United States General Accounting Office

GAO

Report to the Chairmen, Senate and House Committees on Armed Services

July 1994

MEDICAL ADP SYSTEMS

Defense's Tools and Methodology for Managing CHCS Performance Need Strengthening

GAO/AIMD-94-61

GAO

United States
General Accounting Office
Washington, D.C. 20548

Accounting and Information
Management Division

B-220732

July 15, 1994

The Honorable Sam Nunn
Chairman, Committee on Armed Services
United States Senate

The Honorable Ron Dellums
Chairman, Committee on Armed Services
House of Representatives

The Composite Health Care System (CHCS) is an automated medical information system being developed by the Department of Defense. CHCS is intended as a state-of-the-art, integrated system. The ultimate purpose of the system is to improve the quality and reduce the cost of providing medical care to beneficiaries of the military health-care system. Defense intends to deploy CHCS at a total of 565 medical treatment facilities worldwide. We are legislatively required to monitor Defense's development, testing, evaluation, and deployment of CHCS and periodically report to your committees.[1]

Ideally, CHCS should be easy to use, its results should be accurate and complete, and its availability should be high with consistently fast response times. A major point of attention throughout the development, testing, and evaluation of CHCS should be how well the system will meet the needs of clinical users, such as doctors and nurses.

On May 20, 1992, we reported to your committees that clinical users at 8 of 12 CHCS test sites were experiencing system performance problems.[2] Two of the test sites—the Walter Reed Army Medical Center in Washington, D.C., and the U.S. Air Force Medical Center at Keesler Air Force Base in Biloxi, Mississippi—had average response times that were 50 percent above Defense's criteria for acceptable performance levels. This report contains our further evaluation of Defense's management of CHCS performance,[3] focusing on the tools for measuring and analyzing this

[1]The National Defense Authorization Act for fiscal year 1987—Public Law 99-661, Sec. 704, Nov. 14, 1986; as amended by the National Defense Authorization Act for fiscal years 1988 and 1989—Public Law 100-180, Sec. 733, Dec. 4, 1987; and as amended by the National Defense Authorization Act for fiscal year 1991—Public Law 101-510, Sec. 717, Nov. 5, 1990.

[2]Medical ADP Systems: Composite Health Care System Is Not Ready To Be Deployed (GAO/IMTEC-92-54, May 20, 1992).

[3]For purposes of this report, we limited our evaluation of CHCS performance management to those aspects of system performance relating specifically to the execution of instructions on the hardware portion of the system, excluding human response to the system.

performance and the methodology for evaluating the response-time aspect of CHCS performance.

Results in Brief

Because of deficiencies in its CHCS performance management tools and weaknesses in its CHCS performance management methodology, Defense is not managing CHCS performance as effectively or economically as is warranted for a state-of-the-art system.

The performance measurement tools Defense uses at its CHCS sites do not collect all the data Defense needs to detect response-time problems, diagnose their causes, and determine their significance. In addition, Defense lacks state-of-the-art analysis tools, which would help it determine the causes of performance problems and project the impact on response times of changes in workload and/or system configuration. As a result, Defense does not make optimal use of the limited data collected by its performance measurement tools.

In addition, Defense's methodology for managing CHCS performance is weak. The methodology does not require Defense's routine analysis and elimination of extremely long response times that occur sporadically. Rather, Defense must rely on user complaints to generate performance management concerns with such occurrences. As a result, these occurrences will persist and can discourage clinical users from using CHCS to its full potential. Also, Defense's method of providing for reserve CHCS capacity is unreliable and may result in either excessive reserve capacity, thereby incurring unnecessary cost, or in deficient reserve capacity, thereby leading to unsatisfactory system performance.

Background

CHCS is a comprehensive automated medical information system designed and developed to provide support to military medical treatment facilities worldwide. The system is composed of integrated modules that, activated together or independently, will support high-volume workloads and enhance communications within medical treatment facilities.

In November 1991, congressional conferees supported Defense's proposal to deploy CHCS in two phases. The first phase involves deploying and activating a common, integrated CHCS database to Defense medical treatment facilities worldwide. This database will support patient appointment scheduling, pharmacy, laboratory, radiology, patient administration, outpatient order entry by physicians, and

inpatient/outpatient medical test result reports. The second phase relates to activating CHCS physician inpatient order entry functions at specified Defense medical treatment facilities.

Defense tested CHCS between 1988 and 1991 and, in January 1992, published a formal evaluation of the results of its CHCS Operational Test and Evaluation. With congressional approval, in January 1993, Defense began the phase one deployment. This deployment authority does not include the activation of phase two CHCS functions.

Through fiscal year 1993, Defense had obligated over $700 million for CHCS development and initial operating costs at designated medical treatment facilities. Currently, Defense is obligating approximately $13 million per month for (1) continued CHCS development, (2) deployment of CHCS beyond the designated test sites, and (3) overall CHCS operations. CHCS is currently installed at test sites on Virtual Address Extension (VAX) mini-computer platforms[4] using Digital Equipment Corporation's (DEC) Virtual Memory System (VMS) operating system. The medical staff access the system from terminals located throughout the hospitals and clinics. Defense is in the process of installing CHCS on personal-computer platforms using a version of the Unix operating system.

Objective, Scope, and Methodology

The National Defense Authorization Act for fiscal year 1987, as amended, requires that GAO (1) monitor Defense's development, testing, evaluation, and deployment of CHCS and (2) periodically report to the Senate and House Committees on Armed Services prior to the worldwide deployment of CHCS.

Our objective was to determine the adequacy of Defense's management of CHCS performance. Of the different aspects of system performance, such as system response time and ease-of-use, we focused largely on the response-time aspect because 8 of the 12 CHCS test sites had reported response-time problems. We have another assignment currently in process that will address problems relating to the ease-of-use aspect of CHCS performance.

In conducting our review, we reviewed Defense's October 20, 1992, CHCS Performance Management Plan; reviewed a copy of Defense's October 22, 1992, CHCS Operational Test and Evaluation report to the Senate and House Armed Services Committees; reviewed Defense's

[4]A computer platform is the hardware configuration on which an operating system resides.

December 16, 1991, system sizing algorithms; evaluated monthly progress reports provided to Defense by the CHCS contractor through March 7, 1994; and modeled the CHCS systems at the Walter Reed Army Medical Center, Washington, D.C., and Naval Hospital Charleston, Charleston, South Carolina, using a commercially available analytic modeling tool called BEST/1-VMS from BGS Systems Inc., Waltham, Massachusetts.

Since our December 15, 1992 report,[5] we have conducted an analytical review of CHCS performance at 2 of the 12 operational CHCS test sites: the Naval Hospital Charleston and the Walter Reed Army Medical Center. Both serve as beta test sites—operational test sites to which system capabilities are deployed for the purpose of conducting operational test and evaluation. We also examined response-time measurements taken at the Walter Reed Army Medical Center between October 18, 1992, and November 10, 1992; measurements taken at Naval Hospital Charleston between October 12, 1992, and November 19, 1992; and measurements taken at the 12 primary CHCS test sites for the period June 27, 1993, through July 27, 1993. We did not verify the accuracy of the tool Defense used to make these measurements. Also, we did not evaluate the response-time goals set by Defense for these measurements. In addition, we met with officials of Science Applications International Corporation (SAIC), the prime contractor, and officials from the CHCS Program Office in Falls Church, Virginia.

We worked closely with senior program management officials at Defense to discuss our concerns as they arose and confirm our understanding of potential problems and their implications for the achievement of performance management objectives. We briefed senior program management officials at Defense during our review.

The Department of Defense provided written comments on a draft of this report. These comments have been incorporated into the report where appropriate and are included in appendix I.

Defense Does Not Have Adequate Tools for Managing CHCS Performance

Effective management of CHCS performance requires the use of two categories of tools—performance measurement and performance analysis tools. Performance measurement tools collect system performance and system utilization data, while performance analysis tools extract information from these data. The performance measurement tools Defense

[5]Composite Health Care System: Outpatient Capability Is Nearly Ready for Worldwide Deployment (GAO/IMTEC-93-11, Dec. 15, 1992).

uses at its CHCS test sites have serious limitations in the data they collect, thereby adversely affecting Defense's ability to manage response time. Also, Defense has not acquired state-of-the-art performance analysis tools in order to make optimal use of the limited data collected by the performance measurement tools.

Performance Measurement Tools Have Serious Limitations

Managing response time effectively requires three types of measurement data: (1) response-time measurements for individual user functions,[6] (2) measurements of system component use,[7] by user function, and (3) measurements of the frequency with which system users employ user functions. The tools used by Defense do not provide these measurement data. Furthermore, the operating system Defense is acquiring for use with CHCS on personal-computer platforms also lacks adequate performance-measurement tools.

Response-Time Data Are Severely Limited

Response-time measurements generally serve as indicators of how well a system is performing. They allow performance analysts to detect and correct developing response-time problems before they generate user complaints. Defense, however, does not have a tool that measures the system response times actually experienced by system users. While it has two DEC system monitors that measure system component utilization, they do not measure system response time for either the CHCS user functions or the transactions of which the functions are composed.[8] As a result, Defense has no comprehensive indicator of how well CHCS is performing.

Because of this deficiency, Defense has developed a tool, called the Performance Monitoring Tool (PMT), which is intended to simulate certain critical CHCS user activities and to capture the response times of these activities. PMT periodically and automatically submits simulated user activity to CHCS from a personal computer and measures system response time for this activity. However, the sample response-time measurements made by PMT are not fully representative of the response times actually experienced by system users. First, PMT submits to CHCS only 14 user

[6]User functions include such activities as entering laboratory orders, retrieving clinical results, entering nursing orders, filing nursing orders, and ordering prescriptions.

