Evaluation of Smoking Related Nursing Home Expenditures in Florida

Prepared by:

Dr. Tim Lynch
Director
Center for Economic Forecasting and Analysis
Florida State University
2035 E. Paul Dirac Dr.
Rm. 130
Herb Morgan Building
Innovation Park
Tallahassee, Florida, 32310
http://www.cefa.fsu.edu

and:

Richard S. Hopkins, M.D. M.S.P.H.
State Epidemiologist
Disease Control
Florida Department of Health Tallahassee, Florida, 32399

January, 1999

Survey and Research Assistance
Dr. Robert Hays
Haysmar and Associates, Inc.
And
Ms. Susan Flaherty
Department of Finance
Florida State University

Funded by:
Robert Wood Johnson Foundation
The Policy Substance Abuse Research Program

Acknowledgements

This two year $100,000 research project was funded by the Robert Wood Johnson Foundation. Substance Abuse Research Program. The Project managers are David G. Altman, Ph.D., National Program Director, and Ms. Andrea L. Ebbers, MA, Department of Public Health Sciences, Wake Forest University, School of Medicine, Medical Center Boulevard, Winston-Salem, N.C., 27157-1063. The study examines tobacco related nursing home costs in Florida with special attention to Medicaid expenditures.

Researchers at the Center for Economic Forecasting and Analysis at Florida State University jointly conducted this study with researchers from the Bureau of Epidemiology, Florida Department of Health. All conclusions, omissions, errors are the responsibility of the authors and not of the Robert Wood Johnson Foundation, reviewers or supporting agencies.

The full text of this report can be found at ***www.cefa.fsu.edu*** the Web site for the Center for Economic Forecasting and Analysis at Florida State University, Tallahassee, Florida. All comments and questions can be forwarded to Thomas (Tim) Lynch, Ph.D., at ***tilynch@garnet.acns.fsu.edu***.

The authors wish to express special thanks to the staff of the State Center for Health Statistics, Agency for Health Care Administration, State of Florida, Tallahassee Florida. Special thanks to Dr. Fred Roberson, Mr. James Tillery and Mr. Randy Baker and the other staff member that supported this effort. Special thanks also goes to Dr. James Howell, former Director of the Division of Health Policy, AHCA, and Secretary of the Florida Department of Health for his encouragement and support throughout this study and to Dolores Katz, Ph.D. and Ursula Bauer, Ph.D., staff in the office of the State Epidemiologist.

Executive Summary

This two-year $100,000 research project funded by the Robert Wood Johnson Foundation examines tobacco related nursing home costs in Florida with special attention to Medicaid expenditures. Researchers at the Center for Economic Forecasting and Analysis at Florida State University jointly conducted the study in cooperation with the Chief of the Bureau of Epidemiology, Florida Department of Health.

While national models have been developed to estimate state-specific costs of smoking, they estimate nursing home costs only indirectly, generally assuming they are the same as the proportion of hospital costs associated with smoking. Nursing home costs cannot currently be estimated directly because no information exists on the smoking histories of nursing home residents. All national and state smoking surveys done to date have included non-institutionalized adults only, as has the (earlier) National Medical Expenditures Survey.

This study helps resolve this knowledge gap by interviewing a representative sample of Florida nursing home residents about their smoking habits and related demographic data and evaluated these factors (and their smoking status) relative to each client's length of stay. From a stratified random sample of Florida nursing homes, we selected 721 residents who had their first nursing home admissions during the first half of 1995 and interviewed them about their smoking histories before admission and their education, marital status, former occupations, and household income. Using the results of the survey and data from the Florida Behavioral Risk Factor Surveillance Survey, we calculated odds ratios for nursing home admission, comparing current and former smokers to never-smokers by sex and five-year age group cohorts.

From each nursing home we also obtained information on each patient's length of stay over the first thirty months after admission, and measured nursing home length of stay for each surveyed resident. The central question this analysis evaluates is the nature of the relationship between smoking status (former, current and never and then ever and never) and nursing home length of stay.

A. Summary and Conclusions

Section 1 Summary

Section 1 overviews the current status of tobacco damage, health research in the US and set forth the need for the analysis and the null hypothesis this we focused on in the research that follows in Sections 2 and 3 of this report.

Section 2 Summary

Section 2 overviews the survey sample data secured for this analysis and a preliminary comparison of the sample data to the population from which it was drawn. Section 2 thereafter completes an extensive profile of the age, gender, regional dispersion, and length of stay for the sample population. Comparisons and contrasts of the average lengths of stay of ever and never smokers by gender and age are then completed. This is followed by a case study evaluation of the rate of admission to nursing homes (from our sample) of ever and never smokers to the Florida population at risk.

This initial analysis comparisons between both the population at risk and the nursing home sample indicates that:

- Never and ever smokers tend to enter nursing homes in comparable proportions to their numbers in the population at risk.
- There are more ever smokers in the relatively younger age cohorts and more never smokers in the older cohorts in both the population at risk and the nursing home sample
- There are fewer ever smokers among females than males in our sample and the population at risk
- Higher numbers of females are resident in both the population at risk and in the nursing homes sample in virtually every age group evaluated

Conclusions based on the Nursing Home Sample Exclusively

- On average, ever smokers tend to die at much younger age cohorts then never smokers among those nursing home residents that die.
- Never smokers tend to experience longer lengths of stay among the older age cohorts
- Ever smokers tend to have longer average and total lengths of stay in nursing homes in age cohorts up through age 79.
- Survival analysis testing of differences between the ever and never smokers average length of stay in different cohort is only statistically significant in the 55-59 (ever smokers stayed longer) and 95-105 age cohorts (never smokers stayed longer).
- The average difference in LOS for virtually every other age cohort, (except 85-89) while not statistically significant, is still considerable with ever smokers staying far longer than never smokers.
- When the nursing home survey clients are segmented by age into three and then two large cohorts, the ever smokers clearly have longer lengths of stay then never smokers. However, these differences are not at a statistically significant level.

Section 3 Summary

Section 3 examines both the linear and Cox Regression models developed in this study to evaluate the relationship between ever smoking and lengths of stay and age at discharge. The conclusions of this analysis indicate that the following key independent variables are statistically associated in linear regression models with longer lengths of stay.

Regression Model Conclusions

- Longer are associated with:
  1. The average number of cigarettes smoked per day for both those over and under 80 years of age
  2. The average number of cigarettes smoked and number of years smoking.

- Shorter lengths of stay are associated with:
  1. A higher number of years married for all ages.

- Younger age at discharge (dead or alive) is associated with:
  1. A larger number of years not smoking
  2. A larger number of years married

3. The average number of cigarettes smoked per day (perhaps most associated with the residents discharge dead)
4. A larger number of years not smoking.

Cox Model Conclusions

- The hazard of discharge for residents who have never been married is higher than that of residents who have been married residents (i.e. ever married patients tend to have a higher probability of discharged).
- Residents in large facilities experience a higher hazard of discharge relative to residents in smaller facilities.
- The hazard of discharge for male patients is about 48% of the hazard of discharge for female patients (i.e. males are less likely to be discharged relative to females and thus incur longer lengths of stay).
- The hazard of discharge is slightly reduced for residents that have consumed a greater number of cigarettes in their lifetimes.
- The hazard of discharge is slightly increased for patients who have smoked for a longer number of years.

Medical theory would indicate that people who had ever smoked should be more ill on admission to the nursing home, and therefore die sooner after admission than never-smokers. We would expect survival till death to be shorter in ever-smokers than in never-smokers.

- Results from our analysis indicate ever smokers (that died in nursing homes in our sample) had a higher hazard of discharge, or shorter length of stay, than never-smokers. The level of statistical significance is .18 (less than the traditional level of statistical significance of .01, .05 or .1). However the value of the hazard function for ever-smokers is 1.31 times as high as for never-smokers – that is, ever-smokers are more likely to die in nursing homes.

In the second analysis, persons known to have died while in the nursing home are excluded from the analysis. Time is measured till discharge. If the subject is still in the nursing home at the time of the study, his/her observation is considered to be censored. This measures the relationship between smoking status and hazard of discharge, or length of stay, for patients who did not die while in the nursing home.

- The negative sign of the coefficient for the smoking variable indicates that ever-smokers have a reduced hazard of discharge, or longer average length of stay, relative to never-smokers. The level of statistical significance is .26 (less than the traditional level of statistical significance of .01, .05 or .1). The value of the hazard function for ever-smokers is 0.89 times as high as for never-smokers – that is, ever-smokers are *less likely to be discharged, and thus incur longer lengths of stay*.

Cox Analysis Summary

In conclusion, neither of these analyses indicates a strong or statistically significant effect of smoking. However in both analyses, the directions of the observed effects are those predicted by the known health effects of tobacco use: ***an increased risk of early death after admission to nursing home and a longer length of stay for tobacco-users who survive***. As these two effects

tend to cancel each other out, the relatively small magnitude of the observed effects of smoking on length of stay is not unexpected, when persons discharged dead or alive or still living in the nursing home are considered together.

The Estimated Higher Medicaid Nursing Home Expenditures for Ever Smokers than for Never Smokers and Future Research Needs

Summary of Findings of Higher Tobacco Related Medicaid Nursing Home Expenditures in Florida

Table 1 provides a profile of the number of ever and never smokers by age cohort and their respective average length of stay secured from our survey. The average LOS for *ever* smokers is longer in every age category except for the 85 to 89 and the 95+ age group. These are at the age when never smokers tend to out number ever smokers two to one. Ever smokers in NH tend to be younger on admission, experience longer lengths of stays and die at a younger age than never smokers.

Our survey secured meaningful length of stay data from 349 never smokers and 327 ever smokers. This count includes all individuals that were still in the nursing home at the time of the survey. The never smokers reported a total length of stay of 76,079 days while the many fewer ever smokers report 73,552 days. Although there are 6.7% more never smokers than ever smokers, their respective total lengths of stay is virtually identical. Table 1 provides an analysis that weights the relative lengths of stay to the actual proportion of persons admitted to nursing homes by age. This analysis demonstrates that the longer LOS for smokers among the younger age groups, who are more numerous, outweighs the longer LOS for the ever smokers in some of the older groups (where ever smokers are less numerous). As Table 1 indicates the cumulative LOS for the *numerically fewer* ever smokers is higher across each age cohort until the very final 95+ age group. The final comparison demonstrates that in this oldest age grouping (where never smokers out number ever smokers 2.6 to 1) the larger population of never smokers finally dominates the reported length of stay totals.

The bottom of Table 1 presents an estimate of the lower lengths of stay that ever smokers would have reported if they had experienced the average EVER smokers average length of stay. Restated, if ever smokers had stayed in the nursing home, on average, the same number of days as never smokers, the total length of stay for the entire sample population would be 10,781 days fewer in Florida nursing homes. These excess days amount to 7.21% of all the days incurred by both ever-smokers and non-smokers in our population. Our population is highly representative of all Florida nursing home residents with respect to age and length of stay, so we believe these estimates can be generalized to all nursing home admissions.

In 1997, the Florida Medicaid system expended $1.43 billion for Florida nursing home expenditures. If one assumes that length of stay is a direct measure of expenditure, then ***ever smokers generate excess expenditures of $103.2 million, or 7.21% of the total $1.4 billion in nursing home expenditures during 1997***.

Table 1

Estimated 1997 Smoking Related Nursing Home Costs to State Medicaid Budget

Never Smokers - Number of Residents and Length of Stay

Age Cohort	Average LOS Never	Numbers of Never Smokers	Cumulative Number Never Smokers	Cumulative Percent Never Smokers	Total LOS Never Smokers	Total Cumulative LOS Never Smokers
Age Cohort: 1-54		1	1	0.3%	-	-
Age Cohort: 55-59	31	4	5	1.4%	124	124
Age Cohort: 60-64	343	6	11	3.1%	2,058	2,182
Age Cohort: 64-69	31	18	29	8.3%	558	2,740
Age Cohort: 70-74	151	36	65	18.6%	5,436	8,176
Age Cohort: 75-79	189	60	125	35.7%	11,340	19,516
Age Cohort: 80-84	162	74	199	56.9%	11,988	31,504
Age Cohort: 85-89	241	70	269	76.9%	16,870	48,374
Age Cohort: 90-94	273	57	326	93.1%	15,561	63,935
Age Cohort: 95+	506	24	350	100.0%	12,144	76,079
		350			**76,079**	

Ever Smokers - Number of Residents and Length of Stay

Age Cohort	Average LOS Ever	Ever	Cumulative Total Ever	Cum Percent	Total Ever LOS by Age Group	Total Cumulative LOS Ever Smokers	Cumulative Difference Never-Ever Smokers
Age Cohort: 1-54	642	4	4	1.2%	2,568	2,568	2,568
Age Cohort: 55-59	493	3	7	2.1%	1,479	4,047	3,923
Age Cohort: 60-64	422	14	21	6.4%	5,908	9,955	7,773
Age Cohort: 64-69	125	22	43	13.1%	2,750	12,705	9,965
Age Cohort: 70-74	239	59	102	31.2%	14,101	26,806	18,630
Age Cohort: 75-79	208	67	169	51.7%	13,936	40,742	21,226
Age Cohort: 80-84	172	71	240	73.4%	12,212	52,954	21,450
Age Cohort: 85-89	224	54	294	89.9%	12,096	65,050	16,676
Age Cohort: 90-94	328	24	318	97.2%	7,872	72,922	8,987
Age Cohort: 95+	70	9	327	100.0%	630	73,552	(2,527)
		327			**73,552**		

Length of Stay of Ever Smokers If they Stayed as Long As Never Smokers

Age Cohort	Ever Smokers With Never Smoker LOS Avg	Ever Smokers With Their LOS Avg	Difference	
Age Cohort: 1-54	-	2,568	2,568	
Age Cohort: 55-59	93	1,479	1,386	
Age Cohort: 60-64	4,802	5,908	1,106	
Age Cohort: 64-69	682	2,750	2,068	
Age Cohort: 70-74	8,909	14,101	5,192	
Age Cohort: 75-79	12,663	13,936	1,273	
Age Cohort: 80-84	11,502	12,212	710	
Age Cohort: 85-89	13,014	12,096	(918)	
Age Cohort: 90-94	6,552	7,872	1,320	
Age Cohort: 95+	4,554	630	(3,924)	
TOTALS	62,771	73,552	10,781	
				10,781
PERCENT OF TOTAL LOS FOR ALL SURVEYED NH RESIDENTS			**7.21%**	

MEDICAID NURSING HOME SPENDING FOR 1997	$	1,432,000,000
ESTIMATED 1997 SMOKING RELATED NURSING HOME COSTS TO STATE MEDICAID BUDGET	**$**	**103,176,427**

Table of Contents

List of Tables

List of Figures

Section 1. Introduction Literature Review and Research Design

A. Introduction

Much of the estimated $6 billion spent annually on tobacco-related illnesses in the state of Florida is borne directly by the people of Florida in the form of higher Medicaid costs. Since at least 20% of Medicaid expenses are for nursing home care, it is important to clarify the impact of cigarette smoking on Medicaid (and non-Medicaid) related nursing home expenses.

Smoking could increase Medicaid and general nursing home expenditures in several ways: smokers may (1) enter nursing homes earlier than non-smokers; (2) have more complications and more complex illnesses than non-smokers and thus incur more expenditures.

While national models have been developed to estimate state-specific costs of smoking, they estimate nursing home costs only indirectly, generally assuming they are the same as the proportion of hospital costs associated with smoking. Nursing home costs cannot currently be estimated directly because no information exists on the smoking histories of nursing home residents. All national and state smoking surveys done to date have included non-institutionalized adults only, as has the (earlier) National Medical Expenditures Survey.

In this study helps resolve this knowledge gap by interviewing a representative sample of nursing home residents about their smoking habits. From a stratified random sample of Florida nursing homes, we selected 721 residents who had their first nursing home admissions during the first half of 1995 and interviewed them about their smoking histories before admission and their education, marital status, former occupations, and household income. Using the results of the survey and data from the Florida Behavioral Risk Factor Surveillance Survey, we calculated odds ratios for nursing home admission, comparing current and former smokers to never-smokers by sex and five-year age group cohorts.

From each nursing home we also obtained information on each patient's length of stay over the first thirty months after admission, and measured nursing home length of stay for each surveyed resident. The central question this analysis evaluates is the nature of the relationship between smoking status (former, current and never and then ever and never) and nursing home length of stay. Essentially researchers posited the questions:

- Are nursing home lengths of stay (costs) for current and ever smokers longer than for never smokers?
- Controlling for age, are ever smokers more likely to be admitted to nursing homes than never smokers?
- Do current or ever smokers tend to die sooner after admission in nursing homes?

We employed a number of statistical techniques to extensively evaluate the survey data to answer these questions and evaluate the relationship between smoking status and residents respective length of stay where this variable is viewed as a proxy for cost. The evaluation and discussion that follows responding to these inquiries present a systematic analysis of the data, the results of analysis at various stages in the evaluation and the final conclusions as well as study limitations and recommendations. Ultimately we believe the results of this work will assist other evaluators to develop more accurate models of tobacco-related health care costs for the elderly and expand our knowledge about the impact of smoking on the health of older Americans.

B. Review of the Literature

Almost 2.5 million (23.7%) adult Floridians smoke more than 1.3 billion packs of cigarettes per year, resulting in 28,000 deaths.[1] One of every five deaths among Florida residents is related to tobacco use. The overall economic costs of smoking-related death and dying are well documented: annual costs attributable to smoking in Florida are close to $6 billion.[2] Floridians bear a large proportion of this cost is directly in the form of Medicaid payments for treatment of smoking-related illnesses. The overall costs of smoking to society (number of days of illness, lost wages, lost productivity and so forth) are well known. Researchers have also been successful in evaluating and documenting a number of smoking related health care costs.

For example researchers are able to accurately estimate smoking related health care costs for health services provide by doctors, acute care hospitals, and for drug related expenditures.[3] These cost estimates rely heavily on national surveys of the *non-institutionalized*. Therefore, by definition the institutionalized populations, such as this nation's nursing home residents, are systematically excluded from any national or regional tobacco related health care cost evaluation. This leaves a large gap in our knowledge about tobacco related damage costs to a relatively large segment of the American population- the nursing homebound elderly.[4]

This issue is especially problematic since these residents are usually quite old and typically demand considerably more medical services support. These residents therefore accrue much higher medical costs than either the average American or their non-institutionalized age cohort peers. This suggests the possibility that as tobacco related health care costs are considerably underestimated for these nursing home clients these costs are not accurately captured in existing national models.

---

[1]Joyner-Sims, J. Opposition to tobacco use: a nonpartisan public health issue. J Fla Med Assoc.; 1996; 83: 79-80

[2] Lynch T., Hopkins R. Estimating tobacco-related health care and mortality costs in Florida. J Fla Med Assoc.; 1996; 83: 123-132.

[3] See for example expanded descriptions of the currently available SAMMAC models summarized later in this report and more detailed summaries described in Hopkins, R.,-Lynch,T., "Final Florida Estimates of Tobacco Related Medicaid Costs 1995-1997", Provided to the Florida First District Court, May,1997 (Unpublished)

[4] Current estimates indicate that over 2 million Americans are projected to be nursing home residents in 1999 with almost 95,000 residents in Florida alone Source: "Facts and Trends: The Nursing Facility Sourcebook, American Health Care Association (AHCA), 1996.

This may especially be the case in Florida with its very large population of older retirees. Initial State estimates indicate a minimum of 20% of Medicaid expenditures are for nursing home care.[5] Smoking could increase Medicaid nursing home costs in three ways:

- Smokers may enter nursing homes at a younger mean age than non-smokers, and therefore require more Medicaid resources over their lifetimes since they start earlier.
- Because smoking may interfere with earning capacity, smokers may exhaust their resources faster than non-smokers, and consequently become eligible for Medicaid earlier in their illnesses.
- Smokers may have more complications and more complex illnesses that require higher levels of care.

While national models have been developed to estimate smoking-related health care costs for various client groups (such as Medicaid recipients), they do a poor job of capturing nursing-home related costs. The existing models use the National Medical Care Expenditures Survey (NMES) and the National Health Interview Survey (NHIS) to link smoking status (current, former, and never smokers) with health status and then with costs associated with hospitalization, physician visits, and prescription drugs.[6]

Researchers at the University of California-Berkeley School of Public Health have combined these models with state-specific estimates of smoking prevalence and Medicaid expenditures to produce estimates of smoking-related Medicaid expenditures for each state[7]. Analysts from the Florida Department of Health have worked with the California researchers to refine these models, which they agree do not adequately capture nursing home cost differences between smokers and non-smokers. The models have three primary deficiencies.

[5] Personal communication, Tony Swinson, Agency for Health Care Administration, March 1996
[6] Ibid., Hopkins-Lynch, 1997
[7] Miller, L., Zhang,X.,Rice,D.,Max,W.,Novotny,T: State Estimates of Direct Medical-Care

First, we do not know what proportion of nursing home residents were smokers or former smokers at the time of admission. The NHIS, which surveys a stratified random sample of non-institutionalized U.S. residents, does not include nursing home residents, and neither do the state surveys that provide state-specific estimates of smoking prevalence.

