116037 ~~14663~~

UNITED STATES GENERAL ACCOUNTING OFFICE
WASHINGTON, D.C. 20548

HUMAN RESOURCES DIVISION

AUGUST 5, 1981

B-204011

116037

Dr. Arthur Hull Hayes, Jr.
Commissioner, Food and
Drug Administration

Dear Dr. Hayes:

Subject: Software Used in Medical Devices Needs Better Controls To Avoid Compromising Patient Safety (AFMD-81-95)

Computer technology is becoming more and more an integral part of many medical devices and is providing many benefits in the diagnosis and treatment of patients' illnesses. We have looked into the use of computerized medical devices to determine if they are dependable enough to avoid compromising patient safety and, if not, how the controls being developed by the Bureau of Medical Devices will address this area of computer technology. As a result of our inquiry we believe controls are needed to ensure that the software used in medical devices operates as expected.

We understand the goals and objectives of the Bureau's recently formed Task Group on Computers and Associated Software as Medical Devices will be established later, but that part of the group's efforts will include assessing the performance of software used for medical purposes. We believe this is a step in the right direction because it will address the type of problems our review has indicated are of emerging significance. Therefore, we plan no further work in this specific area at this time.

To assist the task group in formulating its objectives and approach, we are summarizing below the information we developed and offering our suggestions for addressing software reliability.

During our review, we looked at various reports from the Bureau's medical device reporting system, legislation, regulations, publications, and Bureau procedures. We interviewed Food and Drug Administration (FDA) and Bureau officers, other Government agency officials, representatives from various medical organizations, health care professionals, and medical device manufacturers.

COMPUTER SOFTWARE IN MEDICAL DEVICES NEEDS TO BE VALIDATED

Although computerized medical devices provide many benefits, the software used in them can sometimes be troublesome. We found

(913460)

017821/116037

that the software in medical devices did not always function as expected. We also found agreement among several professional groups that software is a problem requiring attention. If controls are not developed to ensure the reliability of software, many people may be exposed to undue risks by improper diagnosis or treatment.

Computer software has caused devices to fail

We identified 78 cases involving unreliable computerized medical devices that occurred from June 1976 to August 1979. These cases were obtained from health care professionals, FDA recall files, and the Bureau's medical device problem reporting system. In each case problems had been encountered with the reliability of a computerized medical device. In all, we estimate that about 30 different types of computerized medical devices were involved. As you know, medical device manufacturers, hospitals, and health care professionals do not always report failures to the Bureau primarily because they fear litigation. The Freedom of Information Act can open a case to public scrutiny. Therefore, we believe the cases identified may be only the "tip of the iceberg." A Bureau official who is responsible for the medical device problem reporting system estimates that for every case reported, nine are not.

We selected for review 24 cases and found 13 of them had software problems. The problems involved seven different devices. The remaining cases involved various hardware component failures, manufacturing deficiencies, or design defects. Sufficient detailed information was not available to permit us to perform our own assessment of these problems. We relied upon the assessments made by manufacturers, FDA representatives, or health care professionals.

The exposure of patients to risk is potentially significant. In four of these seven devices with problems, manufacturers voluntarily recalled nearly 1,500 units. The magnitude of risk exposure from these devices alone is large because each unit can be used for hundreds, sometimes thousands, of patients.

Discussed below are two examples of software problems.

Blood gas analyzer

In January 1978, a manufacturer of a blood gas analyzer determined that 885 devices were defective because the software they used included calculations that were not acceptable to the scientific community. According to an FDA report, reliance on this device could lead to too much or too little medication being given to a patient and thus adversely affect recovery.

In August and September 1977, the manufacturer began receiving complaints from the scientific community because the software

in the instrument made a correction for the patient's temperature when it calculated the values of certain blood gases such as oxygen. In October 1977, the manufacturer, concerned about these complaints, had two consultants evaluate the validity of the software calculation. Both considered that temperature correction for measuring certain gases was inappropriate. Because of these opinions the manufacturer deleted this feature from the software.

The manufacturer decided to recall the devices. Of the 885 units, 434 had been distributed in the United States and Canada; the rest had been sent to other foreign countries. As of October 1978, all but one of the units distributed in North America had been fitted with new integrated circuit boards containing a revised software program.

Computerized electrocardiogram interpretation software package

According to some doctors, this type of software cannot always be relied on to produce correct interpretations of electrocardiograms. For example, in February 1979 a doctor at one hospital complained that the software package was "unreliable and full of diagnostic errors." Other doctors at the same hospital made a study to evaluate the merits of the complaint. They concluded that the diagnoses made by the software contained minor errors 30 percent of the time and major errors 5 percent of the time. Their report did not define minor or major errors.