[7]System components include such things as a computer's processors, its input/output devices, its memory, and any networks connecting computer terminals to the computer system.

[8]A transaction begins when a user hits the enter key and ends when the user receives a response to what was entered. A CHCS user function, such as entering a laboratory order, is composed of one or more transactions, such as entering the patient's name and designating the lab work to be done.

functions—only a sample of the thousands of CHCS functions actually employed by system users. For instance, PMT submits no user function for entering or canceling radiology orders. Second, the way in which PMT periodically submits the 14 user functions does not necessarily reflect the frequency with which they are employed by actual users. For example, there is no reason to believe that clinical users employ the function "cancel laboratory order" with the same frequency that they employ the function "enter laboratory order." Yet, PMT executes both of these functions at the same frequency.

Data on System Component Use Do Not Relate to CHCS User Functions

The system response time for a user function is composed of (1) time spent waiting for service and (2) time spent receiving service at various system components, such as processors, disks, and networks. The time that a function spends receiving service at each component is fairly constant from one execution of that function to the next. Any exceptionally long response time is usually due to a function's having to wait for service at some unusually busy system component. To resolve exceptionally long response times, a performance analyst needs to determine which component is delaying the execution of a user function.

However, the measurement tools in place at the CHCS test sites do not enable performance analysts to do this quickly and easily. For example, while the system monitors indicate excessively high utilization of certain disks, they do not indicate which CHCS user functions are using these disks. Therefore, the performance analyst is unable to trace a response-time problem with a particular user function to the specific disk causing the delay, using the measurement data provided by the DEC system monitors alone. Currently, individuals with a high level of technical knowledge must engage in lengthy and labor-intensive efforts to trace such response-time problems.

The two DEC VMS system monitors mentioned above collect comprehensive information relating to system component utilization at a system-wide level and somewhat less comprehensive information relating to system component utilization at the process level.[9] Because processes do not correspond to user functions, these two monitors cannot show system component use by user function. As a result, the VMS monitors do not provide the data performance analysts need to determine which components are causing delays in the execution of user functions, if and when they occur.

[9]A process begins when a user signs on to the system and ends when that user signs off.

Defense has developed a tool, called the Enhanced Option Audit, which holds the promise of eventually relating system-component use to CHCS-function use. The tool is an enhancement of a software monitor that Defense originally developed for CHCS security purposes. Defense enhanced the monitor to show system component use by option (i.e., a menu item that a user selects, such as "Enter/Maintain Lab Orders" or "Lab Order Entry/Login"). However, since the tool measures component use at the option level rather than at the user-function[10] level and the correspondence between CHCS options and CHCS user functions is not one-to-one, Defense still cannot link system resource use directly with CHCS user functions. As a result, Defense cannot readily determine which system components are responsible for the delays.

CHCS Measurement Tools Provide No Information on the Frequency of Function Use

To determine the seriousness of a response-time problem, an analyst needs to know how often CHCS users experience the problem. Defense's PMT detects some response-time problems, but provides no measure of how often users experience the same problems. As a result, Defense performance analysts are unable to make reliable inferences from the sample response-time measurements as to what users actually experience.

Defense's Enhanced Option Audit collects data on how often users select each menu option; however, as pointed out above, the correspondence between CHCS menu options and CHCS user functions is not one-to-one. Therefore, like the VMS monitors and PMT, Defense's Enhanced Option Audit cannot provide performance analysts with measurement data on the frequency with which system users employ various CHCS user functions.

The Newly Acquired Operating System for Use With CHCS on Personal Computer Platforms Lacks Adequate Performance Measurement Tools

While CHCS is currently installed at test sites on VAX mini-computer platforms using DEC's VMS operating system, Defense is in the process of installing CHCS on personal-computer platforms using a Unix-based operating system. However, the performance measurement tools that come with Unix-based operating systems are generally inadequate for CHCS purposes.

For example, because they are designed for on-line analysis of performance problems, the performance measurement tools that come with Unix have very limited archival capabilities. As a result, these tools are not practical for investigating past performance problems. While

[10]A user function, such as "Enter a Laboratory Order" or "Retrieve a Single Laboratory Result," may require the use of more than one menu option, and several different user functions may execute under the same option.

commercial vendors have developed performance-measurement tools for some versions of Unix, they have not developed such tools for the version being acquired by Defense for CHCS. Currently, there are no adequate performance-measurement tools available for the version of Unix being acquired by Defense.

Defense's CHCS Performance Analysis Tools Are Not State-of-the-Art

In addition to the limitations of the above performance measurement tools, Defense has not acquired state-of-the-art performance analysis tools that would allow it to make optimal use of the limited data collected by the performance measurement tools. Such state-of-the-art tools incorporate sophisticated analysis techniques, such as analytic and simulation modeling, which would help Defense determine the causes of performance problems and project the impact of workload growth and system-configuration changes on response times.

State-of-the-Art Modeling Tools Would Help in Determining the Causes of Response-Time Problems

In determining the causes of CHCS performance problems, analysts need to interpret data collected by the system monitors. Without special tools, it is very difficult to interpret the voluminous monitor data collected. An analytic modeling tool developed specifically for the VMS operating system currently exists that could assist Defense performance analysts in the interpretation of performance data collected by the VMS system monitors. Defense performance analysts could use this tool to (1) extract needed data directly from files created by the system monitors, (2) aggregate these data to facilitate analysis, and (3) calculate response-time characteristics. There are also other commercially available analytic and simulation modeling tools that could enhance Defense's performance management capability. While Defense has been looking into the acquisition of such a tool since August 1992, it has not yet acquired one.

Our use of the VMS-specific analytic modeling tool showed that some excessively long response times occurring in the October-November 1992 time frame at the Walter Reed Army Medical Center were caused, in part, by greatly over-utilized input/output devices.[11] Specifically, the general rule of thumb for disks used in interactive processing is that they should not exceed 30 percent utilization over any extended period of time, such as a 15-minute interval. We found that five disks regularly exceeded this criterion—one showing in excess of 90-percent utilization over a 15-minute interval. Over-utilized disks can result in excessively long response times.

[11]Input/output devices are data storage devices such as tapes and disks.

State-of-the-Art Modeling Tools Would Help in Projecting the Impact on Response Times of Workload Growth and Configuration Changes

Analysts need to project the impact of workload growth and configuration changes on performance. Analytic and simulation modeling tools are particularly helpful in doing this. Using such tools, a performance analyst can change the workload or system configuration in the model and have the model calculate the impact of the change on system response times. Such modeling would have been helpful to Defense in its recent upgrades of computer processors.

Over the past 2 years, Defense has replaced old computer processors with new, faster processors at most of its CHCS test sites. While these upgrades have improved overall CHCS response time, Defense could have done a better job of estimating the number of new computer processors it needed by using analytic or simulation modeling to evaluate the proposed upgrades. Our analytic modeling of the upgrade at one test site illustrates this point.

In April 1993, Defense replaced 20 old processors at the Walter Reed Army Medical Center test site with 9 new, faster processors. We used an analytic modeling tool to predict the impact on CHCS response times of this processor upgrade, as well as several alternative processor upgrades Defense could have made. We found that Defense could have obtained essentially the same benefit from 6, rather than 9, new processors. The model showed that the 9 new processors would reduce CHCS response times by an average of 49.8 percent, while 6 new processors would reduce them by 49.1 percent. In terms of a CHCS function, this means that a function that would take 5 seconds of system response time with 9 processors would take 5.07 seconds with just 6 processors. The approximately 1.3 percent response-time improvement resulting from three extra processors came at a cost of $180,000. The inability to optimize CHCS configurations may become economically significant as Defense deploys CHCS worldwide.

Defense's Methodology for Managing CHCS Performance Is Weak

Defense sets forth a CHCS performance evaluation methodology in its CHCS Performance Management Plan. For each of the 14 CHCS user functions periodically and automatically submitted to CHCS by PMT, the plan lists a response-time goal, in seconds, and defines a performance index—the ratio of that function's measured 95th-percentile response time to its 95th-percentile response-time goal. The plan also defines, for each CHCS site, a composite performance index, which is the arithmetic average of the performance indices at that site for the 14 individual functions

submitted by PMT. The plan then sets an explicit objective of 0.8 for the composite performance index at each site.

The evaluation methodology prescribed in the plan is not appropriate for a state-of-the-art system. Specifically, the methodology (1) requires no routine analysis of exceptionally long response times that occur sporadically and (2) erroneously assumes that increases in system utilization ordinarily result in proportional increases in system response time.

Performance Evaluation Methodology Sets No Goal for Reducing the Severity or Frequency of Exceptionally Long Response Times

Defense's CHCS performance evaluation methodology, as presented in the plan discussed above, is based entirely upon the 95th-percentiles of response-time measurements made by PMT. The methodology has no provision for reducing the severity or frequency of exceptionally long response times that occur sporadically and fall well beyond the 95th percentile. This omission could cause Defense to rely on user complaints to generate performance management concern with such occurrences.

The concept of statistical process control, first set forth by Walter A. Shewhart,[12] is an integral component of quality control. Its fundamental thesis is that all processes generate small, unavoidable random variations in outcome, but that where variations are large and due to identifiable, controllable causes rather than uncontrollable random variation, these causes should be investigated, their origin determined, and their effect eliminated. Good performance management would focus on investigating and correcting such variations early, before they become a major irritant to system users.