Second, the NMES survey estimates only hospital expenditures, not nursing home expenditures. The national model that estimates Medicaid costs uses the hospital expenditure estimates for ages 65 and over to estimate the *proportion* of nursing home costs that are attributable to smoking. Florida economists believe that the use of hospital-based estimates may significantly underestimate actual smoking-related Medicaid expenditures. They point out that nursing home costs have been rising very fast in recent years; per diem expenses for nursing home care have increased by more than 80% in the past six years.[8]

Moreover, Florida has the highest proportion of nursing home beds in for-profit institutions of any state (85.4% compared to 57.5% nationwide) and costs tend to be higher in for-profit facilities. In addition, the structure of nursing home costs is very different from that of hospital costs. Because of a growing emphasis on restorative care in nursing homes, nursing home residents have a high proportion of their costs in ancillary services such as physical therapy and other rehabilitative services. Total ancillary expenses per nursing home patient increased from 7.1% of patient care expenses in 1987 to 16.7% in 1992 and have increased by over 300% in the past five years. [9]

Finally, assuming that smokers become ill earlier than non-smokers (a reasonable assumption since smokers do not live as long as non-smokers), the models do not take into account the higher value of earlier expenditures for smokers compared to later expenditures for non-smokers.

Expenditures Attributable to Cigarette Smoking, University of California-Berkeley School of Public Health, July, 1997

[8] Florida Nursing Home Annual Financial Report, 1997, Florida Agency for Health Care Administration (AHCA)

[9] Ibid., AHCA, 1997

C. Literature Review: Estimating the Economic Costs of Smoking

The first widely used approach to estimating the direct and indirect economic impact of smoking was the Smoking-Attributable Mortality, Morbidity and Economic Cost (SAMMEC)[10] model originally developed under a CDC contract, by the Minnesota Department of Health in 1984. The U.S. Center for Disease Control and Prevention (CDC) has adopted this model and provides software and technical support to states to annually calculate national and state specific estimates of financial and economic costs attributable to smoking.

This model estimates the fraction of deaths attributable to smoking calculated for 27 separate medical diagnoses that have been clearly shown to be associated with smoking. This fraction, the Smoking-Attributable Fraction or SAF, is the proportion of deaths in the population that is attributable to smoking. In general, the higher the prevalence of smoking among people with a particular disease and the higher the smoking-associated relative risks, the higher the SAF. For example, the SAF for death from lung cancer is in the range of 90%, meaning that about 90% of all lung cancer deaths in the United States are attributable to smoking. The SAMMEC model then applies these smoking-attributable fractions to nationwide estimates of deaths (by age, sex and cause) to calculate a number of smoking-attributable deaths and related medical costs.[11]

This same SAF has been used on Florida data to estimate tobacco related costs of hospitalization and physician visits, pharmacy and home health care costs, days lost from work, and years of life lost. These are then used to calculate direct and indirect costs of smoking. Direct costs are costs for medical care; indirect costs are costs of lost productivity for those who die prematurely. For example Florida's projected direct and indirect losses from smoking-related illnesses in 1995 exceeds $5.8 billion[12] (Table 1-1).

[10] Smoking and Health: National Status Report, 2nd ed., DHHS Publication No. 87-8396 (revised 02/90)

[11] When the authors ran SAMMEC for 1996 in Florida just over 20% of deaths were smoking-attributable (about 30,000 per year).

Table 1-1. Direct and Indirect Costs of Smoking, Florida, 1995

Type of Cost	Costs for Males (Millions)	Costs for Females (Millions)	Costs for All Floridians (Millions)
Direct*	$1,279.20	$607.70	$1,887.00
Mortality**	$2,560.80	$995.30	$3,556.10
Morbidity**	$279.60	$117.40	$397.00
Total	$4,119.60	$1,720.40	$5,840.10

* Direct costs = physician, hospital, nursing home, prescription and other direct expenditures for preventing, diagnosing, and treating smoking-related illnesses

** Mortality and morbidity costs = indirect costs, including losses to the economy, lost wages, and losses in services by homemakers. Source: 1992 Florida SAMMEC Data

More recently, CDC has refined these models using the National Medical Care Expenditure Survey, the National Health Interview Survey, and other data sources. The goal is still to calculate a smoking-attributable fraction, but now the model includes a component that measures the effects of smoking on overall health status. Therefore, the SAF is estimated based on the increased likelihood of hospitalization, physician visits, and other medical expenses for all causes in smokers compared to non-smokers[13]

The CDC estimated that in 1987 the nation's $326.5 billion in total medical expenditures included more than $21.2 billion, or 6.5% of the national total, in smoking related expenditures.

[12] Smoking-Attributable Costs in Florida, 1992, Office of Family Health Services, State Health Office, Department of Health and Rehabilitative Services, 1993
[13] Phone communications with Dr. Ann Malarcher, Center for Disease Control, Atlanta Georgia, November, 1998

In Florida, estimated annual medical expenditures in 1987 exceeded $19.5 billion of which almost $1.4 billion, or 7.11%, was smoking related (Table 1-2):[14]

Table 1-2 - Estimated Total and Smoking Related Medical Costs for Florida and the U.S. for 1987, in Millions of Dollars

	FLORIDA		UNITED STATES	
Type of Cost	Total Cost (Millions)	Smoking-Related Costs (Millions)	Total Cost (Millions)	Smoking-Related Costs (Millions)
Outpatient	$4,995.00	$381.90	$93,271.80	$6,649.60
Prescription	$2,117.40	$53.30	$31,692.80	$650.30
Hospital	$11,179.80	$844.40	$182,340.90	$12,491.60
Total	$19,548.80	$1,389.10	$326,536.50	$21,227.50

Source: Miller, L., Zhang,X.,Rice,D.,Max,W.,Novotny,T: State Estimates of Direct Medical-Care Expenditures Attributable to Cigarette Smoking

These estimates are based on health care expenditures in Florida from ten years ago. Over each of these years, medical costs have increased dramatically. Cost increases and population-client growth alone have driven Medicaid costs up by as much as 25% per year for a number of those years.

To develop more up-to-date estimates, the Florida Medicaid staff used a published methodology similar to the SAMMEC technique modified for Florida referred to as the Modified SAMMEC Estimates.[15] This analysis, updated to 1995 dollars, indicates that total Florida medical

[14] Miller, et al., 1997

[15] Smoking-Related Deaths and Financial Costs, U.S. Office of Technology, 1993 referenced in "Clearing the Air About Smoking- A Briefing on the Health Problems Associated with the Use of Tobacco", Agency for Health Care Administration, May, 1994

expenditures for 1995 exceeded $45 billion, of which $1.428 billion, or 3.2%, are smoking related (Table 1-3).

Table 1-3. Modified SAMMEC Estimates of Total and Smoking-Related Health Care Costs in Florida, in Millions of Dollars, 1995

	Total	Smoking[16]
Type of Cost	Cost (Millions)	Related Cost
Hospitalization	$18,738.7	$95.5
Physician Care	$11,195.2	$18.5
Nursing Home	$1,912.6	$101.8
Medications	$2,200.5	$20.4
Other Professional	$1,841.5	$21.9
Other	$9,483.4	$107
Total	$45,371.9	$365.2
Source: 1992 Florida SAMMEC Data		

In Florida, as elsewhere, smoking is more common among low income residents; in 1994, smoking was reported for 26% of Florida residents with incomes less than $10,000, compared with 17.8% of those with incomes of $50,000 or higher.[17] Poor people also have poorer health; low-income residents report higher prevalence of hypertension and more days of poor health and disability days.[18] The combination of poorer health and higher smoking prevalence in the population from which Medicaid clients are drawn suggests that smoking-related costs should account for an even greater proportion of the Medicaid budget.

[16] Ibid, Hopkins-Lynch, 1997
[17] Florida Behavioral Risk Factor Surveillance System, State of Florida Department of Health and Rehabilitative Services. "Health and Lifestyle of Florida Residents," 1995.
[18] Florida Behavioral Risk Factor Surveillance System, State of Florida Department of Health and Rehabilitative Services. "Health and Lifestyle of Florida Residents: State Health office 1986-92 Fact Sheet."

Table 1-4 provides a summary of Florida Medicaid 1992-93 expenditures on diseases that have scientifically established linkages to tobacco smoking. The table includes the number of Medicaid enrollees treated for each disease, total Medicaid expenditures for each disease, and the proportion of cases and expenditures attributable to smoking expressed in 1995 dollars.

In 1992-93, the Florida Medicaid program paid almost $400 million (1995 dollars) for smoking-related illnesses. This is approximately 6% of its total budget or twice the estimated proportion of all Floridians' health care expenses that are due to smoking. This is probably an underestimate since it does not include the effect of tobacco smoking on overall health.

Table 1-4. Florida Medicaid Costs Attributable to Smoking
1995-1996 Cost Projections*

Disease	Number of Medicaid Patients	Total Cost (Millions)	% Smoking Related	Number Smoking Related	Smoking Related Cost (Millions)
Cancers of the oral cavity	702	$9.8	77.8	546	$7.6
Cancers of the digestive system	675	$12.6	43.0	290	$5.4
Cancers of the respiratory tract	1,941	$32.4	82.6	1,604	$26.8
Cancers of the uterus*	1,225	$9.5	30.0	368	$2.9
Cancers of the urinary tract	832	$12.5	36.3	302	$4.5
Coronary artery disease	8,821	$94.4	28.8	2,540	$27.2
Cerebrovascular disease	11,621	$171.5	34.7	4,036	$59.5
Pneumonia and influenza	59,071	$430.4	24.7	14,590	$106.3
Emphysema and bronchitis	20,129	$196.9	79.7	16,043	$156.9
Total	105,017	$970.0	38.4	40,319	$397.1

* These projections are based on FY 1992-93 Medicaid Program Expenditures. Note that cancer of the uterus is included inappropriately in this listing and would be excluded if in strict conformance with the CDC SAMMEC model structure but was included in the 1993 AHCA analysis.

This data can be used to estimate the discounted net present value (NPV) of all Medicaid expenditures attributable to smoking-related illnesses over the recent past and foreseeable future.

from 1990 through 1995 ranged from 9.2% to 25.3% with an average annual rate of increase of 19.1%. Using a 9.2% growth rate and discount rates of 3% to 6%, the NPV of Medicaid

from $55 billion for discount rates of 6% to $154 billion for discount rates of 3% (Table 1-5).

Table 1-5. Estimated Medicaid Smoking-Related Cost Increases over 50

Number of Years				
	3%	4%	5%	6%
6	$2,095	$2,078	$2,061	$2,044
7	$2,695	$2,654	$2,614	$2,574
8	$3,337	$3,264	$3,192	$3,122
9	$4,022	$3,908	$3,797	$3,689
10	$4,754	$4,590	$4,431	$4,277
15	$9,237	$8,629	$8,064	$7,541
20	$15,474	$13,964	$12,619	$11,421
30	$36,226	$30,326	$25,490	$21,522
40	$76,398	$58,880	$45,719	$35,805
50	$154,161	$108,713	$77,514	$55,998

With a 25.2% growth rate in smoking related expenditures in Florida, the NPV of discounted values grows at a very accelerated rate. Over 50 years, the NPV ranges from $37.6 to $16.2 trillion for discount rates of 3% and 6%, respectively.

Other states have calculated similar Medicaid losses. Most recently, Mississippi estimated its annual smoking-attributable Medicaid costs to range between least $40 to $60 million per year.[19]

[19] "Medicaid Related Smoking Costs for the State of Mississippi", Wendy Max, Ph.D., Associate Professor of Medical Economics, University of California, May, 1997 (Unpublished)

Mississippi's estimated costs are lower than Florida's because Florida has an older population, higher income cut-off for Medicaid eligibility, and higher Medicaid reimbursement rates.

The surveys on which these analyses are based include only non-institutionalized persons living on their own in the community. These surveys are not designed to measure nursing home costs, which typically account for at least 20% of Florida's Medicaid budget. The most accurate way to calculate the full impact of smoking on Medicaid requires the specific inclusion of smoking-attributable nursing home costs.

The strategy that Florida researchers developed to evaluate the prospects of securing this missing information is to develop a survey of nursing home residents to determine their smoking histories before admission and their primary payment source (Medicaid and non-Medicaid primary or secondary payer) since admission. Researchers are hopeful that the resulting estimates of smoking-related Medicaid nursing home expenditures will help enhance Florida-specific estimates of total Medicaid costs and more realistically estimate the substantial public health care subsidies associated with tobacco-related illnesses. To date the most accurate statewide estimates indicate that annual Medicaid smoking related costs in Florida exceed $365 million over the FY 1995-96 time period.[20]

[20] Ibid, Hopkins-Lynch, 1997

D. Development of a Research Design and Analytical Methodology

This study was initially intended to correct these deficiencies of under-estimation of tobacco related costs by surveying a representative sample of nursing home residents about their smoking histories and calculating costs for smokers and non-smokers in the sample. Florida is a particularly appropriate state for such a survey because it has one of the most comprehensive nursing home resident databases in the nation. [21]

Development of the Problem Statement

This project is designed to measure two hypothesized components of excess nursing home costs for smokers compared to non-smokers:

Hypothesis 1:

At any age, smokers are more likely than former smokers or never-smokers to be admitted to nursing homes.

Hypothesis 2:

Once admitted to nursing homes, smokers incur more total costs (measured in length of stay) than former smokers or non-smokers.

To test these hypotheses, we used data from our survey and from the Florida Behavioral Risk Factor Survey to:

1. Calculated odds ratios for nursing home admission, comparing current and former smokers to never-smokers by sex and five-year age group cohorts
2. Calculate the difference in average nursing home length of stay costs between patients with smoking histories and those who never smoked, controlling for differences in age, race, sex, and socio-economic status.
3. Provide an estimate of the proportion of nursing home costs attributable to smoking.

[21] As the later analysis will demonstrate the data collection contains (age, health and discharge status- many had died) and ultimate data limits (limitations of accurate memory of respondents and information limitations of response of next of kin) preclude completing the full scope of cost analysis originally envisioned.

4. Estimate first the hypothesized increased probability of dying soon after NH admission – i.e. the hazard ratio using the proportional hazards model; and then the hypothesized decreased probability of early discharge alive after NH admission.

We believe the conclusions of this study will provide important new information to help clarify some of the adverse impacts of tobacco smoking on Florida's older citizens. Additionally, we believe the conclusions of this study should help expand the existing national model on smoking-related medical (and nursing home) expenditures and will enhance our collective understanding of the institutionalized nursing home client smoking related costs. For example, the results from this analysis will be provided to researchers at the University of California and elsewhere to assist in their research evaluating the total economic cost of smoking across the U.S. [22]. The methodology (and estimates) developed by this study may also be used by other states to develop more precise estimates of smoking-related nursing home health care costs in their respective states.

[22] Private phone discussions with Vincent Miller, Ph.D., Miller and Associates indicate that he is currently under CDC contract to expand existing SAMMEC models and hopes to use results from this research to expand the model's nursing home cost estimates as appropriate. October, 1998

Section 2 . Methodology and Data

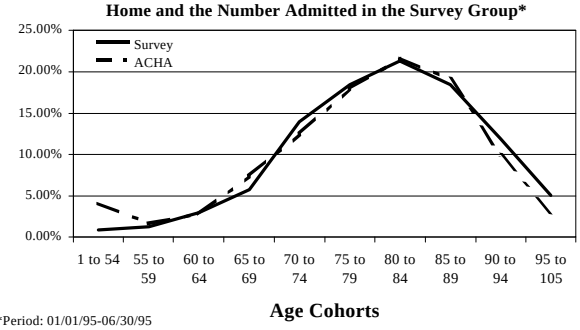
A. Comparison of the Nursing Home Patient Survey Sample and The Florida Nursing Home Population Admitted to Nursing Homes Between January 1 and June 30, 1995

We believe it is important to compare the sample of nursing home patients surveyed to the total Florida nursing home clients at risk to verify the representative nature of the sample. This following section provides a summary of comparisons used to determine if the basic characteristics of the sample group approximate those of the general population. These comparisons include age cohorts, age distribution of new admissions during the time period of the study, length of stay in days, payer mix, and geographic location. All general, Florida nursing home patient data used in these comparisons were provided by the Agency for Health Care Administration (AHCA), State Center for Health Statistics, Nursing Home [1]

Age Cohort Comparison

By dividing patients into 5-year age groups, the age distribution of the sample could be compared to that of the general nursing home population in the state of Florida. The research design specified only clients over 55 would be surveyed. Almost 73% of the patients surveyed are concentrated between ages 70 and 89. Only 11% fall in the four age cohorts younger than 70 while the remaining 16% of the sample includes the three age cohorts from 89 to 105. The curves from the sample and the general nursing home population are virtually superimposed.

FIGURE 2-1
Comparison of the Number of Patients Admitted to a Florida Nursing Home and the Number Admitted in the Survey Group*



*Period: 01/01/95-06/30/95

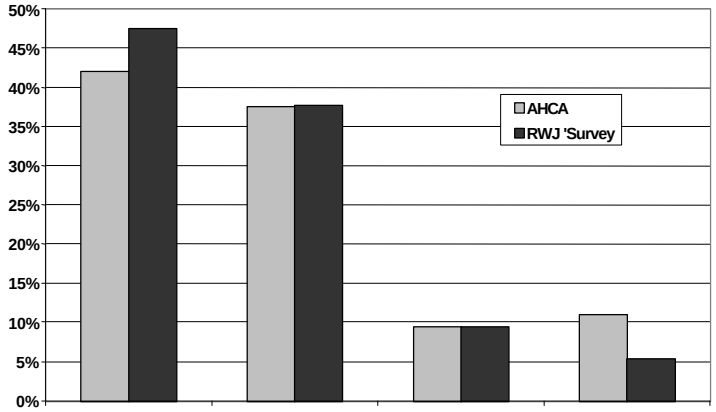
[1] Special thanks too to those supportive staff that provided us with technical and data support over the period of our analysis and without whom this analysis would not have been possible.

These results indicate that the age distribution for the sample data is fairly representative of the age distribution in the general population of nursing home patients in Florida (Figure 2-1).

Length of Stay Comparison

Figure 2-2 provides a comparative profile of the average length of stay (LOS) for the total Florida nursing home population[2] and our nursing home sample survey population. Both groups of patients were admitted to nursing homes during the January 1, 1995 to June 30, 1995 time frame.[3] These distributions demonstrate that the sample and the population from which it was drawn are virtually identical across each LOS category. It is also interesting to note that almost 43% of the sample and 48% of the total Florida nursing home residents stay less than 20 days in nursing homes. Similarly 73% and 80% of the general population and the sample populations respectively involve relatively short stays of less than 100 days. Florida nursing homes, like nursing homes across the country, have converted a considerable amount of their capacity to service relatively short-term, sub-acute or recuperation patients. This trend indicates that nursing homes no longer rely on the traditional long term care patients as their only source of income.[4] However, while the number of long term residents are a small percentage of annual admissions, they still represent the largest fraction of nursing home revenue and costs via their longer total length of stay.[5]

FIGURE 2-2
Comparison of Nursing Home Length of Stay for Total and RWJ Sample Nursing Home



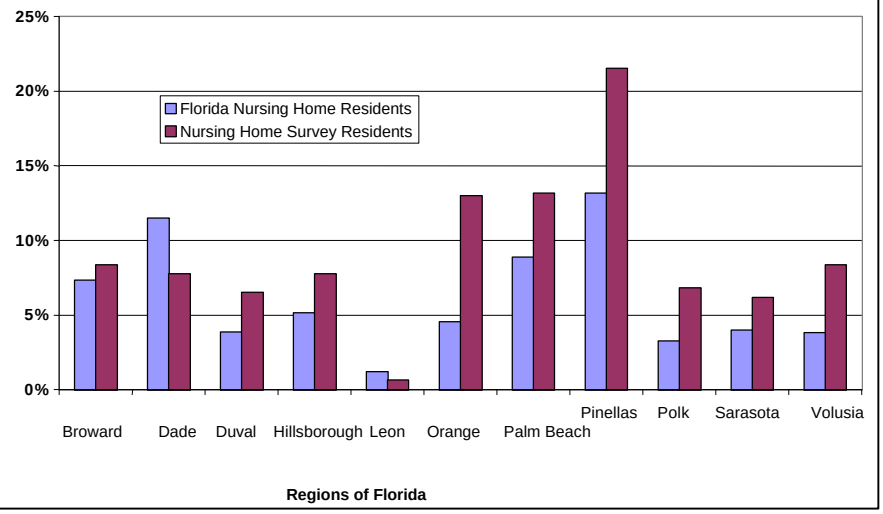
[2] All statewide nursing home information was provided to researchers by, " Agency for Health Care Administration (AHCA), summary statistics were often derived from Florida Nursing Home Reporting System, 1994-1996, Resident Data Report", Agency for Health Care Administration, December, 1997
[3] The LOS stay from AHCA – was it calculated as of approximately the date of the survey

[5] Ibid., AHCA, 1997

Regional Comparison

Figure 2-3 indicates the percentage of patients drawn from various geographic regions for the study sample and how well the study sample represents the general population of Florida nursing home residents admitted over the January 1 through June 30, 1995 time frame. Notice that approximately 40% of total Florida nursing home residents are served in four counties; Pinellas, Dade, Palm Beach and Broward. Despite this large concentration, the remainder of Florida nursing home residents are dispersed over a number of other counties. The design of the survey ensured that a representative sampling would be captured not only from the counties with the largest concentrations but also from the north (Duval, Leon), west (Sarasota), and relatively rural, central Florida (Polk) counties. Figure 2-3 provides a bar graph reflecting this diversity of counties represented in the final nursing home survey sample.

