RATES OF REPORTED PULMONARY DISORDER AMONG THE ADULT POPULATIONS OF SOUTH BOSTON and THE BOSTON METROPOLITAN AREA:*

Boston Area Pulmonary Health Survey, 1983

Charles F. Turner   Mary Ellen Colten
and
Peggy Collins

CENTER FOR SURVEY RESEARCH

a facility of the
UNIVERSITY OF MASSACHUSETTS
and the
JOINT CENTER FOR URBAN STUDIES
MIT and HARVARD UNIVERSITY

* Throughout this report, "Boston metropolitan area" and Boston SMSA denote the federal government's Standard Metropolitan Statistical Area for Boston; see Section I for precise definitions. Where comparisons are reported between South Boston and the Boston SMSA, tabulations contrast the SMSA excluding South Boston to South Boston.

The term pulmonary disorder denotes a class of conditions that includes both pulmonary diseases and pulmonary symptoms.

CONTENTS

*

PREFACE

The following pages report our analysis of the results of the Boston Area Pulmonary Health Survey. This analysis and the survey it reports were carried out under unusually tight time constraints.[1] This tight schedule implied that not every analysis that might have been of potential value could be performed. Thus readers may find that we do not always present every analysis they might have desired and that some results appear to be "loose ends." While we cannot remedy these problems in the present report, readers who wish to conduct further explorations of these data can obtain copies of the data tape and related materials from the Center for Survey Research (CSR).

Nomenclature. Throughout this report we make continual use of one phrase that deserves explanation at the outset. We have chosen to refer to the major dependent variables in our analysis as "pulmonary disorders." We have chosen this somewhat vague term because some of our dependent variables are diseases (e.g., asthma, tuberculosis), and some are symptoms (e.g., Chronic Cough, Phlegm). In this report we use the phrase "pulmonary disorder" to denote both types of observation.

Sponsorship. Funds for this survey were provided by Stone & Webster Engineering Corporation acting on behalf of the Boston Edison Company.

Acknowledgements. Many people made important contributions to this work. Most importantly, this research was made possible by the generous contributions of our respondents who volunteered their time to take part in the survey. The operation of the survey also required the diligent work of fifty interviewers who worked through one of the hottest summers in recent memory. Throughout the survey Margaret Raisty, Dorothy Cerankowski, and their staff cheerfully and competently solved the myriad of problems that inevitably arise in a complex survey operation.

Our other colleagues at CSR provided all sorts of professional advice when we needed it. Steven Dubnoff and Jack Fowler provided particularly valuable help as consultants on sample design and as reviewers of our draft report.

----

[1] 2,991 survey interviews were conducted in a six-week period ending August 28th, and the first (unweighted) computer tabulations of the data were made on September 12, 1983. On October 18th, a draft of the summary of our findings (Section IX) together with relevant tables was presented to the survey sponsors and a complete draft report (minus the abstract) was submitted on October 31st.

The expert advice of Douglas Dockery, Benjamin Ferris, and Frank Speizer of the Harvard School of Public Health and of our statistical consultant, Robert Fay, was invaluable. We learned so much from them that we feel frankly embarrassed by the little we could offer in return.

Needless to say, none of those who helped, critiqued, sponsored, or advised us, is in any way responsible for whatever errors or infelicities may appear in the following pages.

Boston, Massachusetts
29 November 1983

C.F.T., M.E.C., and P.C.

I. SUMMARY

SUMMARY

Survey

During July and August of 1983, the Center for Survey Research conducted a telephone survey of the pulmonary health of adult residents of South Boston and the Boston metropolitan area (i.e. the Boston SMSA)[1]. 2,991 telephone interviews were conducted with a probability sample of English-speaking adults; approximately equal numbers of respondents were drawn from South Boston and from the rest of the Boston metropolitan area.[2] Survey respondents were asked about their pulmonary health and their exposure to various environmental factors that affect pulmonary health such as cigarette smoking. (The survey questionnaire used in this study was a modified version of the American Thoracic Society Questionnaire: ATS-DLD-78A.)

Analyses

The survey was designed to test the claim that there is an excess of pulmonary disorders[3] in South Boston. Seven key

----

1. Throughout this report, we use "Boston metropolitan area" and Boston SMSA to denote the federal government's Standard Metropolitan Statistical Area for Boston; see Section I for precise definitions. Where comparisons are reported between South Boston and the Boston SMSA, our tabulations contrast the SMSA excluding South Boston to South Boston.

2. See Section 2 and Appendix A for details of sampling procedures.

3. The term pulmonary disorder is used to denote a class of conditions that includes both pulmonary diseases and pulmonary symptoms.

measurements of pulmonary health were initially identified for intensive analysis: Chronic Cough, Chronic Phlegm, Chronic Wheeze, Shortness of Breath, Asthma, Chest Colds, and Severe Chest Colds. In addition, because a preliminary tabulation of the data indicated that South Boston residents reported Pulmonary Tuberculosis at 2.7 times its rate in the Boston SMSA, this disorder was also included in our analyses.[4]

The basic strategy of our analyses involved testing for differences between the prevalence of reported pulmonary disorders in South Boston and the Boston SMSA after adjustments had been made for differences in the age, smoking habits, and sex composition of the two samples. This strategy is appropriate since the relevant question is not merely whether the population of South Boston reports more pulmonary disorders, but whether these reports are "excessive" given differences in the smoking history, age, and sex composition of the two populations. (Later more refined analyses tested for differences after adjustments for survey design effects, educational level, work exposure to

----

4. Ratios for other disorders not chosen for intensive analysis were Emphysema (1.65), Chronic Bronchitis (1.32), Pneumonia (1.05), Attacks of Bronchitis (0.74), Hay Fever (0.72), and Sinus Trouble (0.74). Ratios of 1.0 or higher indicate a relative excess of reports in South Boston; ratios less than 1.0 indicate a relative shortage of reports in South Boston. The ratios are defined as:

$$\frac{\text{(Odds of South Boston Respondents Reporting Disorder)}}{\text{(Odds of Boston SMSA Respondents Reporting Disorder)}}$$

dust, type of home heating, and use of stoves for home heating.)

We briefly describe below the results of each stage of our analysis. In addition, Figure A-1 provides an illustration of the outcome of each stage of the analysis. Figure A-1 shows the disorders which were included in the analysis, and the estimated ratios of reporting in South Boston vs. the Boston SMSA after various statistical controls were introduced. When the ratios shown in Figure A-1 exceed 1.0, they indicate that a disorder was more likely to be reported in South Boston than in the Boston SMSA. For example, a ratio of 2 would indicate that adult residents of South Boston were estimated to be twice as likely to report the disorder as adult residents of the Boston SMSA. Conversely, when the ratios shown in Figure A-1 are less than 1.0, they indicate that a disorder was less likely to be reported in South Boston than the Boston SMSA. For example, a ratio of 0.50 would indicate that adult residents of South Boston were only one-half as likely to report the disorder as residents of the Boston SMSA. The notation "n.s." is shown in Figure A-1 when the ratio is not different from 1.0 by a statistically significant margin ($p<.05$), which is to say that statistically one could not discern a difference between the levels of disorder reported by residents of South Boston and residents of the remainder of the Boston metropolitan area.

Stage 1. Our initial analysis (controlling for age, sex, and smoking) indicated statistically significant differences in the reported prevalence of four pulmonary disorders: Chronic Wheeze (adjusted ratio: 2.10), Chronic Phlegm (1.57), Shortness

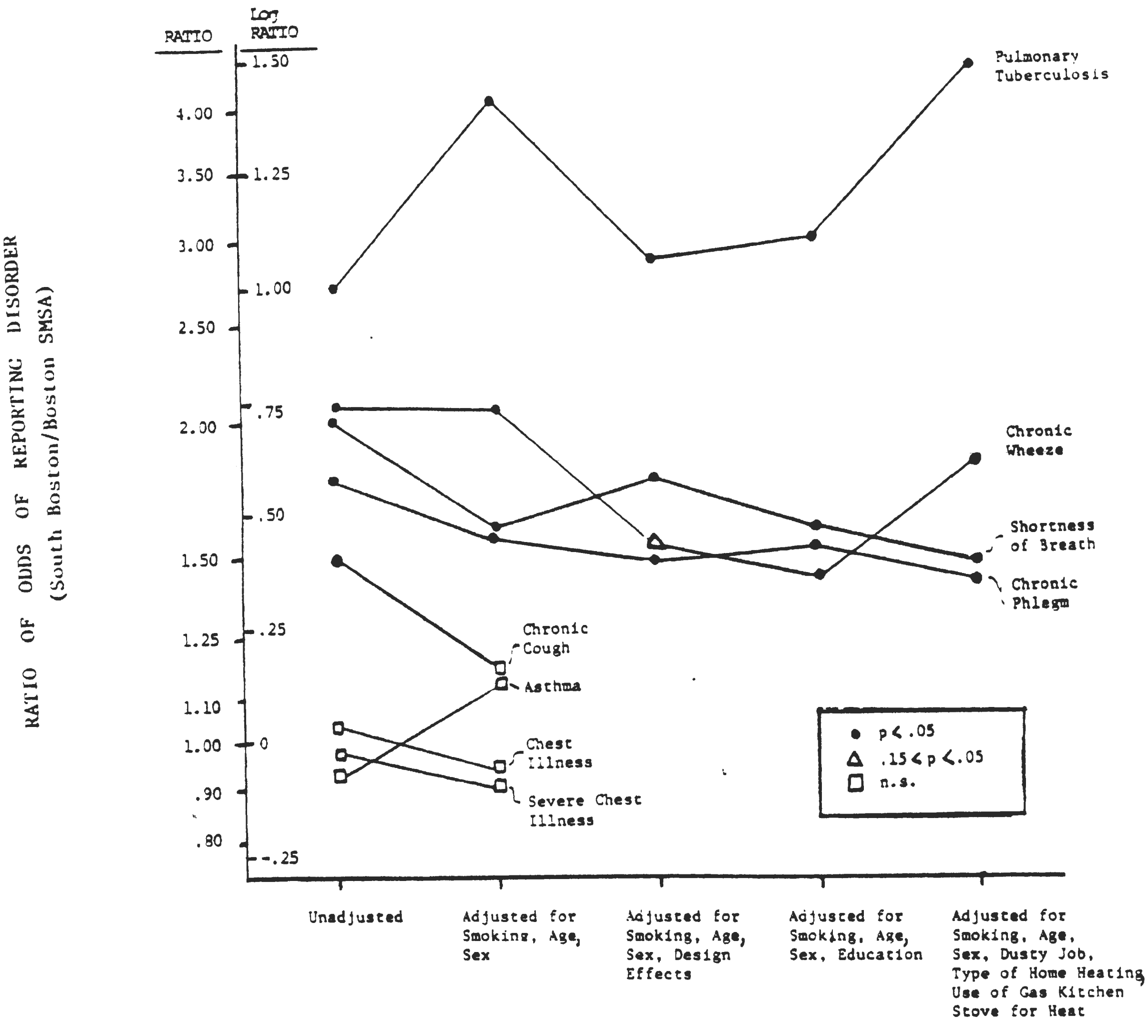


FIGURE A.1 Ratio of Expected Odds (on logarithmic scale) of reporting of Eight Pulmonary Disorders in South Boston versus the Rest of the Boston SMSA, after various Analytic Adjustments. Note that a ratio of 1.0 indicates that the odds of reporting are equal in the two areas; odds of 2.0 indicate that the disorder was twice as likely to reported in South Boston; odds of 0.50 indicate that the disorder was only one-half as likely to be reported in South Boston as in the rest of the Boston SMSA.

Notes. Refer to Tables 1, 9, 12, 13, 22b, 24, and 28 for exact values. Points marked n.s. (i.e. square symbols) indicate that observed ratios were not significantly different from 1.0 (i.e., no significant difference between reporting in South Boston and the Boston SMSA). Ratios shown for Shortness of Breath adjusted for Age, Sex, and Smoking include additional adjustment for heart trouble (see Table 24b).

of Breath (2.26)[5] and Pulmonary Tuberculosis (4.12). For three other pulmonary disorders (Asthma, Chest Colds, Severe Chest Colds, adjusted ratios: 0.96 to 1.08), no statistically significant difference was detected between the reported prevalence rates in South Boston and the Boston metropolitan area.[6] (For the eighth variable, Chronic Cough, a puzzling 5-way interaction was found, indicating that differences in reporting between South Boston and the Boston SMSA varied by sex, age, and smoking behavior.)

Stage 2. A series of methodological analyses were subsequently conducted to assess the potential impact of survey design factors upon the foregoing findings. When these factors were taken into account, the foregoing findings that higher levels of prevalence were reported in South Boston held, except in the case of Chronic Wheeze. For Chronic Wheeze, the excess of symptom reporting in South Boston (adjusted ratio = 1.51) declines below commonly-accepted levels of statistical significance ($.15 < p < .05$).

Stage 3. Two further analyses were conducted introducing additional control variables into the analysis [1: Educational Level; 2: Work in a Dusty Job, Type of Home Heating, and Use of

----

5. Ratio of odds for Level 2 Shortness of Breath (report that respondent had to walk more slowly on level ground than others the same age because of Shortness of Breath) vs. no Shortness of Breath.

6. The results for Chronic Cough showed a puzzling set of inter-area differences that varied by age, sex and smoking in a manner that we found difficult to interpret.

Gas Stove for Heating). Even when these variables were introduced as control factors, the prevalence rates for reported Shortness of Breath, Chronic Phlegm, and Pulmonary Tuberculosis were found to be higher in South Boston than in the rest of the Boston metropolitan area ($p < .05$).

Conclusions. There are limits to the conclusions one can reach based solely on telephone surveys. The results strictly apply only to that fraction of the population that has telephones and is willing to be interviewed. Moreover, the answers that people give in surveys provide imperfect (though quite suggestive) measures of pulmonary problems. We thus cannot rule out the possibility that South Boston residents may overstate their pulmonary symptoms because of a heightened awareness of lung problems, just as we cannot rule out the possibility that South Boston residents may understate their pulmonary symptoms because, on average, they have less education and lower incomes and thus might have less access to medical care[7] or be less aware of medical conditions.

Nonetheless, to the extent that survey reports of pulmonary health conditions can provide useful indicators of actual pulmonary disorders, and in the absence of conflicting evidence, the data from this study support the conclusion that some pulmonary disorders are more prevalent in the South Boston population than in the rest of the Boston metropolitan area.

----

7. See, for example, C. K. Riessman, Improving the use of health care services by the poor. In P. Conrad and R. Kearn (eds.). Sociology of Health and Illness, New York: St. Martin's Press.

II. INTRODUCTION

II. INTRODUCTION

This report presents the major findings from the Boston Area Pulmonary Health Survey. The goal of the survey was to collect data to test the claim that rates of pulmonary disorder[1] in South Boston are abnormally high.

Control Group. Work on this study began with the selection of an appropriate control group against which the South Boston prevalence rates might be meaningfully compared. Although it might seem appropriate to choose another community which is "just like South Boston," this strategy was rejected because it is vulnerable to several severe problems. Most importantly,

1. We do not have sufficient theoretical knowledge of the causes of pulmonary disorders to identify all of the factors which need to be equivalent for two populations to be judged "similar;"
2. Even with adequate theoretical knowledge of all potential risk factors, community-specific population data are not available for even the most obvious ones. For example, we do not know the rates of cigarette smoking in Charlestown, or Revere, etc. Indeed, prior to this study we did not know of any firm estimate of the prevalence of cigarette smoking in South Boston. Without reliable data on the distribution of the various risk factors across different communities, it would

----

1. Throughout this report we use the term pulmonary disorder to denote a class of conditions that includes both pulmonary symptoms and pulmonary diseases.

> be impossible to make an a priori judgement of "similarity," even if we had adequate theoretical knowledge of all the risk factors that induce pulmonary disorders.

Because of these considerations, an attempt to compare South Boston with any other community would be vulnerable to an infinite series of critiques of the sort: "(Unmeasured) factor X does (or may) vary between the two communities. This factor does (or is likely) to cause variations in symptom rates and may account for the differences found in the survey."

To provide a more satisfactory test, we selected the entire Boston metropolitan area for the control area and have posed the question in the following terms:

> Do the residents of South Boston have a different rate of pulmonary disorder than residents of the Boston metropolitan area?

To define the Boston metropolitan area we used the federal government's definition of the Standard Metropolitan Statistical Area (SMSA) for Boston. The federal government's definition of the Boston SMSA includes all of Suffolk, and parts of Norfolk, Middlesex, Plymouth and Essex counties. The map on the following page marks this area with a thick black border.[2]

Survey. From July 16 to August 28, 1983, the Center for Survey Research conducted 2,991 telephone interviews as part of a pulmonary health study of the Boston area. The purpose of the study was to compare the self-reported pulmonary health of residents in South Boston with that of residents of the remainder of the Boston Standard Metropolitan Statistical Area (SMSA). The survey instrument (see Appendix E) is primarily comprised of a pulmonary health questionnaire (ATS-DLD-78A) that was developed by the American Thoracic Society[3] and adapted in the present study for telephone interviewing.

The probability samples for South Boston and the SMSA were drawn separately. A simple random sample design was used in South Boston and a Waksberg two-stage probability sample[4] was drawn in the SMSA. Both methods produce samples with a known

----

2. "The general concept of a metropolitan area is one of a large population nucleus together with adjacent communities which have a high degree of economic and social integration with that nucleus. The standard metropolitan statistical area (SMSA) classification is a statistical standard, developed for use by Federal agencies in the production, analysis and publication of data on metropolitan areas. The SMSA's are designated and defined by the Office of Management and Budget following a set of official published standards developed by the interagency Federal Committee on Standard Metropolitan Statistical Areas." (U.S. Bureau of the Census, 1980 Census of Population: Characteristics of the Population. Standard Appendix A: Area Classifications.)

3. See Ferris, B.G. et al. Epidemiology Standardization Project II: recommended respiratory disease questionnaires for use with adults and children in epidemiological research. American Review of Respiratory Diseases. 1978: 118 (Supplement, 7-53).

4. See, J. Waksberg, Sampling methods for random digit dialing. Journal of the American Statistical Association, 1978, 73:40-46.

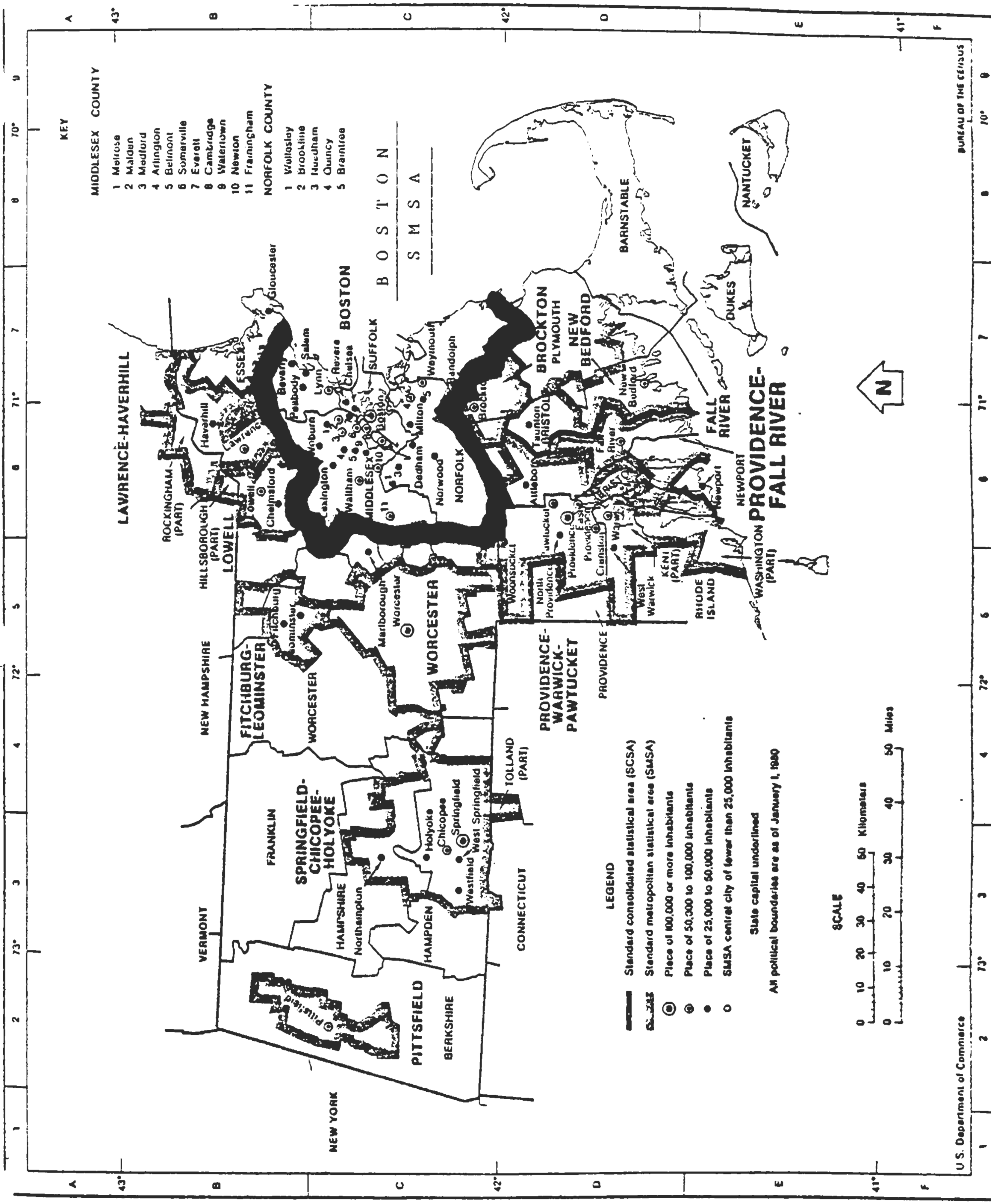


FIGURE 1: Map of Boston Standard Metropolitan Statistical Area (SMSA). [SMSA is marked with wide black border.]

probability of inclusion of each household in the sampled universe. With the exclusion of sampling method, all procedures for South Boston and the SMSA were identical.

The overall response rate for the survey was 70.1 percent: 70.8 percent in the SMSA and 69.3 percent in South Boston. 1,444 interviews were conducted in South Boston and 1,547 in the SMSA.

Qualifications. Because we conducted a survey, certain qualifications must be made to the question we will be answering. Most importantly, we can only speak about reported pulmonary disorders, and thus all of our inferences from the survey data have to do with symptom and disease reports (i.e., what the residents of each area told us) and not clinical diagnoses (although some of our measurements asked respondents to report diagnoses doctors had made). Since we have used a telephone survey, all our inferences must be restricted to those persons living in households with a working residential telephone (95 percent of the households in the Boston SMSA and 86 percent of the households in South Boston have telephones). This restriction in our sample does result in some noteworthy limitations to our inferences; most importantly, we will be missing some of the very poorest households in each area.[5] In addition to the loss of this segment of the population, an additional

---

5. Nationally, in-person surveys conducted between 1972-82 found approximately seven percent of surveyed population did not have a telephone; lack of a telephone was more common among those of lower socio economic status. For example, fewer than 2% of college graduates are without a phone, while 13% of those with a grade-school education lack a phone.

   Among households reporting incomes of less than $10,000, 11.3% lacked phones while the corresponding figure for $30,000+ households was 5.4%

30 percent of the persons selected for interview could not be interviewed because they refused, were unable to be interviewed for reasons of health or language, or were not reachable.[6] The resulting qualifications to our inferences may be succinctly stated as follows:

> All of our inferences are limited to <u>reports</u> of pulmonary disorders given by persons residing in households with telephones. Moreover, rigorous generalizations can only be made to that 70 percent of the population which is willing to be interviewed in such telephone surveys.

While we will investigate in a subsequent section (VI) the impact of this and other survey design issues on generalizations to the whole population, readers should keep these qualifications in mind when generalizing from our results.

<u>Organization of Report</u>. This report discusses a series of statistical analyses which seek to answer our basic question with growing levels of statistical rigor and increasingly long lists of "control" variables which are thought to have a potential influence on pulmonary health. We have attempted in sections III and IV to present our findings in a manner that will be accessible to all readers. In these sections we avoid (for the most part) formal hypothesis testing, but statistically trained readers will find more detailed information in our tables and subsequent sections. Sections V through VIII present more rigorous analyses. We have restricted the prose in these sections to the minimum necessary to inform readers of our methods. For the most part, the tables in this section will be

----

6. The rate of refusal and the rate of non-completion due to ill health were virtually identical in the South Boston and the Boston metropolitan samples; see Table 2 of Appendix A.

of the most interest; they present the results of formal tests of our hypotheses.[7]

In section IX we summarize the findings of our statistical analyses in prose that (we hope) will be accessible to all readers.

---

7. Appendices A to H present further details of our analyses as well as documentation of our fieldwork and coding procedures and a copy of the survey questionnaire.

III. UNADJUSTED ESTIMATES OF PREVALENCE OF PULMONARY DISORDERS

III. UNADJUSTED ESTIMATES OF PREVALENCE OF PULMONARY DISORDERS

Before we delve into formal statistical tests, it may be useful for some readers if we consider what a reasonable approach might be (without using formal statistics) to the question of whether or not the population of South Boston report pulmonary disorders at a different rate than other residents of the Boston metropolitan area (i.e. the Boston SMSA).[1] One reasonable approach might consist of several steps, each providing somewhat more detailed information than the preceding step. We sketch this approach out in sections III and IV before we undertake our formal statistical analyses (sections V to VII).

A reasonable first step is to look at differences in the unadjusted rates of symptom reporting by residents of South Boston and residents of the SMSA. Table 1 presents estimates of the percentage of the population in each of the two areas who

----

1. Throughout this report, we use "Boston metropolitan area" and Boston SMSA to denote the federal government's Standard Metropolitan Statistical Area for Boston; see Section II for precise definitions. Where comparisons are reported <u>between</u> South Boston and the Boston SMSA, our tabulations contrast the SMSA <u>excluding South Boston</u> to South Boston.

   The term pulmonary disorder is used throughout this report to denote a class of conditions that includes both pulmonary diseases and pulmonary symptoms.

report each of the seven primary pulmonary disorders.[2] (These disorders were chosen for intensive analysis prior to the availability of the survey data.)

