## Methods

The Data
The National Health Interview Survey (NHIS) is a representative household survey conducted annually since 1957 by the U.S. Department of Health and Human Service's National Center for Health Statistics (NCHS). In addition to a core questionnaire that measures the self-reported prevalence of various medical conditions, the use of health services, general health status, disability, and demographic characteristics, NCHS adds supplemental questionnaires on specific topics that vary from year to year. In 1982, the NHIS contained supplemental questionnaires on the use of preventive health services and the types and degree of health insurance coverage. Although NCHS has published some data from the preventive services supplement in tabular form (75), no published work to date has attempted to use these data to understand what factors are associated with the use of preventive services by the elderly. OTA decided to conduct such an analysis.

OTA obtained magnetic tapes of the core and supplemental questionnaire data for the 1982 NHIS from NCHS. These data files contained 103,923 observations reflecting respondents of all ages with 11,434 observations for individuals 65 years or older. In addition to responses to survey questionnaires, each observation contained a unique identifier, variables identifying a stratum and cluster from which it was drawn, and the weights necessary to produce representative estimates.

Among the variables on the data set, OTA was interested in:

- the amount of time elapsed since the respondent last received each of five screening services (glaucoma screening, blood pressure measurement, eye examination, breast examination, and Pap smear); and
- factors potentially associated with the use or nonuse of these five services.

OTA converted each of the five variables measuring elapsed time since use of preventive services to a binary variable that measures whether or not the individual used the service within a specified period of time.

These periods of time are based on the recommendations of expert groups presented in table 2 in the text of this paper. Because there is some variation across the different sets of expert guidelines listed in table 2, OTA summarized the published recommendations in the composite measures presented in table 8. These composite measures do not represent a set of recommendations themselves; rather, they are merely one benchmark for comparing actual use to what is generally considered adequate by recommending groups. Where there is disagreement among recommending groups, the composite measures tend toward longer intervals between screenings in order to measure compliance with minimal recommended levels of prevention. These composites of expert recommendations pertain only to the primary analyses conducted by OTA. As a source of comparison, table 8 also includes the periods used by two other studies of the use of preventive procedures. In one of these papers, the authors measured recommended periodicity as the mean of published recommendations (92). The other paper formed a consensus based on their own review of relevant literature (45).

Because of the coding scheme of the NHIS, use within an ' $x$ " year period really means that the individual had used the service within a period of less than but not including " $\mathrm{x}+1$ " years. For example, consider the case of breast examinations. Table 2 suggests that one should measure use within the previous year. However, under the NHIS coding scheme, one would consider an elderly
woman who had her last breast examination 21 months ago to have been adequately screened; a woman whose last breast examination was 24 months ago would not. Table 9 lists all variables used in the analysis and also includes appropriately weighted descriptive statistics.

## Methods of Analysis

OTA used PC-CARP ${ }^{\mathrm{R}}$ software (on an 80286 personal computer) to analyze the data. PC-CARP ${ }^{\mathrm{R}}$ was developed by the Statistical Laboratory at Iowa State University especially for analyzing data from surveys with complex designs like the NHIS. It makes use of the sampling information in the data files to produce appropriate point and variance estimates. OTA performed two separate analyses
on the data set:

- a descriptive summary of the percentages of elderly individuals who reported using each of the five screening services within the specified time; and
- a multivariate weighted logistic regression analysis of the use or nonuse of these services.