Others have also expressed concern over this type of software. A doctor told us that an Indian Health Service hospital used such a software package but discontinued its use because it misdiagnosed cases. The American Heart Association shares the concern about these problems and is pursuing ways to increase the accuracy of the diagnostic readouts produced by software packages.

Software may be a problem for devices on priority list for standards

The Bureau has placed about 200 classes of medical devices, of which 34 percent may be computerized, on its priority list of devices requiring the development of performance standards. Standards are considered necessary to guard against unreasonable risk to patients.

In developing its priority list, Bureau officials relied upon recommendations of its many panels of medical experts as well as its own assessment of the following factors:

--Hazards. (The possible consequences of using or misusing a device, or the failure of the device or its components. The consequences range from fatal to no hazard.)

--Hazard probability. (The probability that a hazard will occur. The probabilities range from over 20 percent to less than 1 percent.)

--Device usage. (The current and projected usage of the device within a 2-year period. Estimated usage ranges from over 2 million to less than 100.)

Although we did not look into it, software could be a significant source of reliability problems with these 200 items. As a matter of fact, the list includes one of the seven devices our sample showed as experiencing software problems.

Health care professionals believe greater software reliability is needed

Many health care professionals believe that the unreliability of software is exposing patients to unnecessary risk. In some selected areas, health care professionals are working on ways to increase reliability.

During 1978 congressional hearings on the use of computers in health care, an American Medical Association representative stated that the reliability of software packages varies and that controls are needed to ensure the safety and effectiveness of complex computer techniques. This concern was echoed during an October 1979 National Institutes of Health Consensus Development Conference on "The Use of Microprocessor-Based 'Intelligent' Machines in Patient Care."

Approximately 350 members of the medical research community--practitioners, consumers, and others--attended the conference to reach consensus on the major benefits and problems encountered in using this technology in the delivery of health care. Conferees concluded that the unreliability of software is a problem and that benchmark programs for software validation and comparison are needed.

Health care professionals continue to meet and discuss the need for ensuring the reliability of software packages. In April 1981, the Bureau of Radiological Health cosponsored two such meetings to generate a dialog about quality assurance techniques applied to radiation treatment software. In November 1981, the Fifth Annual Symposium on Computer Applications in Medical Care will include a workshop to discuss quality control procedures necessary to ensure safe and effective computer programs.

In some selected areas, health care professionals are actively seeking ways to increase the reliability of software packages:

--The Biomedical Computing Technology Information Center, located at the Vanderbilt Medical Center in Nashville,

Tennessee, is developing a data base to measure the reliability of nuclear cardiology software applications. The Center's Codirector told us that the Center, as well as many members of the nuclear medicine community, is concerned about the reliability of computer software.

--The American Heart Association is developing a data base consisting of ventricular arrhythmias readings (abnormal rhythm of that part of the heart that pumps blood into the arteries). The data base will be used to stimulate efforts to improve the reliability and accuracy of software used in arrhythmia monitors. These are devices that scan electrocardiogram rhythms and alert medical personnel when changes in rhythms indicate a threat to a patient's life.

--The Radiation Therapy Scientific Committee, American Association of Physicists in Medicine, has established a task group to work on problems associated with computer applications that calculate the dosage for patient radiation therapy. The chairman of the group told us its concern about validation is increased by a certain mystique surrounding computers. There is " * * * too much tendency in the U.S. for everyone to depend on what comes out of the computer as correct." In July 1979, this task group met and agreed that its primary objective is to develop techniques for validating the accuracy of computer programs that plan radiation therapy treatment. The group has developed some preliminary tests; however, a complete test method for validating this type of software has not yet been developed. Significantly, the preliminary tests identified a "bug" in one software package being used.

ALTERNATIVES TO PERFORMANCE STANDARDS ARE NEEDED

Under the Medical Device Amendments of 1976 to the Federal Food, Drug, and Cosmetic Act, the Bureau can legally rely upon performance standards and good manufacturing practice regulations to address the reliability of computer software. In our opinion, however, performance standards are not likely to provide effective control in the near future; other alternatives, possibly under the aegis of good manufacturing practices, need to be explored.

The 1976 amendments require the Bureau to classify medical devices into three categories:

Class I — Devices which do not present an unreasonable risk of illness and injury and can be monitored by general controls which include adherence to good manufacturing practices.

Class II — Devices for which general controls are insufficient to provide reasonable assurance of safety

and effectiveness and for which there is sufficient information to establish a performance standard to provide such assurance.

Class III Devices which present a potential unreasonable risk of illness or injury and are used in supporting or sustaining life and must be controlled by the Bureau's pre-market approval process.