Table 1 shows substantial differences between South Boston and the SMSA in the rates of reporting for four symptoms: Chronic Cough, Chronic Phlegm, Chronic Wheeze, and Shortness of Breath. The South Boston population is 1.49 times more likely than the SMSA population to report Chronic Cough, 1.78 times more likely to report Chronic Phlegm, and more than twice as likely to report Chronic Wheeze.[3] All of these differences are large enough that chance variation due to sampling is an unlikely explanation. We note from Table 1 that there is also a large difference between the two areas in rates of reporting for shortness of breath. Twenty-eight percent of the population of South Boston reports some symptoms of Shortness of Breath, compared with 18% in the SMSA. The more severe symptoms of Shortness of Breath (e.g., walking more slowly on level ground than others the same age because of Shortness of Breath; having to stop for breath when walking on level ground, etc.) are, when

----

2. These percentages were obtained by weighting the sample data by the reciprocal of the sampling probability; this weighting provides a correction for people who have more than one phone line in their house and who thus have an inflated chance of being interviewed and for people residing in households with several "eligible adults (each of whom will have a reduced chance of being sampled).

3. Readers should note that in our discussion we continue to refer to ratios of odds ratios; thus the statement that a disorder is twice as likely to be reported in South Boston means that the odds ratio (not the rate) for the reporting of this disorder is twice as high in South Boston as in the SMSA.

taken together, reported by the South Boston population at a rate about twice that of the SMSA population.

For the remaining three major symptoms, there is no discernible difference between the reporting in the two areas (ratios=0.93 to 1.04). Respondents from South Boston and from the SMSA report with essentially equal frequency symptoms of asthma, recent chest illness and recent severe chest illness.[4]

In the aggregate, then, significantly more respondents in South Boston than in the SMSA report symptoms of Chronic Cough, Chronic Phlegm, Wheeze, and Shortness of Breath. These four disorders will be a major focus for the remainder of our analyses, although results will be presented for all seven disorders in many instances.

Inter-Area Differences in Demographic Characteristics, Smoking, and Other Factors

A reasonable second step is to consider whether there are ways in which the residents of South Boston as a group may differ from the residents of the remainder of the SMSA that might account for the observed differences in the prevalence rates of reported pulmonary disorders. Table 2 compares South Boston and SMSA residents on a variety of demographic characteristics and

----

4. (1) Chest illness or cold in the past three years that kept respondent at home for less than one week;

   (2) Chest illness or cold with phlegm in the past three years that kept respondent at home for a week or more.

risk factors such as cigarette smoking and exposure to dust or fumes at work. The first three variables in Table 2, cigarette smoking, age, and sex are important "control" variables because of their association with pulmonary health and thus their potential to account for observed differences in the pulmonary health of two populations. Table 2 shows that South Boston sample differs from the SMSA sample in cigarette smoking and age, and that there is a borderline difference between the sex ratios in the two areas.[5] (Table 2 also shows that the South Boston population reports significantly less income and education and different exposure to environmental risk factors than the population of the Boston SMSA -- a consideration to which we will return in Section VIII.)

Smoking differences between the two areas are substantial. More residents of the SMSA have never smoked (38.0 vs. 31.8 percent); furthermore the South Boston population falls into the two heaviest smoking categories (16-30 and 31+ cigarettes daily) at more than one and half times the rate of the SMSA population. The age structure of the sampled populations also differs somewhat. The South Boston sample includes more younger and older persons and has proportionally fewer individuals in the 35-49 year old group than the SMSA sample.

A third major control variable in our analyses is sex. Sex has been used as a control in previous studies of pulmonary

----

5. The observed difference in the sex composition of the two areas is confounded with the age distribution.

health for several reasons. Women have lower mortality but higher morbidity, which means that for any age grouping it would be expected that women would outlive men or that they have outlived their counterparts, but that they might still be reporting more sickness and symptoms for a variety of social and cultural, as well as physiological reasons. So, sex cannot be ruled out as a factor to which some of the observed differences might be attributed.

Relationship of Prevalence of Reported Pulmonary Disorders to Cigarette Smoking, Age, and Sex

Table 3 presents the relationship between each disorder and smoking. This relationship is notably strong for each of the four disorders with differing prevalences in South Boston and the SMSA, and the direction of the relationship is clear. As smoking increases from those currently not smoking to the presently heavy smokers, symptoms increase monotically and dramatically. These striking relationships between smoking and reported pulmonary disorders together with the difference in the rate of smoking in the two areas make evident the necessity of controlling for smoking history when comparing pulmonary disorders in South Boston and the SMSA.

As can be seen in Table 4, age is also related to reporting of some symptoms although not so clearly nor in such an orderly fashion as cigarette smoking. Shortness of Breath increases montonically with age, while chest colds and severe chest colds decline with age. Chronic Cough, Wheeze, Phlegm, and Asthma, however, show no simple trend with age.

Table 5 shows the relationship between sex and the seven disorders. Women report more Chest Colds, Severe Chest Colds, and Shortness of Breath, while men report more Phlegm. There is no sex difference for Chronic Cough, Wheeze, or Asthma.

Inter-Area Differences in Prevalence of Reported Symptoms Controlling Separately for Smoking, Sex, and Age.

The next step is to compare the prevalence of symptoms in South Boston and the SMSA within levels of smoking, of age, and of sex. If, for example, there is no discernible difference in symptom rates between residents of South Boston and the SMSA when differences in smoking beween the two areas are taken into account, then we might say that the difference in the prevalence of pulmonary disorders between South Boston and the SMSA could be "explained" by the differences in the smoking behavior of the two populations. If, however, differences between areas persist within categories of smoking, then the difference in the prevalence rates is not entirely attributable to smoking.

Table 6 compares the reported prevalence of pulmonary disorders in South Boston and the SMSA samples within levels of cigarette smoking. This table suggests that the differences in prevalence rates between South Boston and the SMSA cannot be accounted for solely by the higher smoking rate in South

Boston. For example, Chronic Phlegm is reported more frequently in South Boston by non-smokers, former smokers, and by each category of present smokers.[6] Differences in Wheeze are not apparent for the never and former smokers, but persist in the smoking groups. Chronic Cough is also reported with less frequency in the SMSA by light and moderate smokers than in South Boston. Similarly, for each category of smoking, we find more SMSA respondents reporting no symptoms of Shortness of Breath.

For age (Table 7), some of the differences in prevalence rates between South Boston and the SMSA are seen within age groups, and some diminish or disappear. The two middle-aged groups exhibit differences in Chronic Cough, Chronic Phlegm, and Wheeze, while the youngest and oldest age groups in South Boston and the SMSA look more alike. So, as with smoking, the introduction of age as a control variable suggests that some of the observed difference in reported symptom prevalence in South Boston and the SMSA may be accounted for by the different age compositions of the two populations.

Table 8 shows the symptom frequencies for males and females. There is little hint from this table that differences in the sex distribution of the two samples accounts for the differences observed between South Boston and the SMSA in the reported prevalence of our seven key pulmonary disorders.

----

6. We postpone formal statistical tests until later.

Other Pulmonary Disorders

In addition to the seven key pulmonary disorders, the survey asked respondents about other pulmonary conditions including: Emphysema, Pulmonary Tuberculosis, Pneumonia, Bronchitis, Sinus Trouble, and Hay Fever. While these disorders were not included in our original plan for the analysis of the survey,[7] we did tabulate the responses to these questions. The resulting tabulation is shown in Table 9.

Table 9 indicates that of the seven additional pulmonary disorders, an excess of reporting appears in South Boston for Pulmonary Tuberculosis, Emphysema and Chronic Bronchitis (however only the Tuberculosis finding is statistically reliable). Non-chronic Bronchitis together with Hay Fever and Sinus Trouble are reported more frequently in the SMSA sample. Pneumonia is reported by almost identical proportions of respondents in the two samples.

With the exception of Pulmonary Tuberculosis and Emphysema, the ratios of the odds for reporting the pulmonary disorders shown in Table 9 are all in the range 0.7 to 1.3. These ratios did not seem large enough to require a revision of our basic analysis strategy. For emphysema, the observed ratio is rather higher (1.65), but the finding is unreliable even if we were to assume that our estimates were derived from a simple random sampling (SRS) design.

----

7. C.F. Turner and M.E. Colten, Basic Analysis Plan: Boston Area Pulmonary Health Survey. August 12, 1983. 32 pp.

The Pulmonary Tuberculosis finding, on the other hand, is both "reliable" ($p<.01$ under SRS assumption), and the ratio is quite large: Pulmonary Tuberculosis is 2.75 times more likely to be reported in South Boston than in the SMSA. Because of these two considerations, we have included Pulmonary Tuberculosis in our subsequent analyses. [Because this decision was taken at a rather late point in our work, we have postponed until Section V further consideration of the Pulmonary Tuberculosis data.]

TABLE 1: Unadjusted Estimates of Prevalence Rates for Reported Pulmonary Disorders in South Boston and Boston SMSA: Major Dependent Variables.

Disorder	Boston SMSA	South Boston	Ratio of Odds SB/SMSA		X^2	df
Chronic Cough (Most days for 3 months in a row; Q.A8)	10.6%	15.0%	1.49		12.8	1
Chronic Phlegm (Most days for 3 months in a row; Q.A15)	10.0	16.5	1.78		27.3	1
Chronic Wheeze (most days and nights; Q.A20a)	1.8	3.7	2.10		10.6	1
Asthma (Diagnosed by physician; Q.B5d)	6.1	5.7	0.93		0.3	1
Recent Chest Illness (Kept at home by illness: last 3 years; Q.A24)	34.4	35.2	1.04		0.2	1
Recent Severe Chest Illness (Kept at home & Phlegm for 1 week: during last 3 years; Q.A24b)	18.1	17.6	0.97		0.2	1
Shortness of Breath (Highest Level; QA26-30)[b]						
0 - None	81.7	71.9	(a)		53.8	5
I - When hurrying or hill	13.5	18.8	1.48			
II - Have to walk slowly on level	2.8	4.3	1.56			
III- Breathless walking own pace	0.6	2.1	3.55			
IV - Breathless after few minutes walking on level	1.2	1.5	1.25	2.03		
V - Too breathless to go out or breathless dressing or undressing	0.2	1.3	6.57			
SAMPLE SIZES	1547	1444				

NOTE: Chi-square statistics do not account for non-SRS (simple random sampling) design; hence, common inferences, e.g., p = .05 at 3.84 with 1 d.f., 11.07 with 5 d.f. will not be precisely accurate. Estimates of sample design and interviewer effects upon the variance of these measurements are presented in section VI.

Ratios of odds indicate excess of pulmonary disorders in South Boston when ratios are greater than 1.0; ratios less than 1.0 indicate excess of disorders in SMSA. For example, a ratio of 2.0 would indicate that the odds that a (randomly selected) respondent in South Boston would report a given disorder were twice as high as the odds in the SMSA; a ratio of 0.50 would indicate that the odds were half as high in South Boston as in the SMSA. [Note that the ratio of the odds is not equivalent to ratio of the rates, since, for example, a rate of 1 in 10 (i.e. .10) corresponds to an odds ratio of 1 to 9 (1/9 = .11).]

Base sample sizes vary slightly between measurements due to missing data. Estimates were obtained by weighting sample data by reciprocal of sampling probabilities (see footnote 1, page III-1).

(a) This ratio (0.57) is not shown because it has a different interpretation than other ratios in the table; it indicates the ratio of the odds that a respondent would report no breathlessness.

(b) In subsequent Tables and analyses we use either (A) a 3-category dependent variable: Level 0 vs. Level I vs. Level II or higher; or (B) simple dichotomy: Level 0 or Level I versus Level II or higher.

Table 2

Percentage Distributions for Social and Demographic Variables and Environmental Risk Factors in South Boston and Boston SMSA

Variable	Boston SMSA	South Boston	X^2	df
MAJOR CONTROL VARIABLES				
Smoking				
Never smoked	38.0%	31.8%	56.8	5
Ex. Smoker	22.1	20.0		
Cigar or Pipe Smoker	7.7	4.6		
Cigarette Smoker: 0-15/day	10.7	11.4		
Cigarette Smoker: 16-30/day	15.7	21.4		
Cigarette Smoker: 31+	5.8	10.8		
Age				
18-34	41.7	44.6		
35-49	26.6	20.8	15.1	3
50-64	18.4	21.3		
65+	13.3	13.4		
Sex:				
Male	46.8	43.3	3.8	1
Female	53.2	56.7		
OTHER SOCIAL & DEMOGRAPHIC VARIABLES				
Education				
0-11 Years	10.4	23.9	218.7	2
12 Years	30.8	43.2		
13+ Years	58.7	32.9		
Family Income				
0 - 4,999	4.7	9.5	184.6	6
5,000 - 9,999	10.3	17.7		
10,000 - 14,999	13.7	20.1		
15,000 - 19,999	11.2	14.4		
20,000 - 24,999	12.5	14.0		
25,000 - 29,999	10.6	8.8		
30,000 or more	36.9	15.6		

NOTE: Chi-square statistics do not account for non-SRS sample design; hence, common inferences, e.g., p = .05 at 3.84 with 1 d.f., 5.99 with 2 d.f, 7.81 with 3 d. 11.07 with 5 d.f., and 12.59 with 6 d.f. will not be precisely accurate. Estimates of sample design and interviewer effects upon the variance of these measurements are presented in Tables 16 - 18.

Base sample sizes vary slightly between measurements due to missing data. Estimates were obtained by weighting sample data by reciprocal of sampling probabilities (see footnote 1, page III-1).

Variable	Boston SMSA	South Boston	X^2	df
OTHER SOCIAL & DEMOGRAPHIC VARIABLES				
Work 30 Hours a Week: Yes	62.9%	56.6%	12.3	1
Self-employed: Yes	10.7	5.1	29.5	1
Marital Status				
Married	56.4	43.1	60.0	4
Widowed	7.9	11.5		
Separated	2.0	3.8		
Divorced	6.6	9.8		
Never Married	27.0	31.8		
Race/Ethnicity				
White	93.1	98.5	72.2	4
Black	4.8	0.1		
Hispanic	0.7	0.2		
Amer. Indian	0.4	0.5		
Asian	1.0	0.7		
Ever Lived in South Boston: Yes	2.4	na		
ENVIRONMENTAL AND RISK FACTORS				
Ever Work on Dusty Job for 1 Yr.: Yes	17.9	26.6	30.7	1
Exposed to Gas/Fumes outside Work: Yes	24.6	24.2	0.1	1
Type of Heating System				
Hot Water or Steam	61.3	60.4	111.5	3
Forced Hot Air	27.7	18.9		
Electric Units	6.6	5.6		
Other	4.4	15.1		
Ever Use Kerosene Heater: Yes	2.0	1.5	1.1	1
Cooking Fuel				
Gas	49.9	87.5	485.7	4
Electric	49.8	12.4		
Coal or Coke	0.0	0.0		
Wood	0.2	0.0		
Other	0.1	0.1		
Have a Gas Log: Yes	14.0	39.8	147.3	1
(Use for Heat as % of those having one)	(45.9)	(77.5)	(44.4)	(1)
Ever Use Gas Kitchen Stove for Heat: Yes	6.5	11.5	23.3	1
SAMPLE SIZES	1547	1444		

Table 3

Prevalence of Reported Pulmonary Disorders by Cigarette Smoking
(South Boston and SMSA Samples Combined)

	Cigarette Smoking						
Disorder	Never Smoked	Ex-Smoker	1-15 Daily	16-30 Daily	31+ Daily	X^2	df
Chronic Cough	5.3% (1216)	5.0% (622)	9.7% (327)	25.0% (545)	47.0% (241)	424.0	4
Chronic Phlegm	8.3 (1214)	9.0* (622)	11.1 (327)	20.6 (545)	35.4 (241)	165.6	4
Chronic Wheeze	.9 (1213)	1.3 (623)	2.7 (327)	4.8 (545)	11.1 (241)	91.7	4
Shortness of Breath							
None	85.0	81.0	76.1	66.2	49.0	200.0	8
Level I	11.0	12.2	18.6	23.9	32.2		
Level II	4.0	6.8	5.3	9.8	18.8		
	(1210)	(622)	(326)	(544)	(238)		
Asthma	5.6 (1219)	6.1 (623)	7.7 (327)	5.4 (546)	4.9 (242)	2.8	4
Chest Colds	34.4 (1216)	38.1 (623)	27.1 (327)	37.6 (544)	32.7 (242)	14.1	4
Severe Chest Colds	16.4 (1216)	18.9 (624)	16.2 (327)	20.3 (546)	19.4 (242)	5.4	4

Notes: 1. Base Ns for percentages are shown in parentheses. Base Ns vary slightly due to non-response on particular items.

2. Chi-square statistics do not account for non-SRS sample design; hence, common inferences, e.g., p = .05 at 9.49 with 4 d.f. and 15.51 with 8 d.f. will not be precisely accurate. Estimates of sample design and interviewer effects upon the variance of these measurements are presented in Section VI.

3. These analyses are not intended to provide prevalence estimates for SMSA; sample strata have been combined without re-weighting strata (to reflect oversampling of South Boston residents). Table 21 presents estimates of prevalence rates for the SMSA.

Table 4

Prevalence of Reported Pulmonary Disorders by Age
(South Boston and SMSA Samples Combined)

Disorder	Age 18-34	35-49	50-64	65+	X^2	df
Chronic Cough	11.7% (1260)	14.0% (694)	15.6% (578)	10.2% (389)	8.5	3
Chronic Phlegm	12.7 (1259)	13.1 (692)	15.8 (578)	10.7 (390)	6.0	3
Chronic Wheeze	2.3 (1258)	3.2 (694)	3.4 (578)	2.2 (390)	3.1	3
Shortness of Breath						
None	82.9	77.6	69.4	66.6	83.0	6
Level I	13.6	15.9	18.9	21.4		
Level II	3.6	6.6	11.7	12.0		
	(1259)	(693)	(573)	(385)		
Asthma	6.6 (1260)	7.2 (696)	5.6 (579)	2.9 (390)	9.2	3
Chest Colds	40.5 (1260)	35.5 (696)	28.8 (576)	23.9 (388)	47.3	3
Severe Chest Colds	21.6 (1260)	17.1 (696)	14.5 (579)	12.3 (390)	24.8	3

Notes:

1. Base Ns for percentages are shown in parentheses. Base Ns vary slightly due to non-response on particular items.

2. Chi-square statistics do not account for non-SRS sample design; hence, common inferences, e.g., p = .05 at 7.81 with 3 d.f., and 12.59 with 6 d.f. will not be precisely accurate. Estimates of sample design and interviewer effects upon the variance of these measurements are presented in Section VI.

3. These analyses are <u>not</u> intended to provide prevalence estimates for SMSA; sample strata have been combined <u>without</u> re-weighting strata (to reflect oversampling of South Boston residents). Table 21 presents estimates of prevalence rates for the SMSA .

Table 5

Prevalence of Reported Pulmonary Disorders by Sex
(South Boston and SMSA Samples Combined)

Disorder	Sex: Male	Sex: Female	X^2	df
Chronic Cough	12.3% (1348)	13.1% (1637)	0.4	1
Chronic Phlegm	15.4 (1345)	11.3 (1639)	10.9	1
Chronic Wheeze	2.5 (1348)	2.9 (1635)	0.3	1
Shortness of Breath				
None	83.0	72.0	49.9	2
Level I	12.0	19.4		
Level II	5.0	8.6		
	(1345)	(1628)		
Asthma	6.2 (1349)	5.7 (1640)	0.3	1
Chest Colds	30.4 (1348)	38.5 (1636)	21.4	1
Severe Chest Colds	14.7 (1350)	20.4 (1641)	16.9	1

Notes: 1. Base Ns for percentages are shown in parentheses. Base Ns vary slightly due to non-response on particular items.

2. Chi-square statistics do not account for non-SRS sample design; hence, common inferences, e.g., p = .05 at 3.84 with 1 d.f., and 5.99 with 2 d.f., will not be precisely accurate. Estimates of sample design and interviewer effects upon the variance of these measurements are presented in Section VI.

3. These analyses are not intended to provide prevalence estimates for SMSA; sample strata have been combined without re-weighting strata (to reflect oversampling of South Boston residents). Table 21 presents evidence of prevalence rates for the SMSA.

TABLE 6

Prevalence of Reported Pulmonary Disorders in South Boston and SMSA Samples by Smoking

DISORDER	SMOKING					
	Never Smoked	Ex-Smoker	0-15 daily	16-30 daily	31+ daily	All
CHRONIC COUGH						
SMSA	4.4%	3.1%	10.2%	28.0%	41.7%	10.6%
	(697)	(337)	(163)	(240)	(88)	(1525)
South Boston	6.5	7.2	9.2	22.7	50.0	15.1
	(519)	(285)	(164)	(306)	(153)	(1427)
CHRONIC PHLEGM						
SMSA	7.7	6.6	7.1	18.4	25.1	10.1
	(694)	(337)	(163)	(240)	(85)	(1522)
South Boston	9.0	11.9	15.1	22.3	41.3	16.6
	(520)	(285)	(164)	(306)	(153)	(1428)
CHRONIC WHEEZE						
SMSA	0.8	1.3	2.4	3.8	5.2	1.8
	(696)	(337)	(163)	(240)	(88)	(1524)
South Boston	1.1	1.3	3.1	5.5	14.5	3.8
	(518)	(286)	(164)	(305)	(153)	(1426)
SHORTNESS OF BREATH						
SMSA						
None	86.9	85.3	78.6	68.8	64.9	81.5
Level 1	9.8	8.8	18.7	22.4	28.1	13.6
Level 2	3.3	5.8	2.7	8.8	7.1	4.9
	(694)	(337)	(163)	(240)	(88)	(1522)
South Boston						
None	82.5	75.9	73.6	64.2	39.8	71.7
Level 1	12.5	16.2	18.5	25.1	34.6	19.0
Level 2	5.0	7.9	7.9	10.6	25.6	9.3
	(516)	(285)	(163)	(304)	(150)	(1419)
ASTHMA						
SMSA	5.7	5.7	9.4	5.2	4.4	5.9
	(697)	(338)	(163)	(240)	(88)	(1526)
South Boston	5.3	6.6	5.9	5.5	3.2	5.8
	(521)	(286)	(164)	(306)	(154)	(1431)
CHEST COLDS						
SMSA	34.0	35.6	25.3	40.3	31.6	34.3
	(697)	(337)	(163)	(239)	(88)	(1524)
South Boston	35.0	41.0	28.8	35.5	33.3	35.4
	(519)	(286)	(164)	(305)	(154)	(1428)
SEVERE CHEST COLDS						
SMSA	17.2	17.6	13.0	23.0	22.3	18.1
	(697)	(338)	(163)	(240)	(88)	(1526)
South Boston	15.3	20.4	19.4	18.1	17.7	17.7
	(522)	(286)	(164)	(306)	(154)	(1431)

Note. Base Ns for percentages are shown in parentheses. They vary slightly between disorders due to non-response.

TABLE 7

Prevalence of Reported Pulmonary Disorders in South Boston and SMSA Samples by Age Group

	AGE				
DISORDER	18-34	35-49	50-64	65+	All
CHRONIC COUGH					
SMSA	10.2%	10.6%	12.7%	10.0%	10.8%
	(631)	(403)	(278)	(202)	(1513)
South Boston	13.2	18.7	18.2	10.4	15.0
	(629)	(291)	(300)	(187)	(1407)
CHRONIC PHLEGM					
SMSA	11.3	7.9	10.7	10.5	10.2
	(630)	(401)	(278)	(202)	(1511)
South Boston	14.1	20.2	20.6	10.9	16.3
	(629)	(291)	(299)	(188)	(1408)
CHRONIC WHEEZE					
SMSA	2.1	0.8	2.5	2.1	1.8
	(631)	(403)	(278)	(202)	(1513)
South Boston	2.5	6.5	4.3	2.2	3.7
	(627)	(291)	(300)	(188)	(1406)
SHORTNESS OF BREATH					
SMSA					
None	87.9	82.3	77.2	66.1	81.5
Level 1	10.5	14.2	13.1	23.3	13.6
Level 2	1.7	3.6	9.7	10.6	4.8
South Boston	(631)	(403)	(276)	(200)	(1510)
None	77.8	71.1	62.1	67.1	71.7
Level 1	16.7	18.2	24.4	19.4	19.0
Level 2	5.5	10.7	13.5	13.6	9.3
	(628)	(290)	(297)	(185)	(1400)
ASTHMA					
SMSA	6.5	6.9	6.2	4.0	6.2
	(631)	(403)	(278)	(202)	(1514)
South Boston	6.6	7.5	5.1	1.6	5.8
CHEST COLDS					
SMSA	41.3	34.7	26.4	22.5	34.3
	(631)	(403)	(278)	(200)	(1513)
South Boston	39.7	36.7	31.1	25.4	35.4
	(629)	(293)	(298)	(187)	(1408)
SEVERE CHEST COLDS					
SMSA	23.9	15.1	14.5	11.7	18.2
	(631)	(403)	(278)	(202)	(1514)
South Boston	19.3	19.9	14.5	13.1	17.6
	(629)	(293)	(301)	(188)	(1411)

Note. Base Ns for percentages are shown in parentheses. They vary slightly between disorders due to non-response.

TABLE 8

Prevalence of Reported Pulmonary Disorders in South Boston and SMSA Samples by Sex

DISORDER	Men	Women	Both
CHRONIC COUGH			
SMSA	9.7%	11.3%	10.6%
	(724)	(821)	(1545)
South Boston	15.2	14.8	15.0
	(624)	(816)	(1440)
CHRONIC PHLEGM			
SMSA	11.5	8.7	10.0
	(721)	(821)	(1543)
South Boston	19.8	13.9	16.5
	(623)	(818)	(1441)
CHRONIC WHEEZE			
SMSA	1.8	1.8	1.8
	(724)	(820)	(1544)
South Boston	3.4	4.0	3.7
	(624)	(815)	(1438)
SHORTNESS OF BREATH			
SMSA			
None	86.8	77.2	81.7
Level 1	9.7	16.8	13.5
Level 2	3.5	5.9	4.8
	(724)	(818)	(1543)
South Boston			
None	78.5	66.8	71.9
Level 1	14.7	22.0	18.8
Level 2	6.7	11.2	9.3
	(621)	(809)	(1431)
ASTHMA			
SMSA	6.0	6.2	6.1
	(724)	(823)	(1547)
South Boston	6.4	5.2	5.7
	(625)	(817)	(1442)
CHEST ILLNESS			
SMSA	28.3	39.8	34.4
	(724)	(821)	(1545)
South Boston	32.7	37.1	35.2
	(624)	(815)	(1439)
SEVERE CHEST ILLNESS			
SMSA	14.0	21.7	18.1
	(724)	(823)	(1547)
South Boston	15.4	19.2	17.6
	(625)	(819)	(1444)

Note. Base Ns for percentages are shown in parentheses. They vary slightly between disorders due to non-response.

