Navy Personnel Research and Development Center
San Diego, California 92152-7250 TN-93-6 May 1993

Markov Chains for Random Urinalysis II: Age-Test Model with Absorbing State

James P. Boyle
Douglas J. Hentschel
Theodore J. Thompson

NPRDC-TN-93-6 May 1993

Markov Chains for Random Urinalysis II: Age-Test Model With Absorbing State

James P. Boyle
Douglas J. Hentschel
Theodore J. Thompson

Reviewed by
Mark Chipman

Approved and released by
Murray W. Rowe
Director, Manpower Systems Department

Navy Personnel Research and Development Center
San Diego, California 92152-7250

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. AGENCY USE ONLY *(Leave blank)*	2. REPORT DATE	3. REPORT TYPE AND DATE COVERED
	May 1993	Final—Oct 92-Feb 92

4. TITLE AND SUBTITLE
Markov Chains for Random Urinalysis II: Age-Test Model With Absorbing State

5. FUNDING NUMBERS
Program Element: (1) Reimbursable
(2) 0305889N
Work Unit: (1) 92PODD911
(2) 0305889N.R2143.DR001

6. AUTHOR(S)
James P. Boyle, Douglas J. Hentschel, Theodore J. Thompson

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Navy Personnel Research and Development Center
San Diego, California 92152-7250

8. PERFORMING ORGANIZATION REPORT NUMBER
NPRDC-TN-93-6

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
Chief of Naval Personnel (PERS-63)
Washington, DC 20370-5000

10. SPONSORING/MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES
Functional Area: Manpower Systems
Product Line: Information Support
Effort: Drug Testing Algorithm

12a. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release; distribution is unlimited.

12b. DISTRIBUTION CODE

13. ABSTRACT *(Maximum 200 words)*

This is the second of a series of papers on the use of Markov chains to model random urinalysis programs. Previous work introduced the age-test Markov chain. This chain was used to model random urinalysis strategies stratified by time since last tested. This paper extends the age-test model by including an absorbing state for detection of drug users. We quantify the extent to which age-test urinalysis strategies can be gamed by drug users. The theory and formulas for calculating the distribution of time to absorption (e.g., detection of drug users), the mean time to absorption and the expected number of visits to intermediate states (e.g., number of tests) prior to detection are developed. This model is used to analyze three age-test strategies.

14. SUBJECT TERMS
Drug urinalysis, Markov chains, stratified random sampling, age-test model

15. NUMBER OF PAGES
27

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
UNCLASSIFIED	UNCLASSIFIED	UNCLASSIFIED	UNLIMITED

Foreword

This report was prepared as part of the Markov Models of Random Urinalysis Sampling Procedures project (Reimbursable, Work Unit 92PODD911) and the Statistical Methods for Drug Testing project (Program Element 0305889N, Work Unit 0305889N.R2143.DR001), both sponsored by the Chief of Naval Personnel (PERS-63). The objectives of the projects are to determine if urinalysis strategies based on time since last test can be used to improve the Navy's drug deterrence program and to develop a unified set of statistical methodologies for the analysis of drug testing programs and data. The work described here was performed during FY92 and FY93.

This is the second in a series of reports on the use of Markov chains for the analysis of random urinalysis programs. The first report is *Markov Chains for Random Urinalysis I: Age-Test Model* (NPRDC-TN-93-5). Related work also includes *Probability of Detection of Drug Users by Random Urinalysis in the U.S. Navy* (NPRDC-TN-93-2).

MURRAY W. ROWE
Director, Manpower Systems Department

Summary

Background

This is the second of a series of papers on the use of Markov chains to model random urinalysis programs. Previous work (Thompson, Boyle & Hentschel, 1993) introduced the age-test Markov chain. This chain was used to model random urinalysis strategies stratified by time since last tested. This paper extends the age-test model by including an absorbing state for detection of drug users.

The Nuclear Regulatory Commission (NRC, [1988]) proposed a urinalysis testing strategy based on time since last tested. That is, the probability of a person being tested depends on the amount of time since the person was last tested. Southern California Edison (SCE) has implemented a variation (Murray & Talley, 1988) of the NRC proposal at the San Onofre Nuclear Generating Station in Southern California. We have shown (Thompson, et al., 1993) that age-test urinalysis strategies trade off predictability of an individual being tested for fewer people not tested within a year and fewer people tested excessively during a year.

Objective

The objective of this work is to quantify the extent to which age-test urinalysis strategies can be gamed by drug users.

Approach

The approach includes (1) extending the age-test model to include an absorbing state for detection of drug users, (2) developing the needed theory, and (3) analyzing a number of specific age-test urinalysis strategies.

Results

We define a class of age-test models extending the models presented in Thompson, et al. (1993). The theory and formulas for calculating the distribution of time to absorption (e.g., detection of drug users), the mean time to absorption and the expected number of visits to intermediate states (e.g., number of tests prior to detection) are developed.

This model is used to analyze three age-test strategies. These strategies include an age-test model of the NRC proposal, an age-test model of the SCE process, and an age-test model meeting Navy requirements and assuming a 15% monthly testing rate. The mean times to detection for the NRC proposal are 58 and 200 months for nongaming and gaming drug users, respectively. The mean times for the SCE process are 46 and 84 months. The mean times for the Navy model are 33 and 50 months. These results are based on age-test models assuming drugs are detectable in the user 20% of the time.

Conclusions

Markov chains provide a framework for the systematic analysis of drug testing strategies stratified by time since last tested. Age-test strategies do not change detection time for nongaming drug users. As long as the annual testing rate is fixed, average time to detection is unchanged by using age-test strategies. Age-test strategies do allow gaming drug users to significantly increase time to detection. Constant testing strategies. defined by equal testing probabilities for all states, are resistant to gaming.

Contents

List of Tables

List of Figures

1 Introduction

This is the second in a series of papers on the use of Markov chains to model random urinalysis programs. Previous work (Thompson, Boyle and Hentschel, 1993) introduced the age-test Markov chain. This chain was used to model random urinalysis strategies stratified by time last tested. This paper extends the age-test model by including an absorbing state for detection of drug users.