In addition, OTA:
-examined whether observations dropped from the multivariate analysis because they contained some missing information differ in any important ways from observations included in the analysis, potentially biasing our estimated parameters;

Table 8--- Periods of Time Used by OTA and Two Studies to Measure Older Adults' Use of Preventive Services

| Servi ce | Period of time to neasure use employed by: |  |  |
| :---: | :---: | :---: | :---: |
|  | OTA ${ }^{\text {a }}$ | Lillard, et al., | noo, et al., 1985 ${ }^{\text {c }}$ |
| ■Initial or periodic physi cal exam | 1 year | '''' | 1.4 years |
| $\square$ Bl ood chol esterol l evel | 5 years | ! | 4.5 years |
| ■ Fecal occult blood test | 1 year | 1 year | 1.0 year |
| - Pap snear | 3 years* | 3 years | 4. 1 years |
| ■ G aucoma screeni ng | 2 years* | . $\cdot$. | ---- |
| ■ Opt onet ry/ opht hal nol ogy exam | 2 years* | --- | --- |
| ■ Pneunococcal i mmunization | L i fet i ne | Lifetine | --- |
| ■ Infl uenza i mmunization | 1 year | 1 year | ' ' ' ' |
| $\square$ Tet anus i mmuni zation | 10 years | 10 years | -., $\cdot$ |
| - Hypert ensi on screeni ng | 1 year* | ..-. | 1. 4 years |
| - Breast exami nation | 1 year* | . ${ }^{-}$ | ... |

Abbrevi ation: HMD= health nai ntenance organization.
${ }_{b}$ Intervals listed in this column represent composites of the expert recommendations summarized in table 2.
$\mathrm{b}_{\text {L.A.Lillard, W.G. Manning, C. Peterson, et al., Preventive Medical Care: Standards Usage and Efficacy (Santa }}$ Monica, CA: The Rand Corporation, 1986).
${ }^{c}$ B. Woo, B. Woo, E.F. Cook, et al., "Screening procedures in the Asymptomatic Adult: Comparisons of Physi ci ans' Recommendations, Patients' Desires, Published Guidelines, and Actual Practice, ${ }^{\text {ti }}$ J.A.M.A. 254(11):1480-1484, 1985.
*As described in greater detail in appendices $E$ and $F$, OTAestimated the Use of preventive services among the elderly with two different data sources-a single HMO and the 1982 National Health Interview Survey (NHIS). The asterisk indicates services included in the NHIS analysis. Because of the coding scheme of the NHIS, use within an "x" year period really means that the individual has used the service within a period of less than but not including " $x+1$ " years. Under this scheme, we would consider an elderly woman who had her last breast examination 21 months ago to have been adequately screened; a woman whose last breast examination was 24 months ago would not. For the HMO data, use within "x" years carries a literal definition.
Sources: Office of Technology Assessment, 1989.

- considered whether multicollinearity in the models might reduce precision; and
- examined how the probabilities of use predicted by the models varied with each significant variable.

The multivariate analysis uses Taylor series techniques to estimate a weighted logistic regression model of the form:
$\mathrm{P}_{\mathrm{ij}}=1 /$

## where

$\mathbf{P}_{\mathrm{ij}}$ is the probability of elderly person i using service j . Use is measured by binary variable $\mathrm{Y}_{\mathrm{ij}}(=1$ if person i had service j within the specified period of time; $\mathrm{Y}_{\mathrm{ij}} \mathrm{O}$ otherwise).
$D_{i}$ is a vector of predisposing characteristics describing elderly person i.
$\mathrm{E}_{\mathrm{i}}$ is a vector of enabling characteristics describing elderly person i.
$H_{i}$ is a vector of health status characteristics describing elderly person i.
$\alpha_{\mathbf{j}}$ is an estimated parameter, and $\beta_{\mathbf{j}}$, $\delta_{\mathrm{j} *} \gamma_{\mathrm{j}}$ are vectors of estimated parameters for service j .
$\varepsilon_{\mathbf{j}}$ is an individual, service specific error term.

PC-CARP ${ }^{\mathrm{R}}$ produces estimated coefficients that are consistent and appropriately weighted. Estimates of asymptotic variances also appropriately reflect the complex survey design.

PC-CARP ${ }^{\mathrm{R}}$ requires that no observation in the data matrix contain missing values. For the logistic regression analysis, OTA used SPSS-PC+ to create two data files. OTA purged both files in a listwise fashion of observations containing missing data on any
variable in the models. ${ }^{\text {. }}$ OTA used one of the data sets to estimate models for preventive services potentially used by both sexes--glaucoma screening, eye examinations, and blood pressure measurement. OTA used the other data set, which contained only the observations for women, to estimate models for breast examinations and Pap smears.