Before marketing a class III device, manufacturers must submit to the Bureau data supporting the device's safety and effectiveness. The data must show whether the device fully meets any performance standards that apply to any aspect of the device; noncompliance must be justified.

The Bureau, assisted by panels of medical, scientific, and engineering experts, has tentatively placed 1,200 medical devices into class II. Some of these devices are computerized; others are not. Five of the seven devices with software problems covered by our review are in class II or III.

Timely performance standards are unlikely

In developing a performance standard, the Bureau must follow complex, burdensome, and time-consuming procedures mandated by the amendments. As a result, each standard could take as much as 5 to 10 years to develop.

It has been over 5 years since the Congress passed the Medical Device Amendments of 1976. During this time, the Bureau has been working to classify medical devices. Although the classification process is not complete, the 1,200 items that have been tentatively placed into class II require performance standards. The Bureau also has been working to formulate and publish rules and regulations pertaining to the procedures for developing performance standards. These rules and regulations were published in the Federal Register on February 1, 1980, and became effective on July 30, 1980.

The performance standards development process is complex and burdensome. The process requires the Bureau to publish five Federal Register documents to obtain comments from the public and other interested parties before a proposed standard becomes final. In addition, the process can be time consuming for other reasons. Bureau officials believe that for medical devices the process could take 10 years before a standard would be published.

To begin the development process, the Bureau is required to provide the public, manufacturers, and other interested parties with another opportunity to present arguments on whether a device should be placed in class II. The public has already been given this opportunity during the classification process. If the Bureau

decides against a change in classification, any person may appeal the decision. If the device remains in class II after an appeal, the standard development process will continue.

The Bureau is required to keep the public informed during the development process. It must publish in the Federal Register whether it plans to have a standard developed or to use an existing standard. It must provide other information such as the parties involved, the procedures being used to develop a standard, and the requirements of Federal, State, and local environmental regulations.

The Bureau is also required to give the public the opportunity to participate in the development of a standard. This participation can come during public meetings held by the developer of a standard or it can come when the Bureau publishes a draft in the Federal Register for comments. The Bureau must also establish advisory committees to which proposed standards can be referred. When a proposed standard is referred, the committee must submit within 60 days a report and a recommendation to the Bureau for its resolution.

Other alternatives to mandatory performance standards are possible

Because of the constraints surrounding the development of performance standards, the Bureau needs to look at other alternatives for ensuring that the software in medical devices works as expected. We have not fully investigated alternatives available to the Bureau, but some possibilities appear promising and should be explored. One possibility is a software certification procedure in which a manufacturer would certify to the Food and Drug Administration that the software has been validated. The validation could be performed by the manufacturer or a third party--for example, a group similar to the Biomedical Computing Technology Information Center, which performs a similar service for developers of nuclear medicine applications, or validation and certification could be performed by the Food and Drug Administration.

In conjunction with these alternatives, or separate from them, the Bureau could establish certain labeling requirements under the authority of the amendments. For example, the labeling standards might require manufacturers to disclose the general logic and assumptions included in the design of the software. Such standards could also require manufacturers to disclose whether the software has been validated or certified.

CONCLUSIONS

Advances in computer technology have brought about far more reliable hardware. However, software has been and remains a problem area, regardless of whether it is used in medical or business applications. We believe the use of software in medical devices

is emerging as a troublesome area and requires the attention of the Bureau.

The use of performance standards, as authorized by the Medical Device Amendments of 1976, is a possible mechanism to help control the performance of software in computerized medical devices. Unfortunately, the time-consuming process for developing standards together with the large number of standards to be developed makes it very unlikely that any standards will be available soon. This, coupled with the relatively fast pace at which computer technology changes, makes it unlikely that the standards when developed will be timely enough to validate software in medical devices. Therefore, we believe the Bureau needs to explore other alternatives for validating and certifying that the software in medical devices works as expected.

RECOMMENDATIONS

We believe the Bureau's recent establishment of the task group is a wise and timely initiative. We recognize the goals of the group are still being formulated. We recommend that you direct the Bureau to establish, as one of the task group's primary goals, the identification and evaluation of alternatives to performance standards to ensure that software in medical devices operates as it is expected to operate.

To develop workable alternatives and to capitalize on existing ones, we also recommend that you direct the Bureau and its task group to enlist the support and participation of the medical device industry, the medical profession, and interested Government agencies.

\- - - - -

We have discussed this report with officials from your office and the Bureau. Their comments have been incorporated into the report. We appreciate the courtesies and cooperation extended to our representatives during the review. We would appreciate your comments and advice as to actions you take or plan to take on these matters.

Sincerely yours,

Carl R. Fenstermaker

Carl R. Fenstermaker
Group Director