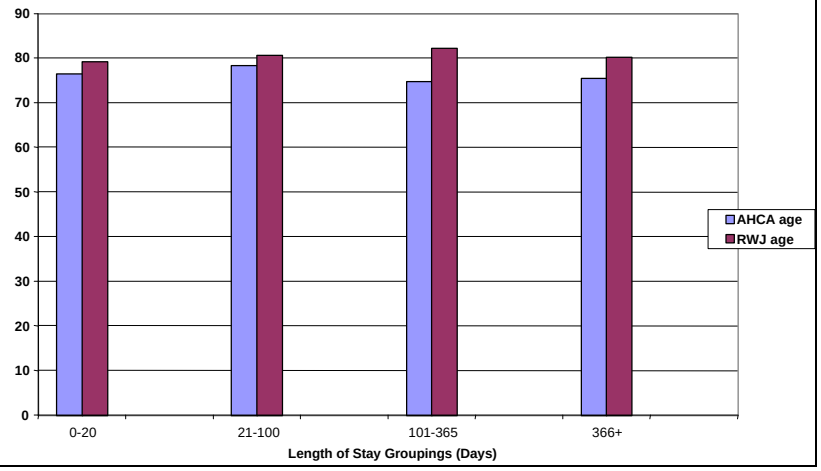
FIGURE 2-3
Regional Comparison of Florida and Survey Nursing Home Population Distributions



Comparison of Average Age of Residents and Length of Stay

Another informative comparison of the study sample and the general Florida nursing home population is by average age and length of stay. Figure 2-4 shows that the sampled population was on average several years older than the resident population. However, these differences were relatively small and ranged from 3% to 6%.

FIGURE 2-4
Comparison of Nursing Home Average Age to Length of Stay
AHCA & Survey Population



The close match of the length of stay, geographic dispersion and age profiles described in this section suggest that the sample drawn is reasonably representative of the state wide nursing home population.

.

B. COMPARISON OF EVER AND NEVER SMOKER LENGTH OF STAY

The goal of the study was to compare the costs (of nursing home care) of smokers versus nonsmokers. Smoking status (former, current, never) and other demographic characteristics of the patient were determined via survey information as well as exposure to smoke. The variable length of stay (LOS) was determined by subtracting the admittance and discharge dates and was used as a proxy for measuring costs.

Description and Characteristics of the Data

As described in the Methods Section of this report, the original data set contained survey results from 721 nursing home residents that entered a Florida nursing home for the first time between January 1 and June 30, 1995. The independent variables consisted of 18 survey questions as well as gender, discharge age, discharge status, date admitted, and date discharged.[6]

The variables used in the analysis are defined in Table 2-1 as follows:

TABLE 2-1

Variable	Description
smoking status	
-current	smoked within one year of admittance
-ever	former smoker
-never	less than 100 cigarettes in lifetime
los	length of stay in nursing home
totcigk	total number of cigarettes smoked in lifetime / 1000
homesize	1=small (< 180 beds) , 2=large (≥ 180 beds), (excluded if ≤ 10)
q7	1 = ever married
discliv	0 = still there, 1 = living at discharge, 2 = deceased at discharge
gender	male or female
avgcigday	average number of cigarettes consumed per day
ddis	discharge status: 0 = deceased, 1= living
myrs	years married
smoyrs	number of years smoking
yrsnt	number of years not smoking
disage	age at discharge
married	ever married = 1, never married = 0

Following the age cohorts used by the BRFSS and ACHA, patients were grouped into 9 five-year age groups, from age 55 to 105.

[6] The original data set also contained some personal, patient information, such as, patient address, phone number, etc., which was not used in the analysis stage. The survey instrument and detailed list of the data is included in Appendix A of this report for the interested reader.

Limitations of the Data and Introduction to the Analysis

There were several problems with the data set. The main problem was the number of missing values in the more important smoking related variables. This proved to cause considerable problems in the analysis stage and prohibited standard regression analysis. The first attempt at ameliorating the problem was through means analysis.

A commonly used method for replacing missing data is to impute means from other observations. This process involves determining the pertinent characteristics of cases that have actual data and developing a mean value based on the available random data associated with those characteristics. This value is then used to fill missing data of cases with similar characteristics. This is standard procedure when variables of interest are missing up to 10 –15% of their data. However statistical theory discourages imputation of means when the amount of missing values exceeds 10-15% especially for the more critical variables in the data set.[7] This process, if abused, leads to biased (toward the mean) regression estimates and of course, inaccurate means tables and comparisons. Since most of the critical variables in this survey data set were missing from 25-40%, mean imputation and replacement could not be employed.

Another option is to delete cases from the study provided the missing data is not characteristic of specific types of patients. This is important because it could indicate a bias in the survey. For example if the oldest smokers could not provide data on length of time smoking, then valuable information about long time smokers would be excluded from the analysis. The final results would say nothing about smokers with those specific characteristics.

To overcome this problem, we first ascertained the severity of the problem and the level of randomness of the missing data. As mentioned above, the missing data problem was severe for several variables. However, it was also evident that there was no relationship between cases with missing data and specific characteristics of these cases.

The sample size was large enough to withstand losing a fair portion of the data while maintaining its validity for statistical testing. However, the missing data was more prominent in the smoking group so case elimination was not the best option.

[7] This occurs because the researcher is essentially tricking the computer into calculating a larger number of degrees of freedom (DF). When imputing means, the degrees of freedom should not increase because the amount of new information has not increased. Any test results must be noted and carefully interpreted.

The structure of the survey did allow for the merging of several similar variables into one variable. The survey was designed to ask the same question in multiple forms. For example, the question regarding the amount of smoking by the patient was asked in two ways. The patient was first asked how many packs per day they smoked and then how many individual cigarettes per day they smoked. Because both variables were missing a substantial amount of data, combining the two variables into one provided more information that could be legitimately included in the analysis.

It should also be noted that there is some degree of measurement error in the variables. Some of the information gathered came from a secondary source, such as, next of kin or from patients with limited coherency. This not only affects the quality of the data but the findings as well.

Profile of the Gender Differences within the Sample

The gender splits within age groups is also interesting to note. Figure 2-5 provides a profile of the number of male and female nursing home residents surveyed by age cohort. The entire survey involved 721 respondents. Female residents represented by 429 or 60% of the survey respondents while male residents were represented by 281 males or 40% of this sample population.[8] The AHCA nursing home data sets contain no gender information that can serve as a basis of comparison. However these gender splits are consistent with national and state estimates provided by the nursing home industry.[9] Notice that only 77 individuals (or 14%) enter nursing homes prior to age 69 and that the gender splits are about equal. Part of this gender equity among the earliest age cohorts may be attributed to the relatively lower death rate of both genders up through the age of 70. As expected, there is a divergence between the number of male and female residents after age 70, as mortality rates are markedly higher for men than women in this age group

FIGURE 2-5
Number of Male and Female Nursing Home Residents Surveyed by Age Cohort

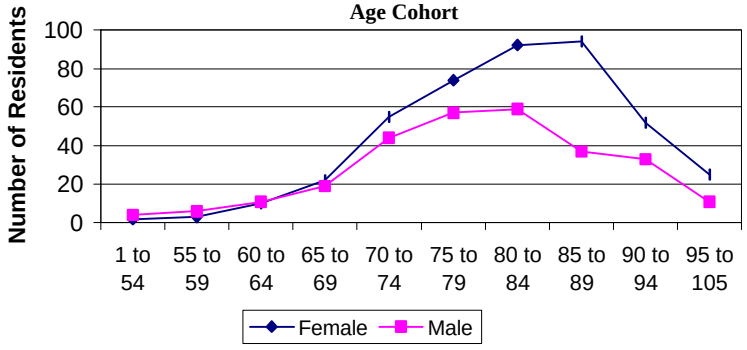
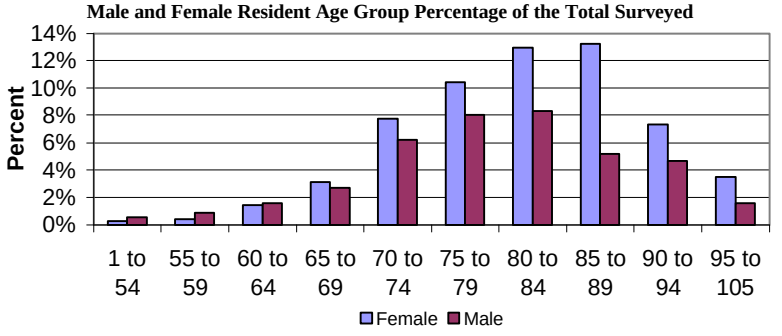


FIGURE 2-6
Male and Female Resident Age Group Percentage of the Total Surveyed



The largest difference occurs in the 85 to 89 age group where females outnumber males by almost three to one (37 males to 94 females).

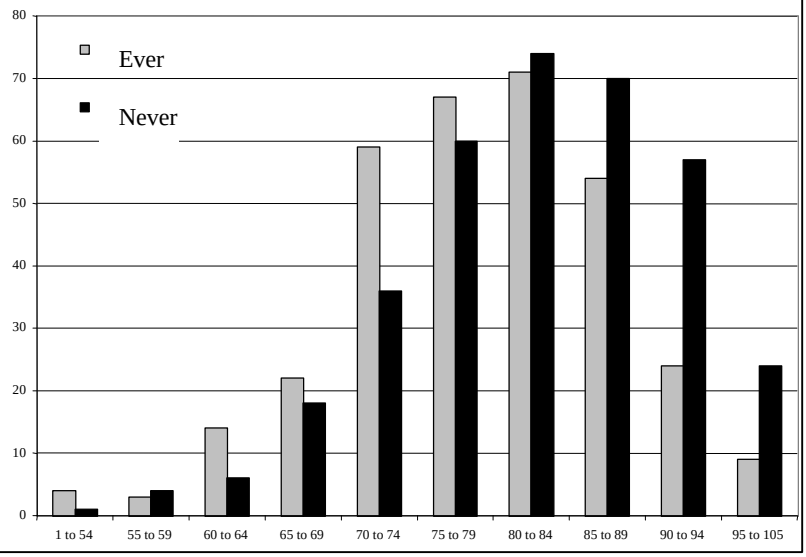
Figure 2-6 presents the relative percentages of the total survey of male and female surveyed residents by age cohort. Again note that the largest percentages (72% of the total) of the surveyed population are in the four age cohorts between ages 70 to 89. Females dominate each cohort and represent 44% while males only sum to 28% of the residents within these age groups.

[8] 11 cases contained missing data and therefore were excluded from this profile

Smoking Status

Almost 52% of the nursing home clients we surveyed were never smokers while 48% fall into the ever smokers category. As Figure 2-7 indicates, ever smokers were a far greater percentage of the younger persons in our sample. This may indicate that ever smokers are far more likely to enter nursing homes at an earlier age than never smokers. Figure 2-7 also indicates that ever smokers out number never smokers in virtually every age group up through and including the 75 to 79 cohort. The minor exception to this is the 55 to 59 age groups where they report virtually identical populations of 3 and 4 clients respectively in the ever and never smokers groupings. By comparison, over 52% of all smokers and only 36% of all never smokers are represented in these relatively younger nursing home cohorts of 79 years of age and younger.

FIGURE 2-7
Profile of the Number of Nursing Home Clients Surveyed that are Ever Smokers and Never Smokers, by Age



Where are all of the never smokers? A review of Figure 2-7 indicates that the never smokers dominate the four oldest cohorts in our survey. The remaining age groups between 80 to 105 contain 64% of the never smokers and only 48% of smokers. These differences are most pronounced in the very old (90 and over) where never smokers out number smokers 2.5 to 1.

Where have all of the smokers gone? Initial indications begin to suggest that a considerable number of smokers have died in larger numbers at earlier ages than never smokers and incurred higher lengths of stay during these younger years, for example see figure 2-12a. Ever smokers are more numerous in the earlier age groups but are still proportionate to their numbers in the general population. They tend to

[9] "Facts and Trends - The Nursing Home Facility Sourcebook," 1997, American Healthcare Association, Washington, D.C.

dominate the nursing home population (in these younger age groups only) since they were admitted in larger numbers in the younger years.

Just as in the general population, the proportion of never-smokers and former smokers rises with increasing age. Also current smokers are not as numerous in the older cohorts since smokers die at higher rates and as people have a chance to move from current smoker to former smoker (it is very rare for adults to start smoking, in the age range we are examining).

Gender Profile of the Distribution of Ever and Never Smokers

The ever and never smoking experience of each gender (within various age cohorts) is also interesting to review. Figure 2-8 provides a profile of the number of ever and never male, nursing home residents surveyed by age cohort. Notice that the ever smokers make up the bulk of the male nursing home respondents in every age category. The absolute number of male smokers peaks in the 80 to 84 age group and declines thereafter. This absolute and relative decline in the number and percentage of male smokers is likely related to the reduced survival rate of smokers in the older cohorts.

FIGURE 2-8
Number of Ever and Never Smokers Among the Male Nursing Home Survey Respondents

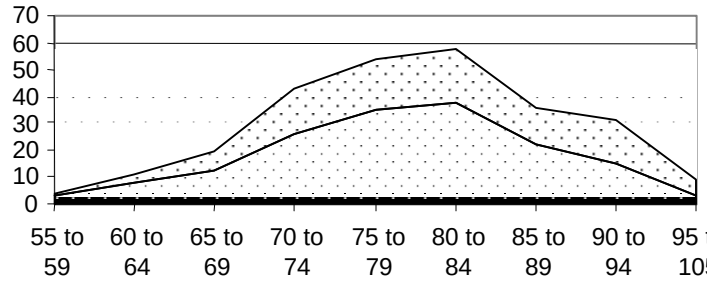


Figure 2-9 provides the number of female ever and never smokers surveyed in the nursing home. A very different pattern of ever and never smokers emerges among female nursing home residents. Never smokers dominate all age cohorts after age 65 by a wide margin. Only 161 (or 40%) of the surveyed females residents were ever smokers while 267 (or 60%) of were never smokers. As Figure 2-9 demonstrates this split is particularly noticeable within the age cohorts over age 65 which accounts for more than 80% of all females in the sample population.

FIGURE 2-9
Number of Ever and Never Smokers Among the Female Nursing Home Survey Respondents

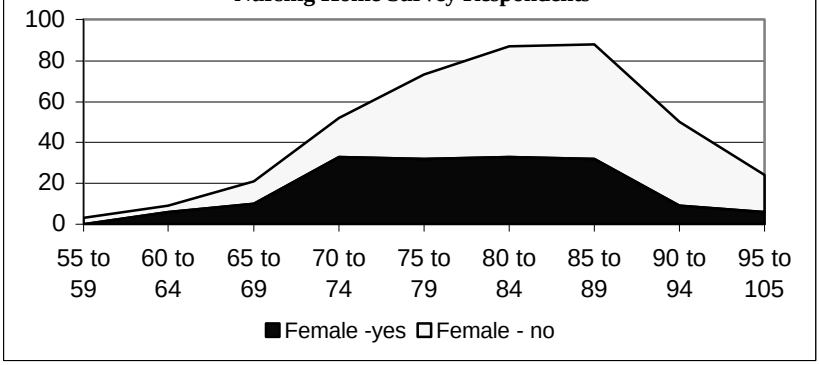
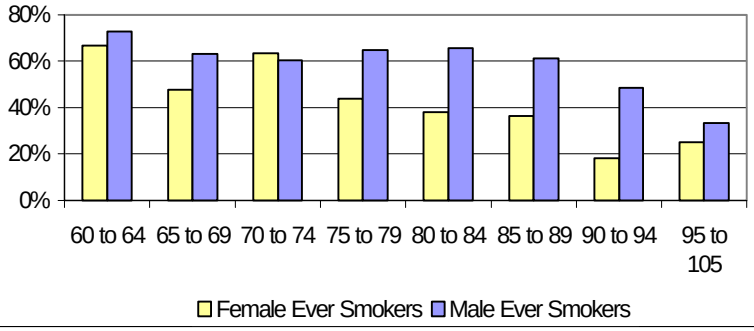


Figure 2-10 provides a comparative bar graph of ever smokers among both male and female residents by age cohorts. Ever smokers constitute 66% of all male nursing home residents surveyed. The number of nursing home residents split dramatically and grows to its largest difference in the 85 to 89 age group

where females outnumber males by almost three to one (37 males to 94 females). Again these differences are consistent with state and national trends and is likely attributable to the longer life expectancy of women and lower smoking rates in these older women, who came to early adulthood at a time (1930-1950) when smoking by women was uncommon compared to smoking by men.

FIGURE 2-10
Percentage of Male and Female Ever Smokers by Age Cohort



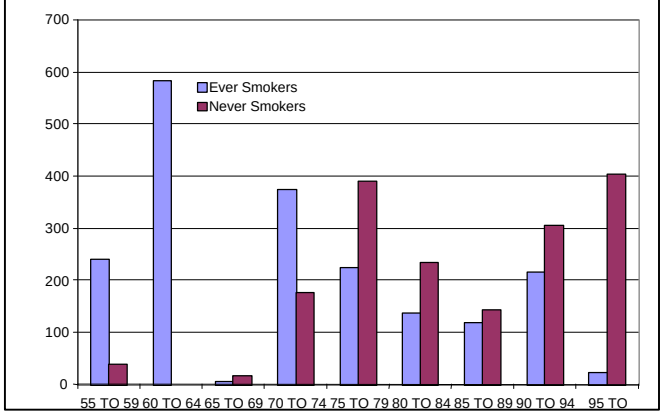
Within the 75 to 79 year age cohort 44% of the females surveyed were ever smokers compared to 65% of the surveyed male nursing home residents. As described in a later section of this report these relative percentages are consistent with similarly aged non-nursing home dependent Floridian male and female residents.

Notice that the percentage of ever-male smokers is considerably higher than the female levels in virtually every age cohort. Also note that the both the male and female rate of ever smokers declines in the older age cohorts. This phenomenon is again likely related to the higher death rate in earlier years of ever smokers over never smokers (see figure 2-13).

Profile of Ever and Never Smokers Average Length of Stay

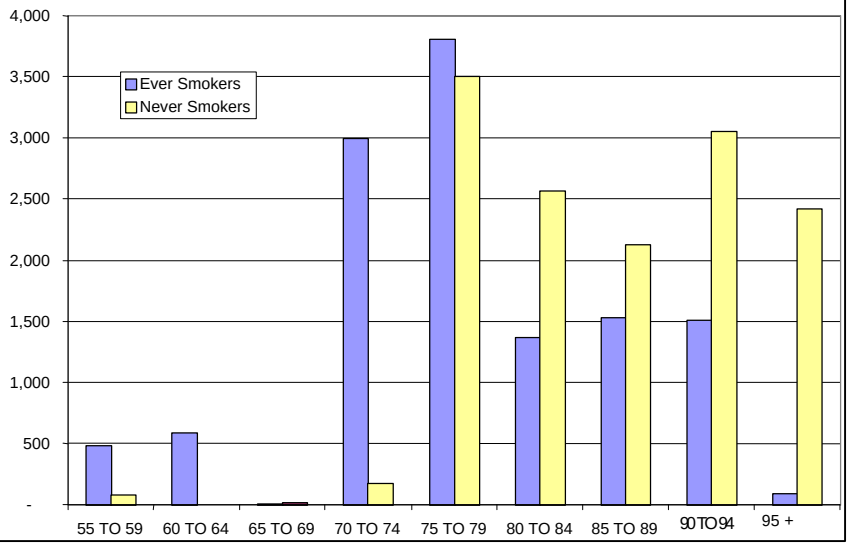
Figure 2-11 provides a profile of the average length of stay for ever and never smokers by age cohort that are discharged dead from the nursing homes surveyed. The 55 to 74 age groups are dominated by the ever smoker population. This trend is reversed permanently at the 75 to 79 age cohort. In essence the shorter LOS here tends to indicate that ever smokers experience earlier discharge (dead or alive). after NH admission in the age cohorts after age 79.

FIGURE 2-11
Comparison of Average Length of Stay of Ever and Never Smoker Discharged Dead



Smokers tend to stay in nursing homes between 2 to 4 times longer than never smokers up through age 75 (i.e. the younger never-smokers die sooner after admission than the younger ever-smokers, the 60 to 64 age cohort is not included in this comparison since there are no never smokers discharged dead in this age group). The never smokers begin to dominate the ever smokers average length of stay from 75 on by a ratio of 1.2 to 18 in the final aged 95 to 105 cohort.

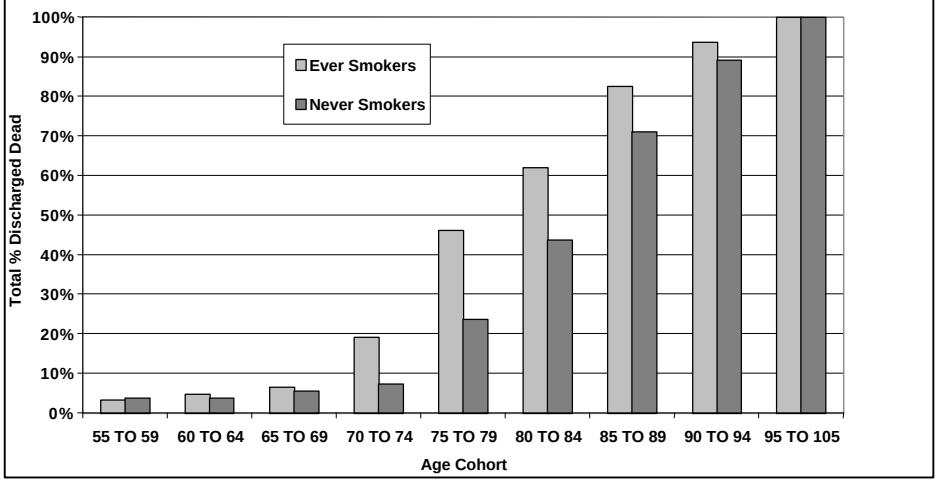
FIGURE 2-12
Comparison of Total Length of Stay of Ever and Never Smokers Discharged Dead



The same trend is evident when examining our nursing home survey results for total length of stay. Figure 2-12 profiles the total length of stay of ever and never smoking nursing home clients *discharged dead* by age group. In this case, the *total* length of stay in the 55 to 75 age cohorts is again dominated by the ever smoker population. Even the 75 to 79 age cohort LOS is dominated by ever smokers in this grouping. While ever smokers may have shorter average lengths of stay (as identified in Figure 2-11 above), the total number of ever smokers entering nursing homes at this

Figure 2-13

Cumulative Comparison of Percent of Ever and Never Smokers Discharged Dead



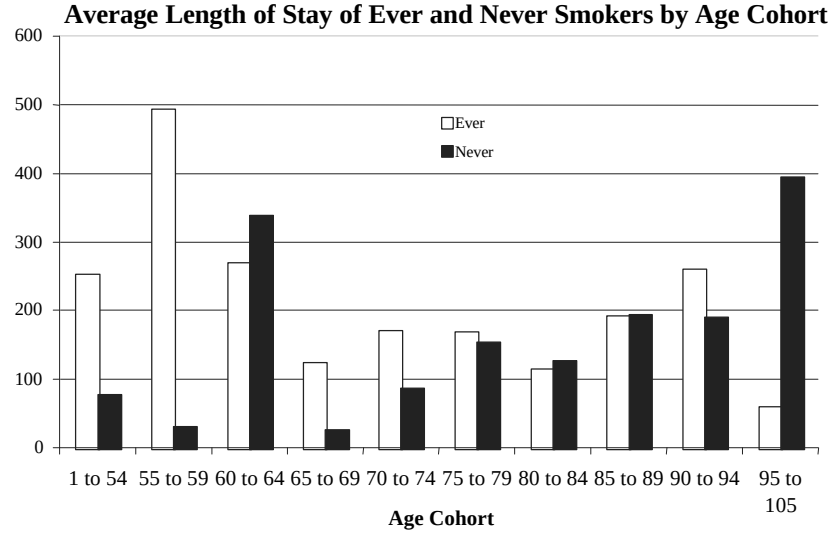
point exceeds the never smokers by a factor of almost two (17 ever versus 9 never smokers- see Figure 2-12) to one with the count of patients.