Table 9

Estimates of Prevalence of Other Reported Pulmonary Disorders in South Boston and Boston SMSA

Disorders	Percent Reporting Disorder		Ratio of Odds SB/SMSA	X^2	d.f.
	Boston SMSA	South Boston			
Pulmonary Tuberculosis	0.7%	1.9%	2.75	7.8	1
Emphysema	1.1	1.8	1.65	2.8	1
Pneumonia	22.0	22.8	1.05	0.2	1
Bronchitis					
Chronic	3.6	4.7	1.32	2.6	1
Acute ("Attacks of")	21.7	17.8	0.78	7.0	1
Hay Fever	23.6	18.2	0.72	12.8	1
Sinus Trouble	31.9	25.7	0.74	14.0	1

NOTE: Chi-square statistics do not account for non-SRS (simple random sampling) sample design; hence, common inferences, e.g., p = .05 at 3.84 with 1 d.f. will not be precisely accurate. Estimates of sample design and interviewer effects upon the variance of these measurements are presented in Section VI.

Ratios of odds indicate excess of pulmonary disorders in South Boston when ratio is greater than 1.0; ratios less than 1.0 indicate excess of pulmonary disorder in SMSA. For example, a ratio of 2.0 would indicate that the odds that a (randomly selected) respondent in South Boston would report a given disorder were twice as high as the odds in the SMSA; a ratio of 0.50 would indicate that the odds were half as high in South Boston as in the SMSA. [Note that the ratio of the odds is not equivalent to ratio of the rates, since, for example, a rate of 1 in 10 (rate of .10) corresponds to an odds ratio of 1 to 9 (1/9 = .11).]

IV. ESTIMATES OF PREVALENCE OF PULMONARY DISORDERS CONTROLLING FOR AGE AND SEX AND FOR SMOKING AND SEX

IV. ESTIMATES OF PREVALENCE OF PULMONARY DISORDERS CONTROLLING FOR AGE AND SEX AND FOR SMOKING AND SEX

We have seen in the preceding section the effects which are obtained when inter-area differences in reported prevalence rates are "controlled" for age or sex or smoking. The question that logically arises is: What if we control for both smoking and sex <u>or</u> age and smoking <u>or</u> sex and age at the same time?

Such an analysis requires that we now compare not only women in South Boston to women in the Boston SMSA,[1] for example, but that we further refine our strategy of comparison so that we compare non-smoking women in South Boston to non-smoking women in the SMSA, light-smoking women in South Boston to their counterparts in the SMSA, and so forth. To use the nomenclature of the later sections, our analysis strategy requires that we build 4-way tables of, for example: Sex by Age by Area by Disorder. This table will have 2 X 4 X 2 X 2 = 32 cells since there are: two sexes, four age groups, two areas (South Boston and SMSA), and two possible values of the disorder variable (i.e., has disorder <u>or</u> does not have disorder).

----

1. Throughout this report, we use "Boston metropolitan area" and Boston SMSA to denote the federal government's Standard Metropolitan Statistical Area for Boston; see Section I for precise definitions. Where comparisons are reported <u>between</u> South Boston and the Boston SMSA, our tabulations contrast the SMSA <u>excluding South Boston</u> to South Boston.

   The term pulmonary disorder is used throughout this report to denote a class of conditions that includes both pulmonary diseases and pulmonary symptoms.

Each respondent in the survey will fit into one of these cells,[2] and each of 2 X 4 = 8 combinations of Sex+Age groupings in the 2 areas would have an average of about 200 respondents. Thus the estimated rates of disorder for each cell would be based, on average, on a relatively small sample, and we might expect that some rates might be based on samples that were considerably smaller than this average. This fact has an important implication for the use of such "intuitive" analytic techniques; it results in very unstable estimates of the rates for any particular cell.[3] Thus, it is important for readers to bear in mind that estimates of individual cells of our 4-way tables may vary considerably (due to chance sampling fluctuations) from the values which would have been obtained if we did a census of the entire population rather than using a sample.

The 4-way Tables 10 and 11 which breakdown the rates of reported pulmonary disorder for each area by Sex and Age (Table

----

2. A small number of respondents (typically about 30 in each sample) will be excluded from the analysis because they did not provide one of the needed bits of information, e.g., they did not give their age or answer the question about the pulmonary disorder.

3. For example, a measured rate of 10 percent for chronic phlegm would have a standard error of 2.12 if it were based upon a simple random sample (SRS) of 200 persons. If we were to allow for a design effect of, say, 1.5 to represent the combined impact of interviewer contributions to the variance and deviations of our sample from an SRS design, the standard error would inflate to 3.18. With a standard error this large meaningful inference becomes extremely difficult, since we must allow for the fact that sampling fluctuations alone would be expected 5 times out of 100 to cause our rate to increase (or decrease) by as much as 6.2 percentage points (i.e., if we repeated the survey 100 times, keeping all factors including the "true" rate of pulmonary disorder constant).

10) and by Sex and Smoking (Table 11) should be examined with this caveat in mind.[4] Although sampling fluctuations do disturb some of the patterns, we again find many familar patterns, e.g., the substantial association between smoking and Chronic Cough, Phlegm, Wheeze and Shortness of Breath (Table 11). We also note that there are several instances in which almost every group of South Boston residents shown in these tables evidences more of a particular symptom than residents of the SMSA. This is particularly the case for Chronic Phlegm and Shortness of Breath. Chronic Cough and Chronic Wheeze[5] evidence a somewhat more unstable pattern. Reports of Asthma, Chest Colds and Severe Chest Colds do not show a consistent pattern.

In the next section we will begin a more formal statistical analysis of these rates.

----

4. We purposely do not present the 4-way table of Symptom by Area by Smoking by Age. Since Smoking has 5 categories and Age has 4, rates within the 2 areas would be based on exceedingly small samples. There are 20 Smoking+Age groupings in each of the 2 areas which yields an average expected sample size of approximately 1500/20 = 75.

5. While we do not wish to belabor our discussion of these tables (given the large sampling errors that are involved), we do nonetheless note some oddities. Chronic wheeze, in particular, shows an unexpected "jump" in its rate of reporting among middle aged men in South Boston. The rate rises from 1.9 percent for 18-34 year olds to 9.0 percent for men aged 35-49. It then recedes to 0.8 and 2.6 percent for men aged 50-64 and 65+. (No similar "jump" is observed for females.) We also note that the rate of chronic wheeze shows an unusual variation in its relative size compared to the rates of reporting for asthma (a pulmonary disorder which manifests itself in wheezing, among other symptoms). In a number of instances the rate of asthma for specific groups is many times that of chronic wheezing, while in other instances the rates of chronic wheezing exceed by a noteworthy margin the rates of reported asthma.

TABLE 10

Estimates of Prevalence of Reported Pulmonary Disorders in South Boston and Boston SMSA Samples by Sex and Age Group

	Males				Females				
DISORDER	18-34	35-49	50-64	65+	18-34	35-49	50-64	65+	ALL
CHRONIC COUGH									
SMSA	7.7% (301)	9.2% (198)	14.9% (139)	10.9% (79)	12.6% (330)	11.9% (205)	10.5% (139)	9.4% (122)	10.8 (1513)
So. Boston	11.9 (291)	20.3 (137)	21.4 (129)	6.1 (58)	14.3 (338)	17.2 (155)	15.7 (170)	12.4 (128)	15.0 (1407)
CHRONIC PHLEGM									
SMSA	12.8 (301)	9.6 (196)	13.1 (139)	9.7 (79)	9.9 (330)	6.3 (205)	8.3 (139)	11.0 (122)	10.2 (1511)
So. Boston	16.1 (291)	20.1 (137)	27.9 (129)	16.7 (58)	12.4 (338)	20.2 (155)	15.1 (170)	8.3 (130)	16.3 (1408)
WHEEZE									
SMSA	1.1 (301)	0.7 (198)	3.8 (139)	3.6 (79)	2.9 (331)	0.9 (205)	1.2 (139)	1.2 (122)	1.8 (1514)
SO. BOSTON	1.9 (290)	9.0 (137)	0.8 (130)	2.6 (58)	3.0 (337)	4.3 (154)	6.9 (170)	2.0 (130)	3.7 (1407)
ASTHMA									
SMSA	7.0 (301)	5.6 (198)	6.6 (139)	3.0 (79)	5.9 (331)	8.2 (205)	5.9 (139)	4.7 (122)	6.2 (1514)
So. Boston	8.5 (291)	7.9 (137)	2.8 (130)	0.9 (58)	5.0 (338)	7.2 (156)	6.9 (169)	2.0 (130)	5.8 (1411)

(Cont'd)

Note. Base Ns for percentages are shown in parentheses. They vary slightly between disorders due to non-response.

TABLE 10 (Cont'd)

Estimates of Prevalence of Reported Pulmonary Disorders in South Boston and Boston SMSA by Sex and Age Group

	Males				Females				
DISORDER	18-34	35-49	50-64	65+	18-34	35-49	50-64	65+	ALL
CHEST COLD									
SMSA	37.1% (301)	30.1% (198)	16.6% (139)	11.5% (79)	45.1% (331)	39.1% (205)	36.1% (139)	29.8% (121)	34.3 (1513)
So. Boston	38.2 (291)	33.8 (137)	21.8 (129)	27.2 (58)	41.0 (338)	39.2 (156)	38.3 (169)	24.7 (129)	35.4 (1408)
SEVERE CHEST COLD									
SMSA	21.7 (301)	11.2 (198)	6.9 (139)	4.8 (79)	25.8 (331)	18.9 (205)	22.0 (139)	16.1 (122)	18.2 (1514)
So. Boston	18.9 (291)	17.4 (137)	9.4 (130)	7.0 (58)	19.7 (338)	22.1 (156)	18.4 (170)	15.8 (130)	17.6 (1411)
SHORTNESS OF BREATH									
SMSA									
None	93.0	87.6	82.7	67.3	83.2	77.1	71.6	65.3	81.5
Level I	6.5	10.0	7.6	25.5	14.0	18.2	18.7	21.9	13.6
Level II	0.5	2.4	9.7	7.3	2.8	4.7	9.8	12.7	4.8
	(301)	(198)	(139)	(79)	(331)	(205)	(138)	(120)	(1510)
So. Boston									
None	88.2	70.5	68.5	69.6	68.9	71.5	57.2	65.9	71.7
Level I	10.2	16.1	22.7	18.8	22.3	20.1	25.7	19.6	19.0
Level II	1.6	13.3	8.8	11.6	8.8	8.4	17.1	14.4	9.3
	(290)	(137)	(129)	(57)	(338)	(153)	(168)	(128	(1400)

TABLE 11

Estimates of Prevalence of Reported Pulmonary Disorders in South Boston and Boston SMSA
by Sex and Cigarette Smoking

	Males					Females					
DISORDER	Never Smoked	Ex-Smoker	1-15 Daily	16-30 Daily	31+ Daily	Never Smoked	Ex-Smoker	1-15 Daily	16-30 Daily	31+ Daily	ALL
CHRONIC COUGH											
SMSA	4.1% (342)	2.2% (155)	20.0% (46)	18.9% (112)	44.1% (49)	4.9% (345)	3.9% (172)	6.7% (112)	36.5% (125)	41.7 (36)	10.8 (1493)
So. Boston	7.4 (243)	5.9 (122)	12.4 (46)	22.6 (129)	48.0 (70)	5.4 (264)	8.5 (156)	8.4 (112)	22.9 (174)	51.5 (79)	15.1 (1396)
CHRONIC PHLEGM											
SMSA	8.5 (340)	6.5 (155)	15.8 (46)	19.7 (112)	30.4 (49)	7.3 (344)	6.9 (173)	3.9 (112)	17.3 (125)	19.9 (36)	10.3 (1491)
So. Boston	11.7 (242)	14.8 (122)	19.1 (46)	25.0 (129)	46.5 (70)	5.4 (266)	9.5 (156)	13.8 (112)	20.6 (174)	37.1 (79)	16.5 (1397)
WHEEZE											
SMSA	0.4 (342)	2.8 (155)	3.2 (46)	3.0 (112)	4.9 (49)	1.3 (345)	0.0 (173)	2.2 (112)	4.6 (125)	6.0 (36)	1.8 (1494)
So. Boston	0.8 (242)	0.8 (122)	4.5 (46)	4.0 (130)	14.3 (70)	1.0 (264)	1.6 (157)	2.7 (112)	6.8 (173)	14.3 (79)	3.7 (1395)
ASTHMA											
SMSA	5.2 (342)	4.3 (155)	11.6 (46)	6.4 (112)	6.9 (49)	6.4 (345)	6.9 (173)	9.0 (112)	4.2 (125)	0.0 (36)	6.0 (1495)
So. Boston	7.4 (243)	8.9 (122)	10.1 (46)	2.0 (130)	5.5 (71)	4.1 (266)	5.2 (157)	4.6 (112)	8.2 (174)	5.2 (79)	5.9 (1399)

(Cont'd)

Note. Base Ns for percentages are shown in parentheses. They vary slightly between disorders due to non-response.

TABLE 11 (Cont'd)

Estimates of Prevalence of Reported Pulmonary Disorders in South Boston and Boston SMSA by Sex and Cigarette Smoking

	Males					Females					
DISORDER	Never Smoked	Ex-Smoker	1-15 Daily	16-30 Daily	31+ Daily	Never Smoked	Ex-Smoker	1-15 Daily	16-30 Daily	31+ Daily	ALL
CHEST COLDS											
SMSA	31.7% (342)	25.2% (155)	15.8% (46)	27.5% (112)	24.5% (49)	35.1% (344)	44.9% (173)	29.7% (112)	52.8% (125)	39.7% (36)	34.1% (1493)
So. Boston	38.4 (242)	37.6 (122)	24.7 (46)	29.2 (130)	19.6 (71)	32.3 (265)	44.3 (157)	30.5 (112)	39.8 (173)	46.3 (79)	35.6 (1396)
SEVERE CHEST COLDS											
SMSA	16.5 (342)	11.8 (155)	7.4 (46)	10.7 (112)	16.7 (49)	18.2 (345)	22.4 (173)	15.9 (112)	34.5 (125)	27.8 (36)	18.1 (1495)
So. Boston	15.5 (243)	18.1 (122)	16.9 (46)	14.2 (130)	13.1 (71)	15.2 (266)	21.8 (157)	20.5 (112)	21.2 (174)	21.5 (79)	17.6 (1400)
SHORTNESS OF BREATH											
SMSA											
None	92.6	87.6	78.9	76.4	69.6	81.3	82.8	77.4	61.2	58.3	81.3
Level I	6.5	6.5	15.8	15.5	27.5	13.2	11.4	20.9	29.2	28.5	13.8
Level II	1.0	5.9	5.3	8.2	2.9	5.5	5.8	1.7	9.8	13.2	4.9
	(342)	(155)	(46)	(112)	(49)	(341)	(173)	(112)	(125)	(36)	(1490)
So. Boston											
None	88.0	77.1	77.3	73.9	54.7	77.2	74.1	71.7	56.8	25.6	71.5
Level I	9.5	16.5	13.6	17.0	29.6	15.6	16.4	20.5	31.2	39.9	19.2
Level II	2.4	6.4	9.1	9.1	15.7	7.2	9.5	7.8	11.9	34.6	9.4
	(242)	(121)	(45)	(130)	(68)	(263)	(156)	(112)	(172)	(77)	(1389)

V. ESTIMATION OF LOG-LINEAR MODELS TO TEST FOR INTER-AREA DIFFERENCES IN PREVALENCE OF PULMONARY DISORDERS

V. ESTIMATION OF LOG-LINEAR MODELS TO TEST FOR INTER-AREA DIFFERENCES IN PREVALENCE OF PULMONARY DISORDERS

Further analysis of the prevalence rates may continue through even more elaborate tables, but ultimately we must confront two constraints:

(1) The number of cells in our tables grows exponentially with the number of variables we "control," thus taxing both our ability to do crude analysis "by eye," and the limited ability of our sample sizes to provide reasonable precision in estimating the prevalence rates for particular subgroups defined by the control variables (e.g., estimates of the rate of Chronic Phlegm for women aged 65+ who smoke 31+ cigarettes a day).

(2) We inevitably must address in a more formal manner the question of whether the totality of the data warrant the conclusion that particular symptoms are reported more frequently (than otherwise expected) in the target area.[1] In particular, given that all of our data are derived from samples of the populations of the Boston

----

1. We may, of course, wish to be specific about population subgroups that are particularly deviant or that run counter to the general pattern, but we shall leave that question for later. It also may be the case that no general pattern exists although there are substantial differences across areas for selected groups. Indeed, we shall see one such result in the following analyses.

SMSA[2] and the South Boston community, there is a non-zero probability, regardless of whatever numbers we find,[3] that the result could have been reversed if we had interviewed every member of the two populations and not contented ourselves with samples of 1,547 and 1,444.

While non-technical readers doubtlessly appreciate the notion that we should place greater faith in "big" differences found in such sample surveys, there are technical details to the formal hypothesis testing we must employ that cannot be briefly stated in everyday language. Even such a simple issue as "What is a big difference?" has a special meaning in the case of sample statistics. Consider, for example, the difference between prevalence rates of 50 and 52 percent in samples the size of ours

----

2. Throughout this report, we use "Boston metropolitan area" and Boston SMSA to denote the federal government's Standard Metropolitan Statistical Area for Boston; see Section II for precise definitions. Where comparisons are reported between South Boston and the Boston SMSA, our tabulations contrast the SMSA excluding South Boston to South Boston.

   The term pulmonary disorder is used to denote a class of conditions that includes both pulmonary diseases and pulmonary symptoms.

3. This is a slight overstatement. Exceptions are of the following sort: Imagine that we sample .1% of the population and find that not a single respondent says that they smoke cigarettes. While there is an infinitesimal but nonetheless non-zero probability that almost every member of the population (i.e., 99.9%) would have admitted to cigarette smoking in our survey (if we had sampled them), there is no possibility that this percentage could exceed 99.9%. In a similar manner, reversals of inferences that a<b become impossible when the number of people saying "a" in our sample equals or exceeds 50% of the total population.

and between a prevalence rate of 1 percent and 2 percent in the same samples. While the former difference of two percentage points might appear larger to some people, the latter difference (of one percentage point) would be more likely to stand up if one were to do a census and measure every person in each of the two populations.[4]

We have tried in the preceding sections to sketch out some of the considerations which motivate our approach to the question posed. At this point, however, we must apply a variety of formal statistical procedures which are, in our judgement, appropriate to eliciting from the survey data answers to the basic question of whether the prevalence of reported pulmonary disorders in South Boston is different (and indeed, greater) than that in the rest of the Boston metropolitan area (i.e, the Boston SMSA). In the concluding section, (section IX) we state in non-technical language the conclusions of these statistical analyses.

----

4. The ratio of the odds in these two cases provides a more useful comparison than the raw percentage point difference. The ratios are: (52/48)/(50/50) = 1.08 in the first case, and (2/98)/(1/99) = 2.02 in the second case.

Estimation of Models

For each of the dependent variables, a hierarchical set of log-linear models were fit to the 5-way table of weighted frequency counts[5] for: Symptom Report by Age by Smoking by Sex by Area. Model estimation was done using an iterative proportional scaling algorithm to obtain maximum likelihood estimators for the parameters of each log-linear model.[6] Models were specified with

---

5. As before, sampling weights (see footnote 1 of section 3) were applied and weights were divided by their average to insure that total weighted sample sizes equalled the unweighted sample sizes.

6. Following Goodman (1970: Sec. 3.1) we use P_{ijklm} to denote the probability (in a population table) that an observation will fall in cell i,j,k,l,m, where i=1,2 (report disorder or not), j=1,2 (SMSA or South Boston), k= 1,2 (male or female), l=1,4 (age: 18-34, 35-49,50-64, or 65+), m=1,5 (smoking: never, ex-smoker, 0-15, 16-30, or 31+ daily). Letting V_{ijklm}=log P_{ijklm} (where log refers to the natural logarithm), we may decompose V_{ijklm} for a model, say Model No. 2 ([Sy][Ar][Ag][Sx][Sm][SySm]), as follows:

$$V_{ijklm} = \lambda_i^{Sy} + \lambda_j^{Ar} + \lambda_k^{Sx} + \lambda_l^{Ag} + \lambda_m^{Sm} + \lambda_{im}^{SySm}$$

Models more complex than No. 2 add additional 2-, 3-, and 4-way effects, such as those the joint effect of Smoking and Age on Symptom Reporting: λ_{ilm}^{SyAgSm} etc.

It should be noted that in the representation of the log-linear model used by Goodman and Fay (i.e., ECTA and CPLX) the parameters for the interaction between two variables represent a system of parameters, one for each combination of outcomes for the two variables. The parameters are constrained to sum to zero both across rows and down columns. When both variables have two levels (such as 1: South Boston, vs. 2: SMSA; and, 1: report disorder, vs. 2: no report of disorder) the single estimated by ECTA or CPLX actually represents the system:

	So. Boston	SMSA
disorder	$+\lambda$	$-\lambda$
no disorder	$-\lambda$	$+\lambda$

All four lambda parameters work in the same direction in the computation of the expected log odds ratio: $[+\lambda - (-\lambda)] - [-\lambda - \lambda] = 4\lambda$.
We have adopted the convention in our tables of reporting tau-squared and 2 to represent the "effects" (where the base category would have, by definition, an effect of zero). The expected odds ratios and log odds ratios thus are twice the and the square of the tau values shown in our tables.

increasing degrees of specificity, constraining various combinations of 2-way, 3-way, and 4-way marginals. Appendix C presents a complete set of results of this model-fitting; it will be seen by inspection of these results that expected patterns of association (e.g. Cigarette smoking producing increased rates of Chronic Cough, etc.) will be found in the data from our survey.

Controls for age, sex, and smoking imposes a constraint on one 4-way marginal: [Sex by Age by Smoking by Area], similarly, the 4-way marginal for the "genesis"[7] of symptomatology [Symptom

---

7. That is, our model testing begins by allowing for associations between such presumed causal factors as cigarette smoking and the dependent variable (Reported pulmonary disorder). These factors which may generate pulmonary disorders are incorporated in our basic model <u>prior</u> to our testing for inter-area differences. Persons interested in testing for an association between symptom reporting and such factors as cigarette smoking may use the results shown in Appendix C. For example, one possible test of the effect of cigarette smoking on symptom reporting would be to contrast models 1 and 2, which provides the following tests of the symptom by smoking parameter:

	Model 1 minus Model 2	
Symptom	L^2	df
Chronic Cough	346.0	4
Chronic Phlegm	142.0	4
Chronic Wheeze	67.2	4
Asthma	2.8	4
Chest Cold	13.3	4
Severe Chest Cold	5.6	4
Shortness of Breath	173.7	8
Pulmonary TB	4.1	4

Under SRS assumptions $L^2 > 9.5$ has $p < .05$ with 4 d.f., and $L^2 > 15.5$ has $p < .05$ with 8 d.f. These results thus suggest highly significant associations of cigarette smoking with reported Chronic Cough, Phlegm, Wheeze and Shortness of Breath, a borderline association with reported Chest Colds, and no (simple) association between cigarette smoking and Severe Chest Colds, Asthma or Pulmonary Tuberculosis.

by Age by Smoking by Sex] must be constrained for the purpose of testing for differences across areas in health status. The model which incudes these constraints (Model 108 in Appendix C) provides the relevant base model. A Model (No.109 in Appendix C) which introduces a Symptom by Area [SyAr] constraint provides the test model. [A comparison of these two models provides the foundation for our analysis (although we will subsequently examine some additional models that stipulate more complicated effects, e.g., inter-area differences that vary by age.)]

Differences between South Boston and the Boston SMSA can be formally assessed by testing the improvement in model fit which is obtained by fitting (to an otherwise appropriately constrained base model) a parameter constraining the two-way marginal [Area by Symptom]. The increment in the model's fit (represented by the decrement in the likelihood ratio chi-square statistic, L^2) provides a test for the existence of a difference in the reported prevalence of a disorder in South Boston and the Boston SMSA, net of differences in all of the control variables. When this parameter significantly increases the fit of the log-linear model to the observed frequency counts, the null hypothesis of "no inter-area difference" might tentatively be rejected (see below for further considerations). The nature of any such "effect" might be further specified through investigation of higher-order interactions involving Area and Symptom plus one or more demographic factors. (Findings from such an analysis might lead to "restricted conclusions," such as: the difference in health status between the areas is specific to particular age groups, or

----

7. Chronic Cough, Phlegm, Wheeze and Shortness of Breath, a borderline association with reported Chest Colds, and no (simple) association between cigarette smoking and Severe Chest Colds, Asthma or Pulmonary Tuberculosis.

to men, or to a specific age group of smoking men, etc.)

Test of null hypothesis. Models fit in our analyses were evaluated using the Likelihood-ratio chi-square statistic to test the hypothesis that the deviations of actual frequency counts for interior cells of the 5-way table from those predicted by a given log-linear model could have arisen through sampling fluctuations. The degrees of freedom for this test is a function only of the number of parameters which are estimated and the number of cells in the table. Excellent discussions of this model testing and estimation strategy can be found in the recent statistical literature.[8]

Rationale. Our choice of this analysis strategy is dictated by two considerations: (1) Since the key dependent variables are discrete, the general linear model cannot be employed directly; and (2) at least one dependent variable, Shortness of Breath, is truly polytomous by virtue of its survey definition.[9]

----

8. For discussion of these specific techniques, see L.A. Goodman, The multivariate analysis of qualitative data. Journal of the American Statistical Association, 65, 1970, 226-256; and, L.A.Goodman, The analysis of multivariate contingency tables: stepwise procedures and direct estimation methods for building models for cross-classifications. Technometrics, 13, 1971, 33-61. For general treatments, see S.J. Haberman, Analysis of Cross-Classified Categorical Data. Two volumes. New York: Academic Press; and, Y.M.Bishop et al, Discrete Multivariate Analysis: Theory and Practice. Cambridge, Mass.: MIT Press, 1975. For a theoretically intensive and rigorous treatment see, S.J. Haberman, Analysis of Frequency Data. Chicago: University of Chicago Press.

9. By virtue of the skip patterns used in the survey, shortness of breath questions are by design interdependent: all persons who are positive on Level II Breathlessness must also be positive on Level I. Thus, while two dichotomous variables can be constructed and logit models fit, use of the polytomy with a multinomial response formulation is preferable.

Results

Tests for improvement in model fit will be found in Table 12. It will be seen from these results that differences between South Boston and the Boston SMSA are found for Chronic Phlegm, Chronic Wheeze, and Shortness of Breath. In addition, as will be discussed in a later section, Pul onary Tuberculosis also shows an elevated rate in South Boston. Of the remaining disorders, Chest Colds, Severe Chest Colds, and Asthma show no inter-area difference (once age, sex, and smoking are controlled). While prevalence rates for Chronic Cough do not show a simple difference, they do pose a special problem since a 5-way interaction must be fit in order to adequately explain the data (see Appendix G for estimates of these 5-way interaction parameters).