The Nuclear Regulatory Commission (NRC, [1988]) proposed a urinalysis testing strategy based on time last tested. That is, the probability of a person being tested depends on the amount of time since the person was last tested. Southern California Edison (SCE) has implemented a variation (Murray and Talley, 1988) of the NRC proposal at the San Onofre Nuclear Generating Station in Southern California. Urinalysis testing strategies based on time last tested are defined by (1) a *high* testing rate for personnel not yet tested in a given time period and (2) a *low* testing rate for personnel found to have negative results in a given time period. The NRC's adopted rules and regulations (Nuclear Regulatory Commission, 1989) for urinalysis do not require a *time last tested* strategy. However, SCE continues to use their variation of this strategy with NRC approval.

We have shown (Thompson, et al., 1993) that age-test urinalysis strategies trade off predictability of an individual being tested for reduced *tail area* of the distribution of the number of tests in a fixed time period. Age-test strategies provide fewer people not tested within a year and fewer people tested excessively during a year than simple random sampling. Age-test strategies are also more predictable, have lower variance in the number of tests, and as a result are subject to gaming by drug users.

Age-test models of both the NRC proposal and the SCE process were shown to have undesirable properties. The NRC proposal required that at least 90% of the individuals be tested each year and that testing rates for individuals already tested with negative results be at least 2.5% per month. A minimum cost age-test model was developed for this proposal. SCE states a 5% annual chance of not being tested and a 130% average annual testing rate. An equal cost age-test model similar to this process was developed. These strategies involve large differences in the testing rates between people tested within the past year and those who were not tested. This implies that, once tested, an individual has a high probability of not being tested again within a year. These probabilities are 0.74 for the age-test model of the NRC proposal and 0.48 for the age test model of the SCE process. The age-test model of the SCE process is such that almost half the tests every month are given to people who know they will be tested.

A related use of Markov chains, modeling classes of drug users, is given in Evanovich (1985). Previous work (Thompson and Boyle, 1992) at NPRDC includes models of detection and gaming of drug users under simple random sampling.

This paper reports on our work to quantify the extent to which age-test urinalysis strategies can be *gamed* by drug users. This paper (1) extends the age-test model to include an absorbing state for detection of drug users, (2) develops the needed theory, and (3) analyzes a number of specific age-test urinalysis strategies. The NRC proposal, the SCE process, and a model meeting Navy requirements are analyzed.

2 Markov Chains

This section develops the theory and notation that will be used in the remainder of the report. Here, we develop the formulas for calculating the distribution of time to absorption (e.g., detection of drug users), including mean time to absorption and for calculating the expected number of visits to intermediate states. It can be considered a continuation of Section 2 of Thompson, et al. (1993). Most of the material and notation is taken from Taylor and Karlin (1984) and Hoel, Port and Stone (1972).

Again, we consider finite state Markov chains, but with the following special structure. The states are labeled $0, 1, \ldots, N$ with transient states $0, 1, \ldots, r-1$ and absorbing states $r, r+1, \ldots, N$. A state a is absorbing if, once at a, the chain stays at a for all time. Also, starting at any transient state, the chain must be eventually trapped by one of the absorbing states. The transition matrix for such a Markov chain can be partitioned as

$$\mathbf{P} = \begin{bmatrix} \mathbf{Q} & \mathbf{R} \\ \mathbf{O} & \mathbf{I} \end{bmatrix}$$

where $\mathbf{Q}$ is an $r \times r$ matrix, $\mathbf{R}$ is an $r \times (N-r+1)$ matrix, $\mathbf{O}$ is an $(N-r+1) \times r$ matrix of zero entries, and $\mathbf{I}$ is the $(N-r+1) \times (N-r+1)$ identity matrix.

Using matrix algebra it can be shown that the n-step transition matrix is

$$\mathbf{P}^n = \begin{bmatrix} \mathbf{Q}^n & (\mathbf{I} + \mathbf{Q} + \cdots + \mathbf{Q}^{n-1})\mathbf{R} \\ \mathbf{O} & \mathbf{I} \end{bmatrix} \tag{1}$$

for $n \geq 2$. Also, from page 20 of Hoel, et al. (1972), $\mathbf{Q}^n \to 0$ as $n \to \infty$. Thus, the infinite series $(\mathbf{I} + \mathbf{Q} + \cdots + \mathbf{Q}^{n-1}) \to (\mathbf{I} - \mathbf{Q})^{-1} = \mathbf{W}$ as $n \to \infty$, and from Equation 1

$$\lim_{n \to \infty} \mathbf{P}^n = \begin{bmatrix} \mathbf{O} & \mathbf{WR} \\ \mathbf{O} & \mathbf{I} \end{bmatrix}. \tag{2}$$

Let $u_{ia}(n)$ be the ith row, ath column entry from the $r \times (N-r+1)$ submatrix $(\mathbf{I} + \mathbf{Q} + \cdots + \mathbf{Q}^{n-1})\mathbf{R}$ in the n-step transition matrix, Equation 1. From page 15 of Hoel, et al. (1972), we have

$$u_{ia}(n) = P[T_a \leq n, X_T = a | X_0 = i] \tag{3}$$

where T_a is the time to absorption by state a, T is the time to absorption for the process, and i is any of the transient states. Equation 3 states that the (i, a) entry in the submatrix from Equation 1 is the probability that, starting at state i, the process is absorbed by state a within n time periods. By basic probability laws,

$$\lim_{n \to \infty} u_{ia}(n) = P[X_T = a | X_0 = i] = u_{ia}.$$

Thus, u_{ia} is the (i, a) entry of the submatrix $\mathbf{U} = \mathbf{WR}$ in Equation 2, and we have observed that the entries in the ith row of $\mathbf{U}$ represent the probabilities, starting at state i, that the process will be trapped by the various absorbing states.

Another way to establish that $\mathbf{U} = \mathbf{WR}$ is to employ what Taylor and Karlin (1984) call first step analysis. Note

$$\begin{aligned} u_{ia} &= P[X_T = a | X_0 = i] \\ &= P[X_T = a, X_1 = a | X_0 = i] + \sum_{j=0}^{r-1} P[X_T = a, X_1 = j | X_0 = i] \\ &= P[X_T = a | X_1 = a, X_0 = i] P[X_1 = a | X_0 = i] \\ &\quad + \sum_{j=0}^{r-1} P[X_T = a | X_1 = j, X_0 = i] P[X_1 = j | X_0 = i] \\ &= p_{ia} + \sum_{j=0}^{r-1} u_{ja} p_{ij}. \end{aligned}$$

In matrix notation the above relations become

$$(\mathbf{I} - \mathbf{Q}) \begin{bmatrix} u_{0a} \\ u_{1a} \\ \vdots \\ u_{r-1\,a} \end{bmatrix} = \begin{bmatrix} p_{0a} \\ p_{1a} \\ \vdots \\ p_{r-1\,a} \end{bmatrix}; \quad a = r, \ldots, N.$$

These equations can be summarized in the single matrix equation $(\mathbf{I} - \mathbf{Q})\mathbf{U} = \mathbf{R}$ and we see that $\mathbf{U} = (\mathbf{I} - \mathbf{Q})^{-1}\mathbf{R} = \mathbf{WR}$.