The data set for both sexes contained exactly 9,000 out of the original 11,434 observations. The remaining 2,434 observations, which had missing data, represented a weighted 21.5 percent of the over-65 population. The single variable with the most missing observations was family income. This variable alone had missing observations representing 15.3 weighted percent of the elderly population. Of each of the other variables containing missing values, none lacked data on observations representing more than 4 weighted percent of individuals over-65. The data set containing only women had 5,040 observations out of a possible 6,655 . The 1,615 observations with missing data represented 19.6 weighted percent of all women over 65.

## Results

The estimates of the national proportions of elderly using each of the five screening services within the specified time are presented in table 3 in the text of the paper. Additional descriptive statistics are presented in table 9. Table 10 below presents the parameters of the estimated logistic regressions that attempt to explain the use or nonuse of each service. Table 7 summarizes these results, and the text of the paper discusses their significance.

[^0]Table 9--- 1982 National Health Interview Survey: Selected Descriptive Statistics for Persons Over 65


SOURCE: Office of Technol ogy Assessnent, 1989.

Table 10. --Elderly Use of Five Screening Services: Logistic Regression Results


Asymptotic standard errors are in parentheses below each estimated coefficient.
Varíable neans were calculated from data natrices used to estimate logit nodel s (i.e., purged of observations with missing val ues).
CAverage "effect" of complex survey design on variances of estimated coefficients. This "effect" is neasured as the number of tines greater the variance from the complex design is than the variance from a simple random desi gn.
*Estimated paraneter si gnificant at the 0.05 level, two-tailed test.
**Estimated parameter significant at the 0.01 level, two-tailed test.
Vari able Key:

| Dependent Vrriables: |  |
| :---: | :---: |
| GLAUCOMA | - $1=$ screened for glaucom within previ ous 2 years, 11 nonths; 0=otherwise |
| EYE | --I =recei ved eye exam nation within previ ous 2 years, 11 nonths; 0=otherwise |
| BLOODP | $\cdots$ - $=$ had blood pressure neasured within previ ous 1 year, 11 nonths; 0=otherwise |
| BREAST | --1=had clinical breast examination within previ ous 1 year, 11 nonths; 0=0therwise |
| PAP | --I had Pap snear within previ ous 3 years, 11 nonths; 0=otherwise |
| Independent Variablea |  |
| MALE | --I =nal e~ Ofrenal e |
| AGE | --respondent's age in years |
| NONWH TE | --I =nonwhite; O=white |
| EDLEVEL | -- hi ghest educational level attai ned; O=none or ki ndergarten; 1=1 to 8 years (el enentary); 2=9 to 11 years ( $\mathrm{sp}, \mathrm{e}$ high school); $3=12$ years (high school graduate); $4=1$ to 3 years college; $5=$ college graduate; $6=$ post-graduate education |
| I NCOME | --fanily incone in dol lars |
| SMEA | --I=resides in a Census Bureau Standard Metropolitan Statistical Area (urban area); O=otherwise |
| PREPAI D | --I =enrolled in HMD or sone other prepaid heal th plan; O=otherwise |
| HLTH NSR | $--I$ has sone heal th insurance coverage or heal th benefits beyond Medi care incl uding prepaid, Veterans', military, or neans tested public assistance health benefit; $0=0$ therwise |
| ALONE | $-\mathrm{I}=\mathrm{i}$ ves al one; $0=0$ herwi se |
| BEDDAYS | -- days in bed during previ evs 12 nonths; O=none; 1=1 to 7 days; 2=8 to 30 days; 3=31 to 180 days; $4=181$ to 365 days |
| LI M TED | --İlimited in some activity; 0=otherwi se |
| SORCE: | i ce of Technol ogy Assessment, 1989 |