Table 13 presents estimates of the lambda parameters for the log-linear models (and the corresponding tau parameters for the multiplicative model). It will be seen that in the cases where we do find statistically significant differences between areas, all differences indicate an elevated level of reported pulmonary disorder in the South Boston population (i.e. tau>1.0).[10] The taus indicate that Chronic Wheeze was 2.10 times more likely to be reported in South Boston than in the SMSA, Chronic Phlegm was 1.57 times more likely to be reported in South Boston, and

----

10. As described in Footnote 6, the square of the tau parameters shown in Table 13 are analogous to the ratio of odds shown in Table 1 after controls have been introduced for age, smoking, and sex. Readers should note that in our discussion we continue to refer to ratios of odds ratios; thus the statement that a disorder is twice as likely to be reported in South Boston means that the odds ratio (not the rate) for the reporting of this disorder is twice as high in South Boston as in the SMSA.

Pulmonary Tuberculosis was four times more likely to be reported in South Boston. Finally, Shortness of Breath level 0 (i.e. no symptoms) is "under" reported in South Boston and level 2 Shortness of Breath is more likely to be reported in South Boston than in the SMSA. (For this 3- level variable, we may, for example, consider the frequency of level 2 Shortness of Breath relative to level 0. Level 2 Shortness of Breath appears 2.26 times [i.e. 1.513/0.670] more likely to be reported in South Boston than in the Boston SMSA.)

Appendix Tables C-1 through C-8 present tests of fit for all models constrained to fit the [ArAgSxSm] and [SyAgSxSm] marginals. Hierarchical comparisons indicate that the excess of disorders reported in South Boston, where present, is adequately modeled by inclusion of a constraint on the 2-way marginal [SyAr], with two exceptions. Chronic Phlegm requires that the model be constrained to fit the 3-way marginal: [SyArAg].[11] The nature of these age variations in the inter-area differences in Chronic Phlegm are well suggested by the 3-way crosstabulation shown in Table 7.

The second instance in which a more complex model is required is for Chronic Cough. No model, including a model which constrains all 4-way marginals (No. 160 in Appendix C), provides an adequate fit to the observed data. (Estimates of the saturated model [No. 161] produce estimates of the 5-way interaction parameters as shown in Appendix F. Examination of this table

---

11. L^2 for difference of Models 109 and 110 (in Appendix C) is 9.5 with 3 d.f. which would have $p<.05$ under SRS assumptions.

suggested that the failure to fit the model constraining all 4-way marginals does not arise from one or two "deviant" cells, but rather from discrepancies involving a number of cells. Five of the 20 lambda parameters had estimates which exceeded their standard errors.)

Additional Analyses

In addition to the foregoing analyses, which were proposed in our original analysis plan, we undertook two exploratory analyses in response to suggestions made during discussions of the preliminary results of the survey. We briefly report those analyses here.

Composite of Symptoms. The pulmonary disorders we have been studying do not (always) exist in isolation. Some respondents report only one, others may report more than one, and some combinations of symptoms may be particularly "popular" or significant clinically. The three symptoms of Chronic Cough, Phlegm, and Wheeze, are of particular interest in this regard. Table 14 displays the reporting frequency for all eight possible combinations of these three symptoms.[12] It will be seen from this table that:

----

12. Since the 3 symptom variables can have 2 "states" (i.e., respondent has symptom or does not have symptom), our composite, which represents all possible combinations of the 3 symptom variables, has $2^3=8$ values. These values represent the eight possible combinations, e.g. no symptoms at all, only Chronic Cough, only Chronic Phlegm, Chronic Phlegm plus Chronic Cough, etc. Table 14 lists each combination.

(1) There is apparently considerable non-randomness to the configuring of symptoms, e.g., more than three times as many people report having all three symptoms as report having only Phlegm and Wheeze;

(2) For almost every combination of two or more symptoms, the rate in South Boston exceeds that in the SMSA. For Cough with Phlegm and Cough with Phlegm and Wheeze, the ratios (of odds) both exceed two. (The sole exception to this pattern is Cough with Wheeze; for this combination of symptoms the prevalent rate is 0.5% in both South Boston and the SMSA.) While we did undertake some exploratory analyses with this *ad hoc* composite of pulmonary symptoms, it is clear to us that these data would merit a formal analysis of their latent structure.[13] Constraints of time and resources prevented our pursuing such strategies.

Pulmonary Tuberculosis. As discussed in Section II, we did not intend to conduct an analysis of Pulmonary Tuberculosis rates, but preliminary tabulations prompted us to conduct a more intensive analysis. Ordinarily physicians expect to find increased rates of Tuberculosis among the elderly,[14] and so we began our analysis with the 4-way table of: Tuberculosis by Age by Sex by Area (see Table 15). This table produced an unexpected

----

13. See, for example, L. A. Goodman, Exploratory Latent Structure Analysis using both identifiable and Unidentifiable Models. *Biometrika*, 61, 1974, 215-231.

14. See, for example, G.W. Thorn et al., Harrison's *Principles of Internal Medicine*, eighth edition, New York: McGraw Hill, 1977, pages 901 and 912.

result in that the age distribution of those reporting (physician-diagnosed) Pulmonary Tuberculosis in South Boston differs from that found in the SMSA.

As Table 9 showed, Pulmonary Tuberculosis in South Boston is reported at more than twice its rate in the SMSA. Since none of the 65+ year olds report Tuberculosis in South Boston, it follows that the excess of reported Pulmonary Tuberculosis in South Boston must arise from even greater differences among the younger cohorts. Thus we find in Table 15 that Pulmonary Tuberculosis among persons aged 50 to 64 is reported in South Boston at more than 4 times its rate in the SMSA. Among the youngest group (aged 18 to 34), Tuberculosis is reported in South Boston at over seven times its rate in the SMSA.

When the Pulmonary Tuberculosis data are subjected to the same analysis we performed on the other disorders we obtain the results previously shown in Tables 12 and 13. The simple excess of reported Pulmonary Tuberculosis in South Boston is statistically significant ($L^2 = 13.9$, d.f.=1, $p = .002$ under SRS assumptions), and there are two significant 3-way interactions indicating that the excess of reported Pulmonary Tuberculosis varies by sex and age.[15]

----

15. Using Appendix C, the Test for [SyArSx] parameter is Model 109 - Model 111:
$L^2 = 5.4$, d.f. = 1, $p < .05$, under SRS assumptions.

Test for [SyArAg] parameter is Model 111 - Model 113:
$L^2 = 8.2$, d.f. = 3, $p < .05$, under SRS assumptions.

Table 12

Test for Inter-Area Differences in Prevalence of Pulmonary Disorders: Improvement of Fit Obtained by Adding Symptom by Area Parameter

	Test for Improvement in Model Fit (Base Model minus Test Model)		
Disorder	L^2	df	P_{srs}
Chronic Cough	1.7	1	.192
Chronic Phlegm	14.5	1	<.001
Chronic Wheeze	8.3	1	.004
Asthma	0.9	1	.343
Chest Cold	0.5	1	.480
Severe Chest Cold	0.6	1	.439
Pulmonary T.B.	13.9	1	.002
Shortness of Breath (3 levels)	33.0	2	<.001

P_{srs}: The probabilities shown in this table were computed by treating the observed frequencies as if they had been produced from a simple random sample of the population. (Intra-interviewer response correlations have been similarly ignored.)

Table 13

Estimates of Lambda and Tau Coefficients
for Area by Symptom Parameters
in Log-Linear Model No. 109

Pulmonary Disorder	Estimates of Area Parameter	
	Lambda (λ')	Tau (t')
Dichotomous Disorders[a]		
Chronic Cough	.082	1.085
Chronic Wheeze	.370	1.448
Chronic Phlegm	.224	1.251
Asthma	.074	1.077
Chest Cold	-.030	.970
Severe Chest Cold	-.040	.961
Pulmonary T.B.	.708	2.030
Polytomous Disorders[b]		
Shortness of Breath		
None	-.400[b]	.670
Level 1	+.014[b]	1.014
Level 2	+.414[b]	1.513

[a]Dichotomous Variables: While we have used the algorithm of Goodman and Fay for these calculations, we have not used their convention of representing effects involving dichotomous variables as the difference between a specific category and the average of the two categories. Rather effects shown contrast South Boston to the SMSA; lambda >0.0 and tau >1.0 indicate an excess of reported pulmonary disorders in South Boston. (The values shown above λ', t' are defined as: $\lambda' = 2\lambda$, and $t' = t^2$, where t and λ are the default specification used by the Goodman and Fay program.)

[b]Polytomous Variables: As above, we continue the practice of contrasting South Boston to the SMSA, but for the 3-category symptom variable, $\lambda > 0.0$ and $t > 1.0$ indicate an excess of the specified category in South Boston.

Table 14

Estimates of Prevalence for Composite of Symptoms: Chronic Cough, Chronic Phlegm, Chronic Wheeze

Symptoms Reported	So. Boston	SMSA
None	76.2%	83.0%
Cough alone	5.7	6.1
Phlegm alone	7.4	5.6
Wheeze alone	1.0	0.5
Cough plus Phlegm	6.9	3.5
Cough plus Wheeze	0.5	0.5
Phlegm plus Wheeze	0.4	0.2
All three (Cough + Phlegm + Wheeze)	1.7	0.6

Table 15

Estimates of Prevalence Rates for (Doctor Diagnosed) Pulmonary Tuberculosis in South Boston and the Boston SMSA by Age and Sex

		Percent Reporting Pulmonary TB	
Age Group	Sex	So. Boston	SMSA
18-34	Male	2.5% (290)	0.0% (301)
	Female	0.6 (338)	0.4 (331)
	Both	1.5 (628)	0.2 (631)
35-49	Male	0.7 (137)	0.0 (197)
	Female	1.6 (156)	1.2 (205)
	Both	1.2 (293)	0.6 (402)
50-64	Male	6.3 (130)	0.0 (139)
	Female	3.0 (170)	1.7 (139)
	Both	4.4 (300)	0.9 (278)
65+	Male	0.0 (58)	1.2 (79)
	Female	0.0 (130)	1.6 (122)
	Both	0.0 (188)	1.4 (201)

Note: Respondents were counted as having had Pulmonary Tuberculosis only if they reported having had Pulmonary TB and if they also replied "yes" when asked if a doctor had told them so. (Relevant survey questions are B1.e and B3.e.)

Base Ns for each rate are shown in parentheses.

VI. EVALUATION OF IMPACT OF SURVEY METHODOLOGY ON ESTIMATES OF PREVALENCE OF PULMONARY DISORDERS

VI. EVALUATION OF IMPACT OF SURVEY METHODOLOGY ON ESTIMATES OF PREVALENCE OF PULMONARY DISORDERS

In this section we report on the impact of sample design and survey execution factors[1] upon the variance of estimates derived from the survey data. We discuss here the assessment of variance components due to sample clustering, interviewer contributions to the variance, and the effects of sampling frame bias and non-response.

These variance components in survey estimates are typically referred to as "design effects". Design effects have two major sources:

1. Multistage selection and sample clustering. For reasons of cost efficiency, self-weighting simple random samples are not ordinarily used in large-scale surveys; rather, designs incorporating a multi-stage selection process and one or more types of sample clustering are employed. In telephone surveys a limited number of telephone exchanges are selected at the first stage of sampling to decrease the proportion of non-working and non-residential phone numbers which have to be dialed.

----

1. Including variation in sample weights.

(As was done with the SMSA[2] sample in this survey.) While these design considerations introduce considerable cost savings they do typically inflate the variance of estimates (vs. comparable estimates derived from an SRS design of equal size).

2. Interviewer effects. Individual interviewers affect respondents in systematic ways. Since all interviewers have some idiosyncratic characteristics, it is inevitable that the survey responses obtained by any single interviewer will look more like each other than a random sample of responses obtained by all other interviewers. This intra-interviewer response correlation contributes a component to the total variance of the survey estimates (over and above that due to sample design factors).

In addition to these variance components, non-response in the survey introduces the possibility of other systematic errors in survey estimates. Where identifiable population segments are underrepresented, ratio adjustment procedures are sometimes employed to produce "adjusted" estimates. These adjustments are made by post-stratification weighting to bring sample estimates

----

2. Throughout this report, we use "Boston metropolitan area" and Boston SMSA to denote the federal government's Standard Metropolitan Statistical Area for Boston; see Section II for precise definitions. Where comparisons are reported between South Boston and the Boston SMSA, our tabulatins contrast the SMSA excluding South Boston to South Boston.

The term pulmonary disorder is used to denote a class of conditions that includes both pulmonay diseases and pulmonary symptoms.

of important population characteristics (e.g., age, sex, race) into accord with presumed "true" counts such as Census estimates. (Although these adjustment procedures have several desirable characteristics, they are not without problems -- as will be described later).

Analysis

As noted above, the use of a complex sample design affects all variance computations.[3] In particular, the convenient algorithms for computing variances for simple random samples (SRS) are not appropriate. Since variances and their roots figure crucially in all inferential statistics, it is a necessary step to assay the effect of our survey design (and execution) upon the variance estimates to be derived from our survey data. Similarly, interviewers also contribute a variance component which must be allowed for.[4]

----

3. For rigorous treatments of this matter, see: M.H. Hansen, W.N. Hurwitz, and W.G. Madow, Sample Survey Methods and Theory. Two volumes. New York: Wiley; or L. Kish, Survey Sampling. New York: Wiley. 1965. For an elementary introduction, see H. Blalock, Social Statistics. New York: McGraw Hill, pp. 520-530. For a description of the sampling procedure used in the present survey, see J. Waksberg, Sampling methods for random digit dialing. Journal of the American Statistical Association, 73, 1978, 40-41.

4. See, for example, L. Bailey, et al., An interviewer variance study for the eight impact cities of the National Crime Survey Cities Sample. Journal of the American Statistical Association. 73, 16-23; and Volume 1, Chapter 3 ("Measurement and Error"), in C.F. Turner and E. Martin (Eds.), Surveying Subjective Phenomena. Final Report of the National Research Council-National Academy of Sciences' Panel on Survey Measurement of Subjective Phenomena. New York: Basic Books & Russell Sage Foundation, (in press).

This is of great importance at a practical level since most common statistical packages (e.g., SPSS, SAS, BMD) compute variances under the assumption that the data derive from SRS samples, and that interviewer effects can be ignored (or modelled as a simple bias factor rather than as a contribution to the total variance of the estimates). When this is done, serious errors of inference can be made.[5] Use of more appropriate algorithms for the variance computations, however, can be extremely expensive since the iterative calculations they use greatly increase the computer resources which must be expended to produce a given estimate. (Jackknife estimation procedures, for example, typically recompute estimates of all statistics n-1 times, where n for a simple cluster sample would equal the number of primary sampling units. In a sample of 1,500 with an average cluster size of 10, n would equal 149!)

Our strategy for dealing with this question was to begin by assessing the intra-cluster and intra-interviewer response

----

5. This is so because use of SRS formulae to calculate the standard errors from commonly-used cluster designs may considerably understate true standard errors for means, regression coefficients, etc.

   Statistical research on health surveys conducted for the National Center for Health Statistics indicates that for some variables that are geographically clustered (such as race, due to residential segregation) true standard errors (for such things as the unstandardized partial regression coefficient representing the "effect" of race on number of decayed teeth, net of age, sex, and consumption of sugar) may be underestimated by a factor of two or more if SRS formulae are used in one's computations. (See J.R. Landis, et al., A statistical methodology for analyzing data from a complex survey. Vital and Health Statistics Series 2, no. 92, September, 1982.)

correlation to derive estimates of the contribution of these factors to the total variance of our estimates of statistics for univariate distributions. Such design effects (deff) are commonly expressed (see Kish, 1965, op. cit., p. 258ff) as the ratio of the actual variance of a sample estimate to the variance of an SRS sample of equal size. This ratio can play a useful role in "deflating" statistics derived under SRS assumptions. For example, in the case of the t-statistic:

$$deft = (deff)^{0.5}$$

$$t_{true} = t_{srs}/deft$$

Note, however, that a low value of deft for the variance of a univariate mean does not necessarily translate into a similar finding for bivariate and multivariate statistics. For this reason, relatively large design effects (i.e., deft >1.1) for crucial variables would lead to a questioning of the assumption that the multivariate design effects can be well accounted for by any simple model. (This led ultimately to our consideration of the need for providing jackknifed model estimates for the key multivariate analyses.)

Calculations. Variance decompositions were performed to estimate the effects of sample clustering and intra-interviewer response correlation upon estimates of the variance for means of all variables. For non-metric variables other than dichotomies, this analysis was repeated for each of n-1 of the n categories of the variable. Table 16 presents results of the variance decomposition[6] for our key symptom variables. Appendix D presents the results for other variables measured in the survey. (Note that sample clustering effects are only relevant in the SMSA sample, the South Boston sample was drawn as a simple random sample of households.)[7]

Table 16 shows that there are several noteworthy interviewer effects. In particular, the symptom "Chronic Phlegm" evidences an overall design effect for the total sample of 1.432 (indicating that the interviewer variance inflates the standard error of our estimates by over 40 percent). These results are not entirely unexpected. "Chronic Phlegm," involves a survey question which interviewers reported as problematic. The concept of "phlegm" did not appear to be well understood by respondents, and interviewers were repeatedly asked by respondents to "explain what they meant."[8] While the interviewers were instructed not to

---

6. Computation. The requisite information for the computation of the deff ratio is provided by the partitioned sums of squared deviations within and between sample elements (or interviewers in the case of estimating the interviewer variance). In these calculations, interviews obtained after an initial refusal were eliminated (because these estimates were not randomly re-assigned, but rather were given to specially-trained "refusal convertors").

7. The interaction term in the variance analysis (i.e., interviewer by sample element) is theoretically of considerable interest, but it cannot be estimated in the present case since with almost 200 sample elements and 50 interviewers in each sample, there will be a very large number of cells with Ns of 0 or 1 (in the 200 by 50 matrix of interviewer by sample element.)

8. Hindsight and the experience of others who have used similar measurements leads us to regret our decision not to code into the database an indication of whether the respondent asked for clarification of the meaning of this question. Clinical experts suggest that those respondents who need a definition probably don't have chronic phlegm (i.e., if they have it they know it; if they need an explanation, they don't have it). Unfortunately, we do not have the necessary data to try the recoding scheme such an axiom suggests.

supply a definition but rather to use one of the standard responses (e.g., what it means to you), the high incidence of definitional problems provides the opportunity for non-standard behavior on the part of the interviewers. The symptom of "Chronic Cough" in the total sample and Chest Illnesses in South Boston also show noteworthy interviewer effects. Wheeze and Asthma, in contrast, show no major interviewer effects.

Table 17 presents estimates of the interviewer effects for key demographic variables. In interpreting these results it is important that readers keep in mind the fact that interviewer effects occur not only because interviewers affect the way their respondents answer the survey questions, but also because interviewers vary in their ability to obtain interviews with different segments of the population. (For example, elderly women may be more likely to consent to a telephone interview by a female interviewer than by a young male interviewer.)[9]

While the interviewer effects shown in Table 17 summarize both differential "responding" and differential "refusal" rates across levels of the variable being measured, there is good reason to suspect that the latter component is dominant. Table 17 indicates that two key demographic variables, sex and age, show interviewer effects (deft) of 1.37 or larger when the total sample is considered.

----

9. In addition, since interviewers (despite attempts to manage it otherwise) were not perfectly randomized across time periods, some effects of deviations from complete randomization will be confounded in our estimates with the "true" interviewer effects.

In designing the survey we had expected sample clustering effects to be considerably smaller than the interviewer effects for our key measurements of pulmonary disorders. This expectation followed both from the fact that few of the pulmonary disorders we measured were contagious, and from the fact that major determinants of pulmonary functioning (age and smoking) are only moderately clustered geographically in the Boston SMSA. (Remember also that our "clusters" in the SMSA were 4-digit telephone exchanges which should usually yield a more heterogeneous population than, say, blocks or ennumeration districts.)

Our computations (see Table 18) bear out these expectations; no pulmonary disorder variable had a sample design effect greater than 1.144 (deft). For our key control variables, sample design effects were: 1.105 for sex, 1.209 for age, and 1.059 to 1.097 for the various categories of cigarette smoking.

Taken alone, the sample design effects in the SMSA are sufficiently small that one might choose to assume that a simple strategy (e.g., deflation of sample size by deff) would provide a reasonable protection in hypothesis-testing using the multivariate distributions. This, however, does not hold true for the interviewer effects, which are sufficiently large for Chronic Cough and Chronic Phlegm that one must be suspicious of any assumption that is made about the effects of the interviewers upon the multivariate distributions of our key variables.

Appropriate adjustments for such interviewer effects will therefore be presented in Section VII.[10]

Impact of Sampling Frame Bias and Non-Response Upon Estimates

In addition to the above issues, two other sampling concerns need to be explicitly addressed. These are: (1) the effects of use of a telephone sample, and (2) the effects of nonresponse. Both of these factors contribute (potential) biases when survey data are used to make inferences about the adult population of South Boston and the Boston SMSA.

The choice of a telephone sampling frame results in non-trivial exclusions of households from the universe sampled in our survey, particularly in South Boston. The 1980 Census of the

----

10. The impact of weighting factors on sampling variances also needs to be taken into account. Weighting affects variance computations in that stratum specific variances must be computed, weighted and summed to estimate the total variance (strata here are unique values of the weight variable). Since most statistical procedures assume that one's data are derived from a self-weighting sample (an assumption which usually is not true, and which certainly is not in this case) the variances estimated by such programs as SPSS, SAS, etc. are subject to some potential bias.
While this can be a severe problem for highly stratified samples (where considerable variation in sample selection probabilities is built into the design), it is considerably less threatening in the present case. This is so because there is a relatively low level of variation possible in the sample selection probabilities. The overwhelming majority of households have a single telephone line and either 1 or 2 eligible adults. The other potential strata (e.g., 1 adult households with 4 phone lines; 10 adult households with one phone line, etc.) are usually empty or very sparsely populated.

The effects of weighting will be taken into account susbsequently when we recompute our basic 5-way model using jacknife estimation procedures.

Population found that 14% (1,749 households) of the 12,616 households in South Boston did not have telephones.[11] This unusually high level of non-ownership of telephones means that an unusually large proportion of South Boston households were excluded from the sampling frame for this study. (For Massachusetts as whole, the comparable percentage is 4.3%.)

Clearly, people who do not have a telephone were missed in our survey. Similarly, those who refused to be interviewed are not reflected in the survey's estimates. Since our non-interview ratė was 30.percent (not an atypical rate for telephone surveys in urban areas), the potential for serious deviations from "representativeness" is substantial. In Table 19, we present a comparison of the distribution of four important variables (Sex, Race, Age, and Education) in the 1980 Census and the corresponding estimate derived from the Boston Area Pulmonary Health Survey. This comparison indicates that the sex and race distributions obtained in our survey closely match those found in the 1980 Census of the Population.[12] Age and education, however, are not as well represented. The combined impact of the telephone sampling frame and non-response produced a survey

----

11. Source: 1980 Census of Population and Housing, (U.S. Department of Commerce: Bureau of the Census), Summary Tape File STF3.

12. CSR estimates (throughout this report) and Census tabulations shown in Table 14 refer to the population aged 18 and over not the entire population of the SMSA. For education, Census data are presently available only for the population aged 25 and over; CSR estimates for this variable were similarly restricted to this population for this analysis.

sample which was somewhat younger than the population (37.0 percent of the adult population of the Boston SMSA was aged 50+ in 1980 compared to 31.7 percent estimated by our survey). This deviation is sufficiently large that it is unlikely to have been due to sampling fluctuations. Since age is a control variable in all of our formal analyses, this minor deviation of our survey sample from the population distribution is not of great concern.[13]

Education exhibits a more marked deviation, particularly in the reporting of the least-educated and college-educated groups. The population of persons with 0-8 years of schooling is over 3 times the size of the estimate generated from the survey (i.e., 10.8 vs. 3.1 percent). In turn the survey shows a corresponding excess of persons reporting 1 to 4 years of college (45.6 percent in survey versus 28.3 percent in Census). This large deviation is of concern, since it suggests that the survey estimates are based on a population which substantially over-represents the more-educated segment of the Boston population.[14]

----

13. Any effects of age upon comparisons between control and target populations would be controlled by the fact that our all of our comparisons would be taking place within age categories.

14. It should be emphasized that these comparisons apply to the entire SMSA; effects specific to South Boston's demographic characteristics will be affected by the larger number of households without telephones in that community.
In addition we should note that bias in the education distribution is not peculiar to this survey. A review of national surveys conducted by a wide range of survey organizations suggests that such biases are not infrequent and are typically in the direction found here, i.e., an over-representation of the college-educated population. (See Turner and Martin, op cit., Volume 1, chapter 3.)

One must allow that this apparent bias in the educational level of the sample (1) may be a consequence of use of a telephone sample, and/or (2) may reflect a propensity of respondents to overreport their educational attainment. Some data on the first proposition is available from in-person interview surveys that have asked respondents about telephone ownership. Table 20a presents estimates of the proportion of the national population owning telephones by educational level, together with similar estimates for the Northeastern United States. (These data are from the General Social Survey of the National Opinion Research Center, University of Chicago.) It can be seen from these figures that an education bias is built into survey estimates when the telephone is used as the (sole) method of interviewing. While this is a partial explanation of our results, the raw figures can be somewhat misleading. Table 20b presents estimates of the educational distribution using the total population and the telephone-owning segment of the population. It will be seen from this tabulation that, both in the nation as a whole and in the Northeast, the effect of excluding non-telephone owners is relatively minor. This is so because they are a relatively small proportion of the total population.

These analyses lead us, by default, to the conclusion that the deviations in the education distribution are due to either response distortions or to differences in the willingness of different segments of the population to be interviewed.

Ratio-adjustment. While there is no way to "represent" elements of the population who are excluded from the sampling frame or who systematically refuse to be interviewed, various supplementary analyses can provide some insight into the potential impact of these inevitable limitations of the (telephone) survey method. As a general caveat, it is probably useful to always bear in mind that the only truly safe statistical generalization that can be made from a telephone survey sample with a typical 30+% non-response rate is: "The sample estimates are representative (within the calculable margins of error) of that 70% of the telephone-owning adult population which is willing to submit to a survey interview." Inferences about non-telephone owners and the percent of the population refusing or unable to be interviewed involve some (often considerable) element of conjecture.