With the minor exception of 65 to 69, ever smokers tend to stay in nursing homes between 1.1 to 17 times longer than never smokers up through age 79.[10] The never smokers begin to dominate the ever smokers total length of stay from 80 years of age and beyond with a ratio of 1.4 to almost 27.5 in the final aged 95 to 105 cohort (note that this last and oldest age cohort contains only 9 ever and 24 never smokers).

This basic analysis suggests that ever smokers tend to enter nursing homes proportionate to their numbers in the population at risk but incur a longer length of stay relative to the non-smoking group despite the fact that both groups are discharged dead. Conversely, never smokers tend to enter at older ages and survive longer. This is illustrated in the Figure 2-12 above. Non-smokers dominate the older age cohorts (80 and above), and consequently incur longer lengths of stay even if they are discharged dead.

FIGURE 2-14

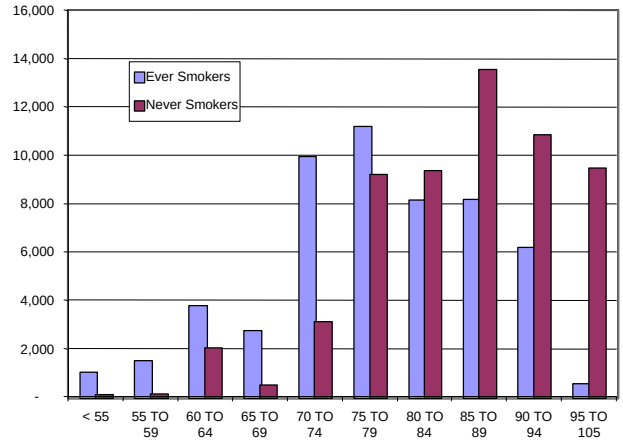
Average Length of Stay of Ever and Never Smokers by Age Cohort



[10] This excludes the 60 to 64 age cohort since there are *no* never smokers discharged dead in this age group.

As Figure 13 shows by age cohort 75 to 79 46% of all those ever smokers compared to only 23.6% of never smokers that will die while in the nursing home have already expired. This represents a 22.4% ***higher absolute cumulative rate of early death of ever to never smokers***. The same holds true for the next several older cohorts (with differences diminishing to 18%, 11.6^ and 4.6% respectively). As the ever smokers continue to die and are almost completely gone in the older cohorts until the very oldest cohort 95+ the difference is erased as all members of both groups that will die in nursing homes have expired.

FIGURE 2-15
Comparison of Total Length of Stay Ever and Never Smokers



The results of the discharged dead sub-group also apply to the sample as a whole. Figure 2-14 provides a comparison of ever and never smokers' average length of stay for all nursing home clients included in the sample. Up through and including age cohort 75 to 79, ever smokers dominate the average length of stay profiles with the exception of the 60 to 64 age cohort. Ever smokers in these cohorts dominate the average length of stay profiles by factors ranging from 1.1 to 18 times that of the never smokers. Non-smokers dominate the average length of stay by factors ranging from 1.1 to 8 with the exception of 90 to 94 cohort.

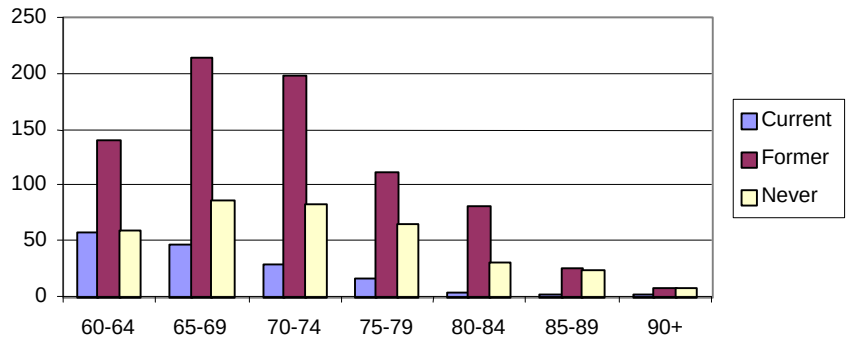
The pattern is even more pronounced for the comparison of *total* lengths of stay between ever and never smokers in the sample. Figure 2-15 portrays the same trend evidenced in Figure 2-14. Ever smokers have higher total lengths of stay up through and including age cohort 75 to 79. After that point, never smokers dominate with sizable leads in the very oldest groupings of 95 to 105. Never smokers exceed ever smokers length of stay by a factor of almost 25 to 1. Again, the results in the oldest cohort should be interpreted with caution, as there are very few individuals in this oldest cohort. In general, the results of the discharged dead subgroup are confirmed in the analysis of the total sample group.

C. Description and Comparison of Rate of Admission to Nursing Homes of Ever and Never Smokers to the Florida Population at Risk

Description of Current Ever and Never Smokers in the Florida Population at Risk

Figure 2-16 shows the number of male respondents in the Florida BRFSS survey by age and smoking status. Notice that current smokers are more numerous in the relatively younger cohorts and become almost negligible by age cohort 80-84. Similar declining numbers with increasing age is evidenced for never and former smokers as well. Among the male residents, former smokers however are the most numerous in every age cohort.

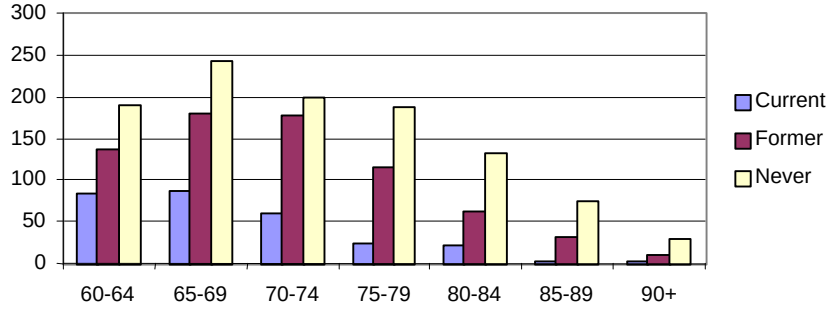
FIGURE 2-16
Number of Male Respondents by Age and Smoking Status, Florida BRFSS



A similar overall pattern exists with the Florida female BRFSS survey respondents. However, since females outnumber males in each cohort they also report higher relative and absolute numbers of current smokers in each age group.

Females present a similar pattern forever smokers with far higher percentages of never smokers and a far lower percentage of former smokers. The significant difference between the genders is in the category of never smokers. Note that female never smokers out number male never smokers in every age group.

FIGURE 2-17
Number of female respondents by age and smoking history, Florida BRFSS



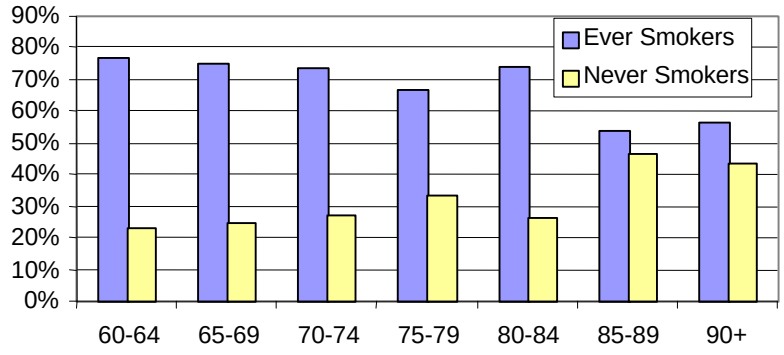
The relative female never smoker category ranges from 46% in the 60-64-age cohort to 70% in the 90+ cohort. The male never smoker percentages range from 23% in the 60 to 64 cohort to 44% in the 90+ cohort.

This relative difference in gender and cohort splits can be more clearly seen in Figures 2-18 and 2-19, which show a distribution by age and gender of BRFSS ever and never smoker respondents.

We undertook this work with the idea that the relationship of nursing home expenditures, at any given age, for smokers and ex-smokers to that for never-smokers would be a product of two relationships:

1. The ratio of the rate of nursing home admission for smokers and ex-smokers to that for never-smokers; and
2. The ratio of average expenditures for smokers and ex-smokers to that for never-smokers among persons admitted to nursing homes.

FIGURE 2-18
Ever Versus Never BRFSS Male Smoker Percentages



If, say for any age group, the product of these two ratios was greater than one for smokers or ex-smokers, we argued, then there would be an excess of nursing home expenditures for smokers or ex-smokers in the population, for that age group.

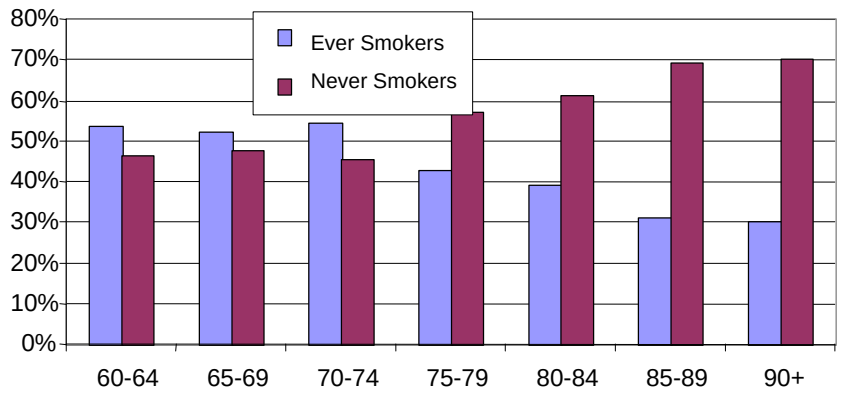
The approach to estimating the first ratio above was to do a case-control study.

For this study, cases were persons aged 55 and older admitted to Florida nursing homes during the period January 1, 1995 to June 30, 1995. A sample of Florida nursing homes was selected with probability proportional to the number of admissions during the previous year. (There was also some geographic

stratification based on urban-versus rural and northern versus southern nursing home concentrations to assure geographic and rural representation.) [11]

In each selected nursing home a proportionate number of subjects were chosen at random from those admitted during the study period. Nursing home administrators assisted with identifying next of kin for interview. Next of kin were contacted by letter with telephone follow-up for the interview. Informed consent was obtained before the interview was begun. Trained interviewers conducted all interviews and almost all were completed over the telephone. A standardized questionnaire and interview script was used. Some demographic items, as well as date of admission to the nursing home, were abstracted from the nursing home records. The study survey team was finally able to finalize a total of 721 completed surveys.

FIGURE 2-19
Ever Versus Never BRFSS Female Smoker Percentages



Controls were respondents to the Florida Behavioral Risk Factor Surveillance System (BRFSS). This is an ongoing random digit dial telephone survey of Florida residents aged 18 and older. Respondents were the 3,339 persons aged 60 and over with completed BRFSS interviews for the years 1993, 1994, and 1995 the two years prior to and the year of the sample survey drawn reported on in this report.

Variables available from *both* data sets included the subject's age in years; sex; smoking history; educational attainment; and marital status. In these analyses, age was grouped in five or ten-year increments. Educational attainment was categorized as: none; 1 to 8 years; 9 to 11 years; 12 years or GED; 1-3 years of college; and 4 or more years of college. Marital status was categorized as ever versus never married. Smoking status was categorized as current (at time of NH admission for cases, at time of interview for controls), former, or never. For some analyses current and former were combined, as ever-smokers; for others, former and never smokers were combined to contrast with current smokers.

[11] See a detailed description of the sampling strategy and methods used in Development of a Research Design and

Comparison of Rate of Admission to Nursing Homes of Ever and Never Smokers to the Florida Population at Risk

The crude odds ratio estimate for nursing home admission comparing current versus never-smokers was 0.66 for males and 0.54 for females. When a Mantel-Haenszel weighted odds ratio was calculated after stratification for age in five-year increments, the odds ratio estimate rose to 1.25 (95% confidence interval 0.73 to 2.15) for men and 1.14 (0.73, 1.79) for women.

When former smokers were compared to never-smokers, crude odds ratio estimates of 0.53 and 0.69 were obtained for men and women; odds ratios weighted for age rose only slightly, to 0.63 and 0.97.

In SPSS logistic regression models comparing any smoking history to none, and including the other four variables in the model, the odds ratio estimate was 0.895, with a p-value of .26. In a similar model comparing current smoking to former or never smoking, the odds ratio estimate was 1.29 with a p-value of .098.

Discussion

From this case-control study, we cannot conclude that, controlling for age, educational attainment, and marital status, current or former smokers are more (or less) likely to be admitted to nursing home than never-smokers. The lower odds ratios for crude estimates than for age-controlled estimates reflect the fact that, within the age range studied, younger persons are more likely to be smokers and also less likely to be admitted to nursing home than older persons.

D. Means Testing Analysis

Evaluation Of The Differences In Ever And Never Smokers Nursing Home Average Length Of Stay

The series of graphs and the preliminary case-control study stresses the need to investigate the data further. It is evident that there is a natural grouping in the data. Current and former Smokers dominate age categories from 1 through 79 while never-smokers dominate age categories from 80 through 105. This phenomenon also holds when comparing smoking status with the number of patients discharged dead and length of stay. From this, four groups were formed for testing purposes: ever smokers, never smokers, patients between the age of 55 and 79, and patients between the age of 80 and 105.

Although the previous tests are suggestive, the validity of these tests is diminished by specific problems of the data. Some of the data deficiencies include missing values (as mentioned above), measurement errors, and some censoring of patient information. The data also exhibits a non-normal distribution. These deficiencies limited the use of traditional analysis. Due to these characteristics of the data, survival analysis techniques were used to generate non-parametric statistics.

The choice of survival analysis techniques is appropriate because we are interested in determining the length of stay in the nursing home. Survival analysis is used to investigate the distribution of times between two events. In this study, the two events are admittance and discharge dates. This also determines the patient length of stay. Therefore the critical event in the study is the discharge of the patient. Some of the sample patients were 'lost to follow up' (i.e. the survey ended before the patient was discharged so length of stay cannot be determined); however, survival analysis techniques incorporate information from all patients in the sample, including those who have not experienced the terminal event, discharge. Those patients who have been lost to follow up are called 'censored' observations and are considered incomplete.

The Kaplan-Meier (KM) method was used to estimate non-parametric sample statistics. This procedure is a method of estimating time-to-event models in the presence of censored cases. This method produces a survival function for each group being tested. These survival functions can then be compared to determine if there is a difference among groups according to a particular variable. This method is advantageous because it includes all information, uncensored and censored. Each patient contributes

information as long as they are in the nursing home, and censored patients contribute to the calculation until they are lost to follow up.

Survival Analysis

This study was interested in determining if a relationship exits between smoking and length of stay. Patient cost was calculated by using the length of stay variable as a proxy. Because patient discharge affects the length of stay, it therefore affects the calculation of patient costs. Thus, it is important to have complete information on the discharge status of each patient as well as their length of stay. These variables were then used to generate statistics, which are used to compare length of stay and discharge status rates of smokers and non-smokers.

The distribution of data (see Table 2-2) for the variable discharge status is as follows:

TABLE 2-2

Discharge Status	Frequency	Percent
Expired	127	17.6
Alive	447	62.0
Still There/ Unknown	147	20.4

Censored Observations

Twenty percent of the sample population is missing data on its discharge status. The censored patients are those patients who were still in the nursing home when the study ended. The censored observations add information to the conventional univariate statistics such as the mean, median, etc., but the estimates are not accurate estimators of the respective parameters. Instead, the parameter estimates are a combination of the survival time distribution (data on discharged patients) and a distribution depending on the survival time (data on patients who 'outlast' the study). This analysis defines the event to be evaluated as the resident *discharge* regardless of the status (dead or alive) of the resident when the event occurs.

KM analysis was used to examine the data in three ways:

1) An overall test of equality of LOS for smokers versus non-smokers was conducted for each age category. Median values are compared as well as the log rank test statistic for each group.

2) Female smokers were compared to female non-smokers for each of 2 age categories. The medians and the log rank test statistics were then compared for each age cohort. This process is repeated for male smokers and non-smokers.
3) Female smokers were compared to female non-smokers for each of 3 age categories. The medians and the log rank test statistics were then compared for each age cohort. This process is repeated for male smokers and non-smokers.

One of the fundamental questions of this study was to determine if the length of stay of nursing home patients was affected by smoking. From earlier results, it is also apparent that gender and age play an important, explanatory role when investigating the effects of smoking and nursing home stays. These points were made apparent in the graphs presented in Section 1. The following tests examine the interactive effects of these parameters by considering when and if the patient was discharged.

The data was stratified by age cohort and gender. Smoking status was then chosen as the factor of interest in the KM procedure. The KM method produces several test statistics including the log rank, Breslow, and Tarone-Ware test statistics. The KM method also generates survivorship distributions. Since the data was split by gender and smoking status, the KM procedure produced a plot for each age cohort which compared survival distributions of smokers and non-smokers. This is a visual representation of the null hypothesis, which is a test of equality among groups. Among any group's age cohort, the null hypothesis is that the two survivorship functions are the same. In other words, there is no difference between the probability of discharge of smokers and non-smokers whether the patient is discharged dead or alive. This idea is so important to the basic idea of the study because discharge determines length of stay and the length of stay determines patient costs.

Test 1: Overall Test of Equality of Likelihood of Discharge

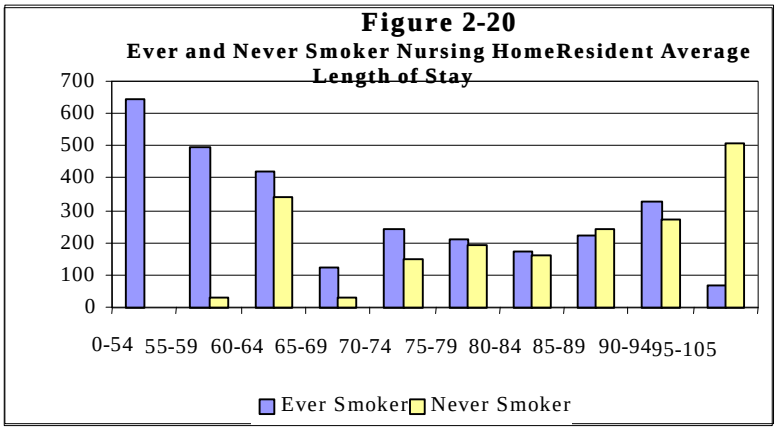
The first test is an overall test of equality of LOS for smokers versus non-smokers. The null hypothesis states that there is no difference between smokers and non-smokers in any age cohort. The data was first stratified by age in 5-year increments. These age categories are the same cohorts that are used in the basic analysis of Sections 1 and 2 and reported on in Table 2-4. It is important to note the relative size of each category. Naturally, any statistics (e.g. mean) formed using these cohorts will be greatly affected by the number of patients in each cohort and smoking status. In fact, the KM method will not produce statistics for cohorts that do not contain a statistically significant sample size. This should be reviewed when evaluating the final results.