As shown in table 1, PMT data from the Walter Reed Army Medical Center for November 1992 show that for 10 of the 14 user functions submitted by PMT, the maximum system response time measured was more than twice as long as the 95th-percentile response time. For example, the maximum system response time measured for retrieving a pharmacy order was 8 times the 95th-percentile, while the maximum for displaying a patient appointment was 9 times the 95th-percentile.

Despite the fact that Defense upgraded its processors at the Walter Reed Army Medical Center in April 1993, the situation with respect to exceptionally long response times was worse in November 1993 than it

[12]Walter A. Shewhart, Economic Control of Quality of Manufactured Product, New York: D. Van Nostrand, Inc., 1931.

had been in November 1992. For example, 13 of the 14 user functions in November 1993 show maximum response times more than twice the 95th-percentile response time, and the maximum system response time for retrieving a pharmacy order was 16 times the 95th-percentile. The causes of these exceptionally long response times that occur sporadically need to be identified and corrected before such long response times become a major irritant to users.

Table 1: Maximum and 95th-Percentile System Response Times Measured by PMT—Composite Health Care System at the Walter Reed Army Medical Center, Washington, D.C., November 1992 and November 1993

	Response time in seconds, November 1992[a]		Response time in seconds, November 1993[b]	
CHCS User Function	Maximum	95th Percentile	Maximum	95th Percentile
Display patient appointment	75.4	8.34	24.9	3.99
Cancel laboratory order	56.2	14.69	311.5	33.35
Cancel nursing order	43.9	14.39	190.8	27.05
Determine eligibility	36.8	24.61	18.8	4.93
File nursing order	18.1	4.15	61.9	6.11
CHCS log-on	90.2	28.12	120.2	24.53
Read mailman message	15.5	6.94	30.2	6.59
Send mailman message	24.9	13.03	66.7	13.28
Get next screen	11.4	2.87	3.8	2.51
Enter laboratory order	84.6	43.15	191.9	25.56
Enter nursing order	50.4	17.52	59.0	14.74
Retrieve laboratory results	19.2	6.13	20.3	6.04
Review clinical results	35.1	23.37	141.8	18.24
Retrieve pharmacy order	296.2	36.44	193.4	11.88
Total	**857.9**	**243.75**	**1,435.2**	**198.80**

[a]Source: SAIC's Progress Report for Reporting Period: November 7 to December 4, 1992, Volume III, page 5-73.

[b]Source: SAIC's Progress Report for Reporting Period: November 6 to December 3, 1993, Volume III, page 7-26.

Defense's CHCS Performance Evaluation Methodology Contains an Erroneous Assumption

In its CHCS Performance Management Plan, Defense adjusts its response-time objective for the composite performance index from 1.0 to 0.8 to provide what it calls a 20-percent reserve system capacity at each CHCS site. To understand this approach, one needs to understand that if each of the 14 simulated user functions was exactly meeting its response-time goal, then each would have a performance index of 1.0 and

the arithmetic average of the 14 indices—the composite performance index—would also be 1.0.

An unanticipated increase in workload (system utilization) could degrade performance to the extent that the composite performance index would exceed 1.0. Defense desires to have enough excess system capacity so that an unanticipated workload increase of up to 25 percent[13] will not degrade performance so much that it causes the composite performance index to exceed 1.0. The approach Defense uses to achieve this 20-percent reserve is to adjust its response-time objective for the composite performance index from 1.0 to 0.8.

Defense's approach, however, is invalid. Since increases in system utilization do not ordinarily result in proportional increases in system response time, meeting a response-time objective that is 20 percent below the response-time goal does not ensure system utilization that is 20 percent below system capacity. This fact was borne out by our analysis at the Walter Reed Army Medical Center.

Using an analysis technique known as analytic modeling, we projected the impact on response times and the composite performance index of changes in CHCS workload at the Walter Reed Army Medical Center. Our analysis indicated that for November 1993, a 100-percent increase in system utilization (from 31 to 62 percent) would have caused less than a 4-percent increase in the composite performance index, showing that changes in response times are not necessarily proportional to changes in system utilization.

While the provision of reserve capacity is a legitimate objective, the example above shows that the composite performance index does not provide a reliable indication of how much reserve capacity there is. Therefore, the composite performance index should not be used for this purpose. The provision of excessive reserve capacity incurs additional costs for little or no improvement in system performance, while the provision of too little reserve capacity can result in poor system performance in the event of an unanticipated increase in workload. A more reliable method of providing for reserve capacity is the use of analytic or simulation modeling, which correctly relates response times to system component utilizations through the use of advanced mathematical techniques, such as queuing theory.

[13]An increase in workload from 80 percent of system capacity to 100 percent of system capacity is a 25-percent increase in the workload.

Conclusions

CHCS is intended to be the backbone for Defense's worldwide medical operations, providing medical personnel with almost instant access to patient information, from medical history to current treatments or vital statistics. As such, the performance and response times of the system will be critical. Defense, however, is placing the performance of the system at risk because it has not provided adequate performance measurement and analysis tools or corrected weaknesses in its performance management methodology.

Without the appropriate performance measurement and analysis tools, Defense cannot diagnose the causes of response-time problems or project how workload growth and configuration upgrades will affect system response times. Moreover, Defense's decision not to focus performance management on routine analysis and elimination of extremely long response times can discourage clinical users from using CHCS. The lack of a reliable methodology to plan for reserve capacity has also led to the unnecessary purchase of excess computer processing power. Unless these problems are corrected, Defense risks continuing to replicate and proliferate system performance problems and adding to the costs of deploying CHCS.

Recommendations

To provide the performance management that a state-of-the-art system warrants, we recommend that the Secretary of Defense direct the Assistant Secretary of Defense for Health Affairs to

- obtain performance measurement tools for each computer operating system under which CHCS runs that can (1) measure system response times at the CHCS user-function level, (2) relate system component utilization with the specific CHCS user functions served, and (3) measure the frequency with which CHCS user functions are employed;
- obtain performance analysis tools for each computer operating system under which CHCS runs that will (1) enable Defense to determine the causes of response-time problems and (2) project the impact of workload growth and system-configuration changes on response times; and
- modify Defense's approach to managing CHCS performance to (1) include objectives for investigating and correcting extremely long response times and (2) provide reliable measures of system reserve capacity.

Agency Comments and Our Evaluation

In commenting on a draft of this report, the Department of Defense stated that it partially concurred with the report. Defense agreed that there

continues to be room for improving CHCS performance management. However, Defense said that we did not recognize some CHCS performance improvements.

While we may not have cited specific system performance improvements relating to CHCS, we recognize that response-time improvements are occurring at some CHCS test sites—as noted previously in table 1. The objective of this report was to determine the adequacy of Defense's management of CHCS performance and to identify areas for improvement, where appropriate.

Defense also claimed that it already has adequate performance measurement and analysis tools to diagnose and resolve system performance problems. We disagree. This report identifies serious limitations in the data collected by the performance measurement tools Defense uses at its CHCS test sites. Further, Defense has not acquired state-of-the-art performance analysis tools needed to make optimal use of the limited data collected by the performance measurement tools. Moreover, based on our review and discussions with Defense officials, we found that other performance monitoring tools used by Defense, such as the Digital Equipment Corporation's DECps performance tool and the Massachusetts General Hospital Utility Multi-Programming System's (MUMPS) tools RTHIST and GLSSTA, contain deficiencies (see comment 2 in appendix I).

Defense also takes the position that deployment of CHCS should proceed as rapidly as possible while incorporating the recommended performance management methodology. We agree. Our position is that to support such deployment, Defense must expeditiously obtain state-of-the-art tools and implement an effective performance management program. Managing system performance at more than 500 medical treatment facilities worldwide will be much more complex than managing performance at the CHCS test sites currently in operation. Also, while Defense's performance management weaknesses may be tolerable when running only CHCS' outpatient functions, these weaknesses, if left uncorrected, may render the system operationally risky and may result in significant cost increases when Defense activates CHCS inpatient functions.

Lastly, Defense concurred with two of our three recommendations. Defense partially concurred with our recommendation that it obtain performance measurement tools for each computer operating system under which CHCS runs that can (1) measure system response times for

each CHCS user function, (2) relate system component utilization with the specific CHCS user function served, and (3) measure the frequency with which each CHCS user function is employed. While Defense concurred that the recommended tools may be of use for a variety of specific purposes, it did not concur that the tools should be used routinely at all operational sites for continuous performance monitoring and measurement.

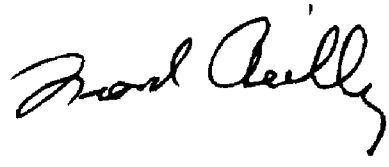
Defense contends that continuously capturing user function response times by function and by user would create a tremendous amount of data, greatly increase storage requirements, and have a negative impact on CHCS performance. We disagree. Typically, performance measurement tools record less than the total of measurements captured. For example, while a software monitor is capable of capturing or measuring the response time of every single transaction, it is usually set to record just the mean, standard deviation, and maximum values over a designated measurement interval, such as 15 minutes. Similarly, such a monitor can show aggregate component use and transaction volume (frequency) by user function over a 15-minute interval. Consequently, performance analysts can obtain useful performance information without making excessive demands on system storage.