One technique which does assist in the process is post-stratification weighting. This technique performs a ratio adjustment of sample estimates to bring specified parameters of the survey sample into agreement with (presumably) true population values for these same parameters. One commonly applied approach involves ratio adjustments to bring "important" population variables, e.g., age, sex, and race into agreement with Census figures. A vector or n-dimensional matrix of ratio-adjustment weights (e.g. for the univariate education distribution or the bivariate education-by-age distribution) can be constructed by dividing the presumed "true values" obtained in the Census by the proportions obtained in a given survey.[15]

----

15. Census estimates might also be adjusted (prior to their use in the post-stratification weighting) to reflect temporal changes in the population since the last Census. Since the

While such procedures have many desirable properties, one can never be certain that the procedure is really "correcting" the values of the relevant variables. For example, if, as one suspects, "alienation" and "distrust" play an important role in a person's refusal to be interviewed, it is not evident that merely matching the Census' sex, age, and race distributions will necessarily result in a better characterization of the population. (Such adjustments are somewhat akin to treating "refusers" as if they were the same as the average person from the same population segment who gave an interview. This may sometimes be a quite erroneous assumption.)

Despite such theoretical problems, it is useful to assay the potential impact of non-coverage and non-response through the use of such a technique, since it does provide at least one way of estimating what the survey results might have been if the survey frame had included the entire universe and if every eligible respondent had consented to be interviewed.

Calculations. The two sets of adjustment factors we used were calculated as the ratio of the proportions of sample and Census which fell into each category of the age and education distribution shown in Table 19, e.g., since the Census found that .165 of the adult population of the SMSA were aged 65+ and the survey found .133, the adjustment factor (applied as a post-stratification weighting factor) for this group would be .165/.133. More generally, adjustment ratios (a) can be directly computed as:

---

15. Census data is only three years old and since the education distribution changes mainly through cohort replacement, there is little reason to consider such a fine point in our present explorations.

$$a_{ij...n} = (p_{ij...n}/p'_{ij...n})$$

where p is the proportion of the total population in cell i,j...n, (e.g., the proportion who are white males aged 65+) using the Census figures, and $p'_{ij...n}$ is the survey estimate of the same proportion.

Impact of Ratio Adjustment on Symptom rates. Table 21 shows the effect of ratio-adjustments for age and education upon our estimates of the prevalence of each of the key pulmonary symptoms. The age adjustment has an extremely modest impact upon our estimates. The education adjustment has a larger impact upon our estimates, but it, too, is not overwhelming. The rate of reporting of Chronic Cough and Chronic Phlegm both increase by 0.9 percentage points (from 10.6% to 11.5% and from 10.1 to 11.0%), Shortness of Breath Level 1 increases from 13.5 to 16.8 percent, Shortness of Breath Level 2 increases from 4.8 to 6.6 percent. Chronic Wheeze and Asthma show smaller changes, and the two Chest Cold measurements decrease when the ratio adjustment for education is performed.

While these variations are noteworthy, and our analysis indicates the difficulties encountered in generalizing from a telephone survey to the total population, it was our judgment that incorporation of ratio-adjustments in our analyses would not be warranted. This decision was based upon the size of the effects we have seen, the previously noted theoretical problems with such adjustments, and the time and resource constraints under which this analysis was conducted.

We will therefore proceed (as noted in the introduction to this report) with the qualification that all of our estimates must be understood to refer to (noninstitutionalized) persons residing in telephone-owning households. Moreover, inferences can only be made about that 70 percent of such households that consent to be interviewed in telephone surveys such as ours. Any inferences readers make about households without telephones, or about that 30 percent of the population who do not participate in telephone interviews are subject to unknown (and potentially tantial) error.

Table 16

Analysis of Interviewer Variance in Univariate Distributions of Key Symptom Variables

Variable	R_I	F	p	deft
Total Sample				
Chronic Cough	0.009	1.562	0.01	1.244
Chronic Phlegm	0.018	2.087	<0.01	1.432
Chronic Wheeze	0.002	1.091	0.32	1.044
Recent Chest Illness	0.004	1.259	0.12	1.120
Asthma	0.004	1.235	0.14	1.109
Severe Chest Illness	0.004	1.254	0.12	1.117
Shortness of Breath	0.001	1.059	0.37	1.029
SMSA Sample				
Chronic Cough	0.004	1.136	0.25	1.064
Chronic Phlegm	0.019	1.611	0.01	1.257
Chronic Wheeze	-0.002	0.926	0.61	0.963
Recent Chest Illness	0.010	1.305	0.09	1.137
Asthma	0.002	1.057	0.37	1.027
Severe Chest Illness	0.008	1.246	0.13	1.112
Shortness of Breath	0.013	1.428	0.04	1.187
South Boston Sample				
Chronic Cough	0.019	1.543	0.01	1.230
Chronic Phlegm	0.033	1.955	<0.01	1.375
Chronic Wheeze	0.005	1.131	0.26	1.061
Recent Chest Illness	0.021	1.598	0.01	1.251
Asthma	-0.001	0.968	0.53	0.984
Severe Chest Illness	0.019	1.536	0.01	1.228
Shortness of Breath	0.005	1.128	0.26	1.060

R_I: intraclass (i.e. intra-interviewer) response correlation

deft: square root of design effect (deff) coefficient. (Deff is ratio of total variance of estimate to the variance of an SRS sample of equal sample size).

F ratio with df = $(N_I-1)/(N_r-1)$. N_I is number of interviewers and N_r number of respondents answering item. They were as follows:

	N_I	N_r
SMSA	44	1386 to 1390
S.Bost.	45	1262 to 1275

In these computations, interviewers completing fewer than 8 interviews were excluded from analysis.

Table 17
Analysis of Interviewer Variance in Univariate Distributions of Key Demographic and Risk Variables

Variable	R_1	F	p	deft
<u>Total Sample</u>				
Sex	0.015	1.928	<0.01	1.378
Education	0.010	1.588	0.01	1.254
Age	0.017	1.975	<0.01	1.394
Income	0.012	1.642	0.01	1.274
Smoking				
1. Never Smoked	-0.005	0.714	0.92	0.847
2. Ex-Smoker	-0.003	0.836	0.77	0.916
3. 0-15 Daily	0.003	1.206	0.17	1.096
4. 16-30 Daily	0.016	1.988	<0.01	1.398
5. 31+ Daily	0.006	1.357	0.06	1.161
<u>SMSA Sample</u>				
Sex	0.020	1.644	0.01	1.269
Education	0.002	1.069	0.35	1.033
Age	0.018	1.579	0.01	1.245
Income	0.010	1.292	0.02	1.131
Smoking				
1. Never Smoked	-0.002	0.944	0.58	0.972
2. Ex-Smoker	-0.004	0.869	0.71	0.934
3. 0-15 Daily	-0.003	0.890	0.68	0.945
4. 16-30 Daily	0.028	1.909	0.01	1.362
5. 31 + Daily	0.037	2.222	0.01	1.463
<u>South Boston Sample</u>				
Sex	0.010	1.300	0.09	1.134
Education	0.016	1.456	0.03	1.197
Age	0.028	1.798	0.01	1.322
Income	0.028	1.717	<0.01	1.292
Smoking				
1. Never Smoked	-0.007	0.809	0.81	0.902
2. Ex-Smoker	-0.003	0.907	0.65	0.954
3. 0-15 Daily	0.024	1.699	<0.01	1.288
4. 16-30 Daily	0.013	1.360	0.06	1.159
5. 31 + Daily	0.014	1.395	0.05	1.173

R_1: intraclass (i.e. intra-interviewer) response correlation

deft: square root of design effect (deff) coefficient. (Deff is ratio of total variance of estimate to the variance of an SRS sample of equal sample size.

F ratio with df = $(N_1-1)/(N_r-1)$. N_1 is number of interviewers and N_r number of respondents answering item. They were as follows:

	N_1	N_r
Total	45	2653 to 2665
SMSA	44	1386 to 1390
S.Bost.	45	1262 to 1275

Except that N_r was 2383, 1252, and 1131 respectively for Income. In these computations, interviewers completing fewer than 8 interviews were excluded from analysis.

Table 18

Analysis of Sample Design Effects Upon Univariate Distribution of Key Symptom and Demographic Variables

Variable	R_i	F	p	deft
Symptoms				
Chronic Cough	0.029	1.140	0.07	1.052
Chronic Phlegm	0.046	1.222	0.01	1.080
Chronic Wheeze	-0.008	0.965	0.64	0.986
Recent Chest Illness	0.075	1.372	<0.01	1.127
Asthma	0.052	1.251	0.01	1.089
Severe Chest Illness	0.062	1.304	<0.01	1.106
Shortness of Breath	0.086	1.431	<0.01	1.144
Demographics and Risk Factors				
Sex	0.061	1.299	<0.01	1.105
Education	0.141	1.757	<0.01	1.228
Age	0.131	1.680	<0.01	1.209
Income	0.119	1.563	<0.01	1.174
Smoking				
1. Never Smoked	0.046	1.220	0.01	1.079
2. Ex-Smoker	0.056	1.273	<0.01	1.096
3. 0-15 Daily	0.034	1.161	0.05	1.059
4. 16-30 Daily	0.041	1.196	0.02	1.071
5. 31 + Daily	0.056	1.276	<0.01	1.097

Note. Sample clustering was not used in the South Boston sample.

R_i: intraclass (i.e. intra-cluster) response correlation

deft: square root of design effect (deff) coefficient. (Deff is ratio of total variance of estimate to the variance of an SRS sample of equal sample size.)

F ratio with df = $302/(N_r-1)$. N_r number of respondents answering item. In the SMSA this varied from 1394 to 1398 for the above variables with the exception of education (1303) and income (1260).

Table 19

Comparison of Demographic Characteristics of Boston SMSA in 1980 Census and Estimates Derived from Boston Area Pulmonary Health Survey

Variable	Percent Distribution	
	Census	CSR
Sex		
Male	46.2	46.8
Female	53.8	53.2
Race		
White	92.3	93.2
Black	5.1	4.8
Other	2.5	2.1
Age		
18-34	41.7	41.7
35-49	21.3	26.5
50-64	20.5	18.4
65+	16.5	13.3
Education (persons aged 25+)		
8th or less	10.8	3.1
1-3 High School	11.9	7.5
High School Graduate	36.3	31.0
1-3 College	16.2	23.3
College Graduate	12.1	22.3
Graduate Work	12.6	12.8

Note. CSR estimates for SMSA were derived from weighted survey data. In addition to the weighting used throughout our analyses, population estimates for the whole SMSA require that the South Boston and (rest of) SMSA strata in sample be weighted in proportion to the actual number of households in each area (in order to correct for the deliberate oversampling of South Boston households). This factor is 12,616/(990,779 - 12,616).

Census estimates were derived from U.S. Bureau of the Census, 1980 Census of Population and housing, Characteristics of the Population,

1. General Population Characteristics: Massachusetts, (Part 23: PC80-1-B23).
2. General Social and Economic Characteristics: Massachusetts, (part 23: PC80-1-C23).
3. Census Tracts: Boston, Mass. Standard Metropolitan Statistical Area, (Report PHC 80-2-98).

Table 20a

Percent of Persons Reporting No Telephone
In Residence for Continental U.S.A. and Northeastern States
1972 - 1982

	Percent Without Phone		Base N	
Educational Level	Northeast	Rest of Nation	N.E.	Rest
0-8 Years	8.0	13.6	343	1,498
9-11 Years	7.7	11.0	439	1,728
12 Years	1.9	5.4	978	3,135
13-15 Years	3.0	4.7	445	1,615
16+ Years	0.2	1.4	462	1,201
ALL	3.6	7.2	2,678	9,204

Table 20b

"Expected" Education Distributions for Telephone and Personal Surveys
In continental U.S.A. and Northeastern States
(Derived fron NORC General Social Survey, 1972-82)

	Northeast		Rest of Nation	
Educational Level	Phone	Personal	Phone	Personal
0-8 years	12.3	12.9	86.4	16.3
9-11 years	15.8	16.5	18.1	18.8
12 years	37.3	36.7	34.8	34.2
13-15 years	16.8	16.7	18.1	17.6
16+ years	17.9	17.3	13.9	13.1

Notes. Estimates were derived by aggregating NORC 1972, 1973, 1974, 1975, 1976, 1977, 1978, 1980, 1982 national samples and the special supplemental black sample drawn in 1982 for the General Social Survey. All data from the 1982 Supplemental are weighted to compensate for the oversampling of blacks in that year.

Characteristics of households are denoted by (h) in the row heading. No weighting, other than that noted above, was required for these estimates (the GSS uses a self-weighting sample of households). Population estimates were derived by weighting the sample in proportion to the number of adults (aged 18+) reported to reside in the respondent's household.

Ns for population estimates are for weighted data, but all household weights were divided by the value the mean weight (to provide reasonable Ns). Ns in 1982 were not adjusted upwards to allow for the oversampling of blacks in that year.

Northeastern states include New England and New York, New Jersey and Pennsylvania.

Table 21

Population Estimates for Prevalence of Key Pulmonary Disorders in SMSA (including South Boston) Derived Using Alternative Ratio-Adjustments of Sample Data

	Estimate		
Symptom	Not Adjusted	Adjusted for Educ.	Adjusted for Age
Chronic Cough	10.6%	11.5%	10.7%
Chronic Phlegm	10.1	11.0	10.2
Chronic Wheeze	1.8	2.0	1.9
Shortness of Breath			
Level 1	13.5	16.8	13.8
Level 2	4.8	6.6	5.2
Asthma	6.1	6.3	6.0
Chest Cold	34.4	33.0	33.9
Severe Chest Cold	18.1	16.8	18.0

Note: All SMSA estimates in this table are weighted to compensate for oversampling of South Boston (see note to Table 29) and to adjust for household size and number of telephone lines in household (see note to Table 1).

In addition, ratio-adjusted estimates weight the sample estimates for each education (or age) strata by $P_{i,Census} / P_{i,csr}$ where $P_{i,csr}$ is the CSR estimate of the proportion of the SMSA population falling into education (or age) strata i, and P_i is the corresponding "estimate" from the Census (see Table 29).

VII. RE-ESTIMATION OF LOG-LINEAR MODELS TAKING ACCOUNT OF INTERVIEWER EFFECTS, SAMPLE WEIGHTING, AND SAMPLE CLUSTERING

VII. RE-ESTIMATION OF LOG-LINEAR MODELS TAKING ACOUNT OF INTERVIEWER EFFECTS, SAMPLE WEIGHTING AND SAMPLE CLUSTERING

Our initial fitting of log-linear models for the 5-way tables used one of the most widely available computing algorithms[1] for fitting log-linear models to data obtained from simple random sampling (SRS) designs. In the preceding section of this report, we discussed in some detail the complexities of design and survey execution which make ours a complex sample (by design), and which through the intra-interviewer correlation of survey responses would require (by execution) the introduction of more complex analytic procedures to cope with the non-independence of survey measurements. The magnitude of these effects (particularly those introduced by intra-interviewer response correlations) is sufficient to warrant some more elaborate analyses that take account of these complexities in survey design and execution.

For this purpose we have repeated our crucial analyses using a successor algorithim developed at the Statistical Methods Division of the U.S. Bureau of the Census. This program generalizes the log-linear model to cases involving complex

----

1 ECTA developed by Robert Fay and Leo Goodman at the Department of Statistics, University of Chicago. The version of ECTA used in these analyses was the current version maintained by the Harvard University Computer Center (OIT).

sample designs.[2] Unlike earlier algorithms, this new procedure does not assume that the multinomial (or other simple) distribution underlies the observed frequency distribution. In place of fixed assumptions about the underlying sampling distribution, this new algorithim uses a jackknife procedure[3] performed on independent replicates of the sample to measure the underlying variablility in the sample estimates.

---

2. R.E. Fay CPLX: Contingency Table Analysis for Complex Sample Design. Statistical Methods Division, U.S. Bureau of the Census, 1982. See also, R.E. Fay, Adjusting Pearson chi-square statistic for clustered samples. Proceedings of the American Statistical Association (Social Statistics Section) 1979: 402-406; R.E. Fay, Contingency table analysis for complex sample designs. Proceedings of the American Statistical Association (Survey Methods section), 1982; see also, J.N.K. Rao and A.J. Scott, Analysis of categorical data from complex sample designs: chi-squared tests for goodness of fit and independence in two-way tables. Journal of the American Statistical Association, 76:221-230.

3. For rigorous treatment of the use of jackknife procedures see: R.G. Miller, A trustworthy jackknife. Annals of Mathematical Statistics, 35, 1964, 1594-1605; and R.G. Miller, The jackknife: a review. Biometrika, 61, 1974, 1-15. For a basic source on replication methods, see, B. Efron, The Jackknife, the Bootstrap, and other Resampling Plans. Philadelphia: Society for Industrial and Applied Mathematics. For an introductory overview, see F. Mosteller and J. Tukey, Data Analysis and Regression: a second course in statistics. Reading, Mass.: Addison-Wesley, 1977, chapter 8. For an application in the survey context, see R. Seltzer and C.F. Turner, Estimating the Effects of Interviewer Characteristics on Survey Response. CSR Working Paper Series, November, 1982.

Computational costs increase markedly when the jackknife algorithm is used.[4] For this reason, we initially planned to make use of this algorithm sparingly, particularly if the results of our analyses of the univariate distrbutions showed small effects. Given the results presented in Section VI, it was decided to re-estimate the base and test models using the jackknife algorithm, but we have excluded disorders (Asthma, Chest Colds, Severe Chest Colds, and Chronic Cough) which showed no "significant" improvement in fit (under SRS assumptions) when the Symptom by Area parameter was added to the base model (i.e., No. 108 in Appendix C).

Table 22 shows the results obtained when the jackknife procedure was employed to test for the improvement in fit obtained by the test Model (No. 109). It will be seen that the conclusion that there is a statistically significant excess of Shortness of Breath reported in South Boston[5] holds (tau=1.34 with p<.005 or better). (This is true whether we use all 3 levels of the symptom or restrict ourselves only to the most severe level.) In addition, we still find a statistically

---

4. Parameters for each model must be re-estimated n times, where n is the number of independent replicates. For the interviewer analyses, n is 43, and n is 303 in the sample design analysis. Since n=303 for the sample design analysis is luxuriously large, clusters were collapsed to n=100 in order to facilitate efficient computation.

5. Throughout this report, we use "Boston metropolitan area" and Boston SMSA to denote the federal government's Standard Metropolitan Statistical Area for Boston; see Section II for precise definitions. Where comparisons are reported between South Boston and the Boston SMSA, our tabulatins contrast the SMSA excluding South Boston to South Boston.

   In addition, the term pulmonary disorder is used to denote a class of conditions that includes both pulmonay diseases and pulmonary symptoms.

significant (tau=1.22 with $p<.025$) excess of reports of Chronic Phlegm in South Boston. Finally, Pulmonary Tuberculosis shows a reliable excess of reporting in South Boston (tau=1.7, with $p<.05$).

In contrast, our finding for Chronic Wheeze ($p = .004$ under SRS assumptions) is reduced below conventional levels of significance. This is an interesting result in that it occurs for a variable that shows a relatively small interviewer effect upon its univariate distribution. We should note, however, that Chronic Wheeze did show an unusual pattern of divergence from the results for Asthma. (See discussion in Section IV.)

Conclusion.[6] When interviewer and sample design effects are accounted for, the elevated levels of reported Shortness of Breath, and Chronic Phlegm and Pulmonary Tuberculosis in South Boston remain. This more rigorous analysis does, however, diminish the finding for Chronic Wheeze below the level which would ordinarily be considered statistically reliable.

----

6. This analysis also does permit an important conclusion about the power of our design. The power of the given sample design may be assessed through the estimated standard errors for the lambda parameters corresponding to the area by symptom effect. For example, since the estimated standard errors for the [Wheeze x Area] parameter is approximately .13 (see Table 22b), then for the given model we may infer that the L_j test of this parameter will reject only about 5 percent of the time (at the .05 level) if the true lambda is in fact 0, about 16 percent of the time if the true lambda is .13; 50 percent for a true value of .26; 84 percent for .39, and 97.5 percent for a true .52. [These power comparisons can be made because under the null hypothesis, and for suitable close alternatives, the two tests (i.e. L_j and comparison of the estimated lambda to its s.e.) will be asymptotically equivalent.]

Table 21

Population Estimates for Prevalence of Key Pulmonary Disorders in SMSA (including South Boston) Derived Using Alternative Ratio-Adjustments of Sample Data

Symptom	Estimate: Not Adjusted	Adjusted for Educ.	Adjusted for Age
Chronic Cough	10.6%	11.5%	10.7%
Chronic Phlegm	10.1	11.0	10.2
Chronic Wheeze	1.8	2.0	1.9
Shortness of Breath			
Level 1	13.5	16.8	13.8
Level 2	4.8	6.6	5.2
Asthma	6.1	6.3	6.0
Chest Cold	34.4	33.0	33.9
Severe Chest Cold	18.1	16.8	18.0

Note: All SMSA estimates in this table are weighted to compensate for oversampling of South Boston (see note to Table 29) and to adjust for household size and number of telephone lines in household (see note to Table 1).

In addition, ratio-adjusted estimates weight the sample estimates for each education (or age) strata by $P_{i,Census} / P_{i,csr}$ where $P_{i,csr}$ is the CSR estimate of the proportion of the SMSA population falling into education (or age) strata i, and $P_{i,Census}$ is the corresponding "estimate" from the Census (see Table 29).

Table 22a

Results of Jackknifed Model Fitting
(5-way table: Symptom by Area by Age by Sex by Smoking)

		Interviewers		Sample Design	
Pulmonary Disorder	df	L_J	p	L_J	p
Shortness of Breath					
2 Levels (a)	1	2.99	<.005	2.92	<.005
Chronic Phlegm	1	2.17	<.025	3.23	<.005
Chronic Wheeze	1	0.82	.15<P<.10	1.04	.10 < p <.05
Pulmonary Tuberculosis	1	1.89	<.025	2.07	<.05

L_J: Jackknifed likelihood ratio chi-square statistic testing the improvement in fit obtained by adding a Symptom by Area parameter [SyAr] to Base Model (108); test statistic is computed as jackknifed difference between fit of Base Model (108) and Test Model (109).

(a) 2-level Shortness of Breath variable is coded as 1 if respondent said he/she had to walk more slowly than other people of the same age because of breathlessness; coded zero otherwise.

Table 22b

Estimates of Lambda and Tau Coefficients
for Area x Symptom Parameter in Analyses Accounting
for Complex Sample Design and for Interviewer Effects

Pulmonary Disorder	Analysis[c]	Lambda (λ')	(s.e.)	Tau (t')
Chronic Phlegm	Int.	.202	(.080)	1.22
	Samp.	.202	(.066)	1.22
Chronic Wheeze	Int.	.206	(.130)	1.23
	Samp.	.208	(.128)	1.23
Pulmonary Tuberculosis	Int.	.542	(.212)	1.72
	Samp.	.530	(.242)	1.70
Shortness of Breath	Int.	.292	(.094	1.34
	Samp.	.292	(.098)	1.34

[a]Dichotomous Variables: While we have used Fay's algorithm for these calculations, we have not used his convention of representing effects involving dichotomous variables as the difference between a specific category and the average of the two categories. Rather effects shown contrast South Boston to the SMSA; lambda >0.0 and tau >1.0 indicate that an excess of symptoms are reported in South Boston. (The values shown above λ', t' are defined as: $\lambda' = 2\lambda$, and $t' = t^2$, where t and λ are the default specification used by the Fay's algorithm.)

[b]Polytomous Variables: As above, we continue the practice of contrasting South Boston to the SMSA, but for the 3-category symptom variable, $\lambda > 0.0$ and $t > 1.0$ indicate an excess of the specified category in South Boston.

[c]"Samp." analyses take account of sample weighting and sample design effects; "Int." analyses take account of sample weighting and interviewer effects.

VIII. RE-ESTIMATION OF LOG-LINEAR MODELS USING ADDITIONAL CONTROL VARIABLES

VIII. RE-ESTIMATION OF LOG-LINEAR MODELS USING ADDITIONAL CONTROL VARIABLES

The survey questionnaire contained questions designed to measure a variety of factors which are thought to affect pulmonary health. We have already in our analysis made use of two variables thought to have the strongest causal association with pulmonary health: cigarette smoking and age. However, it is not only the elderly and the smokers who experience elevated risks of pulmonary disorders; those individuals whose day-to-day life exposes their lungs to acute or chronic insult also may be at greater risk. Even factors which may seem innocuous to most people, such as exposure to a gas cooking stove, have been found to contribute to pulmonary disorders (assumedly via their contribution to "indoor pollution").[1]

In this section we report the results of statistical analyses intended to control for a variety of suspected risk factors. While constraints of time and resources precluded our analysis of all the possible candidates for inclusion as "control variables," we have included those factors which consulting medical experts thought to be particularly important.

Initial Analytic Refinements. We begin this discussion with two preliminary analyses designed to explore potential artifacts in our finding of elevated levels of (some) pulmonary disorders in

---

1. See for example: F. Speizer et al. Respiratory disease rates and pulmonary function in children associated with NO_2 exposure. American Review of Respiratory Disease, 1980, 121, 3-10.

South Boston. One line of explanation that might be proffered is that the excess reporting of pulmonary disorders in South Boston may reflect their socioeconomic disadvantage, in one or more ways. Some people, for example, might argue that South Boston residents evidence more illness (pulmonary and perhaps others) as a result of the greater incidence of poverty in that commmunity (compared to the average for the Boston area). Other people might argue that persons of different social classes may have different propensities to report their symptoms, and thus "apparent" differences between reported symptom rates in South Boston and the SMSA[2] might be artifactual.[3]

----

2. Throughout this report, we use "Boston metropolitan area" and Boston SMSA to denote the federal government's Standard Metropolitan Statistical Area for Boston; see Section II for precise definitions. Where comparisons are reported between South Boston and the Boston SMSA, our tabulations contrast the SMSA excluding South Boston to South Boston.

   In addition, the term pulmonary disorder is used to denote a class of conditions that includes both pulmonary diseases and pulmonary symptoms.

3. There are inferential problems in testing any of these hypotheses with the present data. Were we to find that the excess of symptoms in South Boston was a function of poverty, it would raise some thorny questions. For example, such a finding would prompt one to ask whether the poor have more symptoms because they live in more polluted neighborhoods.

   In the present case, the results of our analyses allow us to forego jousting with such questions, since South Boston residents report more pulmonary disorders even when social class is controlled.

It is true, as Table 23 documents, that a number of pulmonary symptoms are related to educational level.[4] Shortness of Breath for example, shows a strong association with education both in the South Boston and the SMSA samples. Over 40 percent of those with 0-11 years of education in both samples report Shortness of Breath (Level 1 or 2); this figure drops to less than 20 percent among those with some college education.