Table 2-3
Distribution of Patients by Age Categories and Smoking Status

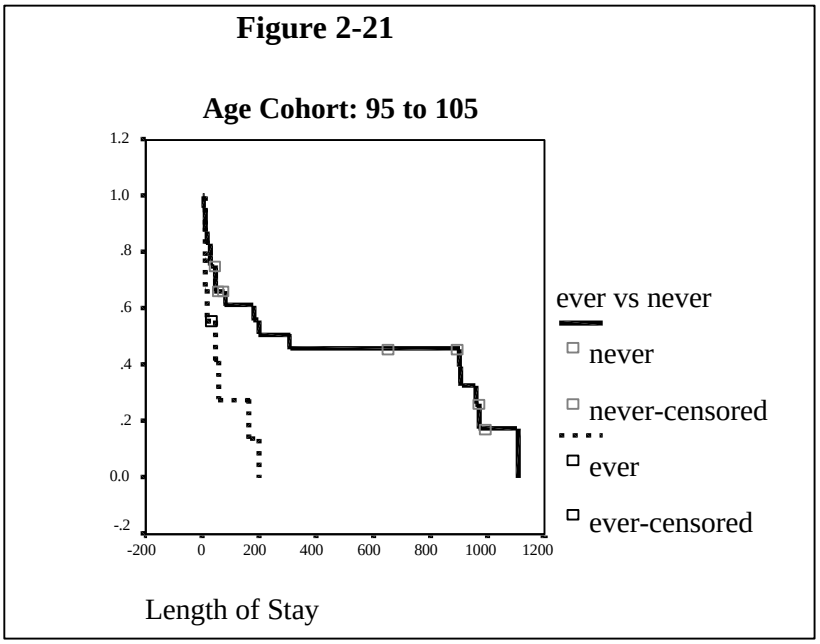
	ever	never	Total
55 to 59	3	4	7
60 to 64	14	6	20
65 to 69	22	18	40
70 to 74	59	36	95
75 to 79	67	60	127
80 to 84	71	74	145
85 to 89	54	70	124
90 to 94	24	57	81
95 to 105	9	24	33
Total	327	350	677

Table 2-4. Results-Test 1: Equality of Survival Distributions for Smoking Status Per Age Cohort

Age Cohort	Mean	Median	Test Statistic		df.	Significance
Age Cohort: 1-54			**Log Rank**	**.33**	1	**.5637**
Smoking Status: Ever	642	.	Breslow	.33	1	.5637
Never	.	.	Tarone-Ware	.33	1	.5637
Age Cohort: 55-59			**Log Rank**	**5.70**	1	**.0169***
Smoking Status: Ever	493	303	Breslow	4.90	1	.0269
Never	31	38	Tarone-Ware	5.31	1	.0212
Age Cohort: 60-64			**Log Rank**	**.09**	1	**.7690**
Smoking Status: Ever	422	31	Breslow	.22	1	.6356
Never	343	17	Tarone-Ware	.17	1	.6838
Age Cohort: 65-69			**Log Rank**	.23	1	**.6329**
Smoking Status: Ever	125	14	Breslow	.00	1	.9449
Never	31	22	Tarone-Ware	.01	1	.9035
Age Cohort: 70-74			**Log Rank**	.01	1	**.6329**
Smoking Status: Ever	239	25	Breslow	.40	1	.9449
Never	151	30	Tarone-Ware	.12	1	.9035
Age Cohort: 75-79			**Log Rank**	.01	1	**.9119**
Smoking Status: Ever	208	43	Breslow	.05	1	.8189
Never	189	43	Tarone-Ware	.04	1	.8389
Age Cohort: 80-84			**Log Rank**	.14	1	**.7117**
Smoking Status: Ever	172	31	Breslow	.44	1	.5082
Never	162	42	Tarone-Ware	.36	1	.5492
Age Cohort: 85-89			**Log Rank**	.01	1	**.9271**
Smoking Status: Ever	224	35	Breslow	.62	1	.4302
Never	241	31	Tarone-Ware	.27	1	.6003
Age Cohort: 90-94			**Log Rank**	.53	1	**.4685**
Smoking Status: Ever	328	80	Breslow	.65	1	.4185
Never	273	42	Tarone-Ware	.59	1	.4432
Age Cohort: 95-105			**Log Rank**	6.21	1	**.0127***
Smoking Status: Ever	70	45	Breslow	3.90	1	.0482
Never	506	308	Tarone-Ware	4.94	1	.0262



The initial results reported on in Table 2-4 are fairly weak under Test 1 specifications. The KM procedure indicates that there are 2 categories in which smokers and non-smokers are significantly different at the .01 level. The age cohorts of 55-59 and 95-105 have significant log rank statistics. The average LOS difference is graphed in Figure 2-20 and the survival functions are for the 95-105 cohort are illustrated below. The line representing each group is called the survival function, and represents the probability that a patient survives beyond the terminal event, discharge. The functions decrease because the probability of *not* being discharged decreases as the number of patients decreases for the longer lengths of stay (Figure 2-21). The pattern of one survivorship function lying above another means the group defined by the upper curve, non-smokers, were less likely to be discharged relative to the lower curve which represents the smokers. This is consistent with the greater mean and median for the non-smokers. The separation of the functions for the two groups indicates that they are significantly different. This result supports our suspicion that non-smokers tend to outlast smokers especially for the very old. However, the rest of the results for Test 1 are weak.



Test 2: Testing the Effect of Gender and Smoking Status with 3 Age Cohorts

This last test was completed to verify the results from Test 2. Because many patients enter the nursing home only for rehabilitative care when they are relatively young, Test 3 was completed to focus on the older cohorts without the influence of younger patients. This test is focused on patients in the second and

third cohorts described below in Tables 2-5 through 2-7 and Figures 2-22 and 2-23. It is evident from the distribution that there are a significant number of patients in the 1-69 age cohort.

Table 2-5

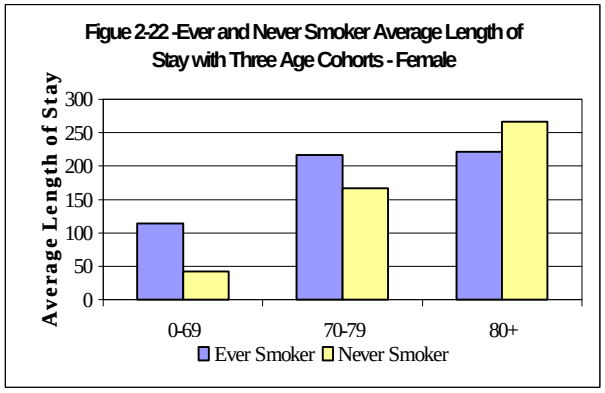
Distribution of Patients by Age Category, Gender, and Smoking Status

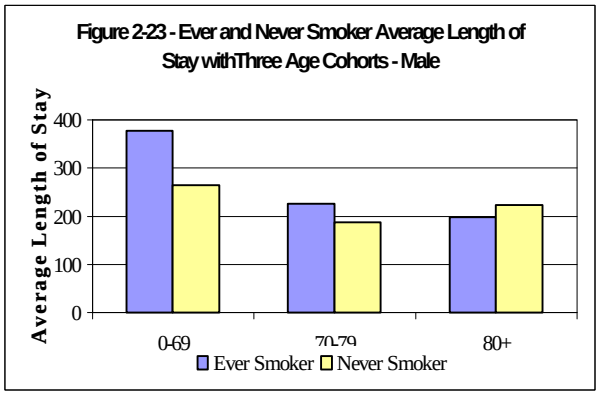
	Smoking Status				Total	
	ever		never			
	female	male	female	male	female	male
1 to 69	16	27	18	11	34	38
70 to 70	65	61	60	36	125	97
80+	80	78	169	56	249	134
Total	161	166	247	103	408	269

- **Results for Test 2: Equality of Survival Distributions for Smoking Status Per Age Cohort and Gender**

Table 2-6 - FEMALES						
Age Cohort	**Mean**	**Median**	**Test Statistic**		**df.**	**Significance**
Age Cohort: 1-69			**Log Rank**	.19	1	**.6633**
Smoking Status: Ever	114	23	Breslow	.10	1	.7521
Never	42	22	Tarone-Ware	.00	1	.9874
Age Cohort: 70-79			**Log Rank**	.01	1	**.9266**
Smoking Status: Ever	217	30	Breslow	.00	1	.9469
Never	167	40	Tarone-Ware	.00	1	.9734
Age Cohort: 80+			**Log Rank**	1.40	1	**.2370**
Smoking Status: Ever	221	39	Breslow	.16	1	.6891
Never	266	44	Tarone-Ware	.47	1	.4908

Table 2-7 - MALES						
Age Cohort	**Mean**	**Median**	**Test Statistic**		**df.**	**Significance**
Age Cohort: 1-69			**Log Rank**	.26	1	**.6085**
Smoking Status: Ever	377	38	Breslow	.67	1	.4117
Never	265	39	Tarone-Ware	.45	1	.5003
Age Cohort: 70-79			**Log Rank**	.01	1	**.9251**
Smoking Status: Ever	226	37	Breslow	.03	1	.8623
Never	188	35	Tarone-Ware	.01	1	.9361
Age Cohort: 80+			**Log Rank**	.08	1	**.7835**
Smoking Status: Ever	198	32	Breslow	.27	1	.6024
Never	223	36	Tarone-Ware	.12	1	.7293





(The graphs of the survival functions are similar to those in the previous selection. They are included in the Appendix.)

Like the results in Test 2, the null hypothesis cannot be rejected for any age cohort. This indicates that there is not statistical difference in the probability of not being discharged between gender specific smokers and nonsmokers for any age cohort. However, never smokers are less likely to be discharged in the older age cohorts.

Test 3: Testing the Effect of Gender and Smoking Status with 2 Age Cohorts

From the preliminary analysis, it is evident that gender and age important. There is a natural break in the data at age 80 and females tend to dominate the older categories. It is evident from the figures below that

females, in particularly non-smoking females, dominate the sample for older cohorts. The analogous graph also illustrated for males and the same result holds for males as for females and again, non-smokers dominate older cohorts.

Table 2-8

Distribution of Patients By Age Categories, Gender, and Smoking Status

Count

		Smoking Status				Total	
		ever		never			
		Gender		Gender		Gender	
		female	male	female	male	female	male
Age Cohorts	Less than 80	*82*	*91*	*84*	*47*	*166*	*138*
	80+	*80*	*78*	*169*	*56*	*249*	*134*
Total		*162*	*169*	*253*	*103*	*415*	*272*

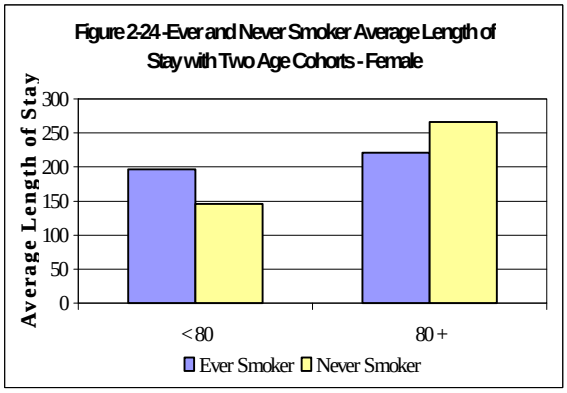
The KM methodology was used to test female (male) smokers in age cohorts 1 through 79 and 80+ against female (male) non-smokers. A priori, we anticipate that female and male nonsmokers will have different survival functions than those of smokers. Again, under the null hypothesis, the smokers and non-smokers within a particular age cohort and gender will have the same likelihood of not being discharged.

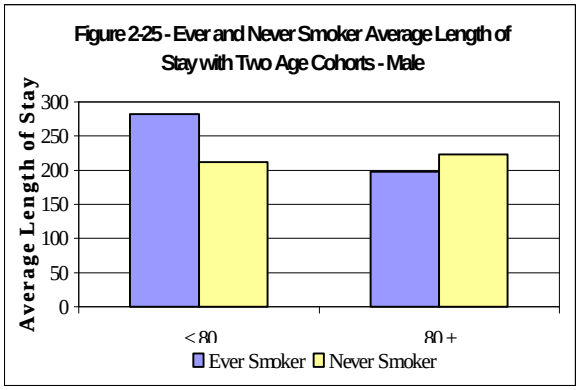
When the data is grouped into 2 large age categories, the mean and median values will have greater explanatory power. The Tables 2-8 through 2-10 and Figures 2-24 through 2-27 below illustrate the distribution of patients in each category and their different LOS of both genders in each category. Again, it is important to recognize the actual number of observations in category when considering the statistical results and their validity. The results of the KM analysis of the difference in survival distributions are presented below.

- **Results for Test 3: Equality of Survival Distributions for Smoking Status Per Age Cohort and Gender**

Table 2-9 – FEMALES

Age Cohort	Mean	Median	Test Statistic		df.	Significance
Age Cohort: <80			**Log Rank**	.06	1	**.7634**
Smoking Status: Ever	197	30	Breslow	.02	1	.8892
Never	146	36	Tarone-Ware	.01	1	.9410
Age Cohort: 80+			**Log Rank**	1.40	1	**.2370**
Smoking Status: Ever	221	39	Breslow	.16	1	.6891
Never	266	44	Tarone-Ware	.47	1	.4908





The KM procedure indicates that there is no significant difference between smokers and non-smokers for either age cohort. We accept the null hypothesis. The survival functions are illustrated below in Figures 2-26 and 2-27.

In the graph of females over 80, the functions flip. This indicates that non-smokers are less likely to be discharged. This result also holds with the original findings that never smokers are less likely to be discharged in the older age cohorts. However, none of the test statistics were significant for the sub-groups. This is also evident in the graphs by the functions lying so close to each other. The results for the males are similar to those of the females.

Figure 2-26

Females

Less than 80

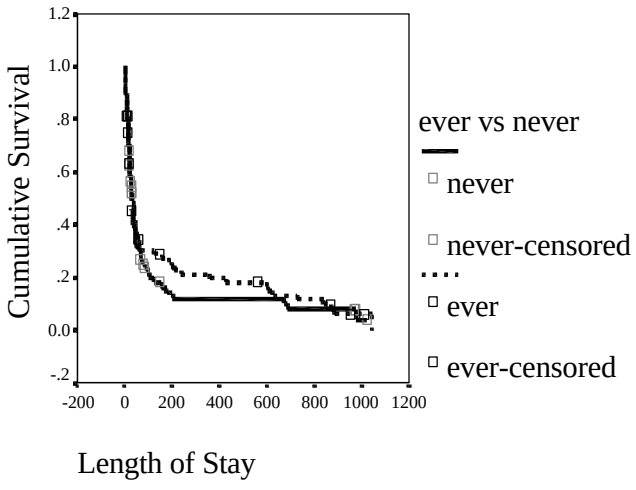


Figure 2-27

Females

80+

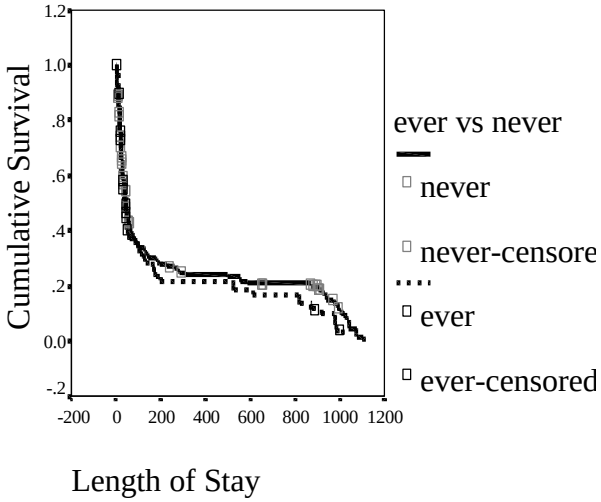


Table 2-10 –MALES						
Age Cohort	**Mean**	**Median**	**Test Statistic**		**df.**	**Significance**
Age Cohort: <80			**Log Rank**	.21	1	**.6440**
Smoking Status: Ever	282	38	Breslow	.17	1	.6796
Never	212	35	Tarone-Ware	.23	1	.6296
Age Cohort: 80+			**Log Rank**	.08	1	**.7835**
Smoking Status: Ever	198	32	Breslow	.27	1	.6024
Never	223	36	Tarone-Ware	.12	1	.7293

The survival functions are illustrated below.

The same phenomenon occurs for the male sub-group (see Figures 2-28 and 2-29). Smokers are less

Figure 2-28

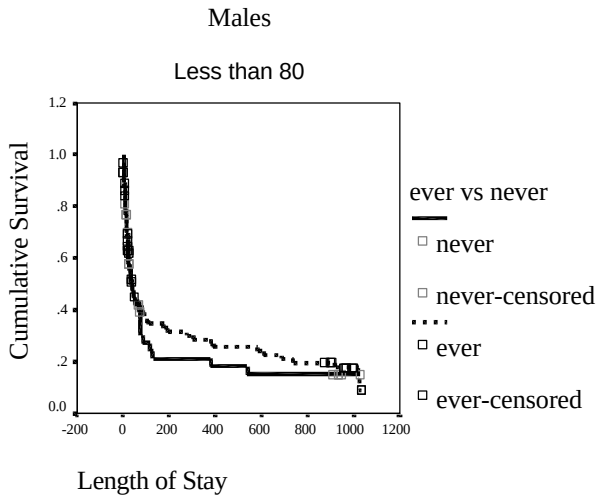
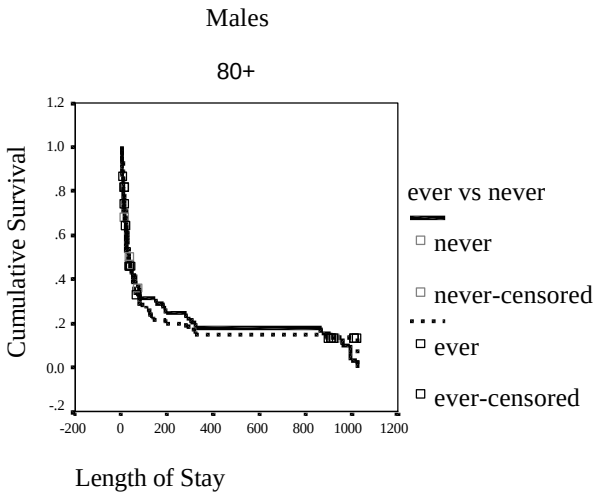


Figure 2-29



likely to be discharged for the less than 80 cohort and are more likely to be discharged relative to the non-smoking group. A problem can occur if the estimated functions cross one another. This means that in some time intervals one group has a more favorable experience and in other time intervals the other group has a more favorable experience. There is some evidence of crossing in the graph in Figure 2-29 of males over 80. However, this is representative of the dwindling numbers of male patients which actually stay in the nursing home for long lengths of stay.

Section 3. Dose Response Linear and Cox Regression Analysis

A. Introduction to the Dose Response Regression Modeling Analysis

The previous analysis demonstrated that ever smokers tend to incur longer lengths of stays in nursing homes. The second phase of the analysis involved evaluating the magnitude and statistical relationship of key variables. More specifically, this analysis evaluates the dose response relationship of the effects of independent smoking variables, such as number of years and number of cigarettes smoked on dependent variables of interest, such as, LOS and discharge status. Standard linear[1] and Survival Analysis Cox Regression techniques[2] are used to examine the dose response relationships.

The results of these analyses tend to support the findings of a statistical association between independent socio-policy variables and the independent variables, LOS, discharge age, and discharge status. The robustness of the results is examined using standard, regression diagnostic statistics. These statistics include but are not limited to the R^2, t and F statistics and their respective *p* values.[3]

Dose Response Regression Modeling of the Smoking Effects on Nursing Home Residents

The medical and social role of nursing homes has shifted dramatically over the past few years from a primarily, long-term caretaker of the chronically ill to that of *short-term*, sub-acute care of temporarily ill or injured elderly. Until recently, this latter function was historically performed by acute care hospitals and is now the domain of the nursing homes because of their overall lower capital and labor costs.
According to the AHCA analysis, currently almost *70%* of all entering nursing home residents will stay in the nursing home for less than 90 days.[4] Despite this significant shift in magnitude to short-term

[1] Standard linear regression was deemed acceptable in this setting during statistical consulting sessions with Dr. Douglas Zahn and his associate Glenn Laird, FSU Statistical Consulting Center, September 11, 1998
[2] Survival Analysis and Cox regression techniques were introduced into this analysis at the recommendation of Richard Hopkins, MD, MPH, State of Florida, Epidemiologist, Florida Department of Health, and by Dr. Douglas Zahn and his associate Glenn Laird, FSU Statistical Consulting Center
[3] Other diagnostic statistics such as residual plots and ANOVA were also examined. Complete results are also available by written request to the authors.
[4] "Nursing Home Reporting System", 1993-1997 Annual Reports, Agency for Health Care Administration, State Center for Health Care Administration, Tallahassee, Florida

residents, long term residents (90+ days) constitute a vast majority of nursing home revenues. Therefore, due to the changing composite of nursing home patients, the data is split into two sections.

This splitting of the data helps to facilitate the study and is an important step in evaluating the questions of interest. The study is interested in determining the costs associated with smoking behavior and uses LOS as a proxy for measuring costs. Splitting the data allows those patients who were admitted for rehabilitative treatment (shorter lengths of stay) to be filtered out. These patients generally do not change their smoking behavior due to nursing home admittance and do not significantly add to nursing home revenues.

The patients of interest in this section of the study are those patients who are Medicaid recipients and those who have already depleted Medicare eligibility and family resources. This group represents a substantial portion of nursing home revenues and costs to the state of Florida. Therefore, the sample data is split into two groups: group 1 contains patients whose LOS exceeded 21 days and the remainder is put into group 2. This bifurcation of the data is deemed especially important in this analysis, which seeks to use LOS as the key dependent, proxy variable for evaluation of smoker and non-smoker cost differences. High volumes of shorter lengths of stay tend to obfuscate this analysis and lend little insight beyond initial descriptive inferential analysis.