We clarified part of that recommendation—that Defense obtain actual response-time and frequency measurements for each CHCS user function. It was not our intention that Defense record statistical representations of response-times, component use, and frequency for each of the more than 2,000 CHCS user functions. However, we believe it is critical that such representations be recorded, on a routine basis, for a representative number of key or critical CHCS user functions.

Detailed Department of Defense comments and our evaluation are contained in appendix I.

We are sending copies of this report to the Chairmen of the House and Senate Committees on Appropriations; the Secretary of Defense; and the Director, Office of Management and Budget. Copies also will be made available to other interested parties upon request.

We conducted our evaluation from August 1992 to December 1993, in accordance with generally accepted government auditing standards. This work was performed under the direction of William S. Franklin, Associate Director, who can be reached at (202) 512-6234. Other major contributors are listed in appendix II.



Frank W. Reilly
Director, Information Resources
Management/Health, Education, and
Human Services

Contents

Abbreviations

CHCS	Composite Health Care System
DEC	Digital Equipment Corporation
DECps	Digital Equipment Corporation Performance Solution
MAISRC	Major Automated Information Systems Review Council
MUMPS	Massachusetts General Hospital Utility Multi-Programming System
PC-CHCS	Personal Computer-Composite Health Care System
PCPS	Polycenter Performance Solution
PMT	performance monitoring tool
SAIC	Science Applications International Corporation
SPM	system performance monitor
VAX	Virtual Address Extension
VMS	Virtual Memory System

Comments From the Department of Defense

Note: GAO comments supplementing those in the report text appear at the end of this appendix.

HEALTH AFFAIRS

THE ASSISTANT SECRETARY OF DEFENSE
WASHINGTON, DC 20301-1200

MAY 19 1994

Mr. Gene L. Dodaro
Assistant Comptroller General
Accounting and Information Management Division
U. S. General Accounting Office
Washington, DC 20548

Dear Mr. Dodaro:

This is the Department of Defense (DoD) response to the General Accounting Office (GAO) draft report entitled -- "MEDICAL ADP SYSTEMS: Defense's Tools and Methodology for Managing Composite Health Care System Performance are Deficient," dated April 25, 1994 (GAO Code 510883), OSD Case 9672. The DoD partially concurs with the report.

The GAO report raises the question of deficiencies in the Composite Health Care System performance management tools. While the Department agrees that there continues to be room for improvement in the process of Composite Health Care System performance management, the GAO did not recognize several significant accomplishments. System performance improvements are evident in Performance Monitoring date evaluation, operational test and evaluation results, and user feedback.

The DOD agrees that, without adequate performance measurement and analysis tools, the Department cannot diagnose the causes of response time problems. Adequate tools, however, already exist to diagnose and resolve system performance problems--as evidenced by numerous instances in which existing tools have been used to identify, analyze, and resolve Composite Health Care System performance problems at operational sites. Although the GAO acknowledged the existence and use of two Virtual Address Extension (VAX)/Virtual Memory System performance monitoring tools, the GAO neglected to mention a number of other performance monitoring tools in use that contribute to the routing performance analysis of Composite Health Care System sites.

The DoD concurs that some performance concerns may be occasionally present at Composite Health Care System sites, due to a need for system tuning or increases in system workload. While DoD agrees that cost realism would be enhanced by the techniques recommended, the benefits of Composite Health Care System are so significant, the cost risk so low, and the time to adopt the new techniques so short, that deployment should proceed as rapidly as possible while incorporating the recommended methodology.

The detailed DoD comments on the report findings and recommendations are enclosed. The DoD appreciates the opportunity to comment on the draft report.

Sincerely,

Edward D. Martin/for
Stephen C. Joseph, M. D., M.P.H.

Enclosure:
As Stated

**GAO DRAFT REPORT — DATED APRIL 25, 1994
(GAO CODE 510883) OSD CASE 9672**

"MEDICAL ADP SYSTEMS: DEFENSE'S TOOLS AND METHODOLOGY FOR MANAGING COMPOSITE HEALTH CARE SYSTEM PERFORMANCE ARE DEFICIENT"

DEPARTMENT OF DEFENSE COMMENTS

* * * * *

FINDINGS

FINDING A: The Composite Health Care System Automated Medical Information System. The GAO reported that the Composite Health Care System is a comprehensive automated medical information system designed and developed to provide support to military medical treatment facilities worldwide. The GAO noted that the system comprises integrated modules that, when activated together or independently, will support high volume workloads and enhance communications within medical treatment facilities.

The GAO further reported that, in November 1991, the congressional conferees supported the DoD proposal to deploy the Composite Health Care System in two phases. The GAO noted that the first phase involved deploying and activating a common, integrated Composite Health Care System database to DoD medical treatment facilities worldwide. The GAO further noted that the database was designed to support: (1) patient appointment scheduling, (2) pharmacy, (3) laboratory, (4) radiology, (5) patient administration, (6) outpatient order entry by physicians, and (7) inpatient/outpatient medical test result reporting. The GAO observed that the second phase was related to activating the Composite Health Care System physician inpatient order entry functions at specified DoD medical treatment facilities. The GAO noted that, with congressional approval to begin phase one only, the deployment began in January 1993.

The GAO found that, through FY 1993, the DoD obligated over $700 million for Composite Health Care System development—and, currently, was obligating approximately $13 million per month, primarily for continued Composite Health Care System development and operations at its test sites. The GAO also observed that the Composite Health Care System was currently installed at test sites on Virtual Address Extension (VAX) minicomputer platforms using the Digital Equipment Corporation Virtual Memory System operating system. The GAO noted that medical staff accesses the system from terminals located throughout the hospitals and clinics. The GAO further noted that the DoD was in the process of installing the Composite Health Care System on personal computer platforms—using a version of the UNIX operating system. (pp. 3–5/GAO Draft Report)

See comment 1.

DOD RESPONSE: Partially concur. Although the DoD agrees with most of the Finding, the GAO states that the Department has obligated $700 million through FY 1993 for *development* and is obligating $13 million per month, primarily for continued development and operations at

its test sites. The $700 million figure reflects all costs through FY 1993 in the Composite Health Care System, not just development cost. It is unclear how the GAO derived the $13 million per month figure. The Composite Health Care System Program has projected $163.4 million for investment and $58.0 million for Operation and Support for FY 1994.

FINDING B: The DoD Does Not Have Adequate Tools for Managing the Composite Health Care System Performance. The GAO asserted that effective management of Composite Health Care System performance requires the use of two categories of tools—(1) performance measurement and (2) performance analysis tools. The GAO explained that performance measurement tools collect system performance and system utilization data, while performance analysis tools extract information from those data. The GAO found that the performance measurement tools the DoD used at its Composite Health Care System test sites have serious limitations in the data collection and, therefore, adversely affect the ability of the DoD to manage response time. The GAO also found that the DoD had not acquired state-of-the-art performance analysis tools to make optimum use of the limited numbers of data collected by the performance measurement tools. (p. 7/GAO Draft Report)

See comment 2.

DOD RESPONSE: Partially concur. Although the DoD concurs with the GAO statements regarding the need for adequate tools, the DoD disagrees that the performance measurement tools used by the Department have *serious* limitations, thereby adversely affecting the ability to manage response time. Further, while the DoD agrees that additional, proven performance monitoring, measurement, and analysis tools potentially could be of benefit to the Composite Health Care System, the GAO did not assess the full spectrum of tools currently used. For example, while the GAO has been exposed to the Performance Monitoring Tool, Option Auditing, and the System Performance Monitor, the GAO did not specifically evaluate those tools in any level of detail, nor did the GAO investigate or evaluate other tools in use by the DoD, including the Virtual Address Extension Performance Advisor or any of the Massachusetts General Hospital Utility Multi-Programming System (MUMPS) tools including the RTHIST and GLSSTA routines used to generate real time histograms and global status reports. In addition, the GAO has not evaluated the pending use of the Digital Equipment Corporation DECps performance tool (which includes a collector, analyzer, and performance advisor and planner) nor the use of new Massachusetts General Hospital Utility Multi-Programming System tools such as ANASYS, an analysis routine. The combination of all those tools—along with the knowledge of the Virtual Address Extension/Virtual Memory System system management and tuning, as well as Digital Standard Massachusetts General Hospital Utility Multi-Programming System and the Composite Health Care System system architecture—does provide The DoD with adequate performance monitoring, management, and analysis capabilities.

See comment 3.

The DoD does not agree that actual response time measurements for all individual user functions are required for managing system response time effectively. While that information could be of benefit in assessing new functionality or in evaluating some specific instances of performance problems, routine collection of actual response time measurements for all Composite Health Care System user functions would not yield significant benefit to Composite Health Care System performance management. The cost to acquire and maintain those tools

may not justify the benefit. In addition, it is possible that collection and storage of such a large volume of information could, in itself, result in a significant, negative impact on system performance and available data storage capacity.

The DoD concurs with items (2) and (3) on page 8 of the draft report regarding data on the system component (system resource) usage by user function (as defined in footnote 5) and with the benefits of capturing data on the frequency with which users employ each function. The benefits derived from the use of those tools may vary, however, between the Virtual Address Extension and Personal Computer–Composite Health Care System configurations.

The DoD also concurs that adequate performance monitoring, measurement, and analysis tools are required for the Personal Computer–Composite Health Care System platform. It should be noted, however, that due to the low cost of the Personal Computer–Composite Health Care System hardware, the acquisition and use of various tools to prevent system oversizing may result in the situation that the cost of tool acquisition, use, and maintenance may exceed the cost of any "excess" system capacity that may be purchased.