Table 11. --Correlation Matrix for Variables in Logistic Regression Models ${ }^{a}$

|  | G_AUCOMA | BLOODP | P EYE | BREAST | PAP | MALE | AGE | NONWH TE | ED. <br> LEVEL | I NCOME | E SMSA | PREPA D | $\begin{aligned} & \text { HLH - } \\ & \text { I NSR } \end{aligned}$ | ALONE | BEDDAYS | LIM TED |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| G_AUCOMA | 1.00 |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
|  | x |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| BLOODP | 0. 19 | 1. 00 |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
|  | xX | xX |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| EYE | 0. 69 | 0. 18 | 1. 00 |  |  |  |  |  |  |  |  |  |  |  |  |  |
|  | x x | xX | xx |  |  |  |  |  |  |  |  |  |  |  |  |  |
| BREAST | XX | XX | XX | XX |  |  |  |  |  |  |  |  |  |  |  |  |
|  | xx | xx | XX | 1. 00 |  |  |  |  |  |  |  |  |  |  |  |  |
| PAP | XX | XX | xx | xX | xx |  |  |  |  |  |  |  |  |  |  |  |
|  | xX | XX | xX | 0.49 | 1. 00 |  |  |  |  |  |  |  |  |  |  |  |
| MALE | -0. 06 | -0. 04 | -0. 05 | XX |  | 1. 00 |  |  |  |  |  |  |  |  |  |  |
|  | x x | x x | xX | xX | xX | 1. 00 |  |  |  |  |  |  |  |  |  |  |
| AGE | -0. 05 | 0. 03 | -0. 03 | XX |  | -0.06 | 1. 00 |  |  |  |  |  |  |  |  |  |
|  | x x | xx | XX | -0. 06 | -0. 15 | 5 XX | 1. 00 |  |  |  |  |  |  |  |  |  |
| NONHH TE | -0. 09 | -0. 02 | -0.03 | xX | xX | 0. 00 | -0.01 | 1. 00 |  |  |  |  |  |  |  |  |
|  | xx | xx | xX | -0. 02 | 0. 00 | xx | -0. 01 | 1. 00 |  |  |  |  |  |  |  |  |
| EDLEVEL | 0. 15 | 0.05 | 0. 10 | xx | xx | 0. 01 | -0.14 | -0.16 | 1. 00 |  |  |  |  |  |  |  |
|  | xx | xx | xx | 0. 11 | 0. 11 | xX | -0.14 | -0. 18 | 1. 00 |  |  |  |  |  |  |  |
| I NCOME | 0. 12 | 0. 04 | 0.07 | xX | xx | 0. 10 | "0.11 | -0.13 | 0. 39 | 1. 00 |  |  |  |  |  |  |
|  | XX | XX | xx | 0.09 | 0. 10 | x | -0. 09 | "0.13 | 0. 32 | 1. 00 |  |  |  |  |  |  |
| SMSA | 0. 07 | 0.01 | 0. 06 | XX | xx | -0. 01 | -0.00 | 0.04 | 0. 09 | 0. 12 | 1. 00 |  |  |  |  |  |
|  | XX | XX | xx | 0. 09 | 0. 05 | XX | 0.01 | 0.02 | 0. 08 | 0. 11 | 1. 00 |  |  |  |  |  |
| PREPAI D | 0.03 | 0. 02 | 0. 02 | XX | xX | 0. 03 | -0.03 | -0. 00 | 0. 04 | 0. 03 | 0. 10 | 1. 00 |  |  |  |  |
|  | xX | xx | xX | 0. 04 | 0. 04 | xX | -0. 03 | -0. 01 | 0. 03 | 0. 01 | 0. 08 | 1. 00 |  |  |  |  |
| HLTH NSR | 0. 12 | 0. 07 | 0. 09 | xX | xx | 0. 02 | -0.09 | -0.12 | 0. 18 | 0. 11 | 0.01 | 0. 08 | 1. 00 |  |  |  |
|  | xX | x x | xx | 0. 10 | 0. 09 | x | -0. 08 | "0. 11 | 0. 18 | 0.07 | 0. 00 | 0.08 | 1. 00 |  |  |  |
| ALONE | -0. 01 | -0.01 | 0. 01 | XX | XX | -0. 26 | 0. 18 | 0.01 | -0. 02 | -0. 31 | -0.00 | -0. 04 | -0.01 | 1. 00 |  |  |
|  | xx | xX | XX | 0.01 | -0.06 | XX | 0.21 | "0. 01 | -0. 00 | -0. 36 | -0.01 | -0.03 | 0.03 | 1. 00 |  |  |
| BEDDAYS | 0. 04 | 0. 13 | 0.02 | XX | xx | -0. 02 | 0.06 | 0.04 | -0. 05 | -0.03- | -0.01 | -0. 02 | -0.01 | 0.01 | 1. 00 |  |
|  | xx | xx | XX | 0. 15 | 0. 08 | $x X$ | 0.06 | 0.04 | -0. 05 | -0.04 | -0.00 | 0. 00 | -0. 02 | 0. 01 | 1. 00 |  |
| LI M TED | -0.00 | 0. 06 | -0.03 | 3 xx | xX | -0.04 | -0.07 | 0.06 | -0. 12 | -0.07 | -0. 04 | -0.00 | -0. 02 | -0.03 | 0.31 | 1. 00 |
|  | XX | xX | xX | 0. 04 | 0. 03 | XX | -0. 04 | 0.07 | -0. 12 | "0.05 | -0. 03 | 0. 01 | -0.01 | -0. 04 | 0.45 | 1. 00 |