The matrix $(\mathbf{I} - \mathbf{Q})^{-1} = \mathbf{W}$ is often called the fundamental matrix. First step analysis can be used to yield a probabilistic interpretation of the elements of $\mathbf{W}$. Let i and k be transient states and set

$$v_{ik} = E\left[\sum_{n=0}^{T-1} 1_k(X_n) | X_0 = i\right] \tag{4}$$

where $1_k(X_n)$ is the indicator function associated with state k. The expression in Equation 4 is the expected or mean number of visits to state k before absorption, given the process starts at state i. Using conditional probabilities we see that

$$\begin{aligned} v_{ik} &= \sum_{a=r}^{N} E\left[\sum_{n=0}^{T-1} 1_k(X_n) | X_1 = a, X_0 = i\right] P[X_1 = a | X_0 = i] \\ &\quad + \sum_{j=0}^{r-1} E\left[\sum_{n=0}^{T-1} 1_k(X_n) | X_1 = j, X_0 = i\right] P[X_1 = j | X_0 = i] \\ &= \sum_{a=r}^{N} 1_k(i) p_{ia} + \sum_{j=0}^{r-1} [v_{jk} + 1_k(i)] p_{ij} \\ &= \delta_{ik} + \sum_{j=0}^{r-1} v_{jk} \end{aligned} \tag{5}$$

for $i, k = 0, 1, \ldots, r-1$. The last equation follows since $\mathbf{P}$ is stochastic and its rows sum to unity. δ_{ik} is the Kronecker delta, which equals 1 when $i = k$ and equals 0 otherwise.

The set of Equations 5 has the matrix representation

$$\begin{bmatrix} 1-p_{00} & -p_{01} & \cdots & -p_{0\,r-1} \\ -p_{10} & 1-p_{11} & \cdots & -p_{1\,r-1} \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ -p_{r-1\,0} & -p_{r-1\,1} & \cdots & 1-p_{r-1\,r-1} \end{bmatrix} \begin{bmatrix} v_{0k} \\ v_{1k} \\ \cdot \\ \cdot \\ \cdot \\ v_{r-1\,k} \end{bmatrix} = e_k \tag{6}$$

where e_k is the kth column of an $r \times r$ identity matrix. Combining all systems in Equation 6 for $k = 0, 1, \ldots, r-1$ leads to the single matrix equation

$$(\mathbf{I} - \mathbf{Q}) \begin{bmatrix} v_{00} & \cdots & v_{0\,r-1} \\ \cdot & & \cdot \\ \cdot & & \cdot \\ \cdot & & \cdot \\ v_{r-1\,0} & \cdots & v_{r-1\,r-1} \end{bmatrix} = \mathbf{I} \tag{7}$$

From Equation 7 it is clear that $v_{ik} = w_{ik}$, and w_{ik} equals the mean number of visits to state k from state i before absorption. When $i = k$ in Equation 4, the expectation counts the starting state i as a visit to state i.

Lastly, define W_k as the number of visits to state k for $k = 0, 1, \ldots, r-1$. Clearly, we must have

$$T = W_0 + \cdots + W_{r-1}$$

where, again, T equals time to absorption. Hence

$$E[T|X_0 = i] = E\left[\sum_{k=0}^{r-1} W_k | X_0 = i\right] = \sum_{k=0}^{r-1} w_{ik}$$

and we may compute the expected time to absorption starting at state i as the sum of the elements in the ith row of the matrix $\mathbf{W}$.

3 Age-Test With Absorption

We now define a class of age-test models extending the models presented in Section 3 of Thompson, et al. (1993). The transition matrix is

$$\mathbf{P} = \begin{bmatrix} (1-\alpha_1)p_1 & q_1 & 0 & 0 & \cdots & 0 & \alpha_1 p_1 \\ (1-\alpha_2)p_2 & 0 & q_2 & 0 & \cdots & 0 & \alpha_2 p_2 \\ (1-\alpha_3)p_3 & 0 & 0 & q_3 & \cdots & 0 & \alpha_3 p_3 \\ \cdot & \cdot & \cdot & \cdot & & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & & \cdot & \cdot \\ (1-\alpha_d)p_d & 0 & 0 & 0 & \cdots & q_d & \alpha_d p_d \\ (1-\alpha_{d+1})p_{d+1} & 0 & 0 & 0 & \cdots & q_{d+1} & \alpha_{d+1} p_{d+1} \\ 0 & 0 & 0 & 0 & \cdots & 0 & 1 \end{bmatrix}. \tag{8}$$

In Equation 8, an individual is in state i ($1 \leq i \leq d$) if the individual tested negative i time periods ago, and in state $d+1$ if tested $d+1$ or more time periods ago. The p_i's are the probabilities of being tested given state i, and the α_i's are the conditional probabilities of a positive result given the individual is tested. State $d+2$ is the single absorbing state defined in Equation 8. A person is absorbed or caught if tested positive. To summarize, an individual residing in state i ($1 \leq i \leq d$) must in the next period test negative and move to state 1, age to state $i+1$, or be absorbed to state $d+2$ by testing positive. A person in state $d+1$ "ages" by staying in state $d+1$.

As in Section 3 of Thompson, et al. (1993), we impose the restrictions $0 < p_1, \ldots, p_d < 1$ and $0 < p_{d+1} \leq 1$. We allow the α_i's to be unrestricted probabilities, i.e. $0 \leq \alpha_1, \ldots, \alpha_{d+1} \leq 1$, except we insist at least one of the α_i's be positive. All of this guarantees that states 1 through $d+1$ must lead to the absorbing state and, by Theorem 2 on page 21 of Hoel, et al. (1972), these states are transient and the results from Section 2 apply.