Any effort to acquire new tools and to integrate those tools into the Composite Health Care System architecture and the Composite Health Care System performance management program must be assessed from a cost/benefit perspective to ensure that the anticipated benefit from use of any tool justifies the cost.

FINDING C: Performance Measurement Tools—Response Time Data are Severely Limited. The GAO reported that response time measurements generally serve as indicators of how well a system is performing and allow performance analysts to detect and correct developing response time problems before they generate user complaints. The GAO found that the DoD does not have a tool that measures the system response times actually experienced by system users. The GAO concluded that, while the DoD had two Digital Equipment Corporation system monitors measuring system component utilization, they do not measure system response time for either the Composite Health Care System user functions or the transactions of which the functions are composed. The GAO further concluded that, as a result, the DoD had no comprehensive indicator of how well the Composite Health Care System was performing. The GAO observed that, because of the cited deficiency, the DoD developed the Performance Monitoring Tool. The GAO observed that the Performance Monitoring Tool periodically and automatically submits simulated user activity to the Composite Health Care System from a personal computer and measures system response time for the activity. The GAO concluded, however, that the sample response time measurements made by the Performance Monitoring Tool are not fully representative of the response times actually experienced by system users. The GAO found that (1) the Performance Monitoring Tool submits to the Composite Health Care System only 14 user functions—which is only a sample of the thousands of functions actually employed by system users—and (2) the way in which the Performance Monitoring Tool periodically submits the 14 user functions does not necessarily reflect the frequency with which they are employed by actual users. (pp. 8–9/GAO Draft Report)

See comment 4.

DOD RESPONSE: Partially concur. The DoD agrees that it does not have a tool that measures the entire spectrum of response times actually experienced by system users. However, the DoD does not concur that detailed response time measurements are required for each system function in order to manage system performance effectively, or with the GAO conclusion that the DoD has no comprehensive indicator of how well the Composite Health Care System is performing.

The GAO report is incorrect in its statement that the DoD developed the Performance Monitoring Tool to provide analysts with some indication of Composite Health Care System response times. While the Performance Monitoring Tool does emulate a Composite Health Care System user through terminal emulation and the use of scripts that simulate a user, and while the Performance Monitoring Tool uses the Composite Health Care System functionality in essentially the same way as a Composite Health Care System user, the Performance Monitoring Tool was not designed to determine actual user response times. The Performance Monitoring Tool cannot, nor was it intended to, execute every user menu option and every user function, and accommodate all variations of the use of the Composite Health Care System. It was instead designed to focus on critical functional components of the Composite Health Care System that were deemed to be the best early warning indicators of system performance problems. (That is evidenced by the fact that 45 specific user functions account for 75 percent of the Composite Health Care System workload during peak periods). In essence, the Performance Monitoring Tool was designed as a performance "monitor," designed to gather relative performance data that are compared to goals and thresholds for each script. As a result, the DoD has never considered Performance Monitoring Tool data to represent the actual response times experienced by a Composite Health Care System user. Instead, Performance Monitoring Tool is a tool that executes 14 specific application "benchmark scripts" each hour and records the results of those "benchmarks." As such, the Performance Monitoring Tool competes with all other user processes for the required system resources (central processing unit, disk input/ output, memory, network services, etc.), and any limitation of those resources will impact the Performance Monitoring Tool response time in essentially the same manner as resource contention will impact a typical user process. While the Performance Monitoring Tool does not simulate every user function in every possible variation, the Performance Monitoring Tool does offer relevant data and is useful as a performance problem "sensor" in identifying potential performance problems. As a result, the data are useful in comparison with established goals and thresholds based on user satisfaction surveys and in trend comparisons for monitoring changes in system performance, due to changes in workload processed through the Composite Health Care System, changes in the way the system is used at a site, or changes in the Composite Health Care System hardware, software, or system tuning.

The DoD concurs that there are a number of enhancements that can be made to Performance Monitoring Tool to improve the quality and breadth of data collected. Those enhancements potentially include incorporating additional scripts to cover more Composite Health Care System functionality, "weighting" scripts or altering relative script frequency to represent better the actual function and transaction mix at a site, increasing the frequency with which scripts run, enhancing Performance Monitoring Tool reporting to focus on the statistical transforms applied to the data (such as distribution analysis and evaluation of very high (Maximum) values), and

continuing ongoing efforts to validate and recalibrate Performance Monitoring Tool goals and thresholds, as required. The DoD does not concur, however, that the Performance Monitoring Tool should be enhanced to capture or represent actual user response times or that the Performance Monitoring Tool should measure 100 percent of the more than 2000 user functions in the Composite Health Care System. Such an expansion likely would flood performance management analysts with excessive, potentially useless data and likely would impact performance of the system under test.

FINDING D: Performance Measurement Tools—Data on System Component Use Do Not Relate to the Composite Health Care System User Functions. The GAO observed that system response time for a user function is composed of (1) time spent waiting for service and (2) time spent receiving service at various system components, such as processors, disks, and networks. The GAO noted that the time that a function spends receiving service at each component is fairly constant from one execution of that function to the next. The GAO further noted that any exceptionally long response time was usually due to a function having to wait for service at some unusually busy system component. The GAO asserted that to resolve exceptionally long response times, a performance analyst needs to determine which component is delaying the execution of a user function. The GAO concluded that the measurement tools in place at the Composite Health Care System test sites do not enable performance analysts to perform that function. The GAO cited the example that, although the system monitors indicate excessively high utilization of certain disks, they do not indicate which Composite Health Care System user functions are using those disks. The GAO concluded, therefore, that the performance analyst is unable to trace a response time problem with a particular user function to the specific disk causing the delay.

The GAO found that the DoD developed a tool, called the Enhanced Option Audit, that holds the promise eventually of relating system component use to the Composite Health Care System function use. The GAO explained that the tool is an enhancement of a software monitor that the DoD originally developed for Composite Health Care System security purposes. The GAO noted that the DoD enhanced the monitor to show system component use by option (i.e., a menu item that a user selects, such as "Enter/Maintain Lab Orders" or "Lab Order Entry/Login"). The GAO concluded, however, that since the tool measures component use at the option level, rather than at the user functions level, and that the correspondence between the Composite Health Care System options and the Composite Health Care System user functions is not one–to-one, the DoD still cannot link system resource use directly with the Composite Health Care System user functions, and as a result cannot determine readily which system components are responsible for the delays. (pp. 10–11/GAO Draft Report)

See comment 5.

DOD RESPONSE: Partially concur. The DoD concurs with the high–level GAO description of system response time and with the GAO statement that in order to resolve long response times, a performance analyst needs to determine which component (i.e., system resource) is delaying execution of a user function. However, the DoD does not concur that the tools necessary to determine which component is delaying a user function (i.e., to identify system resource bottlenecks) are not in place at Composite Health Care System test sites. The GAO is

incorrect in the assertion that the DoD would not know, for example, which user functions would be impacted by an excessively high utilization of a specific disk.

In that particular example, based on the design of the Composite Health Care System, excessively high utilization of specific disk spindles in the Composite Health Care System correlate to specific Composite Health Care System volume sets and specific Massachusetts General Hospital Utility Multi–Programming System globals, which does provide insight into which user processes would be affected. Thus, the tools available, combined with understanding the Composite Health Care System architecture (central processing unit, disks, network, operating system, layered software, Massachusetts General Hospital Utility Multi–Programming System, application software, etc.) do provide adequate insight into Composite Health Care System performance to identify, diagnose, and resolve Composite Health Care System performance problems and to identify system resource bottlenecks.

The DoD concurs that additional tools and capabilities may be useful in reducing the time required to conduct performance analysis or to reduce the level of technical knowledge required to diagnose and resolve problems fully. The DoD also concurs that the availability of tools that provide system resource utilization statistics (i.e., component use) by Composite Health Care System function or by transaction would be useful in identifying areas of the Composite Health Care System application that require performance enhancement, and with the potential for enhancement of Option Auditing tools to relate system resource (i.e., component) usage to Composite Health Care System function usage. While the DoD does not have the tools to quantify system resource utilization by Composite Health Care System function, that ability is not essential to the identification of system resource bottlenecks and the appropriate resolution of Composite Health Care System performance problems through system tuning, configuration changes, or hardware upgrades as appropriate.

FINDING E: The Measurement Tools Provide No Information on the Frequency of Function Use. The GAO asserted that to determine the seriousness of a response time problem, an analyst needs to know how often Composite Health Care System users experience the problem. The GAO found that the DoD Performance Monitoring Tool detects some response time problems, but provides no measure of how often users experience the same problems. The GAO concluded that, as a result, DoD performance analysts are unable to make reliable inferences from the sample response time measurements about what users actually experience.

The GAO reported that, although the DoD Enhanced Option Audit collects data on how often users select each menu option, the correspondence between the Composite Health Care System menu options and the Composite Health Care System user functions is not one–to-one. The GAO concluded, therefore, that as with the Virtual Memory System monitors and the Performance Monitoring Tool, the DoD Enhanced Option Audit also cannot provide performance analysts with measurement data on the frequency with which system users employ the various Composite Health Care System user functions. (p. 12/GAO Draft Report)

DOD RESPONSE: Concur. The DoD agrees that understanding the frequency of problem occurrence, frequency of maximum response time measurements, as well as the distribution of longer response time measurements, would be valuable in the analysis of Composite Health Care System performance problems. The DoD will investigate enhancements of Performance Monitoring Tool and Option Auditing tools to provide additional information on the frequency of Composite Health Care System function usage. (See the DoD response to Recommendation 1.)