Additionally, there seems to be two countervailing forces at work within the nursing home population under examination. The first is that those residents who are smokers tend to experience longer lengths of stay in nursing homes in each of the younger age cohorts. The second is that these same individuals tend to die much more frequently after admittance to Florida nursing homes. Smokers tend to die more frequently and at an earlier age than non-smokers, which reduces their total length of stay (had they continued to lived).[5]

Confounding these factors is the likelihood that never smokers tend to both live longer and stay longer in the older years among those surveyed. These countervailing and offsetting trends tends to mask total nursing home sample length of stay differences between ever and never smokers across all age groups. It is also worth noting that this splitting of the patient data by LOS for analytical purposes is common

[5] Data analysis later presented in this report attempted to examine these differences across age groups by comparing discharge events for living and non-living residents separately.

practice among other researchers and AHCA when evaluating nursing home patient data for a variety of policy purposes.[6]

B. LINEAR REGRESSION MODELS

Length of Stay Regression Models

All ever-smoking residents, those discharged dead, alive or still in the nursing home are included in the following analyses. Cases with missing observations in the discharge status variable are excluded. This analysis, by definition, is restricted to only those nursing home residents admitted who were 80 years of age or younger and excludes relatively shorter lengths of stay. These models primarily evaluate the relative length of stay of those ever smokers who responded to the question of the average number of cigarettes consumed per day on the survey. Thus it is an evaluation of the relative effects on LOS of a dose or level of cigarettes consumed by the patient.

Regression Model Equation 1

Residents Under Age 80 and LOS Over 21 Days

Model 1 uses a subset of the sample data and only includes patients under the age of 80 and with a LOS greater than 21 days. This model evaluates the relationship between the average number of cigarettes consumed over a lifetime and the resident's total length of stay. A priori, we expect the coefficient for the variable representing the average number of cigarettes consumed over a lifetime to have a positive sign and to be significant.

The results in Table 3-1 indicate a strong positive relationship at a *.05 level of significance* with a *t*-statistic of 2.232 and a *p*-value of .034 for the average daily number of cigarettes consumed. The R^2 has a value of .389 and the F value is 4.98 with a significance level of .034.

Specifically, these results indicate that for every additional cigarette consumed per day the average nursing home client's length of stay in the nursing home increases by 8.987 days for patients 79 years of age and younger.

[6] Ibid AHCA, 1993-97

Regression Model Equation 2

Residents Age 80 and Older and LOS Over 21 Days

Model 2 also uses a subset of the sample data and only includes patients who are 80 years of age and older and with a LOS greater than 21 days. This model evaluates the relationship between the average number of cigarettes consumed over a lifetime and the resident's total length of stay. Again, we anticipate the coefficient for the variable representing the average number of cigarettes consumed over a lifetime to have a positive sign and to be significant.

The results in Table 3-1indicate a strong positive relationship at a *.01 level of significance* with a *t*-statistic of 2.597 and a *p*-value of .012 for the average daily number of cigarettes consumed. The R^2 has a value of .316 and the F value is 6.74 with a significance level of .012.

Specifically, these results indicate that for every additional cigarette consumed per day the average nursing home client's nursing home length of stay increases by 5.520 days for patients 80 years of age and older. These models suggest a positive and strong statistical association between the average number of cigarettes consumed per day and nursing home length of stay for individuals in various age groups.

Table 3-1
The effect of a dose or level of cigarettes on patients whose LOS was greater than 21 days
Dependent variable: LOS

	Sample	Independent Variable	Coefficient	*t*-statistic	*p* value	R^2	F statistic	F sig.
Model 1:	Patients whose age < 80 years	Constant	-26.022			.121	4.98	.034
N=30		Avg. Cigarettes Smoked per Day	8.987	2.232	.034			
Model 2:	Patients whose age ≥ 80 years	Constant	49.714			.085	6.74	6.74
N=30		Avg. Cigarettes Smoked per Day	5.520	2.597	.012			

Variables in italics are categorical variables.
'N' - the number of observations for that model.

*Denotes significance at the .01 level.
** Denotes significance at the .05 level.

*** Denotes significance at the .10 level.

Regression Model Equation 3

All Aged Residents and Discharge Age – Average Daily Number of Cigarettes and Years Married and Years Smoked and Number Years Not Smoked

Model 3 examines the relationship between age at discharge and the following explanatory variables: years not smoking, the number of years married, the number of years smoked, and the average number of cigarettes consumed per day. This model uses the full sample and makes no restriction on the length of stay (see Table 3-2).

Each of these four variables has a statistically significant effect on a patient's discharge age. (See Table 3-2). The R^2 is .187 and the F-value is 6.74 with a *p*-value of .012.

Regression Model Equation 4

All Aged Residents and Discharge Age – Years Smoked and Years Not Smoked

Regression model 4 evaluates the relationship between the number of years spent both smoking and not smoking and the age of the patient at discharge. This model suggests that for every year a person smokes their age at discharge should increase by 0.547 years. This seems to contradict earlier findings where smokers tend to die at an *earlier age* compared to never smokers. However, this equation also indicates that for every year a person does ***not*** smoke their age at discharge should increase by 0.657.

Initially these results may seem like a logical contradiction. However, status at discharge is not included in this equation. This contradictory result captures two trends in the sample population. Patients who are smokers tend to die sooner and are therefore released at a younger age. These patients are also often afflicted with chronic diseases where they are admitted on several different occasions over a period of time. The second trend in the data involves the number of years not smoking. The values in this variable are derived by taking the difference between the discharge age and the age when the patient stopped smoking permanently. Previous studies have shown that there are benefits from quitting smoking no matter the age. The results for this variable support these previous studies. (see Table 3-2).

Table 3-2
The effect of a dose or level of cigarettes on patients whose LOS was greater than 21 days
Dependent variable: DISAGE

	Sample	Independent Variable	Coefficient	*t*-statistic	*p* value	R^2	F statistic	F sig.
Model 3:	All Patients	Constant	61.513			.187	6.74	.012
N=103		Number of Years not Smoking	.102	2.245	.027			
		Number of Years Married	.157	3.416	.001			
		Number of Years Smoking	.262	2.745	.007			
		Avg. Cigarettes Smoked per Day	.114	2.446	.015			
Model 4:	All Patients	Constant	44.786			.548	62.853	.000
N=103		Number of Years Smoking	.547	9.931	.000			
		Number of Years not Smoking	.657	11.2	.000			

Variables in italics are categorical variables.
'N' - the number of observations for that model.

*Denotes significance at the .01 level.
** Denotes significance at the .05 level.
*** Denotes significance at the .10 level.

C. Cox Regression Models

A Cox Regression Primer

The following description is a brief introduction to the Cox regression model. It includes the most pertinent terms needed to understand the results in this study. Therefore, the description of the statistical tool is couched in the terms and variables for *this particular study*. (Note: This description is taken directly from *'Chapter 23: Cox Regression Examples'*, **SPSS Advanced Statistics**, **SPSS® Base 7.5 for Windows®**, 1997.)

Why use the Cox regression model as opposed to the standard linear regression model?

> "Multiple linear regression is a technique used to determine the influence of predictor variables on a dependent variable. However, linear regression has no mechanism for handling censored cases. (Censored cases are those cases that were 'lost to follow up', i.e. the survey ended before the patient was discharged so length of stay cannot be determined.) A nonlinear model is also more reasonable as an approximation for a variable, such as survival time or hazard rate, that cannot take on negative values."[7]

Terminology and Interpretation:

Hazard Function, h(t): Estimates the relative risk of discharge given a particular unit of time. If the dependent variable is LOS, the unit of time is one day. If the dependent variable is DISCHARGE AGE, the unit of time is one year. The hazard function is a rate. The hazard function is not a probability and therefore it can exceed 1.

A high hazard function indicates a high rate of discharge.

Survival Function, S(t): Estimates the relative likelihood of *not being discharged* for a given unit of time. This value is a probability. The survival function is the inverse of the cumulative hazard function. $H(t) = -\ln S(t)$

A high survival function indicates a higher probability of not being discharged.

Covariate, X_i: Independent variables which can be categorical or continuous.

[7] Linear regression models were also tested and used in this report as recommended by Dr. Douglas Zahn and Glen Laird, Department of Statistics, Center for Statistical Consultation, September 8, 1998

General Form of the Cox Regression:

$$h(t) = [h_0(t)]e^{(B_1X_1 + B_2X_2 + \ldots + B_pX_p)} \qquad \text{Equation (1)}$$

where B is a regression coefficient, X_I is a covariate, *e* is the base of the natural logarithm, $h_0(t)$ is the baseline hazard function when X equals zero for one dichotomous covariate.[8] An alternate form is:

$$Y = B_1X_1 + B_2X_2 + \ldots B_pX_p \qquad \text{Equation (2)}$$

The Cox Regression

The regression is designed similarly to an ordinary least squares regression as is evidenced in Equation (2) and described in footnote 1.

A Cox Regression must contain at least two variables: a dependent variable that acts as a survival time indicator and an independent variable that acts as a status indicator. The survival time indicator for this study is LOS. Notice that this variable is incremented in days and indicates the days of survival (i.e. not being discharged) in the nursing home. The status indicator for this study is DISCHARGE STATUS, and it records whether the event, discharge, has occurred. The remaining independent variables are also entered with the status indicator as predictors of the dependent variable, LOS.

Univariate Form:

$$Y = B_1X_1$$

Multivariate Form:

$$Y = B_1X_1 + B_2X_2 + \ldots B_pX_p$$

<u>Dependent Variable</u>
LOS

<u>Independent Variables</u>
Gender, Marriage Status, Homesize,
Total number of cigarettes smoked, etc.

<u>Status Indicator (this is entered separately)</u>
Discharge Status

[8] "This equation actually closely resembles an ordinary regression equation through a simple manipulation. Starting with the original equation: $h(t) = [h_0(t)]e^{(B_1X_1 + B_2X_2 + \ldots + B_pX_p)}$ divide both sides by $h_0(t)$ and take the natural log is taken of both sides. This results in: $\ln[h(t)/h_0(t)] = B_1X_1 + B_2X_2 + \ldots B_pX_p$. If $\ln[h(t)/h_0(t)]$ is set equal to Y, the formula now resembles an ordinary regression equation."

Interpreting the Output

Table 3-3: Description of the Results From the Cox Regression Procedure.

	OUTPUT	INTERPRETATION FOR THIS STUDY
B	The estimated coefficient. It is interpreted as the predicted change in log hazard for a unit increase in the predictor.	(-) negative value As the predictor variable is incremented by one unit, the risk of discharge decreases. (+) positive value As the predictor variable is incremented by one unit, the risk of discharge increases.
S.E.	Standard error of the estimated coefficient.	
Wald	A test statistic.	It is used to test whether the estimated coefficient B is different from 0 in the population. The larger the value the less support for the null hypothesis that B = 0.
Sig.	It is the significance level for the Wald statistic.	A value less than .01[†] indicates that the variable is significant. The smaller the value the less support for the null hypothesis that B = 0.
Exp (B)	It estimates the percentage change in risk with each unit change in the covariate. (If there is only one categorical covariate, this is a measure of relative risk.)	*Exp (B)* > 1: Each unit increase in the covariate is expected to increase the risk of discharge, e.g. an *Exp (B)* = 1.50 indicates that the risk of discharge is .50 times more likely. *Exp (B)* < 1: Each unit increase in the covariate is expected to make the chance of discharge less likely, e.g. an *Exp (B)* = .50 indicates that the risk of discharge is .50 times less likely.

[†] There are three acceptable levels of significance: the .01 level (the strongest), the .05 level, and the .10 level (the weakest).

The Cox Regression Analysis Models

Previous linear regression analysis (Section 3-B) indicated that there were several variables that were significant in explaining length of stay in the nursing home, such as marriage and nursing home size. Based on these results, the Cox procedure was initially conducted on single variables and then in a multiple regression format.

Following the earlier results the data was evaluated in three classifications:

(1) The full sample of data is used in the analysis.
(2) The data sample included only those patients who were 79 years of age or younger.
(3) The data sample included only those patients who were 80 years of age or older.

(1) Results from the Full Sample Analysis

Regression: Univariate (individual covariates)
Dependent Variable: Length of Stay
Status Indicator (used for censoring the data): Discharge Status

The results in Table 3-4 are derived by including only one independent variable at a time.

Table 3-4 - FULL SAMPLE -- UNIVARIATE RESULTS

	N	Independent Variable	B	S.E.	Wald Statistic	Sig.	*Exp(B)*
Model 1:	698	*Gender*	-.0035	.0871	.0016	.9681	.9965
Model 2:	698	*Married*	-.3377**	.1644	4.2186	.0400	.7134
Model 3:	698	Approximate Number of Cigarettes Smoked	-.0001	.0001769	.3212	.5709	.9999
Model 4:	698	Number of Years Smoking	.0004918	.0019	.0663	.7968	1.0005

Variables in italics are categorical variables.
'N' - the number of observations for that model.

*Denotes significance at the .01 level.
** Denotes significance at the .05 level.
*** Denotes significance at the .10 level.

From the results, it is evident that one variable is significant of the four tests completed (see Table 3-5). The variable 'Married" is a categorical variable (0- ever married, 1-unmarried), which means that

Exp(B) can be interpreted as the relative risk between unmarried and married individuals. Since *Exp(B)* is less than 1, it indicates that there is a decreased relative risk of discharge when a patient is unmarried. The hazard of discharge for unmarried patients is about 71% that of the hazard of discharge for married patients when no other variables are considered. The negative value of the regression coefficient indicates that the risk of discharge is less for unmarried patients relative to married patients. This result implies that patients that have never been married incur longer lengths of stay. This finding supports the results of the linear regression models described earlier regarding the effect of marriage on length of stay.

The model for Cox regression was then expanded to include more than one covariate.

Regression: Multiple Covariates
Dependent Variable: Length of Stay
Status Indicator (used for censoring the data): Discharge Status
Number of Observations: 698

Table 3-5 -FULL SAMPLE -- MULTIVARIATE RESULTS

Independent Variable	**B**	**S.E.**	**Wald Statistic**	**Sig.**	***Exp(B)***
Gender	-.0017	.0910	.0003	.9854	.9983
Married	-.3311**	.1652	4.0177	.0450	.7181
Approximate Number of Cigarettes Smoked	-.000227	.0002399	.8927	.3447	.9998
Number of Years Smoking	.0021	.0026	.6372	.4247	1.0021
Homesize	.2279*	.0860	7.0129	.0081	1.2559

Variables in italics are categorical variables.

*Denotes significance at the .01 level.
** Denotes significance at the .05 level.
*** Denotes significance at the .10 level.

These results indicate that two variables are significant in this test, married and homesize. The 'married' variable remains significant at the .05 level. The hazard of discharge for unmarried patients is approximately 72% that of the hazard of discharge for married patients when no other variables are considered. This finding is very similar to the result in the univariate test reported above. The 'homesize' variable is strongly significant at the .01 level. This indicates that patients in larger nursing homes are 25% more likely to be discharged relative to those in small nursing homes. This finding is consistent with

the fact that many larger nursing provide rehabilitative programs for the seriously injured and often admits and then re-admits the chronically ill.

The data was then split into two groups, discharge age ≤ 79 and discharge age ≥ 80, and analyzed in the same way as above (see Tables 3-5 though 3-9).

(2) Results from the Discharge Age ≤ 79 Sample Analysis

Regression: Univariate (individual covariates)
Dependent Variable: Length of Stay
Status Indicator (used for censoring the data): Discharge Status
The following results are derived by including only one independent variable at a time.

Table 3-6 -DISCHARGE AGE ≤ 79 SAMPLE -- UNIVARIATE RESULTS

	N	Independent Variable	B	S.E.	Wald Statistic	Sig.	*Exp(B)*
Model 1:	308	*Gender*	-.1958	.1306	2.2471	.1339	.8222
Model 2:	308	*Married*	-.6432**	.2529	6.4662	.0110	.5256
Model 3:	308	Approximate Number of Cigarettes Smoked	-.0002	.0002416	.6865	.4074	.9998
Model 4:	308	Number of Years Smoking	.0013	.0029	.1998	.6549	1.0013

Variables in italics are categorical variables.

*Denotes significance at the .01 level.
** Denotes significance at the .05 level.
*** Denotes significance at the .10 level.

These results are very similar to the Full Sample univariate results. Again, since *Exp(B)* is less than 1 for the married variable, it indicates that there is a decreased relative risk of discharge when a patient is unmarried. The hazard of discharge for unmarried patients is about 53% that of the hazard of discharge for married patients when no other variables are considered. The negative value of the regression coefficient indicates that the risk of discharge is less for unmarried patients relative to married patients. In other words, unmarried patients are less likely to be discharged and are more likely to incur a longer length of stay.

Regression: Multiple Covariates
Dependent Variable: Length of Stay
Status Indicator (used for censoring the data): Discharge Status
Number of Observations: 308

Table 3-7 - DISCHARGE AGE ≤ 79 SAMPLE -- MULTIVARIATE RESULTS					
Independent Variable	**B**	**S.E.**	**Wald Statistic**	**Sig.**	***Exp(B)***
Gender	-.2117	.1356	2.4371	.1185	.8092
Married	-.6383**	.2568	6.1787	.0129	.5282
Approximate Number of Cigarettes Smoked	-.000595***	.000353	2.8306	.0925	.9994
Number of Years Smoking	.0083**	.0041	3.999	.0455	1.0083
Homesize	.1142	.1326	.7424	.3889	1.121

Variables in italics are categorical variables.

*Denotes significance at the .01 level.
** Denotes significance at the .05 level.
*** Denotes significance at the .10 level.

These results indicate that three variables are significant in this test, marriage status, approximate number of cigarettes smoked and the number of years smoking. *Exp(B)* is less than 1 for the variable married indicating that there is a decreased relative risk of discharge when a patient is unmarried for this age group. The hazard of discharge for unmarried patients is about 52% of the hazard of discharge for married patients. Therefore, an unmarried patient is more likely to incur a longer length of stay relative to a married patient. This approximate number of cigarettes smoked indicates that for every additional cigarette smoked the patient is less likely to be discharged and will incur a longer length of stay. The number of years smoking variable indicates that for every additional year spent smoking the patients is 0.83% more likely to be discharged. (It is important to note that 'type of discharge-alive or expired', is not indicated by this test.)

(3) Results from the Discharge Age ≥80 Sample Analysis

Regression: Univariate (individual covariates)
Dependent Variable: Length of Stay
Status Indicator (used for censoring the data): Discharge Status

The following results are derived by including only one independent variable at a time.

Table 3-8 - DISCHARGE AGE ≥ 80 SAMPLE -- UNIVARIATE RESULTS

	N	Independent Variable	B	S.E.	Wald Statistic	Sig.	*Exp(B)*
Model 1:	390	*Gender*	.1497	.1179	1.6124	.2042	1.1615
Model 2:	390	*Married*	-.0233	.2168	.0116	.9144	.9770
Model 3:	390	Approximate Number of Cigarettes Smoked	-.00000911	.0002642	.0012	.9725	1.000
Model 4:	390	Number of Years Smoking	-.000294	.0026	.0127	.9103	.9997

Variables in italics are categorical variables.

*Denotes significance at the .01 level.
** Denotes significance at the .05 level.
*** Denotes significance at the .10 level.

These results indicate that none of the variables are significantly different from zero. This may indicate that the composition of this age group is very different. Variables outside this analysis may be important but their effects are not captured in these variables.

Regression: Multiple Covariates
Dependent Variable: Length of Stay
Status Indicator (used for censoring the data): Discharge Status
Number of Observations: 78

Table 3-9 –DISCHARGE AGE ≥ 80 SAMPLE -- MULTIVARIATE RESULTS					
Independent Variable	**B**	**S.E.**	**Wald Statistic**	**Sig.**	***Exp(B)***
Gender	.1732	.1230	1.9830	.1591	1.1890
Married	-.0281	.2185	.0166	.8975	.9722
Approximate Number of Cigarettes Smoked	.00005823	.0003412	.0291	.8645	1.0001
Number of Years Smoking	-.0019	.0034	.3142	.5751	.9981
Homesize	.4620***	.2614	3.1235	.0772	1.5872

Variables in italics are categorical variables.

*Denotes significance at the .01 level.
** Denotes significance at the .05 level.
*** Denotes significance at the .10 level.

These results indicate that only one variable is significant in this test, 'homesize'. The results for 'homesize' indicate that the risk of discharge is about 58% higher for patients in larger nursing homes. This may be due to the range of services offered by large nursing homes. These homes tend to provide rehabilitative services as well as providing long term care. Patients who are admitted for relatively shorter lengths of stay may be driving this result. It could also be a factor of the sample subset. These individuals are amongst the oldest members in the overall sample. Older patients generally have more health problems, which may cause multiple admits for shorter time periods. These patients may be using the nursing home for short term rehabilitative services.

Effects of Smoking on Risk of Death and on Length of Stay in Patients Who Did Not Die

In order to explore possible reasons for inconsistent and weak results linking smoking history with length of stay, the following models represent two complementary analyses. In one analysis, we asked whether the survival time until death was different for ever-smokers versus never-smokers. In the other, we asked whether the length of stay in the nursing home was different for ever-smokers versus never-smokers, among persons who did not die in the nursing home. If, as we might expect based on the known health effects of cigarette smoking, people who had ever smoked were more ill on admission to the nursing home, and therefore died sooner after admission than never-smokers, then we would expect survival till death to be shorter in ever-smokers than in never-smokers. Similarly, we might expect that ever-smokers who do not die may have a longer length of stay in the nursing home than never-smokers who do not die.

The first effect would tend to give ever-smokers a shorter stay in the nursing home, while the second would tend to give ever-smokers a longer stay in the nursing home. As these effects are in opposite directions, they may tend to cancel each other out in analyses that simply examine length of stay in relation to past and present cigarette use.

In the first analysis, we built a model with the following variables: gender, marital status (ever- versus never-married), size of nursing home, and a series of dummy variables for five-year age groups. The model measured the hazard of discharge dead – a higher hazard for ever-smokers meant a shorter length of stay until death. Persons who were discharged alive, or were still alive and in the nursing home at the time of the study, were considered censored for this analysis – that is, their vital status was known as of a certain date, and at that date they had not yet died.