3.1 Time to Absorption

Using the notation from the previous section we have

$$\mathbf{Q} = \begin{bmatrix} (1-\alpha_1)p_1 & q_1 & 0 & 0 & \cdots & 0 \\ (1-\alpha_2)p_2 & 0 & q_2 & 0 & \cdots & 0 \\ (1-\alpha_3)p_3 & 0 & 0 & q_3 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & & \vdots \\ (1-\alpha_d)p_d & 0 & 0 & 0 & \cdots & q_d \\ (1-\alpha_{d+1})p_{d+1} & 0 & 0 & 0 & \cdots & q_{d+1} \end{bmatrix}, \quad \mathbf{R} = \begin{bmatrix} \alpha_1 p_1 \\ \alpha_2 p_2 \\ \alpha_3 p_3 \\ \vdots \\ \alpha_d p_d \\ \alpha_{d+1} p_{d+1} \end{bmatrix}. \tag{9}$$

The matrix $\mathbf{U} = (\mathbf{I}-\mathbf{Q})^{-1}\mathbf{R} = \mathbf{W}\mathbf{R} = \mathbf{1}$ is a $(d+1) \times 1$ column of 1's, since there is only one absorbing state and the process is absorbed with certainty. Also, for the same reasons, $T = T_a = T_{d+1}$ and $\{X_T = d+1\}$ is the certain event. Thus, from Equation 2 of Section 2, we have

$$u_{i\,d+1}(n) = P[T \leq n | X_0 = i] \tag{10}$$

equals the distribution function of time to absorption, conditional on the process starting at the transient state i. This is precisely the ith element of the $(d+1)\times 1$ column $(\mathbf{I}+\mathbf{Q}+\cdots+\mathbf{Q}^{n-1})\mathbf{R}$.

We can also define the mean time to absorption $E(T|X_0 = i)$ through a limiting argument. From elementary methods, the expected value of a random variable taking on positive integer values is

$$\lim_{n\to\infty} 1 + \sum_{k=1}^{n} [1 - F(k)] \tag{11}$$

where F is the distribution function of the random variable. Applying Equation 11 and using Equation 10 yields

$$E(T|X_0 = i) = 1 + \lim_{n\to\infty} \sum_{k=1}^{n} [1 - u_{i\,d+1}(n)] \tag{12}$$

where we have conditioned on the event $X_0 = i$. The result Equation 12 can be put into matrix terms as

$$\mathbf{1} + \lim_{n\to\infty} n\mathbf{1} - [\mathbf{R} + (\mathbf{I}+\mathbf{Q})\mathbf{R} + \cdots + (\mathbf{I}+\mathbf{Q}+\cdots+\mathbf{Q}^{n-1})\mathbf{R}]$$

$$= 1 + \lim_{n\to\infty} n\mathbf{1} - [(\mathbf{I}-\mathbf{Q})\mathbf{WR} + (\mathbf{I}-\mathbf{Q}^2)\mathbf{WR} + \cdots + (\mathbf{I}+\mathbf{Q}^n)\mathbf{WR}]$$
$$= 1 + \lim_{n\to\infty} n\mathbf{1} - [(\mathbf{I}-\mathbf{Q}) + (\mathbf{I}-\mathbf{Q}^2) + \cdots + (\mathbf{I}+\mathbf{Q}^n)]\mathbf{1}$$

since $\mathbf{WR} = \mathbf{1}$. The above limit reduces to

$$1 + \lim_{n\to\infty} n\mathbf{1} - [n\mathbf{I} - (\mathbf{Q} + \mathbf{Q}^2 + \cdots + \mathbf{Q}^n)]\mathbf{1}$$
$$= 1 + \lim_{n\to\infty} (\mathbf{Q} + \mathbf{Q}^2 + \cdots + \mathbf{Q}^n)\mathbf{1}$$
$$= 1 + \lim_{n\to\infty} [(\mathbf{I} - \mathbf{Q}^{n+1})\mathbf{W} - \mathbf{I}]\mathbf{1}$$
$$= 1 + (\mathbf{W} - \mathbf{I})\mathbf{1}$$
$$= \mathbf{W1}$$

because $\lim_{n\to\infty} \mathbf{Q}^{n+1} = \mathbf{0}$. Again we note that the expected time to absorption, starting at state i, is the sum of the elements in the ith row of $\mathbf{W}$.

3.2 Number of Tests Prior to Absorption

Recall from Section 2 that, in general, w_{ik} equals the mean number of visits to state k prior to absorption, starting at state i. For the age-test model with absorbing state, we will be interested in calculating the mean number of tests prior to absorption. Since testing negative is equivalent to visiting state 1, this amounts to selecting the ith element from the first column of the fundamental matrix $\mathbf{W}$.

3.3 Drug User Gaming

An interesting problem arises in considering the process of Equation 8. Given a fixed set of p_i's, a drug user might be interested in choosing α_i's in such a way that their average usage rate is at a desired level and the mean time to absorption is a maximum. Here we define the usage rate as

$$Y = \frac{\alpha_1 W_1 + \cdots + \alpha_{d+1} W_{d+1}}{W_1 + \cdots + W_{d+1}}$$

where, again, W_k represents the number of visits to state k prior to absorption. The average or expected usage rate, starting in state i, is $E(Y|X_0 = i)$ and represents the average percentage of time that the user would test positive prior to absorption. Since taking the expectation of a ratio of random variables is somewhat intractable, we henceforth approximate the average usage rate with the ratio of expectations, i.e.,

$$E(Y|X_0 = i) \approx \frac{\alpha_1 w_{i1} + \cdots + \alpha_{d+1} w_{id+1}}{w_{i1} + \cdots + w_{id+1}}$$

Formally, the drug user, knowing the p_i's, wishes to find the α_i's that

$$\text{Maximize} \quad E(T|X_0 = i) = w_{i1} + \cdots + w_{id+1} \tag{13}$$

$$\text{subject to} \quad \frac{\alpha_1 w_{i1} + \cdots + \alpha_{d+1} w_{id+1}}{w_{i1} + \cdots + w_{id+1}} = \tilde{\alpha}$$

$$0 \le \alpha_i \le 1 \quad \forall i$$

where $\bar{\alpha}$ is some desired average usage rate. The next section contains a number of examples illustrating Equation 13.