FINDING F: The Newly Acquired Operating System for Use With the Composite Health Care System on Personal Computer Platforms Lacks Adequate Performance Measurement Tools. The GAO reported that, although the Composite Health Care System is currently installed at test sites on Virtual Address Extension minicomputer platforms using the Digital Equipment Corporation Virtual Memory System operating system, the DoD is in the process of installing the Composite Health Care System on personal computer platforms using a UNIX operating system. The GAO concluded, however, that the performance measurement tools that come with UNIX operating systems are generally inadequate for Composite Health Care System purposes. The GAO cited the example that, because UNIX operating systems are designed for online analysis of performance problems, they have very limited archival capabilities. The GAO further concluded that, as a result, the tools are not practical for investigating past performance problems. The GAO observed that, while commercial vendors have developed performance measurement tools for some versions of UNIX, they have not developed such tools for the version being acquired by the DoD for the Composite Health Care System. The GAO asserted that, currently, there are no adequate performance measurement tools available for the version of UNIX being acquired by the DoD. (pp. 12–13/GAO Draft Report)

See comment 6.

DOD RESPONSE: Partially concur. The DoD concurs that performance monitoring, measurement, and analysis tools are desirable for the Personal Computer/UNIX Composite Health Care System platform. Accordingly, research is underway to evaluate and acquire an appropriate suite of performance measurement products for the system configuration. It should be noted that the Performance Monitoring Tool and Option Auditing will continue to be used with the Personal Computer–Composite Health Care System platform.

Based on discussions with the UNIX vendor as well as with third–party UNIX performance tool vendors, the DoD expects that adequate commercial measurement tools are and will be available for use with the Personal Computer–Composite Health Care System. As a result, pending further investigation, the DoD does not concur that there are no adequate performance measurement tools available for the version of UNIX being acquired by the DoD.

In addition to the evaluation of commercial tools from the UNIX vendor, as well as from third–party vendors, the DoD will support a full range of Composite Health Care System performance tools for the Personal Computer–Composite Health Care System, including Performance Monitoring Tool, Option Auditing, plus Massachusetts General Hospital Utility Multi–Programming System tools that are included with the Micronetics Standard Massachusetts General Hospital Utility Multi–Programming System language used with the Personal

Computer–Composite Health Care System. As a result, Performance Monitoring Tool and an enhancement of the Composite Health Care System Option Auditing tool will be available to support performance management of the Personal Computer–Composite Health Care System platform.

FINDING G: Performance Analysis Tools—State-of-the-Art Modeling Tools Would Help in Determining the Causes of Response Time Problems. The GAO asserted that in determining the causes of Composite Health Care System performance problems, analysts need to interpret data collected by the system monitors. The GAO further asserted that without special tools it is very difficult to interpret the numerous monitor data collected. The GAO noted that an analytic modeling tool developed specifically for the Virtual Memory System operating system currently exists—one that could assist the DoD performance analysts in the interpretation of performance data collected by the Virtual Memory System system monitors. The GAO explained that DoD performance analysts could use the tool to (1) extract needed data directly from files created by the system monitors, (2) aggregate the data to facilitate analysis, and (3) calculate response time characteristics. The GAO pointed out that, while the DoD had been looking into the acquisition of such a tool since August 1992, it had not yet acquired one. The GAO noted that its use of the Virtual Memory System–specific analytic modeling tool showed that some excessively long response times occurring in October–November 1992 at the Walter Reed Army Medical Center were caused, in part, by greatly overutilized input/output devices (disks). (pp. 14–15/GAO Draft Report)

DOD RESPONSE: Concur. The DoD agrees that enhanced analytic or simulation modeling tools would enhance Composite Health Care System performance analysis. In that regard, the DoD has tasked and funded an effort to research, select, acquire, and begin use of appropriate modeling tools for the Composite Health Care System for the Virtual Address Extension and Personal Computer–Composite Health Care System architectures, as well as for large regional configurations supporting overlapping catchment areas.

As a result of that effort, the Composite Health Care System contractor has recommended selection of the SES Workbench simulation modeling tool. Since that recommendation was made, the contractor has initiated construction and validation of Virtual Address Extension, Personal Computer, and regional Overlapping Catchment Areas models, and has initiated planning for the insertion of those modeling tools into the Composite Health Care System product development and capacity planning processes. The Virtual Address Extension model was validated in March 1994, and work has begun on the Personal Computer and Overlapping Catchment Areas models, with completion (validation) expected in September 1994 and January 1995, respectively.

Although the DoD has not yet procured the commercial tools, the Composite Health Care System contractor has proceeded under a loan of product for tool licenses in order to expedite model construction and validation, in parallel with the required contract modifications to allow the DoD to procure the tools through competitive acquisition. The contract modification

package to support procurement of the SES Workbench tools and supporting hardware is due in May 1994.

It must be emphasized, however, that the DoD has concerns regarding the routine use of those tools to resolve performance issues at operational sites. While the tools may be useful in certain circumstances, the routine, exclusive use of the referenced analytic and simulation modeling tools to resolve all performance problems potentially could delay problem identification and resolution, and could result in increased cost for performance analysis. For example, the GAO use of the Best/1 analytic modeling tool required an estimated 6 months of analysis from receipt of System Performance Monitor measurement data for 2 sites until the Best/1 models were created and results produced. That length of time for routine performance analysis is unacceptable in an environment with more than 120 operational host facilities and where medical facilities are dependent upon acceptable performance levels in the Composite Health Care System. As a result, incorporation of the tools into the Composite Health Care System performance management program will be a means to assess the appropriate use of the tools.

In addition, the selected modeling tool (SES Workbench) and the simulation models under development will address many problems that cannot be addressed by the more broadly targeted, hardware platform dependent, analytic modeling tool used by the GAO. Since the SES Workbench simulation modeling tool supports model construction that includes all components of the Composite Health Care System architecture, including the Composite Health Care System application and actual Composite Health Care System workload, the tool offers greater utility and flexibility in modeling all Composite Health Care System hardware platforms, changes in the Composite Health Care System functionality, as well as modeling changes in actual functional workload.

FINDING H: Performance Analysis Tools—State-of-the Art Modeling Tools Would Help in Projecting the Impact on Response Times of Workload Growth and Configuration Changes. The GAO asserted that analysts need to project the impact of workload growth and configuration changes on performance. The GAO noted that the analytic and simulation modeling tools are particularly helpful in doing that. The GAO explained that by using such tools, the performance analyst can change the workload or system configuration in the model and have the model calculate the impact of the change on system response times.

The GAO observed that, over the past 2 years, the DoD had replaced old computer processors with new, faster processors at most of its Composite Health Care System test sites. The GAO concluded that, although the upgrades had improved the overall Composite Health Care System response time, the DoD could have done a better job of estimating the number of new computer processors it needed by using analytic or simulation modeling to evaluate the proposed upgrades. The GAO noted that, in April 1993, the DoD replaced 20 old processors at the Walter Reed Army Medical Center test site with nine new, faster processors. The GAO used an analytic modeling tool to calculate the impact of the processor upgrade, as well as several alternative processor upgrades that the DoD could have made, on Composite Health Care System response times. The GAO concluded that the DoD could have obtained essentially the same benefit from

6, rather than 9, new processors. The GAO noted that the approximately 1.3 percent response time improvement resulting from 3 extra processors came at a cost of $180,000. (pp. 15–16/ GAO Draft Report)

See comment 7.

DOD RESPONSE: Partially concur. The DoD concurs that the use of analytic and/or simulation modeling tools could assist in projecting performance change due to change in system configuration changes, tuning changes, or changes in system workload, as well as in other aspects of the Composite Health Care System performance management program.

The DoD does not concur with the GAO analysis presented for the Walter Reed Army Medical Center and with the GAO conclusions regarding the Best/1 analysis of System Performance Monitor data provided to the GAO for the period October 18–November 10, 1992. In addition, the DoD does not concur with the GAO conclusion that the DoD inappropriately sized Walter Reed Army Medical Center and that the DoD is unable to optimize Composite Health Care System configurations. The following specific comments are offered on each of those points:

- The Composite Health Care System Sizing Algorithm sized the Walter Reed Army Medical Center for a configuration of three Virtual Address Extension 6630s in the standard N+1 configuration. That equates to a total of nine central processing units. The N+1 sizing provides the third Virtual Address Extension 6630 system to support system fail over. Thus, two Virtual Address Extension 6630s are predicted by the algorithm for the Walter Reed Army Medical Center system workload (six central processing units) and the third system is added by the N+1 configuration rule (for a total of nine central processing units). It appears, therefore, that the GAO did not account for the N+1 sizing in the Best/1 model. The three additional central processing units were not purchased to provide additional computing power or with the anticipation of an improvement in overall system response time, but were, instead, added to provide system fail-over capability and to reduce the performance impact during a system failure.

- As a normal practice in sizing new sites or in performing major system upgrades, the DoD includes projected system workload for all satellite facilities and also includes a 20 percent increase in system central processing unit capacity (as predicted by the Composite Health Care System sizing algorithm) to accommodate future growth in site workload, peak periods of performance, and changes in system capacity requirements, due to the release of new versions of Composite Health Care System software. In addition, when the DoD sized the Walter Reed Army Medical Center at three Virtual Address Extension 6630s, an allowance was made for the planned addition of users not on the system during the data collection for the GAO analysis in October–November 1992. That may account for some difference between the GAO analysis of the Walter Reed Army Medical Center using Best/1 and the system upgrade that was installed in April 1993.