While this result is provocative, it is should be remembered that education is, in part, serving as a proxy for age and smoking in this 2-way table. (That is to say, the association of reported symptoms with education partly reflects the excess of smokers and old people at the lower educational levels.)

To pursue the analysis more rigorously, we expanded the analysis strategy we employed in Section V to include Education as a further control variable. Analysis of the resulting 6-way table (Symptom by Area by Sex by Age by Smoking by Education) was carried out in the same fashion as the analyses reported in Section V, but we restricted ourselves to pulmonary disorders which we had found to be significantly in excess in South Boston (i.e. Chronic Phlegm, Wheeze, Shortness of Breath, and Pulmonary

----

4. We have chosen to use educational level as our indicator of socioeconomic status. Of the three widely-used indicators (income, education, and occupation), it is the best measured (i.e., highest measurement reliability and least problematic in terms of respondents willingness to supply an answer). It also does not suffer from difficulties of "relevance," such as those that bedevil attempts to classify the occupations of persons who have little attachment to the conventional labor market, e.g., housewives.

Tuberculosis) and the one symptom (Chronic Cough) for which we found a significant 5-way interaction.

The results of this analysis are shown in Table 24a. It will be seen from this analysis that South Boston continues to have an excess of Chronic Phlegm, Shortness of Breath, and Pulmonary Tuberculosis even when education is introduced as a 6th control variable. Chronic Wheeze, however, ceases to show a significant difference when education is controlled.

A second "artifactual" explanation which might be advanced involves the findings for Shortness of Breath. Some respondents who report breathlessness may be reporting a symptom which is a consequence of heart disease rather than a pulmonary disorder. To test for such an artifact, we constructed a 6-way table for: Shortness of Breath by Area by Sex by Age by Smoking by Heart Disease. (The latter variable was measured by Question B7 in the survey: "Has a doctor ever told you that you had heart trouble?") The results of this analysis are shown in Table 24b; it will be seen from these results that one still must conclude that there is a statistically significant excess of Shortness of Breath reported in South Boston even when the presence of heart trouble is controlled.

Environmental Risk Factors. We subsequently expanded our analysis to include a number of other factors which were thought to pose a potential risk and which varied substantially between the South Boston community and the SMSA. Tables 25 to 27 present simple 3-way tables showing the association of reported pulmonary disorder in each sample to: working in a dusty job for one or more years (Table 25), use of a gas kitchen stove to heat home (Table 26), and type of heating system (Table 27). It will be seen from these tables that there are strong associations between some reported pulmonary disorders and some of these factors.

Working in a "dusty" job, for example, shows a "significant" association with Chronic Cough, Phlegm, Wheeze and Shortness of Breath in both South Boston and the Boston SMSA. Use of a gas stove for heating shows a substantial association with Shortness of Breath and a lesser association with Chronic Phlegm in both samples; it also has a relatively strong association with Chronic Cough for the South Boston sample (but not the SMSA sample). A more mixed pattern of associations is found for type of heating system.

To analyze the effect of each of these variables, we initially analysed 6-way tables (as described in the preceding section) controlling for each risk factor in succession. This, however, does not permit a comprehensive statement to be made about inter-area differences when all of these factors are simultaneously controlled (although it does relieve us of the burden of constructing and analysing some very large and sparsely-populated 8-way tables.)

Fears over the sparseness of the 8-way tables are alleviated somewhat, given our limited purpose of testing the effect of the addition of a single [SyAr] parameter. Some relief is obtained by invocation of a theorem stated by Haberman and reemphasized by Fienberg. When comparing two models which differ by only a single parameter, "differences in likelihood-ratio chi-square statistics have approximate chi-square distributions under quite general conditions"[5] and the conditions include those of "large sparse multinomial structures."[6]

We have chosen therefore to spare readers discussions of the series of 6-way tables, and to proceed directly to our final analyses which estimated models for the 8-way table (Symptom by Area by Sex by Age by Smoking by "Dusty Job" by Use of Gas Kitchen Stove for Heat by Type of Heating System (Hot Water-Steam, Hot Air, Electric, or other). Table 28 presents the 1 degree of freedom tests for the [SyAr] parameter. The results we obtained were consistent for all four disorders: a statistically significant excess of reports of Shortness of Breath (Level 2), Chronic Phlegm, Chronic Wheeze, and Pulmonary Tuberculosis are found in South Boston.

Additional Considerations and Qualifications

It is, of course, possible to introduce further controls into our analysis, and it is not difficult to suggest likely candidates, such as passive exposure to cigarette smoke. One might also try (at a substantial computational cost) to obtain

---

5. S. J. Haberman, Analysis of Qualitative Data. Volume 2, New York: Academic Press, p. 421.

   A proof can be found in S. J. Haberman, Analysis of Frequency Data. Chicago: University of Chicago Press, 1974, pp. 372-373.

   I am greatful to Otis Dudley Duncan and Robert Fay for teaching this to me twice. I have here borrowed Duncan's felicitous phrasing. (c.f.t.)

6. S. E. Fienberg, The Analysis of Cross-Classified Categorical Data. Cambridge, Mass.: MIT Press, 1979, p. 63.

jackknifed estimates for the model paramaters of the 8-way table (taking account of interviewer effects, etc.)

That however is an analytic luxury that we are denied by constraints of time and resources. We clearly cannot say what the results would be if such additional analyses were performed. But within the limits of the variables we have considered, the finding that South Boston residents do report some pulmonary disorders at a higher rate than residents of the Boston metropolitan area cannot be escaped.

Bias. It has been suggested that the results we have obtained might reflect a reporting bias (rather than a true difference in pulmonary health). This is certainly a possibility, but we have no scientific basis for making such an inference. Indeed, it should be remembered that it is also possible that reporting biases worked to make South Boston residents appear relatively healthier than they otherwise would have appeared if measurements (in both areas) were made without any reporting bias. (Since the suggestions of reporting bias doubtlessly will arise, we have summarized our position on this important question in Appendix I.)

Table 23

Estimates of Prevalence of Self-Reported Pulmonary Disorder
By Educational Level for South Boston and Rest of Boston SMSA

		Educational Level				
Disorder	Population	0-11	12	13+	N	X^2
Chronic Cough	So. Boston	21.8%	12.98%	12.7%	1432	16.52
	SMSA	16.0	12.6	8.7	1539	10.34
Chronic Phlegm	So. Boston	20.1	14.4	16.5	1433	5.13
	SMSA	14.0	12.2	8.1	1536	9.15
Chronic Wheeze	So. Boston	5.3	3.3	3.0	1432	3.55
	SMSA	3.3	2.2	1.3	1538	3.55
Pulmonary Tuberculosis	So. Boston	1.8	2.0	1.7	1435	0.10
	SMSA	1.2	1.3	0.1	1539	8.84
Shortness of Breath						
None	So. Boston	55.8	72.2	83.0	1023	-
Level 1	So. Boston	25.6	20.6	11.8	269	-
Level 2	So. Boston	18.6	7.2	5.3	132	-
					1255	84.16
None	SMSA	54.0	76.3	89.3	1424	-
Level 1	SMSA	34.3	16.4	8.3	208	-
Level 2	SMSA	11.6	7.2	2.3	74	-
					1536	128.31
Ns (a)	So. Boston	341	621	470	1432	
	SMSA	159	475	905	1539	

Notes: Chi-square statistics do not account for non-SRS sample design (or interviewer effects, etc.); hence, common inferences, e.g., p = .05 at 3.84 with 1 d.f., 5.99 with 2 d.f., 7.81 with 3 d.f., 9.49 with 4 d.f., will not be precisely accurate. Estimates of sample design and interviewer effects upon the variance of these measurements are presented in Section VI.

Chi-square statistics have 2 d.f. except in case of Shortness of Breath, where d.f. = 4.

(a) Sample sizes (Ns) for columns are from chronic cough analysis; Ns for other tabulations may vary slightly due to missing symptom data.

Table 24a

Test of Inter-Area Differences in 6-way Table:
Symptom by Area by Age by Sex by Smoking by Education

		Test of Area-by-Symptoms Parameter		
Pulmonary Disorder	Tau	L^2	df	p
Chronic Phlegm	1.24	11.19	1	<.001
Chronic Wheeze	1.20	1.91	1	ns
Pulmonary Tuberculosis	1.74	7.98	1	<.005
Shortness of Breath (2 Level)	1.26	7.10	1	<.01
(3-level)				
Level 0	0.80	9.32	2	<.01
Level I	0.94			
Level II+	1.33			

Table 24b

Test of Inter-Area Difference in Shortness of Breath
in 6-way Table Controlling for Doctor-Diagnosed Heart Trouble

Shortness of Breath (2 Level)	1.26	5.45	1	<.02
(3-level)				
Level 0	0.75	12.25	2	<.005
Level I	1.02			
Level II+	1.29			

Note: Tests of inter-area differences were computed as difference in likelihood-ratio chi-square statistics for log-linear models constrained to fit two 5-way marginals [SyAgSxSmEd] and [ArAgSxSmEd]; and a model which also fit [SyAr] marginal. In Table 34b the same strategy was used, substituting Ht (Heart Trouble) for Ed (Education).

Table 25

Estimates of Prevalence of Self-Reported Pulmonary Disorder by Employment in a "dusty job" for 1+ Years for South Boston and Rest of Boston SMSA

Disorder	Population	"Dusty" Job Yes	No	N	X^2
Chronic Cough	So. Boston	20.9%	13.4%	1329	11.36
	SMSA	16.2	9.2	1461	11.26
Chronic Phlegm	So. Boston	22.8	14.9	1330	11.46
	SMSA	17.6	8.6	1459	17.96
Chronic Wheeze	So. Boston	6.6	2.7	1328	11.20
	SMSA	3.1	1.4	1460	3.9
Pulmonary Tuberculosis	So. Boston	1.9	2.4	1328	0.42
	SMSA	0.6	0.6	1459	0.01
Shortness of Breath					
None	So. Boston	66.9	74.2	956	-
Level 1	So. Boston	22.6	17.4	248	-
Level 2	So. Boston	10.5	8.4	118	-
				1332	6.97
None	SMSA	76.1	82.9	1193	-
Level 1	SMSA	17.4	12.8	1099	-
Level 2	SMSA	6.4	4.3	69	-
				1460	6.68
Ns (a)	So. Boston	337	991	1328	
	SMSA	249	1210	1459	

Notes: Chi-square statistics do not account for non-SRS sample design (or interviewer effects, etc.); hence, common inferences, e.g., p = .05 at 3.84 with 1 d.f., 5.99 with 2 d.f., 7.81 with 3 d.f., 9.49 with 4 d.f., will not be precisely accurate. Estimates of sample design and interviewer effects upon the variance of these measurements are presented in Section VI.

Chi-square statistics have 1 d.f. except in case of Shortness of Breath, where d.f. = 2.

Table 26

Estimates of Prevalence of Self-Reported Pulmonary Disorder
By Use of Gas Kitchen Stove for Home Heating
For South Boston and Rest of Boston SMSA

Disorder	Population	Use Gas Stove for Heating: Yes	No	N	X^2
Chronic Cough	So. Boston	24.3%	13.7%	1440	12.83
	SMSA	12.4	10.5	1545	0.39
Chronic Phlegm	So. Boston	25.4	15.3	1441	10.90
	SMSA	15.8	9.6	1543	4.01
Chronic Wheeze	So. Boston	5.3	3.5	1438	1.22
	SMSA	2.4	1.7	1544	0.22
Pulmonary Tuberculosis	So. Boston	1.6	4.0	1441	4.57
	SMSA	0.6	1.0	1545	0.25
Shortness of Breath					
None	So. Boston	54.4	74.2	1028	-
Level 1	So. Boston	27.0	17.8	270	-
Level 2	So. Boston	18.6	8.0	132	-
				1431	32.17
None	SMSA	70.3	82.5	1261	-
Level 1	SMSA	19.1	13.1	208	-
Level 2	SMSA	10.5	4.4	74	-
				1543	11.62
Ns (a)	So. Boston	166	1275	1441	
	SMSA	100	1445	1545	

Notes: Chi-square statistics do not account for non-SRS sample design (or interviewer effects, etc.); hence, common inferences, e.g., p = .05 at 3.84 with 1 d.f., 5.99 with 2 d.f., 7.81 with 3 d.f., 9.49 with 4 d.f., will not be precisely accurate. Estimates of sample design and interviewer effects upon the variance of these measurements are presented in Section VI.

Chi-square statistics have 1 d.f. except in case of Shortness of Breath, where d.f. = 2.

(a) Sample sizes (Ns) for columns are from chronic cough analysis; Ns for other tabulations may vary slightly due to missing symptom data.

Table 27

Estimates of Prevalence of Self-Reported Pulmonary Disorder
By Type of Home Heating for South Boston and Rest of Boston SMSA

Disorder	Population	Type of Home Heating: Hot Water or Steam	Hot Air	Electric	Other	N	X^2
Chronic Cough	So. Boston	14.8%	16.9%	7.2%	15.9	1396	4.74
	SMSA	10.4	10.4	10.1	14.5	1505	1.15
Chronic Phlegm	So. Boston	18.6	10.8	13.0	18.4	1397	9.93
	SMSA	10.1	10.1	5.3	12.3	1502	2.85
Chronic Wheeze	So. Boston	4.2	2.5	3.3	3.4	1396	1.63
	SMSA	1.5	2.0	2.9	0.7	1504	1.65
Pulmonary Tuberculosis	So. Boston	1.6	1.0	6.5	2.7	1399	11.0
	SMSA	0.6	0.5	0.5	0.0	1505	0.43
Shortness of Breath							
None	So. Boston	72.6	71.3	85.0	64.8	999	-
Level 1	So. Boston	17.3	23.5	11.7	22.7	264	-
Level 2	So. Boston	10.1	5.2	3.3	12.6	127	-
						1390	21.94
None	SMSA	81.6	81.8	77.2	88.4	1227	-
Level 1	SMSA	13.4	13.8	15.0	9.4	202	-
Level 2	SMSA	5.0	4.4	7.8	2.0	73	-
						1502	4.57
Ns (a)	So. Boston	844	264	79	210		
	SMSA	923	416	100	66		

Notes: Chi-square statistics do not account for non-SRS sample design (or interviewer effects, etc.); hence, common inferences, e.g., p = .05 at 3.84 with 1 d.f., 5.99 with 2 d.f., 7.81 with 3 d.f., 9.49 with 4 d.f., will not be precisely accurate. Estimates of sample design and interviewer effects upon the variance of these measurements are presented in Section VI.

Chi-square statistics have 3 d.f. except in case of Shortness of Breath, where d.f. = 6.

(a) Sample sizes (Ns) for columns are from chronic cough analysis; Ns other tabulations may vary slightly due to missing symptom data.

Table 28

Test For Inter-Area Difference in Prevalence of Self-Reported Pulmonary Disorder using 8-Way Table: Symptom by Age by Sex by Smoking by Dusty Job (1+Yr.) by Use of Gas Kitchen Stove to Heat Home by Type of Heating

		Test of Area by Symptom Parameter		
Pulmonary Disorder	Tau	L^2	df	p
Chronic Phlegm	1.19	6.21	1	$<.02$
Chronic Wheeze	1.40	4.73	1	$<.03$
Pulmonary Tuberculosis	2.10	9.96	1	$<.002$
Shortness of Breath (dichotomy)	1.24	4.50	1	$<.05$

Note. Test is difference in L^2 for:

Model 1: constrained to fit two 7-way marginals (Symptom by all other factors except Area) and (Area by all other factors except Symptom).

Model 2: constrained to fit same 7-way marginals plus (SyAr) Marginal.

IX. SUMMARY OF FINDINGS FROM STATISTICAL ANALYSIS

IX. SUMMARY OF FINDINGS FROM STATISTICAL ANALYSIS

SUBSTANTIVE FINDINGS

A. A number of pulmonary disorders are reported with substantially greater frequency in South Boston than in the Boston SMSA.[1] Chronic Cough, Chronic Phlegm, Chronic Wheeze, Shortness of Breath and Pulmonary Tuberculosis are 1.49 to 2.71 times more likely to be reported in South Boston than in the rest of the Boston metropolitan area.

B. Demographic differences and smoking behavior are responsible for some portion of these findings. Of particular importance in this regard is the fact that the South Boston population contains a higher proportion of smokers (and more heavy smokers) than the Boston SMSA.

C. Further analysis of reports of eight major pulmonary disorders indicates that Shortness of Breath, Chronic

---

1. Throughout this report, we use "Boston metropolitan area" and Boston SMSA to denote the federal government's Standard Metropolitan Statistical Area for Boston; see Section II for precise definitions. Where comparisons are reported <u>between</u> South Boston and the Boston SMSA, our tabulations contrast the SMSA <u>excluding South Boston</u> to South Boston.

   Pulmonary disorder is used to denote a class of conditions that includes both pulmonary diseases and pulmonary symptoms.

Phlegm, Chronic Wheeze, and Pulmonary Tuberculosis are more prevalent in South Boston than in the rest of the Boston metropolitan area even after adjustments are made for smoking, age and sex. The elevated prevalence rates for these disorders appear to be adequately understood in most cases[2] as a simple difference between areas (i.e., a single parameter common to all categories of smoking, age and sex fits the observed data quite well).

Three other pulmonary disorders, Asthma, Chest Colds, and Severe Chest Colds, show no difference across areas. The eighth disorder (Chronic Cough) exhibits a puzzling 5-way interaction.

D. Examination of the impact of sample design and interviewer effects on the (univariate) distribution of reported pulmonary disorders suggested that some of the foregoing findings might have been due to complexities of design that were inadequately modeled in our initial statistical tests.

E. More rigorous testing for inter-area differences in reported pulmonary disorders (i.e., adjusting for age,

---

2. This is true except in the cases of Chronic Phlegm (where South Boston's disadvantage varies with age) and Pulmonary Tuberculosis (where it varies by sex and by age).

sex, and smoking, and accounting for interviewer effects and sample design effects) does not alter the conclusion that there is an elevated prevalence rate in South Boston for reported Shortness of Breath, Chronic Phlegm and Pulmonary Tuberculosis. More rigorous testing does, however, diminish the finding for Chronic Wheeze to a level ($.15<p<.05$) that would not ordinarily be considered "statistically reliable."

F. When our definition of Shortness of Breath is changed to distinguish between two levels (none vs. must walk more slowly than others the same age--or worse), we still conclude that there is an elevated rate of Shortness of Breath in South Boston.

G. When the analysis is expanded to control for socioeconomic status we still find higher rates in South Boston for the reporting of Chronic Phlegm, Shortness of Breath, and Pulmonary Tuberculosis. (This analysis used Simple Random Sampling assumptions rather than the more rigorous and computationally expensive jackknife procedures).

H. When the analysis is expanded to control for Age, Sex, Smoking, Type of Heating, Employment in a "Dusty job for 1 year," and Use of a Gas Kitchen Stove to Heat Home, the higher rates of reported pulmonary disorders in South Boston are still found for Chronic Phlegm, Chronic Wheeze, Shortness of Breath, and Pulmonary Tuberculosis.

Similarly, introduction of a control for doctor-diagnosed heart trouble in the analysis of Shortness of Breath does not alter the statistical conclusion that there is an excess of Shortness of Breath in South Boston.

ISSUES OF SURVEY TECHNIQUE

I. Coverage of the population: The combined impact of non-response and our use of a telephone sampling frame was to skew the demographic composition of our samples upward in Socioeconomic status. The sex and race composition of the samples were unaffected, and only a modest skew was evident in the age distribution.

Ratio adjustments of our sample estimates for age and education were explored (using 1980 Census counts for the SMSA). Generally, these adjustments produced modest increases in our estimates of the prevalence rates (this was particularly true in the case of the adjustment for education).

J. Bias: Across area differences in reporting-bias are a possible contaminant in all our estimates. Nonetheless, we have little scientific basis at present for inferring that such differential bias does (or does not) exist. Nor, indeed, do we know, were it to exist, whether it would affect our estimates by making the South Boston population appear more or less healthy.

APPENDIX A: REPORT ON FIELDWORK PROCEDURES

Report on Fieldwork

Boston Area Pulmonary Health Study, 1983

From July 16 to August 28, 1983, 2991 telephone interviews were conducted as part of a pulmonary health study of the Boston area. The purpose of the study was to compare the self-reported pulmonary health of residents in South Boston with that of residents of the remainder of the Boston Standard Metropolitan Statistical Area (SMSA). The survey instrument (see Appendix E) is primarily comprised of a revised version of the American Thoracic Society/Division of Lung Diseases Respiratory and Cardiovascular Disease Symptoms Questionnaire adapted for telephone interviewing.

The probability samples for South Boston and the SMSA were drawn separately. A simple random sample design was used in South Boston and a Waksberg two-stage probability sample[1] was drawn in the SMSA. Both methods produce samples with a known probability of inclusion of each household in the universe sampled from. With the exclusion of sampling method, all procedures for South Boston and the SMSA were identical.

The overall response rate for the survey was 70.1 percent: 70.8 percent in the SMSA and 69.3 percent in South Boston. 1444 interviews were conducted in South Boston and 1547 in the SMSA.

This report describes procedures and quality control mechanisms employed in each phase of the survey operations.

---

1. See, J. Waksberg, Sampling methods for random digit dialing. Journal of the American Statistical Association, 1978, 73:40-46.

SAMPLE

Sampling was done with the goal of interviewing approximately 1500 English speaking adults in South Boston and 1500 adults in the Boston SMSA (excluding South Boston). The sampling results are displayed in Table 1.

South Boston. A simple random digit dial sampling method was used in South Boston. For random digit dial samples, four digit random numbers are attached to every phone exchange in the sample area. Since the eligible sample is comprised only of working residential phones, business and non-working phones are excluded from the final count of "eligible" units.

Five thousand random numbers were generated for the two exchanges serving South Boston. These were then randomly divided into eleven replicates[2]: six containing 455 numbers and five containing 454 numbers. Numbers from 8 3/4 replicates (3976 numbers) were used for the final South Boston sample frame. Of those, 2099 were working residential numbers; these numbers constitute the "eligible" sample.

SMSA Sample. For the SMSA, the sample was drawn by the Waksberg method. This multistage approach increases efficiency while still retaining the characteristic of a known probability of selection for each household in the sample. The sampling stages and their outcomes are shown in Table 2.

In the first stage of the sampling, all telephone exchanges (the first three digits of a phone number) serving the SMSA (excluding South Boston) were identified. Two hundred eighty-three telephone exchanges serve communities in

----

2 The replicate system is used to divide the sample into random subsets, not all of which need to be used in order to maintain the characteristic of a known and equal probability of selection for each number. This allows the researcher to wait until later in the study to estimate the number of phone numbers needed to obtain the targeted quantity of interviewers, rather than having to guess when the study begins.

the SMSA. Three four digit random numbers were attached to each of these 283 exchanges, resulting in 849 numbers.

These were then called (in what is known as "Primary Screening") to identify the working residential numbers. The 302 (35.6%) working residential numbers which resulted define a set of clusters from which the second-stage sample is generated. The cluster itself is the exchange plus the first two digits of the random four digit number which was found to be a working residential number. A second-stage sample was then created by attaching a series of random two digit numbers (from the one hundred possible numbers between 00-99) to the five digit root. There were eight numbers assigned for each cluster for a total of 2416 numbers. These were subset into eight replicates, of which seven and one third were used for the study, resulting in an average cluster size of 7.33. When, during the course of calling these second-stage numbers, a business phone or non-working number was reached, the number was replaced with another number generated by random assignment of a different two digits to the root. This procedure was followed until a residential number was found.

The final stage of sampling, occurring once contact was made with a household, was identical for South Boston and the SMSA. When a household was contacted, the number of eligible adults in the household was determined and the designated respondent was selected according to a modified Kish randomization table.[3] Thus, every adult in a selected household had an equal probability of being selected as a respondent, and all households have a known probability of selection. However, since each adult in a N adult-household has a smaller probability of selection than adults residing in an N-1 (or

----

3. See L. Kish, Survey Sampling. New York: wiley, 1965.

smaller) household, the sample data must be weighted by the number of adults in the household to derive population estimates. In addition to this weighting to compensate for unequal probabilities of selection in households of differing size, a second weight is required to compensate for over-sampling of households with two or more telephone lines; this second weight is merely the reciprocal of the number of telephone lines used by the household.

SAMPLING FIELD WORK

First-stage sample screening (primary screening), using 83 hours of interviewer time, was done between June 30 and July 7.

Primary screening required contacting the 849 primary numbers (the 3 numbers generated for each exchange) within the SMSA to identify which were residential numbers. A maximum of four calls, two in the day and two in the evening, were made to a number by an interviewer. If contact was made at the number, the person who answered the telephone was asked if the number was residential or business or both (if both, it is included as a residential number). If no contact was made after 4 calls, the phone number was checked with the telephone company to determine the status of the number as residential, non-working, or business. Non-working numbers were dropped from the sample, residential numbers were retained, and numbers designated by the telephone company as business were contacted to determine if they were also used as residential numbers. All working residential numbers found in this process then defined the second stage clusters.

QUESTIONNAIRE DESIGN

Pretests were conducted on July 7 and July 12. Pretest samples were drawn from listed telephone numbers in Worcester and from numbers in South Boston which did not fall into the main study sample. For each pretest, 20 interviews were conducted, half in South Boston and half in Worcester. Four interviewers participated in the first pretest and three in the second. The pretests were used primarily to improve questionnaire flow, question order, skip patterns, and instructions to interviewers. The first pretest uncovered difficulties with question wording. Most problems were with respondents misunderstanding the meaning of questions. Whenever more than one respondent was unable to understand a question, it was reviewed and changes were made, if possible, or clarifying instructions and/or probes were provided for the interviewers.

TRAINING AND BRIEFING

The briefing for all interviewers and supervisors was conducted on July 16. All interviewers, both new and experienced, conducted practice interviews with another interviewer immediately following the briefing. In addition, all new interviewers (who had just received 3 days of general training) spent a minimum of four hours in the telephone room doing practice interviews (with respondents not in the sample) before their first shift working on the study.

Each interviewer also filled out a questionnaire for himself/herself.

DATA COLLECTION

Interviewing was done between July 16 and August 28, 1983. Interviews were an average of 15 minutes long. Forty-three interviewers worked a total of 3427 hours over 65 six-hour shifts during the six week period. The maximum number of interviewers on any shift was thirteen.

Each shift had at least one supervisor present at all times. The first four shifts had two supervisors to perform extra monitoring, answer questions, and work with interviewers who were having difficulties with response rate or with the interview.