Results are shown in the Table 3-10. As ever-smoker is coded 1 and never-smoker is coded 0, the positive sign of the coefficient for the smoking variable tells us that ever-smokers had a higher hazard of discharge (expired), or shorter length of stay, than never-smokers. The value of the hazard function for ever-smokers is 1.31 times as high as for never-smokers – that is, ever-smokers are more likely to die. The standard error for the coefficient is large, however, and the 95% confidence limits for the coefficient overlap zero – the p-value associated with the estimate is .18, well above the conventional cutoff of .05.

The variable, homesize, is significant at the .01 level indicating that patients in large homes are 49% less likely to be discharged dead relative to patients in small nursing homes. This is consistent with the fact that large nursing homes offer rehabilitative services where patients are discharged alive after their program of care is complete. None of the age groups is significantly associated with an increased hazard of death. Being male is significantly associated with a higher risk of death, with a 1.52 fold increase in hazard of death. Having ever been married is associated with increased risk of discharge (death) but this variable is not significant. Finally, being in a smaller home is associated with greater hazard of death.

Regression: Multiple Covariates
Dependent Variable: Length of Stay
Status Indicator (used for censoring the data): Discharge Status-(those still in the home and those discharged alive are censored)
Number of Observations: 663

Table 3-10 MULTIVARIATE RESULTS - Dead at Discharge

Independent Variable	B	S.E.	Wald Statistic	Sig.	*Exp(B)*
Gender	.4191**	.1973	4.5119	.0337	1.5207
Married	.4822	.3769	1.6362	.2008	1.6196
Smoking: Ever vs. Never	.2682	.2004	1.7899	.1809	1.3076
Homesize	-.7094*	.2230	10.1209	.0015	.4920
Age2dum	-.7205	.8265	.7599	.3834	.4865
Age3dum	-.1656	.5367	.0953	.7576	.8473
Age4dum	.5259	.4607	1.3031	.2536	1.6921
Age5dum	.2763	.4756	.3374	.5613	1.3182
Age6dum	.4734	.4684	1.0217	.3121	1.6055
Age7dum	.4157	.4922	.7134	.3983	1.5155
Age8dum	.6496	.5564	1.3630	.2430	1.9148

Variables in italics are categorical/dummy variables.

*Denotes significance at the .01 level.
** Denotes significance at the .05 level.
*** Denotes significance at the .10 level.

In the second analysis, persons known to have died in the nursing home are excluded from the analysis. Time is measured till discharge. If the subject is still in the nursing home at the time of the study, his/her observation is considered to be censored. What we are measuring here, then, is the relationship between

smoking status and hazard of discharge, or length of stay, given that the subject did not die. The model contains the same variables as before.

Results are shown in the following Table 3-12. As ever-smoker is coded 1 and never-smoker is coded 0, the negative sign of the coefficient for the smoking variable tells us that ever-smokers had a lesser hazard of discharge, or longer average length of stay, than never-smokers. The value of the hazard function for ever-smokers is 0.89 times as high as for never-smokers – that is, ever-smokers are less likely to be discharged, and thus have longer stays.

The standard error for the coefficient is large, however, and the 95% confidence limits for the coefficient overlap zero – the p-value associated with the estimate is .26, well above the conventional cutoff of .05. Having ever been married is associated with a higher risk of discharge, with a 1.41 fold increase in hazard of death, but this is not statistically significant at the .05 level. Not surprisingly, younger age was associated with a higher hazard function for discharge alive and thus a shorter average length of stay. Finally, being in a larger home is associated with greater hazard of discharge, or shorter length of stay.

Regression: Multiple Covariates
Dependent Variable: Length of Stay
Number of Observations: 545

Table 3-11 MULTIVARIATE RESULTS - Alive at Discharge

Independent Variable	B	SE	Wald Statistic	Sig.	*Exp(B)*
Gender	-.0376	.1085	.1198	.7293	.9631
Married	.3440***	.1974	3.0363	.0814	1.4105
Smoking: Ever vs. Never	-.1167	.1037	1.2666	.2604	.8898
Homesize	.4004*	.1008	15.7785	.0001	1.4924
Age2dum	1.0135*	.3424	8.7583	.0031	2.7551
Age3dum	.7240**	.3222	5.0493	.0246	2.0628
Age4dum	.5109	.3166	2.6042	.1066	1.6668
Age5dum	.6085***	.3130	3.7794	.0519	1.8378
Age6dum	.4599	.3194	2.0738	.1498	1.5839
Age7dum	.1754	.3361	.2723	.6018	1.1917
Age8dum	-.2186	.3983	.3012	.5831	.8036

Variables in italics are categorical/dummy variables.

*Denotes significance at the .01 level.
** Denotes significance at the .05 level.
*** Denotes significance at the .10 level.

Neither of these analyses produced a strong or statistically significant effect of smoking, *but in both analyses the directions of the observed effects are those predicted by the known health effects of tobacco use*: **an increased risk of early death after admission to nursing home, and then, if surviving, a longer stay for tobacco-user**s. As these two effects tend to cancel each other out, the relatively small magnitude of the observed effects of smoking on length of stay, when persons discharged dead or alive or still living in the nursing home are considered together, is not unexpected.

Section 4. Final Results Summary and Conclusions

A. Summary

Section 1 Summary

Section 1 overviews the current status of tobacco damage, health research in the US and set forth the need for the analysis and the null hypothesis this we focused on in the research that follows in Sections 2 and 3 of this report.

Section 2 Summary

Section 2 overviews the survey sample data secured for this analysis and a preliminary comparison of the sample data to the population from which it was drawn. Section 2 thereafter completes an extensive profile of the age, gender, regional dispersion location, and length of stay for the sample population. Comparisons and contrasts of the average lengths of stay of ever and never smokers by gender and age are then completed. This is followed by a case study evaluation of the rate of admission to nursing homes (from our sample) of ever and never smokers to the Florida population at risk.

This initial analysis comparisons between both the population at risk and the nursing home sample indicates that:

- Never and ever smokers tend to enter nursing homes in comparable proportions to their numbers in the population at risk.
- There are more ever smokers in the relatively younger age cohorts and more never smokers in the older cohorts in both the population at risk and the nursing home sample
- There are fewer ever smokers among females than males in our sample and the population at risk
- Higher numbers of females are resident in both the population at risk and in the nursing homes sample in virtually every age group evaluated

Conclusions based on the Nursing Home Sample Exclusively

- On average, ever smokers tend to die at much younger age cohorts then never smokers among those nursing home residents that die.
- Never smokers tend to experience longer lengths of stay among the older age cohorts
- Ever smokers tend to have longer average and total lengths of stay in nursing homes in age cohorts up through age 79.
- Survival analysis testing of differences between the ever and never smokers average length of stay in different cohort is only statistically significant in the 55-59 (ever smokers stayed longer) and 95-105 age cohorts (never smokers stayed longer).

- The average difference in LOS for virtually every other age cohort, (except 85-89) while not statistically significant, is still considerable with ever smokers staying far longer than never smokers.
- When the nursing home survey clients are segmented by age into three and then two large cohorts, the ever smokers clearly have longer lengths of stay then never smokers. However, these differences are not at a statistically significant level.

Section 3 Summary

Section 3 examines both the linear and Cox Regression models developed in this study to evaluate the relationship between ever smoking and lengths of stay and age at discharge. The conclusions of this analysis indicate that the following key independent variables are statistically associated in linear regression models with longer lengths of stay.

Regression Model Conclusions

- Longer are associated with:
  1. The average number of cigarettes smoked per day for both those over and under 80 years of age
  2. The average number of cigarettes smoked and number of years smoking.

- Shorter lengths of stay are associated with:
  1. A higher number of years married for all ages.

- Younger age at discharge (dead or alive) is associated with:
  1. A larger number of years not smoking
  2. A larger number of years married
  3. The average number of cigarettes smoked per day (perhaps most associated with the residents discharge dead)
  4. A larger number of years not smoking.

Cox Model Results

- The hazard of discharge for residents who have never been married is higher than that of residents who have been married residents (i.e. ever married patients tend to have a higher probability of discharged).
- Residents in large facilities experience a higher hazard of discharge relative to residents in smaller facilities.
- The hazard of discharge for male patients is about 48% of the hazard of discharge for female patients (i.e. males are less likely to be discharged relative to females and thus incur longer lengths of stay).
- The hazard of discharge is slightly reduced for residents that have consumed a greater number of cigarettes in their lifetimes.

- The hazard of discharge is slightly increased for patients who have smoked for a longer number of years.

Conclusion of the Cox Analysis

Medical theory would indicate that people who had ever smoked should be more ill on admission to the nursing home, and therefore die sooner after admission than never-smokers. We would expect survival till death to be shorter in ever-smokers than in never-smokers.

- Results from our analysis indicate ever smokers (that died in nursing homes in our sample) had a higher hazard of discharge, or shorter length of stay, than never-smokers. The level of statistical significance is .18 (less than the traditional level of statistical significance of .01, .05 or .1). However the value of the hazard function for ever-smokers is 1.31 times as high as for never-smokers – that is, ever-smokers are more likely to die in nursing homes.

In the second analysis, persons known to have died while in the nursing home are excluded from the analysis. Time is measured till discharge. If the subject is still in the nursing home at the time of the study, his/her observation is considered to be censored. This measures the relationship between smoking status and hazard of discharge, or length of stay, for patients who did not die while in the nursing home.

- The negative sign of the coefficient for the smoking variable indicates that ever-smokers have a reduced hazard of discharge, or longer average length of stay, relative to never-smokers. The level of statistical significance is .26 (less than the traditional level of statistical significance of .01, .05 or .1). The value of the hazard function for ever-smokers is 0.89 times as high as for never-smokers – that is, ever-smokers are *less likely to be discharged, and thus incur longer lengths of stay*.

Cox Regression Analysis Final Conclusion

In conclusion, neither of these analyses indicates a strong or statistically significant effect of smoking. However in both analyses, the directions of the observed effects are those predicted by the known health effects of tobacco use: ***an increased risk of early death after admission to nursing home and a longer length of stay for tobacco-users who survive.*** As these two effects tend to cancel each other out, the relatively small magnitude of the observed effects of smoking on length of stay is not unexpected, when persons discharged dead or alive or still living in the nursing home are considered together.

The Estimated Higher Medicaid Nursing Home Expenditures for Ever Smokers than for Never Smokers and Future Research Needs

This two year $100,000 research project funded by the Robert Wood Johnson Foundation examined tobacco related nursing home costs in Florida with special attention to Medicaid expenditures. The study was jointly conducted by researchers at the Center for Economic Forecasting and Analysis at Florida State University and the Bureau of Epidemiology, Florida Department of Health. While this study is complete, there are a number of critical issues yet left unexamined that still need to be addressed. Some of the noteworthy findings emerging from this analysis are summarized below along with a proposal for future research needs in this area.

Summary of Findings of Higher Tobacco Related Medicaid Nursing Home Expenditures in Florida

Table 4-1 provides a profile of the number of ever and never smokers by age cohort and their respective average length of stay secured from our survey. The average LOS for *ever* smokers is longer in every age category except for the 85 to 89 and the 95+ age group. These are at the age when never smokers tend to out number ever smokers two to one. Ever smokers in NH tend to be younger on admission, experience longer lengths of stays and die at a younger age than never smokers.

Our survey secured meaningful length of stay data from 349 never smokers and 327 ever smokers. This count includes all individuals that were still in the nursing home at the time of the survey. The never smokers reported a total length of stay of 76,079 days while the many fewer ever smokers report 73,552 days. Although there are 6.7% more never smokers than ever smokers, their respective total lengths of stay is virtually identical. Table 4-1 provides an analysis that weights the relative lengths of stay to the actual proportion of persons admitted to nursing homes by age. This analysis demonstrates that the longer LOS for smokers among the younger age groups, who are more numerous, outweighs the longer LOS for the ever smokers in some of the older groups (where ever smokers are less numerous). As Table 4-1 indicates the cumulative LOS for the *numerically fewer* ever smokers is higher across each age cohort until the very final 95+ age group. The final comparison demonstrates that in this oldest age grouping (where never smokers out number ever smokers 2.6 to 1) the larger population of never smokers finally dominates the reported length of stay totals.

The bottom of Table 4-1 presents an estimate of the lower lengths of stay that ever smokers would have reported if they had experienced the average EVER smokers average length of stay. Restated, if ever smokers had stayed in the nursing home, on average, the same number of days as never smokers, the total length of stay for the entire sample population would be 10,781 days fewer in Florida nursing homes. These excess days amount to 7.21% of all the days incurred by both ever-smokers and non-smokers in our population. Our population is highly representative of all Florida nursing home residents

with respect to age and length of stay, so we believe these estimates can be generalized to all nursing home admissions.

In 1997, the Florida Medicaid system expended $1.43 billion for Florida nursing home expenditures. If one assumes that length of stay is a direct measure of expenditure, then ever smokers generate excess expenditures of $103.2 million, or 7.21% of the total 1.4 billion dollars in nursing home expenditures during 1997.

Summary of Proposed Additional Nursing Home Tobacco Related Expenditure Research Needs

This study relies exclusively on nursing home residents' length of stay to measure higher smoking related Medicaid expenditures across the Florida industry. Limited resources, time, confidentiality and gaps in patient and next of kin recollection considerably restrained our survey and nursing home data collected for this analysis.

These and other limits constrain researchers from accessing a number of data categories that could otherwise provide considerably greater insights into the tobacco costs in nursing homes. For example, we had no information on the medical conditions that caused the resident to enter into the nursing home. Were these conditions smoking related or not? We had no information on prior or subsequent conditions and related medical care the resident received over the period of their stay. Were these conditions and or treatments smoking related? Were the levels of care accorded current or ever smokers more or less than never smokers? What about medication levels? Are there any differences between current, ever or never smokers levels of medication? Are there any differences between current, ever or never smokers levels of cognitive and wellness functioning levels and needed levels of support?

All of these and a number of other related issues can and should be more meaningfully evaluated with access to the comprehensive Nursing Home Resident Assessment and Care Screening Full Assessment Form – Minimum Data Set (MDS) and linkage of these data to the AHCA Medicaid client records. The Federal government requires nursing home staff to complete the MDS for each prospective nursing home resident prior to admission. A number of large nursing home chains have maintained these data sets in computerized formats for a number of years while a number of smaller facilities have, until recently retained these data in paper form only. These data are now collected on a systematic basis for *all* nursing home clients entering into nursing homes and will soon be available in a computerized form from the AHCA data center.

Researchers believe that access to these data and linkage of the nursing home client to AHCA Medicaid nursing home could further the insights gained during completion of the research reported on in this study. We believe that these steps will further aid our understanding of the scale and magnitude of the costs of treating tobacco related health care needs among our nations elderly nursing home residents.

Table 4-1

Estimated 1997 Smoking Related Nursing Home Costs to State Medicaid Budget

Never Smokers - Number of Residents and Length of Stay

Age Cohort	Average LOS Never	Number of Never Smokers	Cumulative Number Never	Cumulative Percent Never	Total LOS Never Smokers	Cumulative LOS Never Smokers
Age Cohort: 1-54		1	1	0.3%	-	-
Age Cohort: 55-59	31	4	5	1.4%	124	124
Age Cohort: 60-64	343	6	11	3.1%	2,058	2,182
Age Cohort: 64-69	31	18	29	8.3%	558	2,740
Age Cohort: 70-74	151	36	65	18.6%	5,436	8,176
Age Cohort: 75-79	189	60	125	35.7%	11,340	19,516
Age Cohort: 80-84	162	74	199	56.9%	11,988	31,504
Age Cohort: 85-89	241	70	269	76.9%	16,870	48,374
Age Cohort: 90-94	273	57	326	93.1%	15,561	63,935
Age Cohort: 95+	506	24	350	100.0%	12,144	76,079
		350			**76,079**	

Ever Smokers - Number of Residents and Length of Stay

Age Cohort	Average LOS Ever	Ever	Cumulative Total Ever	Cum Percent	Total Ever LOS by Age Group	Total Cumulative LOS Ever Smokers	Cumulative Difference Never-Ever Smokers
Age Cohort: 1-54	642	4	4	1.2%	2,568	2,568	2,568
Age Cohort: 55-59	493	3	7	2.1%	1,479	4,047	3,923
Age Cohort: 60-64	422	14	21	6.4%	5,908	9,955	7,773
Age Cohort: 64-69	125	22	43	13.1%	2,750	12,705	9,965
Age Cohort: 70-74	239	59	102	31.2%	14,101	26,806	18,630
Age Cohort: 75-79	208	67	169	51.7%	13,936	40,742	21,226
Age Cohort: 80-84	172	71	240	73.4%	12,212	52,954	21,450
Age Cohort: 85-89	224	54	294	89.9%	12,096	65,050	16,676
Age Cohort: 90-94	328	24	318	97.2%	7,872	72,922	8,987
Age Cohort: 95+	70	9	327	100.0%	630	73,552	(2,527)
		327			**73,552**		

Length of Stay of Ever Smokers If they Stayed as Long As Never Smokers

Age Cohort	Ever Smokers w/ Never Smokers LOS Avg.	Ever Smokers With Their LOS Avg.	Difference
Age Cohort: 1-54	-	2,568	2,568
Age Cohort: 55-59	93	1,479	1,386
Age Cohort: 60-64	4,802	5,908	1,106
Age Cohort: 64-69	682	2,750	2,068
Age Cohort: 70-74	8,909	14,101	5,192
Age Cohort: 75-79	12,663	13,936	1,273
Age Cohort: 80-84	11,502	12,212	710
Age Cohort: 85-89	13,014	12,096	(918)
Age Cohort: 90-94	6,552	7,872	1,320
Age Cohort: 95+	4,554	630	(3,924)
TOTALS	62,771	73,552	10,781
			10,781
PERCENT OF TOTAL LOS FOR ALL SURVEYED NH RESIDENTS		**7.21%**	

MEDICAID NURSING HOME SPENDING FOR 1997	$1,432,000,000
ESTIMATED 1997 SMOKING RELATED NURSING HOME COSTS TO STATE MEDICAID BUDGET	**$ 103,176,427**

Bibliography

Facts and Trends: The Nursing Facility Sourcebook, American Health Care Association (AHCA), 1996.

Florida Behavioral Risk Factor Surveillance System, State of Florida Department of Health and Rehabilitative Services. "Health and Lifestyle of Florida Residents," 1995.

Florida Behavioral Risk Factor Surveillance System, State of Florida Department of Health and Rehabilitative Services. "Health and Lifestyle of Florida Residents: State Health office 1986-92 Fact Sheet."

Florida Nursing Home Annual Financial Report, 1997, Florida Agency for Health Care Administration (AHCA)

Florida Nursing Home Reporting System, 1994-1996, Resident Data Report, Agency for Health Care Administration, December, 1997

Hopkins, R., Lynch, T., Final Florida Estimates of Tobacco Related Medicaid Costs 1995-1997, Provided to the Florida First District Court, May, 1997 (Unpublished)

Joyner, Sims, J. Opposition to tobacco use: a nonpartisan public health issue. Journal of the Florida Medical Association. 1996; 83: 79-80

Kleinbaum DG, Kupper LL, Morgenstern H. Epidemiologic Research: Principles and Quantitative Methods. New York; Van Nostrand Reinhold Co., 1982; pp. 350-55.

Kalton G. Introduction to Survey Sampling; Newbury Park:Sage Publications, 1990

Lynch T., Hopkins R. Estimating Tobacco-Related Health Care and Mortality Costs in Florida., J Florida Medical Association.; 1996; 83: 123-132.

Medicaid Related Smoking Costs for the State of Mississippi", Wendy Max, Ph.D., Associate Professor of Medical Economics, University of California, May, 1997 (Unpublished)

Miller, L., Zhang,X.,Rice,D.,Max,W.,Novotny,T: State Estimates of Direct Medical-Care Expenditures Attributable to Cigarette Smoking, University of California-Berkeley School of Public Health, July, 1997

Selvin S. Statistical Analysis of Epidemiologic Data. New York: Oxford University Press, 1991: pp 310-338.

Smoking-Attributable Costs in Florida, 1992, Office of Family Health Services, State Health Office, Department of Health and Rehabilitative Services, 1993

Smoking and Health: National Status Report, 2nd ed., DHHS Publication No. 87-8396 (revised 02/90)

Smoking-Related Deaths and Financial Costs, U.S. Office of Technology, 1993 referenced in "Clearing the Air About Smoking- A Briefing on the Health Problems Associated with the Use of Tobacco", Agency for Health Care Administration, May, 1994

SPSS Advanced Statistics, SPSS® Base 7.5 for Windows®, 1997. *'Chapter 23: Cox Regression Examples'*,

APPENDIX 1

DEVELOPMENT OF THE FLORIDA NURSING HOME SURVEY AND OTHER DATA COLLECTION EFFORTS

I. INTRODUCTION

The goal of the data collection phase of the project is to obtain at a minimum 600 completed surveys from patients admitted to a Florida nursing home sometime between January 1, and June 30, 1995. As will be described below great care was been taken to collect a representative a sample as possible. The survey instrument collects information from either the residents or their guardians or next of kin regarding their history of smoking cigarettes and other pertinent data regarding the nursing home stay. In addition, the resident's social security numbers was obtained from the respondents.

II. INITIAL SURVEY APPROACH

The survey plan called for a stratified random sample of residents who were admitted to a Florida nursing home between January 1, and June 30, 1995, called the admissions time frame. The plan had three steps:

1. select target nursing homes;
2. select patients to interview within the target nursing homes; and,
3. complete the survey instruments.