We conclude this section with a result proved in the Appendix. If the p_i's are equal in Equation 9 or the α_i's are all equal to $\bar{\alpha}$ in Equation 13, the mean time to absorption, starting at state 1, is given by $1/\pi_1\bar{\alpha}$. Equal p_i's means testing rates are the same regardless of state and equal α_i's means drug usage rates are the same regardless of state. Here π_1 represents the expected number of tests per time period for a nonuser starting at steady state. Steady state is defined in Thompson, et al. (1993).

This result is very important for two reasons. First, urinalysis strategies with equal p_i's cannot be gamed. To prevent gaming, age-test urinalysis strategies with unequal p_i's should be avoided. Second, age-test urinalysis strategies with unequal p_i's do not change mean time to detection for nongaming drug users. Mean time to detection for nongaming drug users is only a function of the average test rate. This implies age-test strategies are not problematic if gaming is not a concern.

4 Applications

Three age-test strategies presented in Thompson, et al. (1993) are revisited here. These strategies are the age-test model of the NRC proposal, the age-test model of the SCE process, and the age-test model meeting Navy requirements and assuming a 15% monthly testing rate.

The number of tests prior to detection, and the time to detection for both nongaming and gaming drug users is presented for all three age-test strategies. Average drug use ($\bar{\alpha}$) for all following examples is 0.20. That is, drugs are detectable in a user's system an average of 20% of time or about 6 days per month. This is roughly equivalent to using drugs twice a month.

4.1 Nuclear Regulatory Commission (NRC)

The Nuclear Regulatory Commission (1988) proposed two urinalysis testing alternatives. One alternative required that at least 90% of the individuals are tested each year and that testing rates for individuals already tested with negative results be at least 2.5% per month. The age-test strategy $\{p_1 = p_2 = \ldots = p_{12} = 0.025$ and $p_{13} = 0.6338\}$ is the least cost solution to this NRC proposal. The distribution of the number of tests within 12 months is shown in Figure 1. By contrast, simple random sampling at the 103% annual rate (same rate as the minimum cost age-test model) gives a 34% chance of no tests and a 7.7% chance of 3 or more tests.

Time to detection results are presented in Table 1. Average drug use is the same (20%) for both the nongaming and gaming drug user. The gaming strategy is the obvious one. Drug usage in state 13 is set to zero. To keep average usage at 20%, drug usage in states 1 through 12 is increased to 22%. The average time to detection increases significantly from 58 months to 200 months. Also, the average number of tests prior to detection increases from 5 to 17 tests. If detection of drug users and prevention of gaming are high priorities, this age-test strategy should not be used.

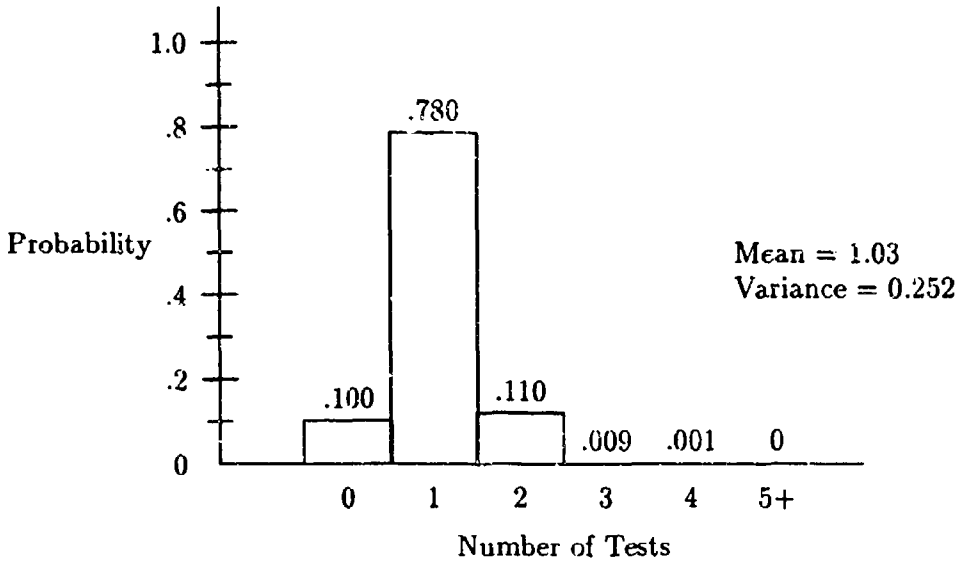

Figure 1. Steady state d' tribution of number of tests within 12 months for age-test model of NRC proposal.

Table 1

Time (in Months) to Detection for Age-Test Model of NRC Proposal

	Drug User	
	Nongaming	Gaining
Mean	58.2	200.2
1st Quartile	20	58
2nd Quartile	41	128
3rd Quartile	79	278

4.2 Southern California Edison (SCE)

SCE has implemented a *composite random sampling* (Murray & Talley, 1988) approach to urinalysis. Their approach is based on a sampling scheme that is part sampling with replacement (individuals are subject to sampling after having been selected) and part sampling without replacement (individuals are not subject to sampling after having been selected). The entire population is sampled at a specified rate with replacement. People who have not been sampled within the past year are sampled at another specified rate without replacement. SCE states a 5% annual chance of not being tested and a 130% average annual testing rate. The age-test strategy $\{p_1 = p_2 = \ldots = p_{12} = 0.0595, p_{13} = 1.0\}$ yields similar results. The distribution of number of tests within 12 months is shown in Figure 2. Simple random sampling at the 130% annual rate gives a 25% chance of no tests and a 13% chance of 3 or more tests.

Time to detection results are presented in Table 2. The gaming strategy is, again, the obvious one. Drug usage in state 13 is set to zero. To keep average usage at 20%, drug usage in states 1 through 12 is increased to 21%. The average time to detection increases significantly from 46 months to 84 months. Also the average number of tests prior to detection increases from 5 to 9 tests. If detection of drug users and prevention of gaming are high priorities, this age-test strategy should not be used.