- While the Best/1 tool may be very accurate in its prediction of system response time improvements, the accuracy of the analytic model has not been determined in a Virtual Address Extension/Virtual Memory System/Digital Standard Massachusetts General Hospital Utility Multi-Programming System/Composite Health Care System environment.

The GAO Best/1 model was completed based on approximately one month of data from 1992, on different hardware and with different versions of Composite Health Care System software than currently are installed. In addition, no analysis was conducted after the Walter Reed Army Medical Center upgrade to the current configuration. Therefore, the GAO Best/1 model requires validation prior to drawing any meaningful conclusions with respect to the accuracy of the response time predictions. Given the lack of model validation and lack of analysis of post–upgrade data, and due to the fact that the GAO only discussed central processing unit capacity, it appears that there are insufficient data to conclude that unnecessary central processing unit capacity was installed at the Walter Reed Army Medical Center.

- Since the DoD has not been provided data or copies of the final modeling and analysis conducted by the GAO, it is not possible to provide additional views concerning the conclusions reached by the GAO. While the DoD did participate in much of the initial modeling effort (data collection, model installation, etc.), the DoD has not been provided any final model analysis, reports, data, or other products resulting from that effort for the Walter Reed Army Medical Center or the Charleston models.

FINDING I: The DoD Methodology For Managing the Composite Health Care System Sets No Goal for Reducing the Severity or Frequency of Exceptionally Long Response Times. The GAO reported that the DoD Composite Health Care System performance evaluation methodology is based entirely upon the 95th centiles of response time measurements made by the Performance Monitoring Tool. The GAO asserted that the DoD methodology had no provision for reducing the severity or frequency of exceptionally long response times that occur sporadically and that fall well beyond the 95th centile. The GAO concluded that omission caused the DoD to rely on user complaints to generate performance management concern with such occurrences.

The GAO observed that the Performance Monitoring Tool data from the Walter Reed Army Medical Center for November 1992 showed that for 10 of the 14 user functions submitted by the Performance Monitoring Tool, the maximum system response time measured was more than twice as long as the 95th centile response time. The GAO asserted that, despite the fact the DoD upgraded its processors at the Walter Reed Army Medical Center in April 1993, the situation with respect to exceptionally long response times was worse in November 1993 than it had been in November 1992. The GAO cited the example that 13 of the 14 user functions in November 1993 showed maximum response times more than twice the 95th centile response time—and the maximum system response time for retrieving a pharmacy order was 16 times the 95th centile. The GAO concluded the causes of the exceptionally long response times that occur sporadically needed to be identified and corrected before such long response times become a major irritant to users. (pp. 17–20/GAO Draft Report)

See comment 8.

DOD RESPONSE: Partially concur. The DoD concurs that effort is required to address and to focus on exceptionally long response times as reported by Performance Monitoring Tool, and action will be taken to address those occurrences. (See the DoD response to Recommendation 3.) However, the DoD does not concur with GAO statements regarding the use of

Performance Monitoring Tool 95th centile as the entire basis of the Composite Health Care System performance evaluation methodology. It is important to note that while the DoD concurs that the high Performance Monitoring Tool values should be investigated, the DoD does not agree that ***all*** occurrences of performance measures that exceed performance goals can be eliminated. It is neither feasible, nor cost effective, to size automated systems to provide levels of system performance that could never be impacted regardless of any load imposed by their user communities.

The Composite Health Care System Performance Evaluation Methodology is not based entirely upon the 95th centiles of response time measurements by the Performance Monitoring Tool. The Performance Monitoring Tool reports 95th centile, mean, and maximum Performance Monitoring Tool readings by script, as well as an overall index for both mean and 95th centile measurements. The performance management methodology does not prevent, nor does it preclude, investigation of high values or investigation of Performance Monitoring Tool data other than 95th centile readings. The DoD concurs, however, that the description of the performance management procedures, as contained within the Performance Management Plan, require enhancement and clarification. That will be accomplished in Version 5.0 of the Composite Health Care System Performance Management Plan, which is currently under development and for which a June 1994 publication is expected.

In addition, the Composite Health Care System Performance Evaluation Methodology also includes review of Virtual Memory System Monitor and System Performance Monitor data for each site, as well as assessment of performance problems reported by Composite Health Care System users and onsite operations and system management personnel. As a result, the GAO is incorrect that the methodology is based entirely upon Performance Monitoring Tool 95th centile results.

FINDING J: The DoD Composite Health Care System Performance Evaluation Methodology Contains An Erroneous Assumption. The GAO reported that, in its Composite Health Care System Performance Management Plan, the DoD adjusts its response time objective for the composite performance index from 1.0 to 0.8 to provide what it calls a 20 percent reserve system capacity at each Composite Health Care System site. The GAO noted that the DoD objective is to have enough excess system capacity so an unanticipated workload increase of up to 25 percent would not degrade performance so much that it caused the composite performance index to exceed 1.0. The GAO explained that the DoD achieves the 20 percent reserve by adjusting its response time objective for the composite performance index from 1.0 to 0.8.

The GAO concluded that the DoD approach is invalid. The GAO further concluded that, since increases in system utilization do not ordinarily result in proportional increases in system response time, meeting a response time objective that is 20 percent below the response time goal does not ensure system utilization that is 20 percent below system capacity. In addition, the GAO concluded that this fact was borne out by its analysis at the Walter Reed Army Medical Center. Using analytic modeling, the GAO projected the impact on response times and the composite performance index of changes in the Composite Health Care System workload at the

Walter Reed Army Medical Center. The GAO determined that, for November 1993, a 100 percent increase in system utilization (from 31 to 62 percent) would have caused a less than a 4 percent increase in the composite performance index—showing that changes in response times are not necessarily proportional to changes in system utilization.

The GAO asserted that, while the provision of reserve capacity is a legitimate objective, it can be seen from the above example that the composite performance index does not provide a reliable indication of how much reserve capacity there is. The GAO further asserted that the composite performance index should not be used for such purpose. The GAO concluded that the provision of excessive reserve capacity incurs additional costs for little or no improvement in system performance, while the provision of too little reserve capacity can result in poor system performance in the event of an unanticipated increase in workload. The GAO further concluded that a more reliable method of providing for reserve capacity was the use of analytic or simulation modeling, which correctly relates response times to system component utilization through the use of advanced mathematical techniques, such as queuing theory. (pp. 20–22/GAO Draft Report)

DOD RESPONSE: Concur. The DoD concurs that increases in system utilization do not result ordinarily in proportional increases in system response time. The most recent versions of the Composite Health Care System Performance Management Plan contain an error that equates a programmatic goal of a Performance Monitoring Tool Composite Performance Index (as an expression of the mean of Performance Monitoring Tool scores divided by their goals) of 0.8 to having a 20 percent system reserve capacity. That is an error in the language of the plan that does not reflect the intent of the Department. The error will be corrected in the next version of the Performance Management Plan.

While the DoD did adopt a programmatic goal of an overall Composite Performance Index of 0.8 (which equates to each Performance Monitoring Tool script running an average of 20 percent below the established goal), the DoD recognizes that the goal does not equal a 20 percent reserve capacity, but is a mere expression of a desire to have all scripts perform well below the established goals. In addition, the DoD also desires to have some amount of excess system capacity at Composite Health Care System sites to accommodate increased demands associated with peak workload periods, growth in site workload, additional capacity demands of new software releases, and the addition of new system users or facilities. The DoD plans to use the simulation modeling tools and models currently under development and validation to pursue and validate further the actual residual capacity present at the Composite Health Care System sites, and to use the simulation modeling tools and models to estimate and provide sufficient residual capacity to accommodate peak workloads and system capacity requirement increases due to new functionality.

* * * * *

See comment 9.

RECOMMENDATIONS

RECOMMENDATION 1: The GAO recommended that the Secretary of Defense direct the Assistant Secretary of Defense (Health Affairs) to obtain performance measurement tools for each computer operating system under which the Composite Health Care System runs that can (1) measure system response times for each Composite Health Care System user function, (2) relate system component utilization with the specific Composite Health Care System user functions served, and (3) measure the frequency with which each Composite Health Care System user function is employed. (pp. 23–24/GAO Draft Report)

DOD RESPONSE: **Partially concur.** The recommended tools may be of use for a variety of specific purposes within the Composite Health Care System to support model and algorithm development and validation and in the resolution of certain system performance issues, particularly in the identification and resolution of function–specific performance problems. However, the DoD does not concur that the tools should be used routinely at all operational sites for continuous performance monitoring and measurement. Continuously capturing user–function response times by function and by user would result in tremendous numbers of data, large increases in storage volume requirements, and potentially could have a negative impact on Composite Health Care System performance.

The DoD concurs that Composite Health Care System Option Auditing may be enhanced to provide system resource (component) utilization and frequency for each system function as described by the GAO, and that there may be some benefit from the enhancements. The DoD currently is assessing the feasibility of those enhancements (cost/schedule) and will weigh that against the potential benefit of the enhancement. The DoD expects analysis to be completed by the end of 1994, at which time adoption of the enhancement will be considered. It should be noted that the utility of the tools may be different for the Virtual Address Extension and the Personal Computer–Composite Health Care System platforms, and that the cost/benefit of the changes will be assessed for each system configuration. The DoD does not assume, however, that the tools should run routinely at all operational sites, and believes that the tools, if acquired, may only be used in a laboratory environment and selectively at operational sites. If run routinely at operational sites, the tools will generate substantial numbers of data, which may be of little routine use, and the collection of these data may cause significant capacity and performance impact on the system under test.