To examine the possibility that multicollinearity among the independent variables in the model might preclude precise estimation, OTA estimated the weighted first-order Pearson correlation matrices for the two data sets. Table 11 presents the correlation statistics. Only three pairs of independent variables had correlations greater than 0.25 (or less than -0.25): EDLEVEL and INCOME ( 0.39 in the two-gender data set and 0.32 in the women only data set), ALONE and INCOME ( -0.31 and -0.36 respectively), and the two measures of health status, LIMITED and BEDDAYS (0.31 and (). 43 respectively). In addition, ALONE and AGE have correlations of 0.18 and 0.21 respectively in the two data sets. However, despite the potential effect of this collinearity on the estimated variances, the conclusions are unlikely to change. In all models except blood pressure measurement, EDLEVEL and INCOME are both already significant predictors of preventive service use. ALONE is significant in three out the five (with blood pressure measurement and Pap smears being the exceptions). Although there is a high degree of correlation between BEDDAYS and LIMITED, at least one of them is statistically significant in all of the models except glaucoma screening, thus supporting the notion that health status is associated with preventive service use among the elderly. ${ }^{\text {² }}$

OTA excluded a substantial proportion of observations because data were missing for one or more variables in the model. In order to examine if these exclusions could have biased the results of the multivariate models, OTA compared the characteristics of the included and excluded groups. In both data sets, the included respondents were significantly different from those eliminated be-

[^1]cause of missing data for only two variables: HLTHINSR (the presence of any health insurance beyond Medicare) and INCOME. ${ }^{3}$ For each of the other variables (including the dependent variables), the mean for the observations with missing data did not differ statistically from the mean for observations included in our analyses. This analysis suggests that the exclusion of observations with missing data is unlikely to introduce bias into the multivariate models, but OTA cannot rule out the possibility.

In order to examine the effect of each significant variable in the estimated models, OTA simulated, one independent variable at a time, how the probability of using each screening service varied with each possible value of the independent variables. In these simulations, all independent variables, except the one whose effect was being simulated, assumed their mean values.