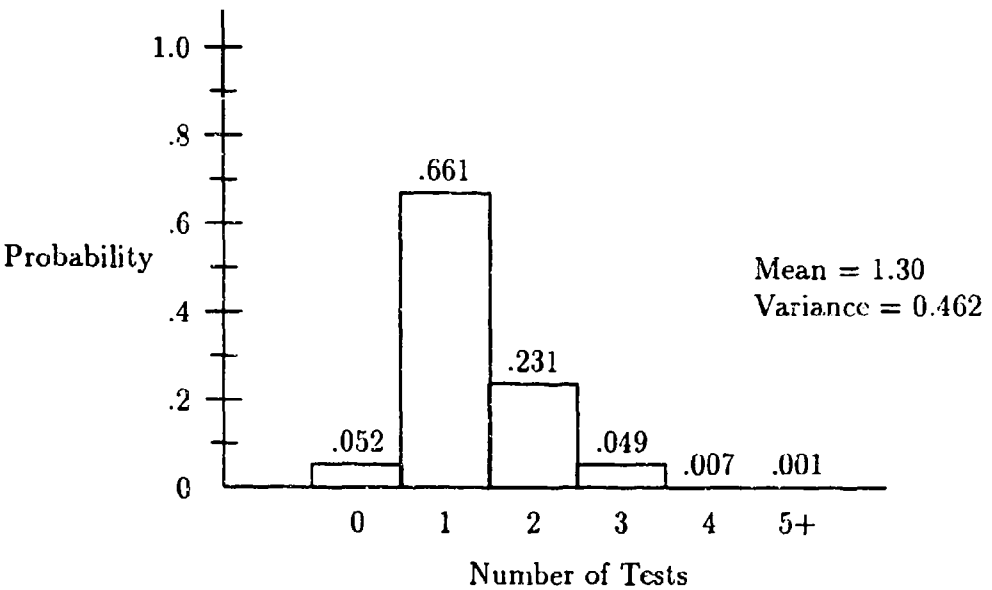

Figure 2. Steady state distribution of number of tests within 12 months for age-test model of SCE's process.

4.3 United States Navy

U.S. Navy policy (Chief of Naval Operations, 1990) directs commands to test 10–20% of their personnel each month. We developed an age-test strategy $\{p_1 = p_2 = p_3 = 0.1, p_4 = 0.1145,$

Table 2

Time (in Months) to Detection for Age-Test Model of SCE Process

	Drug User	
	Nongaming	Gaming
Mean	46.2	83.9
1st Quartile	15	24
2nd Quartile	33	58
3rd Quartile	63	116

$p_5 = \ldots = p_{13} = 0.2$) that maximizes the number of people tested at least once a year given an average monthly testing rate of 15% and given all p_i between 10 and 20%. The midpoint value 15% was chosen. Although the Navy does not use this strategy, the strategy does comply with Navy policy. For this age-test strategy, the distribution of number of tests within 12 months is shown in Figure 3. Simple random sampling at the 180% annual rate gives a 14% chance of no tests and a 8.2% chance of 4 or more tests.

Time to detection results are presented in Table 3. A gaming strategy that seems reasonable is as follows. Drug usage in states 1 to 3 is increased to 30%; drug usage in states 5 through 13 is decreased to 12%. To keep average usage at 20%, drug usage in state 4 is set at 18.5%. The average time to detection increases moderately from 33 months to 38 months. Also, the average number of tests prior to detection increases from 5 to 6 tests. Since this gaming strategy yielded only slight gains, we calculate[1] the optimal gaming strategy. Results from using this strategy are also included in Table 3. This optimal gaming strategy is 100% usage in state 1, 37% usage in state 2, and zero usage for all other states. Using it increases the average time to detection to 50 months. Although this strategy does not seem reasonable, it is presented to show the optimal amount of gaming possible. This age test strategy, which contains no large differences in the values of the p_i's, limits the amount of gaming possible for users.

4.4 Relative Merits

We conclude this section with a comparison of these three age-test strategies. Assume Company XYZ has 1000 employees and the typical drug user has drugs detectable in their system an average of 6 days per month. Also assume that one urinalysis test costs $10.00. A comparison of the relative merits of the 3 age-test strategies is shown in Table 4. The number of tests per year is simply 1000 times the average number of tests per person per year. The average number of tests per person per year is obtained from the steady state distribution of the Markov chain without an absorbing state. The table also includes the number of months until detection with probability 0.10, 0.50 and 0.95. Both gaming and nongaming times are presented. The optimal gaming strategy is used for the age-test model meeting Navy policy.

[1]Excel Solver (Microsoft Corporation, 1991) was used to solve optimization problems.

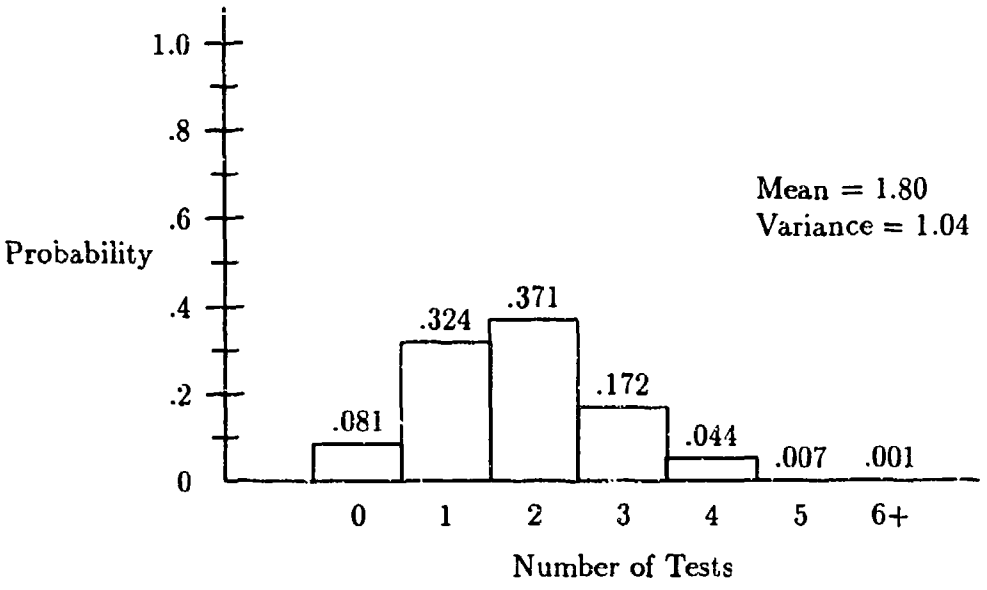

Figure 3. Steady state distribution of number of tests within 12 months for an age-test model meeting Navy requirements.