RECOMMENDATION 2: The GAO recommended that the Secretary of Defense direct the Assistant Secretary of Defense (Health Affairs) to obtain performance analysis tools for each computer operating system under which the Composite Health Care System runs that will (1) enable the DoD to determine the causes of response time problems and (2) project the impact of workload growth and system configuration changes on response times. (p. 24/GAO Draft Report)

DOD RESPONSE: **Concur.** The DoD will proceed with efforts to incorporate simulation modeling tools (SES Workbench and perhaps other tools) into the Composite Health Care System Performance Management Program for capacity testing and planning, for simulation modeling and predictive performance analysis for system changes, and for incorporation into the product development process to minimize and manage better the capacity implications of Composite Health Care System software changes, as discussed in the DoD response to Finding G. The performance analysis modeling results should be available by the end of 1994. At that time, a decision on utilization of additional performance analysis tools will be made.

RECOMMENDATION 3: The GAO recommended that the Secretary of Defense direct the Assistant Secretary of Defense (Health Affairs) to modify the DoD approach to managing the Composite Health Care System performance to (1) include objectives for investigating and correcting extremely long response times and (2) provide reliable measures of system reserve capacity. (p. 24/GAO Draft Report)

DOD RESPONSE: **Concur.** The DoD concurs that action must be taken to address very long response times and to define criteria, objectives, and procedures for investigating and resolving these occurrences. It should be recognized, however, that it is not always cost effective to size automated systems in a manner that precludes the possibility of ***any*** system response time exceeding an established goal. In a system that allows the user community access to system resources on demand (e.g., *ad hoc* reports, etc.), it will be possible always for an authorized user to create a demand that would exceed virtually any available capacity. The DoD does concur that the causes of the long response times must be characterized fully and that all reasonable actions be taken to prevent the response times from impacting the Composite Health Care System user. The Office of the Director, Operational Test and Evaluation, will validate the effectiveness and suitability of those corrective actions once they are formulated. The specific actions will be addressed in Version 5.0 of the Composite Health Care System Performance Management Plan, which should be available by September 1994.

In addition, the DoD concurs with the GAO recommendation to provide reliable measures of system reserve capacity using current and future measurement and analysis tools available within the Composite Health Care System. (See the DoD responses to Recommendations 1 and 2.)

The following are GAO's comments on the Department of Defense's letter dated May 19, 1994.

GAO Comments

1. We have revised the report to say that the more than $700 million figure obligated through fiscal year 1993 includes initial operating costs at designated medical treatment facilities, as well as CHCS development costs. In addition, the $13 million figure in the report represents a monthly average of the amount ($161.9 million) that Defense obligated in fiscal year 1993 for (1) continued CHCS development, (2) deployment of CHCS beyond the designated test sites, and (3) overall CHCS operations.

2. We disagree with Defense's assertion that the limitations of the various performance measurement and analysis tools, currently in use by Defense, do not adversely affect its ability to manage system response time. The tools in use by Defense do not provide sufficient data to effectively manage the system response times experienced by CHCS users. According to Defense officials, a Defense team comprised of members knowledgeable in the use of the various measurement tools can diagnose and resolve response-time problems at any CHCS site. However, because this team approach is time-consuming and labor-intensive, we believe it is unreasonable to expect such an approach to be economically viable for effectively managing CHCS response time on a worldwide scale. Additionally, Defense's current approach does not provide it with the ability to analytically forecast future system requirements to preserve system responsiveness.

We also disagree with Defense's contention that we did not assess the full spectrum of tools currently used by Defense. We reviewed the CHCS performance management tools currently used by Defense and generally found them to be deficient for the following reasons:

(a)The two DEC VMS system monitors (VMS Monitor and System Performance Monitor—SPM) do not provide performance analysts with data needed to determine which system components are causing delays in the execution of user functions when they occur. They also do not measure system-component utilization for individual disks by process, thereby creating a problem when characterizing workloads for use in system modeling.

(b)The Massachusetts General Hospital Utility Multi-Programming System's (MUMPS) tools RTHIST and GLSSTA, according to Defense and

contractor officials, do not provide useful information for analytic modeling.

(c)According to Defense and contractor officials, Defense, at the time of our analysis, had no plans to acquire DECps because of its cost. Nevertheless, in a cursory review of DECps[14] we found that it does not recognize transactions and therefore does not measure transaction response times or transaction resource utilizations.

(d)At the time of our review, Defense had not acquired the Virtual Address Extension (VAX) Performance Advisor—a component of the analysis layer of DECps (PCPS).

(e)Data collected by the Option Auditing tool were not sufficiently refined to be useful in performance analysis. Specifically, the resource-utilization data collected by this tool could not be related to specific CHCS user functions.

Finally, we disagreed with Defense's contention that our exposure to the Performance Monitoring Tool (PMT) was limited. We conducted extensive analysis of the PMT measurements provided us by Defense and discussed the results of this review in the report.

3. We clarified the report to say that actual response-time and frequency measurements should be obtained for a representative sample of key or critical CHCS user functions, and not routinely for every CHCS user function.

4. The report was revised to clarify Defense's intent for developing and using its PMT.

5. The report was revised to recognize that despite the limitations in the performance measurement data generated by the DEC system monitors, it is possible for Defense to trace response-time problems after lengthy and labor-intensive efforts by analysts with a high level of technical knowledge.

6. We do not agree with Defense that there are adequate commercial measurement tools available for use with the Personal Computer-Composite Health Care System (PC-CHCS). Based on our discussions with technical experts, as well as our review of technical

[14]DECps was recently upgraded to Polycenter Performance Solution (PCPS).

publications, there currently are no adequate performance measurement tools available for the version of Unix being acquired by Defense.

7. Contrary to Defense's contention, our analysis of the sizing of the Walter Reed Army Medical Center did take into account the N+1 configuration rule in determining the number of central processing units required. In addition, our modeling provides for a CHCS workload increase of about 70 percent over what Defense measured at the Walter Reed Army Medical Center for October to November 1992.

To test the accuracy of our predictions relating to processor capacity, we compared our forecasted values to Defense's actual processor utilization data, as presented in CHCS monthly progress reports both before the processor upgrade—to help validate our predictive model—and after the processor upgrade—to validate our workload increase projection. It was unnecessary to revalidate the analytic model, since the Best/1 model automatically adjusts the processor characteristics relating to upgrades.

8. While Defense believes that its CHCS performance management methodology does not prevent or preclude it from investigating instances of excessive response time, Defense agrees with us that this is not clearly stated in its CHCS Performance Management Plan. We believe it is critical that Defense's revision of this plan include provisions for routine analysis of exceptionally long response times that occur sporadically.

9. Discussed in the "Agency Comments and Our Evaluation" section of this report.

Major Contributors to This Report

Accounting and Information Management Division, Washington, D.C.

Helen Lew, Assistant Director
John A. Riley, Senior Evaluator
Bruce B. Herbert, Senior Technical Specialist
Shane D. Hartzler, Reports Analyst

Related GAO Products

Composite Health Care System: Outpatient Capability Is Nearly Ready for Worldwide Deployment (GAO/IMTEC-93-11, Dec. 15, 1992).

Medical ADP Systems: Composite Health Care System Is Not Ready To Be Deployed (GAO/IMTEC-92-54, May 20, 1992).

Medical ADP Systems: Changes in Composite Health Care System's Deployment Strategy Are Unwise (GAO/IMTEC-91-47, Sept. 30, 1991).

Medical ADP Systems: Composite Health Care System: Defense Faces a Difficult Task (GAO/IMTEC-90-42, Mar. 15, 1990).

Medical ADP Systems: Composite Health Care System Operational Tests Extended (GAO/IMTEC-89-30, Apr. 10, 1989).

Medical ADP Systems: Analysis of Technical Aspects of DOD's Composite Health Care System (GAO/IMTEC-88-27, July 11, 1988).

Medical ADP Systems: Composite Health Care System Acquisition—Fair, Reasonable, and Supported (GAO/IMTEC-88-26, Mar. 4, 1988).

Medical ADP Systems: Composite Health Care System Operational Test and Evaluation Costs (GAO/IMTEC-88-18BR, Jan. 28, 1988).

ADP Systems: Concerns About DOD's Composite Health Care System Development Contracts (GAO/IMTEC-87-25, June 8, 1987).

ADP Systems: Concerns About the Acquisition Plan for DOD's Composite Health Care System (GAO/IMTEC-86-12, Mar. 31, 1986).

Ordering Information

The first copy of each GAO report and testimony is free. Additional copies are $2 each. Orders should be sent to the following address, accompanied by a check or money order made out to the Superintendent of Documents, when necessary. Orders for 100 or more copies to be mailed to a single address are discounted 25 percent.

Orders by mail:

U.S. General Accounting Office
P.O. Box 6015
Gaithersburg, MD 20884-6015

or visit:

Room 1100
700 4th St. NW (corner of 4th and G Sts. NW)
U.S. General Accounting Office
Washington, DC

Orders may also be placed by calling (202) 512-6000 or by using fax number (301) 258-4066.

Each day, GAO issues a list of newly available reports and testimony. To receive facsimile copies of the daily list or any list from the past 30 days, please call (301) 258-4097 using a touchtone phone. A recorded menu will provide information on how to obtain these lists.

PRINTED ON RECYCLED PAPER

United States
General Accounting Office
Washington, D.C. 20548-0001

Bulk Mail
Postage & Fees Paid
GAO
Permit No. G100