Table 12 and figures 1 through 4 present the results of this analysis for each significant variable in our models. Among the independent variables, holding other factors constant, age, education, and health insurance appear to have the greatest overall effect on the probability of receiving each of these services. This analysis also supports the notion that blood pressure measurement is different from other services. Since almost everyone receives it, there is less variation to explain. Hence, the variables in the model appear less important in predicting its use than they do for the other services.

3 The group of observations excl uded from the anal ysis had a lower mean income ( $\$ 14,475$ versus $\$ 15$, 276 in the two gender data set; $p<0.01$ ) and was less likely to have any insurance coverage beyond Medicare ( $\mathbf{0 . 7 0}$ versus 0.79 in the two gender data set; p<0.05) than was the included group. The in come statistic may not accurately reflect the entire group of observations with missing data since three-quarters of the observations missing any data at al 1 did not have income data.

Table 12--- Effect of Statistically Significant Binary Variables ${ }^{\text {a }}$ in Logistic Regressions on Elderly Use of Five Screening Services: Predicted Probabilities ${ }^{\text {b }}$

${ }^{\text {a Effect of significant non-bi nary variables shown in figures } 1 \text { through } 4 .}$
${ }^{\text {b }}$ Predicted probability is estimed as $1 /\left[1+e^{-X \beta}\right]$ where $B$ is the vector of estimated coefficients and $X$ is the vector of individual characteristics. Of these characteristics (all independent variables included in the estimated model), each takes on its mean value except the one designated in that row of the table above; it takes on the value shown in the row header.

Symbol Key:
XX=1 ndependent variable not included in nodel
-- Estimated coefficient on independent variable not significant at 0.05 level, tuo-tailed test
Variabl e Key:
Dependent Variables:
G_AUCOMA --I =screened for glaucoma within previous 2 years, 11 months; 0=otherwise
EYE $\quad-\quad$ I frecei ved eye examination within previ ous 2 years, 11 nonths; O=otherwise
BLOOOP --I =had blood pressure neasured within previ ous 1 year, 11 nonths; $0=0$ othervise
BREAST --I had clinical breast examination within previ ous 1 year, 11 nonths; $0=0$ herwise
PAP $\quad-\quad$ I had Pap snear within previ ous 3 years, 11 nonths; $0=0$ otherwi se
Independent Variables:
MALE $\quad$--I=male; $0=$ female
NONW TE --I =nonwhite; O=white
SMSA --I=resides in a Census Bureau Standard Metropolitan Statistical Area (urban area); O=otherwise
PREPA D --I =enrolled in HMD or sone other prepaid health plan; O=othervise
HLTH NSR --I has sone health insurance coverage or health benefits beyond Medicare including prepaid, Veterans!, military, or means tested public assistance health benefit; $0=0$ otherwise
ALONE $\quad-I=1$ ives al one $O=o t h e r w i s e$
LIMTED --IIImited in some activity; $0=0$ otherwise
SOURCE: Office of Technol ogy Assessnent, 1989.


$\begin{array}{llll}1-7 \text { Days } & 8-30 \text { Days } \quad 31-180 \text { Days } & \text { 181-385 } \\ \text { Bed Days in Last } 12 \text { Months } & \end{array}$

- olencoma screoning - Blood Prosiure $\rightarrow$ Eyo Exam
alaucoma screoning - Blood Preas
- Brent Exam
Key to Figure 3: Elementary $=1$ to 8 years education; Some HS $=9$ to 11 years education; HS Grad $=12$ years education; Some $\mathrm{Coll}=1$ to 3 years college; Coll Grad = college graduate; Post-Col = post-graduate education.
SOURCE: Office of Technology Assessment, 1989.


[^0]:    1 To estimate the descriptive statistics presented in table 9 and the national percentages of elderly receiving the five screening services, OTA used data sets that contained all observations for which any data is available for the particular variable in question. The Listwise deletion of missing values described here only applies to the logistic regression analysis.

[^1]:    Z However, because multicol 1 i neari ty reduces the precision of the estimator, the standard error of these two variables' estimated coefficients may be biased in the glaucoma screening model.