Extensive attempts were made to contact each number in the sample. Six calls (varying time of day and including both weekday and weekends) were made. If no contact was made after six calls, the phone number was verified with the telephone company. Numbers verified as working residential numbers received, on average, another nine calls (for a total of fifteen before work was terminated on the number.[4] Further, if contact was made with an eligible household, but the designated respondent was not available at the time, at least 12 additional calls were made to reach the respondent before work was terminated on that number.

QUALITY CONTROL

All Center for Survey Research interviewers receive a minimum of 3 days training in general interviewing techniques. Additionally, all interviewers who worked on this study, including experienced ones, received a one-day briefing and training on the study (described earlier).

At least one interview is monitored by the supervisor for each shift an interviewer works, and two additional completed interviews per interviewer are carefully reviewed each shift. Specific areas of interviewing techniques are rated and immediate feedback is given to each interviewer by the supervisor. Additionally, every non-interview and first refusal is reviewed by the supervisor.

Attempts were made to convert all initial refusals into interviews. For each refusal, at least one attempt was made, by a different inteviewer from the one who received the refusal, to convince the respondent to participate. Refusal conversion attempts were made between four and ten days after the

----

[4] Any number deemed to be a working residential number is kept in the sample, and it is counted as a non-interview, even if contact is never made.

initial refusal was received, and at a time of day different from the initial call. For the first several weeks of the study, refusal conversions were assigned only to the most experienced interviewers. In the last 2 weeks of the study, all interviewers worked on refusal conversions.

RESPONSE RATE

The final sample disposition is displayed in Table 2. The interview was conducted only in English, so if the randomly selected adult did not speak English, the number was not counted as an "eligible" sample unit. Final response rate is the total interviews taken as a percentage of the total eligible households in the sample. Total non-response includes all households for which no interview was taken because of refusal, illness, or inability to contact the household or designated respondent. The response rate for this study was 70.1% overall; 69.3% in South Boston and 70.8% in the SMSA. 1536 interviews were taken from the SMSA sample and 1455 from the South Boston sample. The two exchanges serving South Boston have a miniscule spillover into adjoining areas. Eleven respondents from the South Boston sample do not actually live in South Boston, so the final count for analysis is 1547 SMSA residents and 1444 South Boston residents.

TABLE 1

SAMPLING PROCESS IN SMSA

283	Telephone Exchanges
3	Random numbers for each exchange
849	Numbers for primary screening
302	Clusters (working residential phones)
7.33	Numbers per cluster
2215	Numbers in sample
1536	Interviews (70.8% response rate)

Table 2

Field Results

	SMSA	%	South Boston	%	Total	%
AMPLE	2215		3975		6190	
Phone out of service	*		1274		1274	
Business	*		556		556	
Language	45		47		92	
OTAL ELIGIBLE SAMPLE	2170		2099		4269	
INTERVIEWS	1536	70.8%	1455	69.3%	2991	70.1%
NON-INTERVIEWS	634	29.2%	644	30.7%	1278	29.9%
Refusal	436	20.1%	475	22.6%	911	21.3%
Illness	63	2.9%	59	2.8%	122	2.9%
Other	135	6.2%	110	5.2%	153	5.7%

In Waksberg sampling, each number which is a business or a phone out of service is replaced with another random number in that cluster.

APPENDIX B: REPORT ON CODING AND DATABASE PREPARATION

APPENDIX B

REPORT ON CODING AND DATABASE PREPARATION

I. Code Development

A codebook was constructed from the questionnaire. The standard SRC University of Michigan occupation, business/industry codes and a CSR standard SMSA and Massachusetts cities codes were used. The questionnaires were edited. The codes used in most cases were the numbers of the answers in the questionnaire. A 9 was coded for answer not ascertained and a 0 for inapplicable.

II. Coding Changes and Special Issues

There were two forms of the questionnaire, A and B, which differed in the placement of a question on general health (A1 or B12). Form B had the general health question in position B12, therefore a zero was coded for question A1 on form B and a 0 was coded in position B12 for form A.

Variable 479 was not included in the first batch of inteviews, so variable 479 is coded missing (9) for the first several hundred interviews. (In the analysis, these cases were assigned the value of 1 on this variable, since only one percent of the remainder of the sample reported having more than one telephone line.)

Respondent had some difficulty specifying the type of heat they have in their homes (variable 464). In the coding of variable 464, "registers" were recorded as hot air heat, "radiators" as steam, and "gas on gas" was coded as other.

Coders checked each questionnaire to verify that the sample designation on the coversheet (SMSA or South Boston) was in fact where the respondent said that he/she lived. Eleven respondents with South Boston telephone exchanges reported that they actually lived in communities outside of South Boston, so these cases were reassigned to the SMSA group in our analyses.

III. Coding Production and Quality Control

Seven experienced coders worked on this study. All were briefed on the objectives of the study and the codes for the particular responses. Two practice interviews were coded independently by each coder during the briefing session, and then codes were discussed in round-robin fashion. Errors and changes in coding were corrected daily by two means: a coding control log and question cards. Changes or additions were entered into the control log by one research assistant and checked daily by each coder. Final disposition of all questionable responses and codes for which the coder was uncertain was made by the research assistant.

To assure quality, 10% of the questionnaires of the coders were coded by a second coder or the research assistant. The differences were then compared and any errors were corrected. This was done in order to identify coding problems and to calculate error rates for the coding process. Error rates for the seven variables that involved considerable coding were calculated and are listed in Table 1.

Table 1: Error Rates for Seven Variables Coded

V258	General health	4.5%
V322	Height	0.0%
V349	Ounces/pouches pipe tobacco	2.4%
V425	Income	3.8%
V426 - V435	Residential history (So. Boston)	1.7%
V440 - V446	Residential history (SMSA)	1.7%
V479	Number of phones	1.0%

IV. Data Processing

The data were keypunched directly from the questionnaires to tape and 100% key verified. The data were then transferred from tape to on-line disk, sorted into sequential ID and card order, and then checked to see that each ID and card number was correct. Subsequently a series of data editing procedures were used to identify erroneous or inconsistent data codes, errors identified by those procedures were corrected (consulting the original questionnaires as needed).

APPENDIX C: MORE COMPLETE SET OF LOG-LINEAR MODEL RESULTS

TABLE C-1

Tests of Alternative Log-linear Models fit to 5-way Table of Symptom by Area by Age by Sex by Smoking. (Data from Boston Area Pulmonary Health Survey, 1983.)

CHRONIC COUGH

Model	Margins Fit	Test of Model L^2	df	p
1	{ArAgSxSm} {Sy}	476.1	79	<.001
2	{ArAgSxSm} {SySm}	130.1	75	<.001
3	{ArAgSxSm} {SySx}	475.6	78	<.001
4	{ArAgSxSm} {SyAg}	467.9	76	<.001
5	{ArAgSxSm} {SyAr}	461.0	78	<.001
6	{ArAgSxSm} {SySx} {SySm}	129.3	74	<.001
7	{ArAgSxSm} {SyAg} {SySm}	128.5	72	<.001
8	{ArAgSxSm} {SyAg} {SySx}	467.2	75	<.001
9	{ArAgSxSm} {SyAr} {SySm}	128.2	74	<.001
10	{ArAgSxSm} {SyAr} {SySx}	460.7	77	<.001
11	{ArAgSxSm} {SyAr} {SyAg}	452.5	75	<.001
12	{ArAgSxSm} {SyAg} {SySx} {SySm}	127.7	71	<.001
13	{ArAgSxSm} {SyAr} {SySx} {SySm}	127.5	73	<.001
14	{ArAgSxSm} {SyAr} {SyAg} {SySm}	126.8	71	<.001
15	{ArAgSxSm} {SyAr} {SyAg} {SySx}	452.1	74	<.001
16	{ArAgSxSm} {SyAr} {SyAg} {SySx} {SySm}	126.1	70	<.001
17	{ArAgSxSm} {SyArAg}	449.6	72	<.001
18	{ArAgSxSm} {SyArAg} {SySx}	449.2	71	<.001
19	{ArAgSxSm} {SyArAg} {SySm}	125.4	68	<.001
20	{ArAgSxSm} {SyArAg} {SySx} {SySm}	124.7	67	<.001
21	{ArAgSxSm} {SyArSx}	459.9	76	<.001
22	{ArAgSxSm} {SyArSx} {SyAg}	451.2	73	<.001
23	{ArAgSxSm} {SyArSx} {SySm}	125.2	72	<.001
24	{ArAgSxSm} {SyArSx} {SyAg} {SySm}	124.6	69	<.001
25	{ArAgSxSm} {SyArSm}	116.2	70	.001
26	{ArAgSxSm} {SyArSm} {SyAg}	114.7	67	<.001
27	{ArAgSxSm} {SyArSm} {SySx}	115.5	69	.001
28	{ArAgSxSm} {SyArSm} {SyAg} {SySx}	114.1	66	<.001
29	{ArAgSxSm} {SyAgSx}	460.6	72	<.001
30	{ArAgSxSm} {SyAgSx} {SyAr}	445.0	71	<.001
31	{ArAgSxSm} {SyAgSx} {SySm}	117.8	68	<.001
32	{ArAgSxSm} {SyAgSx} {SyAr} {SySm}	116.1	67	<.001
33	{ArAgSxSm} {SyAgSm}	112.9	60	<.001
34	{ArAgSxSm} {SyAgSm} {SyAr}	111.1	59	<.001
35	{ArAgSxSm} {SyAgSm} {SySx}	112.1	59	<.001
36	{ArAgSxSm} {SyAgSm} {SyAr} {SySx}	110.4	58	<.001
37	{ArAgSxSm} {SySxSm}	120.7	70	<.001
38	{ArAgSxSm} {SySxSm} {SyAr}	119.0	69	<.001
39	{ArAgSxSm} {SySxSm} {SyAg}	119.1	67	<.001
40	{ArAgSxSm} {SySxSm} {SyAr} {SyAg}	117.5	66	<.001
41	{ArAgSxSm} {SyAgSm} {SySxSm}	103.8	55	<.001
42	{ArAgSxSm} {SyAgSm} {SySxSm} {SyAr}	102.2	54	<.001
43	{ArAgSxSm} {SyAgSx} {SySxSm}	107.4	64	.001
44	{ArAgSxSm} {SyAgSx} {SySxSm} {SyAr}	105.8	63	.001
45	{ArAgSxSm} {SyAgSx} {SyAgSm}	102.1	56	<.001
46	{ArAgSxSm} {SyAgSx} {SyAgSm} {SyAr}	100.4	55	<.001
47	{ArAgSxSm} {SyArSm} {SySxSm}	106.8	65	.001
48	{ArAgSxSm} {SyArSm} {SySxSm} {SyAg}	105.2	62	.001
49	{ArAgSxSm} {SyArSm} {SyAgSm}	98.9	55	<.001
50	{ArAgSxSm} {SyArSm} {SyAgSm} {SySx}	98.2	54	<.001

(Cont'd)

NOTE. Note. Variables are:
Ar: Area, 2 categories (SMSA and South Boston)
Ag: Age, 4 categories (18-34, 35-49, 50-64, 65+)
Sx: Sex, 2 Categories (Male and Female)
Sm: Smoking, 4 categories (Never Smoked Cigarettes; Ex Cigarette Smoker; 1-15 , 16-30, and 31+ Cigs. per day)
Sy: Symptom, Dependent Variable

L^2 is Likelihood-ratio chi-square statistic for fit of model

TABLE C-1: CHRONIC COUGH
(Cont'd)

Model	Margins Fit						TEST OF MODEL		
							L^2	df	p
51	{ArAgSxSm}	{SyArSm}	{SyAgSx}				104.2	63	.001
52	{ArAgSxSm}	{SyArSx}	{SySxSm}				117.4	68	<.001
53	{ArAgSxSm}	{SyArSx}	{SySxSm}	{SyAg}			115.8	65	<.001
54	{ArAgSxSm}	{SyArSx}	{SyAgSm}				109.3	57	<.001
55	{ArAgSxSm}	{SyArSx}	{SyAgSx}				444.1	70	<.001
56	{ArAgSxSm}	{SyArSx}	{SyAgSx}	{SySm}			114.6	66	<.001
57	{ArAgSxSm}	{SyArSx}	{SyArSm}				114.3	68	.001
58	{ArAgSxSm}	{SyArSx}	{SyArSm}	{SyAg}			112.8	65	<.001
59	{ArAgSxSm}	{SyArAg}	{SySxSm}				116.3	63	<.001
60	{ArAgSxSm}	{SyArAg}	{SyAgSm}				109.8	56	<.001
61	{ArAgSxSm}	{SyArAg}	{SyAgSm}	{SySx}			109.1	55	<.001
62	{ArAgSxSm}	{SyArAg}	{SyAgSx}				441.9	68	<.001
63	{ArAgSxSm}	{SyArAg}	{SyAgSx}	{SySm}			114.2	64	<.001
64	{ArAgSxSm}	{SyArAg}	{SyArSm}				112.5	64	<.001
65	{ArAgSxSm}	{SyArAg}	{SyArSm}	{SySx}			111.8	63	<.001
66	{ArAgSxSm}	{SyArAg}	{SyArSx}				448.5	70	<.001
67	{ArAgSxSm}	{SyArAg}	{SyArSx}	{SySm}			123.4	66	<.001
68	{ArAgSxSm}	{SyAgSx}	{SyAgSm}	{SySxSm}			93.6	52	.001
69	{ArAgSxSm}	{SyAgSx}	{SyAgSm}	{SySxSm}	{SyAr}		92.0	51	.001
70	{ArAgSxSm}	{SyArSm}	{SyAgSm}	{SySxSm}			89.7	50	.001
71	{ArAgSxSm}	{SyArSm}	{SyAgSx}	{SySxSm}			93.7	59	.003
72	{ArAgSxSm}	{SyArSm}	{SyAgSx}	{SyAgSm}			88.1	51	.001
73	{ArAgSxSm}	{SyArSx}	{SyAgSm}	{SySxSm}			100.9	53	<.001
74	{ArAgSxSm}	{SyArSx}	{SyAgSx}	{SySxSm}			103.9	62	.001
75	{ArAgSxSm}	{SyArSx}	{SyAgSx}	{SyAgSm}			99.3	54	<.001
76	{ArAgSxSm}	{SyArSx}	{SyArSm}	{SySxSm}			105.4	64	.001
77	{ArAgSxSm}	{SyArSx}	{SyArSm}	{SySxSm}	{SyAg}		103.7	61	.001
78	{ArAgSxSm}	{SyArSx}	{SyArSm}	{SyAgSm}			97.2	53	<.001
79	{ArAgSxSm}	{SyArSx}	{SyArSm}	{SyAgSx}			102.8	62	.001
80	{ArAgSxSm}	{SyArAg}	{SyAgSm}	{SySxSm}			101.1	51	<.001
81	{ArAgSxSm}	{SyArAg}	{SyAgSx}	{SySxSm}			104.1	60	.001
82	{ArAgSxSm}	{SyArAg}	{SyAgSx}	{SyAgSm}			98.4	52	<.001
83	{ArAgSxSm}	{SyArAg}	{SyArSm}	{SySxSm}			103.1	59	<.001
84	{ArAgSxSm}	{SyArAg}	{SyArSm}	{SyAgSm}			96.9	52	<.001
85	{ArAgSxSm}	{SyArAg}	{SyArSm}	{SyAgSm}	{SySx}		96.2	51	<.001
86	{ArAgSxSm}	{SyArAg}	{SyArSm}	{SyAgSx}			101.2	60	.001
87	{ArAgSxSm}	{SyArAg}	{SyArSx}	{SySxSm}			114.7	62	<.001
88	{ArAgSxSm}	{SyArAg}	{SyArSx}	{SyAgSm}			108.1	54	<.001
89	{ArAgSxSm}	{SyArAg}	{SyArSx}	{SyAgSx}			441.4	67	<.001
90	{ArAgSxSm}	{SyArAg}	{SyArSx}	{SyAgSx}	{SySm}		113.2	63	<.001
91	{ArAgSxSm}	{SyArAg}	{SyArSx}	{SyArSm}			110.5	62	<.001
92	{ArAgSxSm}	{SyArSm}	{SyAgSx}	{SyAgSm}	{SySxSm}		79.6	47	.002
93	{ArAgSxSm}	{SyArSx}	{SyAgSx}	{SyAgSm}	{SySxSm}		90.6	50	.001
94	{ArAgSxSm}	{SyArSx}	{SyArSm}	{SyAgSm}	{SySxSm}		88.6	49	.001
95	{ArAgSxSm}	{SyArSx}	{SyArSm}	{SyAgSx}	{SySxSm}		91.9	58	.003
96	{ArAgSxSm}	{SyArSx}	{SyArSm}	{SyAgSx}	{SyAgSm}		87.1	50	.001
97	{ArAgSxSm}	{SyArAg}	{SyAgSx}	{SyAgSm}	{SySxSm}		90.1	48	<.001
98	{ArAgSxSm}	{SyArAg}	{SyArSm}	{SyAgSm}	{SySxSm}		87.9	47	<.001
99	{ArAgSxSm}	{SyArAg}	{SyArSm}	{SyAgSx}	{SySxSm}		90.7	56	.003
100	{ArAgSxSm}	{SyArAg}	{SyArSm}	{SyAgSx}	{SyAgSm}		85.0	48	.001
101	{ArAgSxSm}	{SyArAg}	{SyArSx}	{SyAgSm}	{SySxSm}		99.9	50	<.001
102	{ArAgSxSm}	{SyArAg}	{SyArSx}	{SyAgSx}	{SySxSm}		102.7	59	.001
103	{ArAgSxSm}	{SyArAg}	{SyArSx}	{SyAgSx}	{SyAgSm}		97.6	51	<.001
104	{ArAgSxSm}	{SyArAg}	{SyArSx}	{SyArSm}	{SySxSm}		101.6	58	.001
105	{ArAgSxSm}	{SyArAg}	{SyArSx}	{SyArSm}	{SyAgSm}		95.3	50	<.001
106	{ArAgSxSm}	{SyArAg}	{SyArSx}	{SyArSm}	{SyAgSx}		100.2	59	.001
107	{ArAgSxSm}	{SyArAg}	{SyArSx}	{SyArSm}	{SyAgSx}	{SySxSm}	89.4	55	.003
108	{ArAgSxSm}	{SyAgSxSm}					74.0	40	.001
109	{ArAgSxSm}	{SyAgSxSm}	{SyAr}				72.3	39	.001
110	{ArAgSxSm}	{SyAgSxSm}	{SyArAg}				70.4	36	.001

(Cont'd)

TABLE C-1: CHRONIC COUGH
(Cont'd)

Model	Margins Fit					TEST OF MODEL L^2	df	p
111	{ArAgSxSm}	{SyAgSxSm}	{SyArSx}			71.1	38	.001
112	{ArAgSxSm}	{SyAgSxSm}	{SyArSm}			61.0	35	.004
113	{ArAgSxSm}	{SyAgSxSm}	{SyArAg}	{SyArSx}		69.6	35	.001
114	{ArAgSxSm}	{SyAgSxSm}	{SyArAg}	{SyArSm}		58.2	32	.003
115	{ArAgSxSm}	{SyAgSxSm}	{SyArSx}	{SyArSm}		59.8	34	.004
116	{ArAgSxSm}	{SyAgSxSm}	{SyArAg}	{SyArSx}	{SyArSm}	57.4	31	.003
117	{ArAgSxSm}	{SyArSxSm}				99.7	60	.001
118	{ArAgSxSm}	{SyArSxSm}	{SyAg}			97.8	57	.001
119	{ArAgSxSm}	{SyArSxSm}	{SyArAg}			95.8	54	.001
120	{ArAgSxSm}	{SyArSxSm}	{SyAgSx}			86.6	54	.003
121	{ArAgSxSm}	{SyArSxSm}	{SyAgSm}			82.5	45	.001
122	{ArAgSxSm}	{SyArSxSm}	{SyArAg}	{SyAgSx}		84.0	51	.003
123	{ArAgSxSm}	{SyArSxSm}	{SyArAg}	{SyAgSm}		80.6	42	<.001
124	{ArAgSxSm}	{SyArSxSm}	{SyAgSx}	{SyAgSm}		72.5	42	.003
125	{ArAgSxSm}	{SyArSxSm}	{SyArAg}	{SyAgSx}	{SyAgSm}	69.7	39	.002
126	{ArAgSxSm}	{SyArAgSm}				72.0	40	.002
127	{ArAgSxSm}	{SyArAgSm}	{SySx}			71.6	39	.001
128	{ArAgSxSm}	{SyArAgSm}	{SyArSx}			70.9	38	.001
129	{ArAgSxSm}	{SyArAgSm}	{SyAgSx}			61.5	36	.005
130	{ArAgSxSm}	{SyArAgSm}	{SySxSm}			62.5	35	.003
131	{ArAgSxSm}	{SyArAgSm}	{SyArSx}	{SyAgSx}		60.9	35	.005
132	{ArAgSxSm}	{SyArAgSm}	{SyArSx}	{SySxSm}		61.6	34	.003
133	{ArAgSxSm}	{SyArAgSm}	{SyAgSx}	{SySxSm}		52.3	32	.013
134	{ArAgSxSm}	{SyArAgSm}	{SyArSx}	{SyAgSx}	{SySxSm}	51.4	31	.012
135	{ArAgSxSm}	{SyArAgSx}				438.3	64	<.001
136	{ArAgSxSm}	{SyArAgSx}	{SySm}			108.0	60	<.001
137	{ArAgSxSm}	{SyArAgSx}	{SyArSm}			96.0	56	.001
138	{ArAgSxSm}	{SyArAgSx}	{SyAgSm}			92.3	48	<.001
139	{ArAgSxSm}	{SyArAgSx}	{SySxSm}			97.3	56	.001
140	{ArAgSxSm}	{SyArAgSx}	{SyArSm}	{SyAgSm}		80.1	44	.001
141	{ArAgSxSm}	{SyArAgSx}	{SyArSm}	{SySxSm}		85.1	52	.003
142	{ArAgSxSm}	{SyArAgSx}	{SyAgSm}	{SySxSm}		83.7	44	<.001
143	{ArAgSxSm}	{SyArAgSx}	{SyArSm}	{SyAgSm}	{SySxSm}	71.5	40	.002
144	{ArAgSxSm}	{SyArAgSm}	{SyArAgSx}			58.9	32	.003
145	{ArAgSxSm}	{SyArAgSm}	{SyArAgSx}	{SySxSm}		48.9	28	.009
146	{ArAgSxSm}	{SyArSxSm}	{SyArAgSx}			80.0	48	.003
147	{ArAgSxSm}	{SyArSxSm}	{SyArAgSx}	{SyAgSm}		66.0	36	.002
148	{ArAgSxSm}	{SyArSxSm}	{SyArAgSm}			54.9	30	.004
149	{ArAgSxSm}	{SyArSxSm}	{SyArAgSm}	{SyAgSx}		44.0	27	.021
150	{ArAgSxSm}	{SyAgSxSm}	{SyArAgSx}			63.9	32	.001
151	{ArAgSxSm}	{SyAgSxSm}	{SyArAgSx}	{SyArSm}		54.4	28	.002
152	{ArAgSxSm}	{SyAgSxSm}	{SyArAgSm}			33.9	20	.027
153	{ArAgSxSm}	{SyAgSxSm}	{SyArAgSm}	{SyArSx}		33.1	19	.023
154	{ArAgSxSm}	{SyAgSxSm}	{SyArSxSm}			53.9	30	.005
155	{ArAgSxSm}	{SyAgSxSm}	{SyArSxSm}	{SyArAg}		51.3	27	.003
156	{ArAgSxSm}	{SyArSxSm}	{SyArAgSm}	{SyArAgSx}		41.6	24	.014
157	{ArAgSxSm}	{SyAgSxSm}	{SyArAgSm}	{SyArAgSx}		32.1	16	.010
158	{ArAgSxSm}	{SyAgSxSm}	{SyArSxSm}	{SyArAgSx}		48.5	24	.002
159	{ArAgSxSm}	{SyAgSxSm}	{SyArSxSm}	{SyArAgSm}		25.6	15	.043
160	{ArAgSxSm}	{SyAgSxSm}	{SyArSxSm}	{SyArAgSm}	{SyArAgSx}	24.4	12	.018
161*	{SyArAgSxSm}					0.0	0	---

TABLE C-2

Tests of Alternative Log-linear Models fit to 5-way Table of Symptom by Area by Age by Sex by Smoking. (Data from Boston Area Pulmonary Health Survey, 1983.)