Table 3

Time (in Months) to Detection for an Age-Test Model Meeting Navy Requirements

	Drug User		
	Nongaming	Gaming	Optimal Gaming
Mean	33.3	38.3	50.0
1st Quartile	11	11	11
2nd Quartile	24	27	35
3rd Quartile	46	53	71

Table 4

Relative Merits of Three Age-Test Strategies

		Age-Test Strategy		
		NRC	SCE	Navy
Number of Tests per Year		1030	1300	1800
Cost per Year		$10,300	$13,000	$18,000
Nongaming User		Time in Months		
Detection Probability	0.10	13	9	5
	0.50	41	33	24
	0.95	163	131	97
Gaming User				
Detection Probability	0.10	21	9	1
	0.50	128	58	33
	0.95	600	250	157

Company XYZ: 1000 employees, typical drug user has drugs detectable in system 20% of the time.

5 Conclusions

Markov chains provide a framework for the systematic analysis of drug testing strategies stratified by time last tested. For nonusers the steady state distribution provides estimates of the number of tests per month and the number of people who have not been tested in the past year. The distribution of the number of tests in a fixed time period (e.g., year), given any initial state, can be calculated. Furthermore, given testing cost estimates, the relative merits of different testing strategies can be easily calculated. For drug users, the distribution of time to detection and the expected number of tests prior to detection can be calculated given any initial state.

Age-test urinalysis strategies trade off predictability of an individual being tested for reduced probability of not being tested and reduced probability of excessive tests. This reduced tail area of the distribution is sometimes perceived as more equitable.

Age-test does not change detection for nongaming drug users. As long as the annual testing rate is fixed, average time to detection is unchanged by using age-test strategies. Therefore, if gaming drug users are not a concern, age-test may be beneficial.

Age-test allows gaming drug users to significantly increase time to detection. The more extreme the age-test strategy, the more gaming is possible. Extreme strategies have large changes in probability tested over time (since last tested).

Constant testing strategies are resistant to gaming. Constant strategies are defined by equal probabilities of testing by state. Since the past and present provide no information about the future, gaming as defined here is impossible.

6 References

Chief of Naval Operations (1990). Alcohol and Drug Abuse Prevention and Control, *OPNAVINST* 5350.4B.

Evanovich, Peter (1985). *A Model for Drug Testing* (CRM 85-33). Alexandria, VA: Center for Naval Analyses.

Hoel, P. G., Port, S. C. & Stone, C. J. (1972). *Introduction to Stochastic Processes*, Boston: Houghton Mifflin Company.

Microsoft Corporation (1991). *Microsoft Excel Solver User's Guide*, Redmond, WA.

Murray, L. C. & Talley, A. E. (1988). *Composite Random Sampling*, Rosemead, CA: Southern California Edison.

Nuclear Regulatory Commission (1988). Fitness-for-Duty Program: Proposed Rules, *Federal Register*, *53* (184).

Nuclear Regulatory Commission (1989). Fitness-for-Duty Program: Rules and Regulations, *Federal Register*, *54* (108).

Taylor, H. M. & Karlin, S. (1984). *An Introduction to Stochastic Modeling*, Orlando: Academic Press, Inc.

Thompson, T. J. & Boyle, J. P. (1992). *Probability of Detection of Drug Users by Random Urinalysis in the U.S. Navy* (NPRDC-TN-93-2). San Diego: Navy Personnel Research and Development Center.

Thompson, T. J., Boyle, J. P. & Hentschel, D. J. (1993). *Markov Chains for Random Urinalysis I: Age-Test Model* (NPRDC-TN-93-5). San Diego: Navy Personnel Research and Development Center.

Appendix

Minimax Gaming

Minimax Gaming

This appendix proves the result mentioned at the end of Section 3.3. A minimax theorem associated with gaming is also established.

Recalling the definition of Q from Equation 9, we have

$$\mathbf{I}-\mathbf{Q}=\begin{bmatrix} \alpha_1 p_1+q_1 & -q_1 & 0 & 0 & \cdots & 0 & 0 \\ \alpha_2 p_2-p_2 & 1 & -q_2 & 0 & \cdots & 0 & 0 \\ \alpha_3 p_3-p_3 & 0 & 1 & -q_3 & \cdots & 0 & 0 \\ \cdot & \cdot & \cdot & \cdot & & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & & \cdot & \cdot \\ \alpha_d p_d-p_d & 0 & 0 & 0 & \cdots & 1 & -q_d \\ \alpha_{d+1} p_{d+1}-p_{d+1} & 0 & 0 & 0 & \cdots & 0 & 1-q_{d+1} \end{bmatrix}.$$

We are interested in deriving the first row of the fundamental matrix $\mathbf{W}=(\mathbf{I}-\mathbf{Q})^{-1}$. The first row of $\mathbf{W}$ provides the mean number of visits to each transient state prior to detection given the system starts in state 1. First we calculate the determinant of $(\mathbf{I}-\mathbf{Q})$. Expanding by elements of the first column yields

$$\begin{aligned} \mathrm{DET}(\mathbf{I}-\mathbf{Q}) &= p_{d+1}\left[\alpha_1 p_1+q_1+\sum_{i=2}^{d+1}(\alpha_i p_i-p_i)\tau_{i-1}\right] \\ &= p_{d+1}\left[\alpha_1 p_1+\sum_{i=2}^{d+1}\alpha_i p_i \tau_{i-1}\right]+p_{d+1}\left[q_1-\sum_{i=2}^{d+1}p_i\tau_{i-1}\right] \end{aligned} \tag{A1}$$

where we define $\tau_i=\prod_{j=1}^{i}q_j$, $i=1,\ldots,d-1$; and $\tau_d=(q_1\cdots q_d)/p_{d+1}$. It can be shown that the last expression in Equation A1 vanishes. Thus

$$\mathrm{DET}(\mathbf{I}-\mathbf{Q})=p_{d+1}\left[\alpha_1 p_1+\sum_{i=2}^{d+1}\alpha_i p_i\tau_{i-1}\right]. \tag{A2}$$

Next, observe that the cofactors of the first column of $(\mathbf{I}-\mathbf{Q})$ are

$$[p_{d+1},\ p_{d+1}\tau_1,\ p_{d+1}\tau_2,\ \cdots,\ p_{d+1}\tau_{d-1},\ p_{d+1}\tau_d]. \tag{A3}$$

Since this row equals the first row of the adjoint of $(\mathbf{I}-\mathbf{Q})$, the first row of $\mathbf{W}$ is expression A3 divided by the expression in Equation A2.