II.1. Selecting Target Nursing Homes

A list of all nursing homes licensed in Florida as of June of 1993 was the starting point. Six of the state's Department of Human Resources Service (HRS) Districts, one rural and five urban were selected for inclusion in the sample. The collective nursing home demographic characteristics such as number of admissions, type of organization, earnings type, (profit/not for profit), bed size, and ownership status, of the target districts was matched to the state's overall values. If the sample's nursing home demographics did not fall within a ±5% range of the state's values, the districts comprising the sample would be modified until the test was passed. If an overall ±5% was unable to be reached, earnings type, ownership type, and admissions would be the deciding characteristics. The criteria required that the target districts would have experienced at least half of the state's admissions for the target time interval.

After the sample districts were chosen, all nursing homes operating within the sample districts where randomly ordered and the first forty nursing homes would

be drawn from the list. The demographic characteristics of these target homes where compared to the overall state values and, as described above, a ± 5% correspondence where sought. If the correspondence level was not achieved, the nursing homes was returned to the sample, the sample was put in a new random order and a new sample drawn until the desired correspondence was achieved.

The nursing homes were then officially contacted. The survey and its purpose would be described and the homes would be offered a copy of the survey results if they would participate by allowing a researcher to come to their institution and either interview selected residents or collect name, address, and telephone data on selected former residents for further contact. If a nursing home declined, it would be replaced by the next nursing home on the random list that had similar demographic characteristics.

II.2 Selecting Patients to Interview

After 40 nursing homes were selected and had agreed to participate, a field worker visited each home and collected background information on all residents admitted to the nursing home during the admission time frame. An interview quota was set for each home based on the nursing home's share of admissions as a percent of the total number of admissions to all homes in the sample. For example, if a home had 8% of all the admissions experienced by the 40 selected homes in the admissions time frame, then that home was expected to yield 48 completed surveys (8% x 600 total interviews). The resident names from each home was put in random order and a sample twice the size of the required interviews for that home was drawn from the list. That list of names became the survey list for each nursing home.

II.3 Completing the Survey Instruments

Field workers then returned to the nursing homes and attempted to complete the surveys. If the person on the survey list was no longer in the nursing home or could not physically or mentally participate in the survey, the field worker would attempt to contact the person's caretaker or next of kin. If that was unsuccessful, the next person on the home's list replaced that person on the survey list. Interviews of persons who had left the nursing home and interviews with caretakers or next of kin where conducted by telephone if they could not be conducted in the nursing home.

Because the confidentiality concerns for the residents and data collected in this survey was to be augmented by data from the AHCA Nursing Home Patient Data Base, it was decided that a consent form would be required. These consent forms might also allow the survey team, subject to the availability of funds at some juncture in the future, to use the state data on the individuals that participated in the survey, for purposes directly related to this research.

It was decided that the consent form would be collected during the interview if the selected residents or the caretakers and next of kin were surveyed face to face. For surveys conducted by telephone, mailing addresses would be obtained from the nursing home records and a consent form mailed to the potential respondent along with a postage paid return envelope. The subject would be requested to complete the consent form and mail it back to the project team. Once the consent form was received, the interview would proceed.

II.4 Test of the Survey Approach

This approach was tested with a sample of eight nursing homes in Dade County, Florida. It was found to be impractical for a number of reasons:

1. State data could not be obtained with sufficient specificity to identify each nursing home's values for the demographic parameters.
2. The admission time frame initially selected, January 1, to March 31, 1994, proved to be too distant from the actual survey time frame, spring and summer of 1997. Few nursing homes had complete records for the time period and only a tiny fraction of the patients that could be identified were still in the nursing homes and able to participate in the survey.
3. The voluntary participation rate of the nursing homes was very low. Of the eight initially selected only two provided full co-operation and a third promised to participate "soon." "Soon" did not arrive during our six week test period.
4. The consent form process proved unworkable. The address information available from the nursing homes was highly inaccurate and the return rate for the "direct mail" nature for the consent forms was very low. Of the approximately 50 consent forms mailed during the test, 22 were returned undeliverable and only three were returned signed.

III. MODIFIED APPROACH

Based on the lessons learned in the test the survey plan was modified. The basic tenets of the initial plan were maintained but the practical application was modified. Again primary emphasis was placed on obtaining as representative a sample as possible. The modifications to the initial plan were:

1. The admissions time frame was changed from the first quarter of 1994 to the first half, January 1 to June 30, of 1995.
2. The stratified random sample approach was modified from the use of HRS Districts to counties as the geographical basis.
3. The demographic characteristic used to collect nursing homes into subgroups within counties was number of admissions within the admissions time frame.

4. The survey completion method was changed from an emphasis on face-to-face interviews to telephone interviews, although direct interviews were still to be made where possible.
5. The consent forms were not sent out until after a person had become a respondent by answering the questions on the survey instrument.
6. A stronger introduction was necessary to encourage the nursing homes to participate in the survey.

The question of response rate was raised at this point, because of concerns about the possible introduction of a non-respondent bias. It was decided that a response rate of 80% would be sought and that the rate would be determined by dividing the number of people who agreed to participate in the survey once they had been contacted and the survey and its purpose explained by the total number of people who were contacted and the survey and its purpose explained. Said another way, the target was that fewer than 20% of the people actually contacted in person or over the telephone would decline to participate once they understood who was conducting the survey and why it was being conducted. Because of the uncertain nature of the contact data supplied by nursing homes, detailed statistics will be kept regarding the disposition of each subject we attempt to contact. Subjects will be tried a total of three times, with each try coming on a different day than the last. If the subject is not reached by the three tries, we changed to the caretaker or next of kin or the subject was deemed unreachable.

III.1. Selecting Target Nursing Homes - modified approach

State data for the number of admissions by LOS within the admissions time frame for all licensed nursing homes was obtained. Twelve counties were selected to provide a broad range of geographic and demographic characteristics to the stratified sample. There were 71,076 total admissions to 613 nursing homes during the admission time frame. There were 48,497 admissions made to 379 nursing homes (with more than 10 admissions) in the 12 sample counties. The sample and state totals matched almost perfectly in percent of admissions by LOS in the four categories provided, e.g., 0-20 day LOS: total=42%; 12 county sample=43%.

The nursing homes in the 12 county sample were then separated into large and small categories based on the total admissions experienced during the six month admission time frame. Nursing homes with 180 or fewer admissions accounted for almost half (24,362 of 48,497) of the total admissions to the 12 county sample, but they constituted almost 75% of the number of homes (286 of 379). Consequently, the nursing homes in each of the sample counties were divided into two groups, large and small.

Next the number of actual contact telephone numbers necessary from each county was determined, based on the county's share of the overall admissions during the admission time period. After calculating the number of contacts necessary from a county, the number of nursing homes necessary to visit to collect the contacts was determined by dividing the number of total contacts necessary in each county/size category by the average number of admissions in the county/size category. For example, if 200 contacts were needed from small nursing homes in county X and the average size of a small nursing home in county X was 70 admissions, three homes would be visited in county X.

After determine the number of homes necessary in each county/size category, the total number of homes in the category was divided to determine a sample choice factor, the homes in each category were put in random order and every n^{th} (the sampling factor) home was chosen. For example if there were 15 small homes in county X, and the required number to visit was three, the homes were put in random order and every 5th (15/3) home was identified as a target home. This was done for all 24 categories (12 counties, large and small homes) to identify the target homes.

III.2 Selecting Patients to Interview - modified approach

Data collection services were provided in nine counties by four professional market research field services, by two local health departments, and, in one county, by members of the project team. These entities will be referred to, collectively, as field services for the remainder of this memo.

Each field service was given a list of all the nursing homes in their county(ies), the target nursing homes were also identified on the list. They were also given general instructions regarding the visits, data collection forms for the required information, and copies of the survey instrument and consent form in case they should encounter a qualifying resident. They were instructed to administer the survey to any qualifying residents they identified and to reduce the number of telephone contacts by 10 for each interview and consent form they could complete. The field services were instructed to move to the next nursing home on the list they were given, if a target home should refuse to participate.

Based on the quality of the data experienced when visiting the test sites, it was decided that 10 resident contact numbers would be collected for each completed survey desired, i.e., to complete 600 surveys, 6,000 initial contact numbers would be necessary. Because the target nursing homes were selected in such a way as to produce approximately 6,000 contact numbers, the field services were instructed to simply go to the identified nursing homes and collect names and telephone

numbers (or complete interviews) of residents and the caretaker or next of kin for all residents admitted in the admission time frame.

While the lists and instructions were being given to the field services, a letter was sent to all nursing homes in the 12 sample counties. The letter was on Florida Department of Public Health stationary and was signed by the State Epidemiologist and by the Executive Director of the State's association of nursing homes, The Florida Health Care Association. It was a short letter that introduced the survey, explained the purposed and encouraged participation.

III.3 Completing the Survey Instruments - modified approach

The field services began contacting nursing homes on September 15, 1997. Telephone calls to the contact numbers began on December 1, 1997. The quota of completed interviews for each county/size category was the target for telephone contacts. The contact numbers provided by the field services where collected into the categories and put in random order. Telephone survey personal using Computer Aided Telephone Interviewing (CATI) hardware and software started at the beginning of each category and begin calling contact numbers, completing interviews and collecting experience data for calculating the response rate as described above.

APPENDIX 2

FINAL NURSING HOME RESIDENT SURVEY

Opening for patients: Good morning (etc.). My name is _____. I have been asked by the State of Florida Department of Health to talk to people who were admitted to a nursing home in Florida during January, February, or March of 1994. We picked your name at random from the approximately 20,000 people who were admitted to a nursing home during that period. We want to learn more about the people who have occasionally had to spend some time in a nursing home in Florida. The answers to the questions we ask will be kept strictly confidential. We will not contact you again and we won't use your name in any report we write. There are only a few questions, it will not take more than 5 or 6 minutes, so lets begin.

Opening for patients' caretakers/next of kin: Good morning (etc.). My name is _____. I have been asked by the State of Florida Department of Health to talk to people who were admitted to a nursing home in Florida sometime during January, February, or March of 1994. We want to learn more about the backgrounds of people who have occasionally had to spend some time in a nursing home in Florida. (name of patient) was picked at random from the approximately 20,000 people who were in a nursing home at that time. It is our understanding that Mr./Ms. (last name)

has passed away

has reached a point where he/she is no longer able to accurately answer questions about his/her background.

On behalf of the State, let me say that

we are sorry for your personal grief that accompanied his/her passing

we are sorry for the additional burden that this puts on you

but, to make sure that the study is accurate we need to get the information about the people who were selected. We were hoping that you would be able to answer a few questions about Mr./Ms. (last name) 's background. The answers to the questions we ask will be kept strictly confidential. We will not contact you again and we won't use Mr./Ms. (last name) 's name in any report we write. There are only a few questions, it will not take more than 5 or 6 minutes, so lets begin.

Participant # _____

1. Were you
Was Mr./Ms. (last name)
admitted to (name of home) between January and March of 1994?

1. ___ YES	CONTINUE
2. ___ NO	PROBE

Are you sure, according to our records you were (Mr./Ms. (last name) was) in (name of home) from _____ to _____ in 1994. Are we wrong?

YES, YOU ARE WRONG	TERMINATE IF NOT IN 1QT 1994
NO, THAT IS CORRECT	CONTINUE TO #2

2. Also according to our records, your (his/her) Social Security Number is (was)

__ __ __ - __ __ - __ __ __ __

Is that correct?

1. ___ YES
2. ___ NO

could you tell me the correct one? __ __ __ - __ __ - __ __ __ __

(Tim, I am not sure that we need this one, what do you think?)

3. What is the highest grade or year of school you (he/she) completed? (READ ONLY IF NECESSARY)

1. ___ never attended school or attended kindergarten only
2. ___ grades 1 through 8 (elementary)
3. ___ grades 9 through 11 (some high school)
4. ___ grade 12 or GED (High school graduate)
5. ___ College 1 year to 3 years (some college or technical school)
6. ___ College 4 years or more (college graduate)
7. ___ refused

4. Have you smoked (did Mr./Ms. ___(last name)___ smoke) at least 100 cigarettes in your (his/her) entire life? (DO NOT READ RESPONSES)

1. ___ YES GO TO 6.
2. ___ NO GO TO 12.
3. ___ DON'T KNOW/NOT SURE GO TO 12.
4. ___ REFUSED GO TO 12.

5. How old were you (was Mr./Ms. ___(last name)___) when you (he/she) started smoking regularly (every day?)

1. ___ years old
2. ___ DON'T KNOW/NOT SURE
3. ___ REFUSED
4. ___ NEVER SMOKED REGULARLY GO TO 10.

6. For how many years did you (Mr./Ms. __(last name)__) smoke regularly?

1. ___ years
2. ___ DON'T KNOW/NOT SURE
3. ___ REFUSED

7. What is the most number of packs per day you (he/she) ever smoked?

1. ___ packs per day

2. ___ DON'T KNOW/NOT SURE
3. ___ REFUSED

8. On average, how many cigarettes per day did you (he/she) smoke when you were (he/she was) smoking?

1. ___ cigarettes per day (one pack equals 20 cigarettes)
2. ___ DON'T KNOW/NOT SURE
3. ___ REFUSED

9. Did you (he/she) smoke cigarettes during the month before you were (he/she was) admitted to the nursing home in early 1994?

1. ___ YES	
2. ___ NO	GO TO 13.
3. ___ DON'T KNOW/NOT SURE	GO TO 13.
4. ___ REFUSED	GO TO 14.

10. On average how many cigarettes a day were you (was he/she) smoking during the month before you (he/she) was admitted to the nursing home in early 1994?

1. ___ number of cigarettes (one pack equals 20 cigarettes)
2. ___ DON'T KNOW/NOT SURE
3. ___ REFUSED

11. During the year before you were admitted to the nursing home, did you quit smoking for 1 day or longer?

1. ___ YES	GO TO 14
2. ___ NO	GO TO 14.
3. ___ DON'T KNOW/NOT SURE	GO TO 14.
4. ___ REFUSED	GO TO 14.

12. About how long before you (he/she) entered the nursing home in early 1994 did you (he/she) stop smoking cigarettes regularly (that is, daily)?

1. ___ 0 to 1 MONTH before entering the nursing home
2. ___ 1 to 3 MONTHS before entering the nursing home
3. ___ 3 to 6 MONTHS before entering the nursing home
4. ___ 6 to 12 MONTHS before entering the nursing home
5. ___ 1 to 5 YEARS before entering the nursing home
6. ___ 5 to 15 YEARS before entering the nursing home
7. ___ 15 or more YEARS before entering the nursing home
8. ___ Don't know/Not sure

9. ___ Never smoked regularly
10.___ Refused

13. Have you ever been married?
Was Mr./MS. __(last name)__ ever married?

1. ___ YES GO TO 13a
2. ___ NO GO TO 14

13a. How many years, all together, have you (was he/she) married?
Number of years married, total __ __?

14. During the last year you (he/she) worked full time, was your (his/her) annual household income from all sources:

1. ___ less than $10,000
2. ___ more than $10,000 but less than $15,000
3. ___ more than $15,000 but less than $20,000
4. ___ more than $20,000 but less than $25,000
5. ___ more than $25,000 but less than $35,000
6. ___ more than $35,000 but less than $50,000
7. ___ more than $50,000 but less than $75,000
8. ___ over than $75,000
9. ___ Don't know/Not sure
10.___ refused

15. Who paid the nursing home's bills for the stay that began in January, February, or March of 1994?
(DO NOT READ THE LIST, PROMPT TWICE: "WERE THERE ANY OTHERS?)

1. ___ the patient
2. ___ personal insurance
3. ___ company provided insurance coverage
(for people employed when they went into the nursing home)
4. ___ company retirement coverage
5. ___ Medicare
6. ___ Medicaid
7. ___ Other _____
(if they choose 5 or 6, repeat the program name, e.g., "Was that MediCARE or MediCAID?"

Thank you for answering our questions. We would also like your permission to review some of your records to get a limited amount of supporting information for

this study. Please read this form and, if it is O.K. with you, sign it at the bottom. I can answer a few questions you might have, and I will give you a paper with the name and address of the person in charge of the study in case you have further questions.

(GO TO THE CONSENT FORM)

Consent Form

Participant #

_____

The Florida Department of Health is performing a study to learn more about why people are admitted to nursing homes. We are particularly interested in learning about any factors in a person's background that may influence whether or not that person is likely to need nursing home care and how those factors affect nursing home expenses.

In order to help this study, we would like to review State records about your stay in the nursing home to identify what conditions you had when you were admitted. We would also like to review the State's records to learn about how much the nursing home charged for your stay and how the bills were paid.

All of the information collected, including the answers you just gave us to the question we asked will be confidential. The information will be entered into a computer without your name. Your name will not appear on any reports. We will not report any information about any individuals. Our interest is only in group results.

This study is directed by Dr. Tim Lynch, Florida State University, Center for Economic Forecasting, 904-644-7357. He will answer any further questions you may have about the study.

You may withdraw your consent at any time for any reason. Any questions you ask, except for questions about other individuals in this study, will be answered.

I have read and understand this information. I consent to allow researchers to access records kept by the State regarding my nursing home stay discussed during this interview.

Signed _____

_____

(signature) (printed)

Witness _____

Date _____

❒ I would like to receive a copy of the final report from this study.

Address: _____City

_____

State: _____ Zip

_____

Consent Form - Next of Kin or Caretaker

Participant # _____

The Florida Department of Health is performing a study to learn more about why people are admitted to nursing homes. We are particularly interested in learning about any factors in a person's background that may influence whether or not that person is likely to need nursing home care and how those factors affect nursing home expenses.

In order to help this study, we would like to review State records about Mr./Ms. (last name) 's stay in the nursing home to identify what conditions he/she had when he/she was admitted. We would also like to review the State's records to learn about how much the nursing home charged for his/her stay and how the bills were paid.

All of the information collected, including the answers you just gave us to the question we asked will be confidential. The information will be entered into a computer without Mr./Ms. (last name) 's name. His/her name will not appear on any reports. We will not report any information about any individuals. Our interest is only in group results.

This study is directed by Dr. Tim Lynch, Florida State University, Center for Economic Forecasting, 904-644-7357. He will answer any further questions you may have about the study.

You may withdrawn your consent at any time for any reason. Any questions you ask, except for questions about other individuals in this study, will be answered.

I have read and understand this information. I consent to allow researchers to access records kept by the State regarding Mr./Ms. (last name) 's nursing home stay discussed during this interview.

Patient name: _____
(printed)

Signed _____
_____
(signature) (printed)

Witness _____

Date _____

❒ I would like to receive a copy of the final report from this study.

Address: _____City
_____

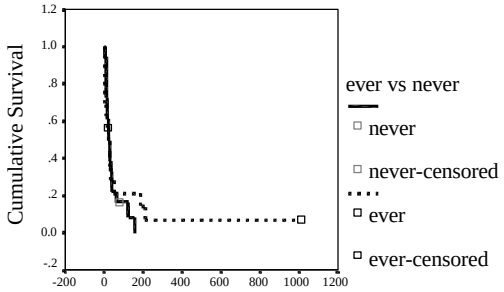
State: _____ Zip
_____

APPENDIX 3

Survival Analysis

Females

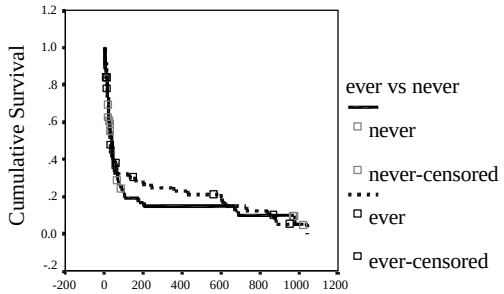
Age Cohort: 1-69



Length of Stay

Females

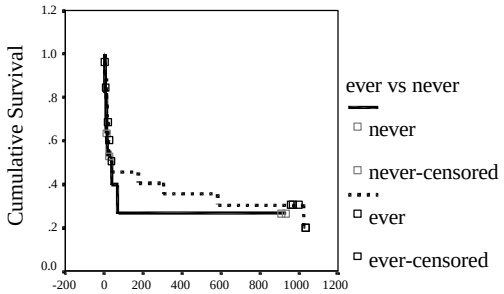
Age Cohort: 70-79



Length of Stay

Males

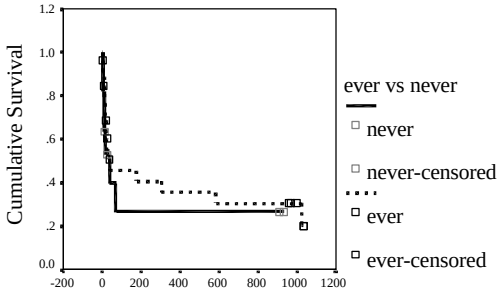
Age Cohort: 80+



Length of Stay

Males

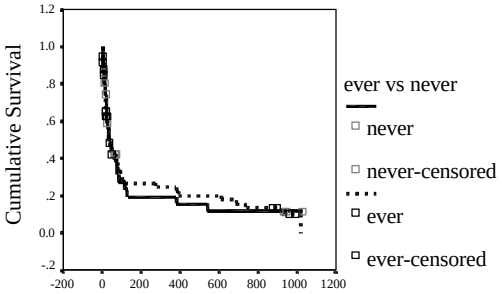
Age Cohort: 1-69



Length of Stay

Males

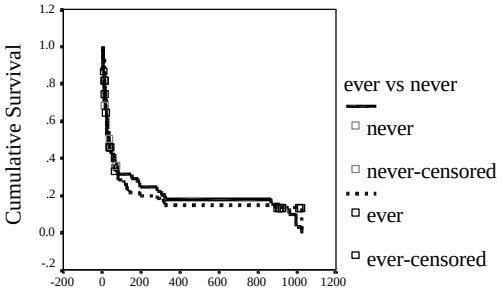
Age Cohort: 70-79



Length of Stay

Males

Age Cohort: 80+



Length of Stay