used for routine therapy. Health Protection Branch of Canada does the same thing. This lasts until August 1987.
- Mar. 5 AECL sends third revision of CAP to FDA
- Mar. Second user's group meeting
- Apr. 9 FDA responds to CAP and asks for additional information
- May 1 AECL sends fourth revision of CAP to FDA
- May 26 FDA approves CAP subject to final testing and safety analysis.
- June 5 AECL sends final test plan and draft safety analysis to FDA
- July Third user's group meeting
- July 21 Fifth (and final) revision of CAP sent to FDA

**1988**

- Jan. 29 Interim safety analysis report issued
- Nov. 3 Final safety analysis report issued.



## References

- [1] Rawlinson, J.A. Report on the Therac-25, OCTRF/OCI Physicists Meeting, Kingston, Ontario, 7 May, 1987.
- [2] Houston, F. "What Do the Simple Folk Do?: Software Safety in the Cottage Industry," *IEEE Computers in Medicine Conference*, Chicago, September 1985.
- [3] AECL memo to FDA, January 4, 1988.
- [4] W.D. (AECL) letter to Fritz Hager (ETCC), January 21 1987
- [5] Breu, G. and Plummer, W. "A Computer Glitch Turns Miracle Machine into Monster for Three Cancer Patients," *People*, Vol. 26 No. 21, November 24, 1986, pp. 48-50.
- [6] Hager, F. Incident Report for ETCC, April 11, 1986.
- [7] Tim Still deposition for ETCC lawsuit.
- [8] Gagne, R.M., memo to E. Miller (FDA), July 11, 1986.
- [9] Saltos, R. "Man Killed by accident with medical radiation," *Boston Globe* June 20, 1986.
- [10] W.D. (AECL QA Manager) deposition in ETCC lawsuit.
- [11] Bowsher, C.A. "Medical Devices: The Public Health at Risk", GAO Report GAO/T-PEMD-90-2, 046987 / 139922, 1990.
- [12] P.R. (Operator) deposition in ETCC lawsuit.
- [13] Hager deposition in ETCC lawsuit.
- [14] Newman, H.F. "The Malfunction '54' Accelerator Accidents," FDA report (Oak Ridge).
- [15] D.N. (AECL Technician) deposition in ETCC lawsuit.
- [16] Gunter, A. W. (AECL), Letter to Schlichtemeier, A. L. (ETCC), September 16, 1985.
- [17] Final Report AECL file Q2.9.3, AECL Complaint No. PHR 135, Access No. R-109611, Medical Device Report, September 16, 1985.
- [18] American Medical News, Jan. 16, 1987

- [19] [\*\*] all this is in the American Medical News, January 1987, in case we can't cite the original.
- [20] Investigative report of FDA, April 25, 1986.
- [21] Davis, A. Initial investigative report memo from FDA, April 25, 1986.
- [22] Gundaker, W.E., Letter from FDA to AECL, May 2 1986
- [23] W.D. (Q.A. Manager, AECL) Letter to W.E. Gundaker (FDA), September 26, 1986.
- [24] American Medical News, Nov. 7, 1986.
- [25] Kivel, M. (ed.), Radiological Health Bulletin, vol. XX, No. 8, Center for Devices and Radiological Health, Food and Drug Administration, Rockville, MD., December 1986.
- [26] Miller, E. (FDA) Memo, June 15, 1987.
- [27] Letter from AECL to users, April 15 1986.
- [28] FDA memo, July 11, 1986.
- [29] Letter from AECL to FDA, September 26, 1986.
- [30] Letter from Ed Miller to W.D. (AECL QA Manager), December 11, 1986.
- [31] Therac-25 Users Group Meeting, AAPM Annual Meeting, Lexington, Kentucky, August 6, 1986.
- [32] Medical Device Regulation, Incident Report #1, AECL Complaint No. PHR #176, February 19, 1987.
- [33] Miller, E. Memo to Director, FDA Office of Compliance, February 18, 1987.
- [34] Letter from FDA to AECL, February 10, 1987.
- [35] *Medical Device Recalls, Examination of Selected Cases*, GAO/PEMD-90-6. October 1989.
- [36] Gagne, R.M., Memo to E. Miller of FDA, June 19, 1987.
- [37] Memo of telephone conversation between E.A. Miller and W.D. (AECL QA Manager), June 23, 1987.
- [38] Ruckelshaus, W.D. "Risk in a Free Society," *Risk Analysis*, Vol. 4, No. 3, 1984, pp. 157-162.

- [39] Ryder, E.A. "The Control of Major Hazards: The Advisory Committee's Third and Final Report," Transcript of Conference on European Major Hazards, Oyez Scientific & Technical Services and Authors, London, 1984.
- [40] Leveson, N.G. "Software Safety: Why, What, and How," *ACM Computing Surveys*, Vol 18, No. 2, June 1986, pp. 25-69.
- [41] Leveson, N.G. "Software Safety in Embedded Computer Systems," *Communications of the ACM*, February, 1991, pp. 34-46.
- [42] Ed Miller, The Therac-25 Experience, *Proc. of Conference of State Radiation Control Program Directors*, 1987,
- [43] American Medical News article, February 6, 1987.
- [44] W.D. (AECL QA Manager) letter to FDA, September 26, 1986.
- [45] FDA memo, October 20, 1986.
- [46] Minutes of Therac-25 User's Meeting of March 30, 1987.
- [47] W.D. and E.M., Therac-25 Safety Analysis, Final Findings Report, (Theratronics/AECL), November 1, 1988.
- [48] Bowman, W.C., Archinoff, G.H., Raina, V.M., Tremaine, D.R., and Leveson, N.G. "An Application of Fault Tree Analysis to Safety-Critical Software at Ontario Hydro," *Conference on Probabilistic Safety Assessment and Management (PSAM)*, Beverly Hills, April 1991.
- [49] Lahar, James, FDA Internal Memo, April 25, 1986.
- [50] Gagne, R.M., Review for the FDA, June 19, 1987.
- [51] Ricky W. Butler, R.W. and Finelli, G.B. "The Infeasibility of Experimental Quantification of Life-Critical Software Reliability", *Proceedings of the ACM SIGSOFT '91 Conference on Software for Critical Systems*, New Orleans, December 1991.