From the above arguments, we can write Equation 13 of Section 3.3 and the associated constraint as

$$\begin{aligned} \text{Maximize}\quad E(T|X_0=1) &= \frac{p_{d+1}\sum_{i=0}^{d}\tau_i}{p_{d+1}\left[\alpha_1 p_1+\sum_{i=2}^{d+1}\alpha_i p_i\tau_{i-1}\right]} \\ &= \frac{1}{\pi_1\left[\alpha_1 p_1+\sum_{i=2}^{d+1}\alpha_i p_i\tau_{i-1}\right]} \end{aligned} \tag{A4}$$

$$\text{subject to} \quad \frac{\sum_{i=1}^{d+1} \alpha_i \tau_{i-1}}{\sum_{i=1}^{d+1} \tau_{i-1}} = \bar{\alpha} \tag{A5}$$

where we set $\tau_0 = 1$. The constraint Equation A5 can be solved for α_{d+1} to yield

$$\alpha_{d+1} = \bar{\alpha} + \sum_{i=1}^{d} \frac{\tau_{i-1}}{\tau_d}(\bar{\alpha} - \alpha_i). \tag{A6}$$

Since π_1 does not involve the α_i's, maximizing $E(T|X_0 = 1)$ with respect to the α_i's is achieved by minimizing the bracketed expression in Equation A4. Substituting the right side of Equation A6 for α_{d+1} in Equation A4 leads to

$$\begin{aligned}
&\alpha_1 p_1 + \sum_{i=2}^{d+1} \alpha_i p_i \tau_{i-1} \\
&= \sum_{i=1}^{d} \alpha_i p_i \tau_{i-1} + \left[\bar{\alpha} + \sum_{i=1}^{d} \frac{\tau_{i-1}}{\tau_d}(\bar{\alpha} - \alpha_i)\right] p_{d+1} \tau_d \\
&= \sum_{i=1}^{d} \bar{\alpha} p_i \tau_{i-1} - \sum_{i=1}^{d} (\bar{\alpha} - \alpha_i) p_i \tau_{i-1} + \left[\bar{\alpha} + \sum_{i=1}^{d} \frac{\tau_{i-1}}{\tau_d}(\bar{\alpha} - \alpha_i)\right] p_{d+1} \tau_d \\
&= \bar{\alpha} \sum_{i=1}^{d+1} p_i \tau_{i-1} + \sum_{i=1}^{d} \tau_{i-1}(p_{d+1} - p_i)(\bar{\alpha} - \alpha_i).
\end{aligned}$$

Recalling the definition of the τ_i's, note that

$$\begin{aligned}
\sum_{i=1}^{d+1} p_i \tau_{i-1} &= p_1 \tau_0 + p_2 \tau_1 + \ldots + p_d \tau_{d-1} + p_{d+1} \tau_d \\
&= p_1 + p_2 q_1 + \ldots + p_d (q_1 \cdots q_{d-1}) + (q_1 \cdots q_d) \\
&= 1
\end{aligned}$$

by an induction argument. Thus we have proved that

$$E(T|X_0 = 1) = \frac{1}{\pi_1 \left[\bar{\alpha} + \sum_{i=1}^{d} \tau_{i-1}(p_{d+1} - p_i)(\bar{\alpha} - \alpha_i)\right]} \tag{A7}$$

and maximizing $E(T|X_0 = 1)$, for fixed p_i's, subject to Equation A5 is equivalent to the optimization problem

$$\text{Minimize} \quad \bar{\alpha} + \sum_{i=1}^{d} \tau_{i-1}(p_{d+1} - p_i)(\bar{\alpha} - \alpha_i) \tag{A8}$$

$$\text{subject to} \quad 0 \le \alpha_i \le 1; \quad i = 1, \ldots, d \tag{A9}$$

$$0 \le \alpha_{d+1} = \bar{\alpha} + \sum_{i=1}^{d} \frac{\tau_{i-1}}{\tau_d}(\bar{\alpha} - \alpha_i) \le 1. \tag{A10}$$

Because of Equation A7, we have shown

$$E(T|X_0 = 1) = \frac{1}{\pi_1 \tilde{\alpha}}$$

whenever $\alpha_1 = \alpha_2 = \ldots = \alpha_d = \alpha_{d+1}$ or $p_1 = p_2 = \ldots = p_d = p_{d+1}$. This gives the first result.

It is not difficult to show that given increasing, not all equal p_i's, one can construct a feasible solution to Equations A9 and A10. This solution takes the following form

$$\alpha_1^*, \ldots, \alpha_{k-1}^* = 1$$
$$0 \leq \alpha_k^* < 1$$
$$\alpha_{k+1}^*, \ldots, \alpha_{d+1}^* = 0.$$

By inspection of all cases for values of $k = d+1, d, \ldots, 1$, it can also be shown that a solution of this form will provide Equation A8 strictly less that $\tilde{\alpha}$. This guarantees that the objective function A8 is strictly greater than $1/\pi_1 \tilde{\alpha}$. Therefore, the maximum expected time to detection for the gaming drug user is minimized when the p_i's are equal. This establishes the minimax theorem.

Distribution List

Distribution:
OCDEP&S, Science and Testing Programs Officer
ODASD, Drug Enforcement Plans and Support
Chief of Naval Personnel (PERS-6), (PERS-63), (PERS-63E), (PERS-6C), (PERS-6E)
Chief, Bureau of Medicine and Surgery (MED-25)
Defense Technical Information Center (4)

Copy to:
Chief of Naval Personnel (PERS-00), (PERS-00B), (PERS-00W), (PERS-01JJ)
Deputy Assistant Secretary of the Navy (Manpower)
Chief of Naval Operations (N091), (N13)
Office of the Assistant Secretary of Defense (FM&P)
Armstrong Laboratory, Operations and Support Directorate (AL/DO), Brooks AFB, TX
AL/HR-DOKL Technical Library, Brooks Air Force Base, TX
President, Naval War College (Code E-112)
Superintendent, Naval Postgraduate School
Director of Research, U.S. Naval Academy