CHRONIC PHLEGM

Model	Margins Fit	TEST OF MODEL L^2	df	p
1	{ArAgSxSm} {Sy}	261.2	79	<.001
2	{ArAgSxSm} {SySm}	119.2	75	.001
3	{ArAgSxSm} {SySx}	251.3	78	<.001
4	{ArAgSxSm} {SyAg}	255.6	76	<.001
5	{ArAgSxSm} {SyAr}	233.4	78	<.001
6	{ArAgSxSm} {SySx} {SySm}	109.3	74	.005
7	{ArAgSxSm} {SyAg} {SySm}	116.2	72	.001
8	{ArAgSxSm} {SyAg} {SySx}	246.3	75	<.001
9	{ArAgSxSm} {SyAr} {SySm}	105.2	74	.010
10	{ArAgSxSm} {SyAr} {SySx}	222.1	77	<.001
11	{ArAgSxSm} {SyAr} {SyAg}	228.3	75	<.001
12	{ArAgSxSm} {SyAg} {SySx} {SySm}	106.1	71	.005
13	{ArAgSxSm} {SyAr} {SySx} {SySm}	94.2	73	.048
14	{ArAgSxSm} {SyAr} {SyAg} {SySm}	102.9	71	.008
15	{ArAgSxSm} {SyAr} {SyAg} {SySx}	217.6	74	<.001
16	{ArAgSxSm} {SyAr} {SyAg} {SySx} {SySm}	91.7	70	.042
17	{ArAgSxSm} {SyArAg}	216.6	72	<.001
18	{ArAgSxSm} {SyArAg} {SySx}	205.7	71	<.001
19	{ArAgSxSm} {SyArAg} {SySm}	93.7	68	.021
20*	{ArAgSxSm} {SyArAg} {SySx} {SySm}	82.3	67	.099
21	{ArAgSxSm} {SyArSx}	222.0	76	<.001
22	{ArAgSxSm} {SyArSx} {SyAg}	217.5	73	<.001
23	{ArAgSxSm} {SyArSx} {SySm}	94.0	72	.042
24	{ArAgSxSm} {SyArSx} {SyAg} {SySm}	91.5	69	.036
25	{ArAgSxSm} {SyArSm}	97.7	70	.016
26	{ArAgSxSm} {SyArSm} {SyAg}	95.1	67	.014
27*	{ArAgSxSm} {SyArSm} {SySx}	86.4	69	.076
28*	{ArAgSxSm} {SyArSm} {SyAg} {SySx}	83.7	66	.070
29	{ArAgSxSm} {SyAgSx}	243.9	72	<.001
30	{ArAgSxSm} {SyAgSx} {SyAr}	214.9	71	<.001
31	{ArAgSxSm} {SyAgSx} {SySm}	102.6	68	.004
32	{ArAgSxSm} {SyAgSx} {SyAr} {SySm}	88.1	67	.043
33	{ArAgSxSm} {SyAgSm}	98.8	60	.001
34	{ArAgSxSm} {SyAgSm} {SyAr}	85.5	59	.014
35	{ArAgSxSm} {SyAgSm} {SySx}	89.8	59	.006
36*	{ArAgSxSm} {SyAgSm} {SyAr} {SySx}	75.6	58	.060
37	{ArAgSxSm} {SySxSm}	107.4	70	.003
38	{ArAgSxSm} {SySxSm} {SyAr}	92.6	69	.031
39	{ArAgSxSm} {SySxSm} {SyAg}	104.4	67	.003
40	{ArAgSxSm} {SySxSm} {SyAr} {SyAg}	90.1	66	.026
41	{ArAgSxSm} {SyAgSm} {SySxSm}	88.4	55	.003
42	{ArAgSxSm} {SyAgSm} {SySxSm} {SyAr}	74.4	54	.034
43	{ArAgSxSm} {SyAgSx} {SySxSm}	100.4	64	.003
44	{ArAgSxSm} {SyAgSx} {SySxSm} {SyAr}	86.0	63	.028
45	{ArAgSxSm} {SyAgSx} {SyAgSm}	86.2	56	.006
46*	{ArAgSxSm} {SyAgSx} {SyAgSm} {SyAr}	71.9	55	.063
47	{ArAgSxSm} {SyArSm} {SySxSm}	84.8	65	.050
48	{ArAgSxSm} {SyArSm} {SySxSm} {SyAg}	82.1	62	.045
49	{ArAgSxSm} {SyArSm} {SyAgSm}	76.5	55	.029
50*	{ArAgSxSm} {SyArSm} {SyAgSm} {SySx}	66.4	54	.121

(Cont'd)

NOTE. Note. Variables are:
Ar: Area, 2 categories (SMSA and South Boston)
Ag: Age, 4 categories (18-34, 35-49, 50-64, 65+)
Sx: Sex, 2 Categories (Male and Female)
Sm: Smoking, 4 categories (Never Smoked Cigarettes; Ex Cigarette Smoker; 1-15 , 16-30, and 31+ Cigs. per day)
Sy: Symptom, Dependent Variable

L^2 is Likelihood-ratio chi-square statistic for fit of model

TABLE C-2 : CHRONIC PHLEGM
(Cont'd)

Model	Margins Fit						TEST OF MODEL L^2	df	p
51*	{ArAgSxSm}	{SyArSm}	{SyAgSx}				79.9	63	.074
52	{ArAgSxSm}	{SyArSx}	{SySxSm}				92.3	68	.026
53	{ArAgSxSm}	{SyArSx}	{SySxSm}	{SyAg}			89.8	65	.022
54*	{ArAgSxSm}	{SyArSx}	{SyAgSm}				75.5	57	.051
55	{ArAgSxSm}	{SyArSx}	{SyAgSx}				214.8	70	<.001
56	{ArAgSxSm}	{SyArSx}	{SyAgSx}	{SySm}			87.9	66	.037
57*	{ArAgSxSm}	{SyArSx}	{SyArSm}				86.0	68	.069
58*	{ArAgSxSm}	{SyArSx}	{SyArSm}	{SyAg}			83.2	65	.063
59*	{ArAgSxSm}	{SyArAg}	{SySxSm}				80.9	63	.064
60	{ArAgSxSm}	{SyArAg}	{SyAgSm}				76.0	56	.039
61*	{ArAgSxSm}	{SyArAg}	{SyAgSm}	{SySx}			66.0	55	.148
62	{ArAgSxSm}	{SyArAg}	{SyAgSx}				202.7	68	<.001
63*	{ArAgSxSm}	{SyArAg}	{SyAgSx}	{SySm}			78.0	64	.112
64	{ArAgSxSm}	{SyArAg}	{SyArSm}				87.1	64	.029
65*	{ArAgSxSm}	{SyArAg}	{SyArSm}	{SySx}			75.7	63	.132
66	{ArAgSxSm}	{SyArAg}	{SyArSx}				205.7	70	<.001
67*	{ArAgSxSm}	{SyArAg}	{SyArSx}	{SySm}			82.1	66	.087
68	{ArAgSxSm}	{SyAgSx}	{SyAgSm}	{SySxSm}			84.6	52	.003
69	{ArAgSxSm}	{SyAgSx}	{SyAgSm}	{SySxSm}	{SyAr}		70.4	51	.037
70*	{ArAgSxSm}	{SyArSm}	{SyAgSm}	{SySxSm}			65.2	50	.073
71*	{ArAgSxSm}	{SyArSm}	{SyAgSx}	{SySxSm}			77.8	59	.051
72*	{ArAgSxSm}	{SyArSm}	{SyAgSx}	{SyAgSm}			62.5	51	.130
73	{ArAgSxSm}	{SyArSx}	{SyAgSm}	{SySxSm}			74.3	53	.028
74	{ArAgSxSm}	{SyArSx}	{SyAgSx}	{SySxSm}			85.8	62	.024
75*	{ArAgSxSm}	{SyArSx}	{SyAgSx}	{SyAgSm}			71.8	54	.053
76	{ArAgSxSm}	{SyArSx}	{SyArSm}	{SySxSm}			84.3	64	.045
[illegible]	{ArAgSxSm}	{SyArSx}	{SyArSm}	{SySxSm}	{SyAg}		81.5	61	.041
[illegible]	{ArAgSxSm}	{SyArSx}	{SyArSm}	{SyAgSm}			66.1	53	.107
79*	{ArAgSxSm}	{SyArSx}	{SyArSm}	{SyAgSx}			79.4	62	.067
80*	{ArAgSxSm}	{SyArAg}	{SyAgSm}	{SySxSm}			65.0	51	.090
81*	{ArAgSxSm}	{SyArAg}	{SyAgSx}	{SySxSm}			76.2	60	.077
82*	{ArAgSxSm}	{SyArAg}	{SyAgSx}	{SyAgSm}			61.4	52	.175
83*	{ArAgSxSm}	{SyArAg}	{SyArSm}	{SySxSm}			74.3	59	.087
84*	{ArAgSxSm}	{SyArAg}	{SyArSm}	{SyAgSm}			69.5	52	.052
85*	{ArAgSxSm}	{SyArAg}	{SyArSm}	{SyAgSm}	{SySx}		59.5	51	.195
86*	{ArAgSxSm}	{SyArAg}	{SyArSm}	{SyAgSx}			71.3	60	.151
87*	{ArAgSxSm}	{SyArAg}	{SyArSx}	{SySxSm}			80.7	62	.056
88*	{ArAgSxSm}	{SyArAg}	{SyArSx}	{SyAgSm}			65.9	54	.128
89	{ArAgSxSm}	{SyArAg}	{SyArSx}	{SyAgSx}			202.7	67	<.001
90*	{ArAgSxSm}	{SyArAg}	{SyArSx}	{SyAgSx}	{SySm}		77.9	63	.098
91*	{ArAgSxSm}	{SyArAg}	{SyArSx}	{SyArSm}			75.3	62	.120
92*	{ArAgSxSm}	{SyArSm}	{SyAgSx}	{SyAgSm}	{SySxSm}		61.1	47	.081
93	{ArAgSxSm}	{SyArSx}	{SyAgSx}	{SyAgSm}	{SySxSm}		70.3	50	.030
94*	{ArAgSxSm}	{SyArSx}	{SyArSm}	{SyAgSm}	{SySxSm}		64.9	49	.063
95	{ArAgSxSm}	{SyArSx}	{SyArSm}	{SyAgSx}	{SySxSm}		77.3	58	.046
96*	{ArAgSxSm}	{SyArSx}	{SyArSm}	{SyAgSx}	{SyAgSm}		62.3	50	.114
97*	{ArAgSxSm}	{SyArAg}	{SyAgSx}	{SyAgSm}	{SySxSm}		60.2	48	.112
98*	{ArAgSxSm}	{SyArAg}	{SyArSm}	{SyAgSm}	{SySxSm}		58.5	47	.122
99*	{ArAgSxSm}	{SyArAg}	{SyArSm}	{SyAgSx}	{SySxSm}		69.5	56	.106
100*	{ArAgSxSm}	{SyArAg}	{SyArSm}	{SyAgSx}	{SyAgSm}		54.8	48	.233
101*	{ArAgSxSm}	{SyArAg}	{SyArSx}	{SyAgSm}	{SySxSm}		64.9	50	.076
102*	{ArAgSxSm}	{SyArAg}	{SyArSx}	{SyAgSx}	{SySxSm}		76.1	59	.066
103*	{ArAgSxSm}	{SyArAg}	{SyArSx}	{SyAgSx}	{SyAgSm}		61.4	51	.151
104*	{ArAgSxSm}	{SyArAg}	{SyArSx}	{SyArSm}	{SySxSm}		73.9	58	.078
105*	{ArAgSxSm}	{SyArAg}	{SyArSx}	{SyArSm}	{SyAgSm}		59.3	50	.173
106*	{ArAgSxSm}	{SyArAg}	{SyArSx}	{SyArSm}	{SyAgSx}		71.1	59	.134
107*	{ArAgSxSm}	{SyArAg}	{SyArSx}	{SyArSm}	{SyAgSx}	{SySxSm}	69.3	55	.094
108	{ArAgSxSm}	{SyAgSxSm}					71.0	40	.002
109	{ArAgSxSm}	{SyAgSxSm}	{SyAr}				56.5	39	.034
110*	{ArAgSxSm}	{SyAgSxSm}	{SyArAg}				47.0	36	.104

(Cont'd)

TABLE C-2: CHRONIC PHLEGM
(Cont'd)

Model	Margins Fit					L^2	df	p
111	{ArAgSxSm}	{SyAgSxSm}	{SyArSx}			56.4	38	.027
112*	{ArAgSxSm}	{SyAgSxSm}	{SyArSm}			47.2	35	.082
113*	{ArAgSxSm}	{SyAgSxSm}	{SyArAg}	{SyArSx}		47.0	35	.085
114*	{ArAgSxSm}	{SyAgSxSm}	{SyArAg}	{SyArSm}		40.4	32	.146
115*	{ArAgSxSm}	{SyAgSxSm}	{SyArSx}	{SyArSm}		46.8	34	.070
116*	{ArAgSxSm}	{SyAgSxSm}	{SyArAg}	{SyArSx}	{SyArSm}	40.3	31	.123
117	{ArAgSxSm}	{SyArSxSm}				80.4	60	.040
118	{ArAgSxSm}	{SyArSxSm}	{SyAg}			77.8	57	.035
119*	{ArAgSxSm}	{SyArSxSm}	{SyArAg}			70.0	54	.070
120	{ArAgSxSm}	{SyArSxSm}	{SyAgSx}			73.9	54	.037
121*	{ArAgSxSm}	{SyArSxSm}	{SyAgSm}			60.9	45	.057
122*	{ArAgSxSm}	{SyArSxSm}	{SyArAg}	{SyAgSx}		65.8	51	.079
123*	{ArAgSxSm}	{SyArSxSm}	{SyArAg}	{SyAgSm}		54.5	42	.094
124*	{ArAgSxSm}	{SyArSxSm}	{SyAgSx}	{SyAgSm}		57.2	42	.059
125*	{ArAgSxSm}	{SyArSxSm}	{SyArAg}	{SyAgSx}	{SyAgSm}	50.2	39	.109
126*	{ArAgSxSm}	{SyArAgSm}				49.9	40	.137
127*	{ArAgSxSm}	{SyArAgSm}	{SySx}			38.1	39	>.5
128*	{ArAgSxSm}	{SyArAgSm}	{SyArSx}			38.0	38	.472
129*	{ArAgSxSm}	{SyArAgSm}	{SyAgSx}			33.8	36	>.5
130*	{ArAgSxSm}	{SyArAgSm}	{SySxSm}			37.3	35	.362
131*	{ArAgSxSm}	{SyArAgSm}	{SyArSx}	{SyAgSx}		33.8	35	>.5
132*	{ArAgSxSm}	{SyArAgSm}	{SyArSx}	{SySxSm}		37.2	34	.323
133*	{ArAgSxSm}	{SyArAgSm}	{SyAgSx}	{SySxSm}		32.9	32	.421
134*	{ArAgSxSm}	{SyArAgSm}	{SyArSx}	{SyAgSx}	{SySxSm}	32.9	31	.375
135	{ArAgSxSm}	{SyArAgSx}				199.7	64	<.001
136*	{ArAgSxSm}	{SyArAgSx}	{SySm}			75.3	60	.089
[illegible]*	{ArAgSxSm}	{SyArAgSx}	{SyArSm}			68.5	56	.122
[illegible]	{ArAgSxSm}	{SyArAgSx}	{SyAgSm}			58.7	48	.139
139*	{ArAgSxSm}	{SyArAgSx}	{SySxSm}			73.6	56	.057
140*	{ArAgSxSm}	{SyArAgSx}	{SyArSm}	{SyAgSm}		52.2	44	.186
141*	{ArAgSxSm}	{SyArAgSx}	{SyArSm}	{SySxSm}		66.7	52	.082
142*	{ArAgSxSm}	{SyArAgSx}	{SyAgSm}	{SySxSm}		57.6	44	.082
143*	{ArAgSxSm}	{SyArAgSx}	{SyArSm}	{SyAgSm}	{SySxSm}	51.0	40	.114
144*	{ArAgSxSm}	{SyArAgSm}	{SyArAgSx}			30.7	32	>.5
145*	{ArAgSxSm}	{SyArAgSm}	{SyArAgSx}	{SySxSm}		30.0	28	.365
146*	{ArAgSxSm}	{SyArSxSm}	{SyArAgSx}			63.7	48	.064
147*	{ArAgSxSm}	{SyArSxSm}	{SyArAgSx}	{SyAgSm}		48.2	36	.084
148*	{ArAgSxSm}	{SyArSxSm}	{SyArAgSm}			33.5	30	.301
149*	{ArAgSxSm}	{SyArSxSm}	{SyArAgSm}	{SyAgSx}		29.7	27	.329
150*	{ArAgSxSm}	{SyAgSxSm}	{SyArAgSx}			44.3	32	.073
151*	{ArAgSxSm}	{SyAgSxSm}	{SyArAgSx}	{SyArSm}		37.7	28	.104
152*	{ArAgSxSm}	{SyAgSxSm}	{SyArAgSm}			18.9	20	>.5
153*	{ArAgSxSm}	{SyAgSxSm}	{SyArAgSm}	{SyArSx}		18.9	19	.464
154*	{ArAgSxSm}	{SyAgSxSm}	{SyArSxSm}			43.1	30	.057
155*	{ArAgSxSm}	{SyAgSxSm}	{SyArSxSm}	{SyArAg}		36.8	27	.099
156*	{ArAgSxSm}	{SyArSxSm}	{SyArAgSm}	{SyArAgSx}		27.3	24	.293
157*	{ArAgSxSm}	{SyAgSxSm}	{SyArAgSm}	{SyArAgSx}		16.8	16	.396
158*	{ArAgSxSm}	{SyAgSxSm}	{SyArSxSm}	{SyArAgSx}		34.7	24	.072
159*	{ArAgSxSm}	{SyAgSxSm}	{SyArSxSm}	{SyArAgSm}		15.1	15	.442
160*	{ArAgSxSm}	{SyAgSxSm}	{SyArSxSm}	{SyArAgSm}	{SyArAgSx}	13.5	12	.333
161*	{SyArAgSxSm}					0.0	0	---

TABLE C-3

Tests of Alternative Log-linear Models fit to 5-way Table of Symptom by Area by Age by Sex by Smoking. (Data from Boston Area Pulmonary Health Survey, 1983.)

SHORTNESS OF BREATH

Model	Margins Fit					TEST OF MODEL L^2	df	p
1	{ArAgSxSm}	{Sy}				591.1	158	<.001
2	{ArAgSxSm}	{SySm}				417.4	150	<.001
3	{ArAgSxSm}	{SySx}				541.5	156	<.001
4	{ArAgSxSm}	{SyAg}				508.1	152	<.001
5	{ArAgSxSm}	{SyAr}				539.3	156	<.001
6	{ArAgSxSm}	{SySx}	{SySm}			365.4	148	<.001
7	{ArAgSxSm}	{SyAg}	{SySm}			325.5	144	<.001
8	{ArAgSxSm}	{SyAg}	{SySx}			462.4	150	<.001
9	{ArAgSxSm}	{SyAr}	{SySm}			385.1	148	<.001
10	{ArAgSxSm}	{SyAr}	{SySx}			492.2	154	<.001
11	{ArAgSxSm}	{SyAr}	{SyAg}			455.2	150	<.001
12	{ArAgSxSm}	{SyAg}	{SySx}	{SySm}		277.7	142	<.001
13	{ArAgSxSm}	{SyAr}	{SySx}	{SySm}		335.3	146	<.001
14	{ArAgSxSm}	{SyAr}	{SyAg}	{SySm}		293.6	142	<.001
15	{ArAgSxSm}	{SyAr}	{SyAg}	{SySx}		412.6	148	<.001
16	{ArAgSxSm}	{SyAr}	{SyAg}	{SySx}	{SySm}	248.7	140	<.001
17	{ArAgSxSm}	{SyArAg}				437.8	144	<.001
18	{ArAgSxSm}	{SyArAg}	{SySx}			394.1	142	<.001
19	{ArAgSxSm}	{SyArAg}	{SySm}			274.8	136	<.001
20	{ArAgSxSm}	{SyArAg}	{SySx}	{SySm}		229.2	134	<.001
21	{ArAgSxSm}	{SyArSx}				491.8	152	<.001
22	{ArAgSxSm}	{SyArSx}	{SyAg}			411.9	146	<.001
[illegible]	{ArAgSxSm}	{SyArSx}	{SySm}			334.7	144	<.001
	{ArAgSxSm}	{SyArSx}	{SyAg}	{SySm}		247.7	138	<.001
25	{ArAgSxSm}	{SyArSm}				369.5	140	<.001
26	{ArAgSxSm}	{SyArSm}	{SyAg}			277.1	134	<.001
27	{ArAgSxSm}	{SyArSm}	{SySx}			320.3	138	<.001
28	{ArAgSxSm}	{SyArSm}	{SyAg}	{SySx}		232.5	132	<.001
29	{ArAgSxSm}	{SyAgSx}				437.5	144	<.001
30	{ArAgSxSm}	{SyAgSx}	{SyAr}			386.2	142	<.001
31	{ArAgSxSm}	{SyAgSx}	{SySm}			258.0	136	<.001
32	{ArAgSxSm}	{SyAgSx}	{SyAr}	{SySm}		228.3	134	<.001
33	{ArAgSxSm}	{SyAgSm}				272.7	120	<.001
34	{ArAgSxSm}	{SyAgSm}	{SyAr}			237.7	118	<.001
35	{ArAgSxSm}	{SyAgSm}	{SySx}			230.1	118	<.001
36	{ArAgSxSm}	{SyAgSm}	{SyAr}	{SySx}		198.2	116	<.001
37	{ArAgSxSm}	{SySxSm}				346.7	140	<.001
38	{ArAgSxSm}	{SySxSm}	{SyAr}			317.0	138	<.001
39	{ArAgSxSm}	{SySxSm}	{SyAg}			259.7	134	<.001
40	{ArAgSxSm}	{SySxSm}	{SyAr}	{SyAg}		230.8	132	<.001
41	{ArAgSxSm}	{SyAgSm}	{SySxSm}			215.0	110	<.001
42	{ArAgSxSm}	{SyAgSm}	{SySxSm}	{SyAr}		183.6	108	<.001
43	{ArAgSxSm}	{SyAgSx}	{SySxSm}			239.1	128	<.001
44	{ArAgSxSm}	{SyAgSx}	{SySxSm}	{SyAr}		209.5	126	<.001
45	{ArAgSxSm}	{SyAgSx}	{SyAgSm}			207.3	112	<.001
46	{ArAgSxSm}	{SyAgSx}	{SyAgSm}	{SyAr}		175.0	110	<.001
47	{ArAgSxSm}	{SyArSm}	{SySxSm}			302.2	130	<.001
48	{ArAgSxSm}	{SyArSm}	{SySxSm}	{SyAg}		214.5	124	<.001
49	{ArAgSxSm}	{SyArSm}	{SyAgSm}			221.8	110	<.001
50	{ArAgSxSm}	{SyArSm}	{SyAgSm}	{SySx}		182.8	108	<.001

(Cont'd)

NOTE. Note. Variables are:
Ar: Area, 2 categories (SMSA and South Boston)
Ag: Age, 4 categories (18-34, 35-49, 50-64, 65+)
Sx: Sex, 2 Categories (Male and Female)
Sm: Smoking, 4 categories (Never Smoked Cigarettes; Ex Cigarette Smoker; 1-15 , 16-30, and 31+ Cigs. per day)
Sy: Symptom, Dependent Variable

L^2 is Likelihood-ratio chi-square statistic for fit of model

TABLE C-3: SHORTNESS OF BREATH
(Cont'd)

Model	Margins Fit						TEST OF MODEL		
							L^2	df	p
51	{ArAgSxSm}	{SyArSm}	{SyAgSx}				213.5	126	<.001
52	{ArAgSxSm}	{SyArSx}	{SySxSm}				316.4	136	<.001
53	{ArAgSxSm}	{SyArSx}	{SySxSm}	{SyAg}			229.6	130	<.001
54	{ArAgSxSm}	{SyArSx}	{SyAgSm}				197.5	114	<.001
55	{ArAgSxSm}	{SyArSx}	{SyAgSx}				384.9	140	<.001
56	{ArAgSxSm}	{SyArSx}	{SyAgSx}	{SySm}			226.9	132	<.001
57	{ArAgSxSm}	{SyArSx}	{SyArSm}				320.0	136	<.001
58	{ArAgSxSm}	{SyArSx}	{SyArSm}	{SyAg}			231.9	130	<.001
59	{ArAgSxSm}	{SyArAg}	{SySxSm}				212.0	126	<.001
60	{ArAgSxSm}	{SyArAg}	{SyAgSm}				222.7	112	<.001
61	{ArAgSxSm}	{SyArAg}	{SyAgSm}	{SySx}			182.4	110	<.001
62	{ArAgSxSm}	{SyArAg}	{SyAgSx}				368.4	136	<.001
63	{ArAgSxSm}	{SyArAg}	{SyAgSx}	{SySm}			209.6	128	<.001
64	{ArAgSxSm}	{SyArAg}	{SyArSm}				258.6	128	<.001
65	{ArAgSxSm}	{SyArAg}	{SyArSm}	{SySx}			213.0	126	<.001
66	{ArAgSxSm}	{SyArAg}	{SyArSx}				393.7	140	<.001
67	{ArAgSxSm}	{SyArAg}	{SyArSx}	{SySm}			228.6	132	<.001
68	{ArAgSxSm}	{SyAgSx}	{SyAgSm}	{SySxSm}			192.2	104	<.001
69	{ArAgSxSm}	{SyAgSx}	{SyAgSm}	{SySxSm}	{SyAr}		160.2	102	<.001
70	{ArAgSxSm}	{SyArSm}	{SyAgSm}	{SySxSm}			168.7	100	<.001
71	{ArAgSxSm}	{SyArSm}	{SyAgSx}	{SySxSm}			194.7	118	<.001
72	{ArAgSxSm}	{SyArSm}	{SyAgSx}	{SyAgSm}			161.3	102	<.001
73	{ArAgSxSm}	{SyArSx}	{SyAgSm}	{SySxSm}			182.9	106	<.001
74	{ArAgSxSm}	{SyArSx}	{SyAgSx}	{SySxSm}			207.9	124	<.001
75	{ArAgSxSm}	{SyArSx}	{SyAgSx}	{SyAgSm}			173.7	108	<.001
76	{ArAgSxSm}	{SyArSx}	{SyArSm}	{SySxSm}			302.0	128	<.001
[illegible]	{ArAgSxSm}	{SyArSx}	{SyArSm}	{SySxSm}	{SyAg}		213.9	122	<.001
[illegible]	{ArAgSxSm}	{SyArSx}	{SyArSm}	{SyAgSm}			182.4	106	<.001
79	{ArAgSxSm}	{SyArSx}	{SyArSm}	{SyAgSx}			212.6	124	<.001
80	{ArAgSxSm}	{SyArAg}	{SyAgSm}	{SySxSm}			168.1	102	<.001
81	{ArAgSxSm}	{SyArAg}	{SyAgSx}	{SySxSm}			191.4	120	<.001
82	{ArAgSxSm}	{SyArAg}	{SyAgSx}	{SyAgSm}			160.2	104	<.001
83	{ArAgSxSm}	{SyArAg}	{SyArSm}	{SySxSm}			195.7	118	<.001
84	{ArAgSxSm}	{SyArAg}	{SyArSm}	{SyAgSm}			208.7	104	<.001
85	{ArAgSxSm}	{SyArAg}	{SyArSm}	{SyAgSm}	{SySx}		168.7	102	<.001
86	{ArAgSxSm}	{SyArAg}	{SyArSm}	{SyAgSx}			194.6	120	<.001
87	{ArAgSxSm}	{SyArAg}	{SyArSx}	{SySxSm}			211.2	124	<.001
88	{ArAgSxSm}	{SyArAg}	{SyArSx}	{SyAgSm}			181.9	108	<.001
89	{ArAgSxSm}	{SyArAg}	{SyArSx}	{SyAgSx}			367.1	134	<.001
90	{ArAgSxSm}	{SyArAg}	{SyArSx}	{SyAgSx}	{SySm}		208.1	126	<.001
91	{ArAgSxSm}	{SyArAg}	{SyArSx}	{SyArSm}			212.6	124	<.001
92	{ArAgSxSm}	{SyArSm}	{SyAgSx}	{SyAgSm}	{SySxSm}		146.7	94	.001
93	{ArAgSxSm}	{SyArSx}	{SyAgSx}	{SyAgSm}	{SySxSm}		158.9	100	<.001
94	{ArAgSxSm}	{SyArSx}	{SyArSm}	{SyAgSm}	{SySxSm}		168.2	98	<.001
95	{ArAgSxSm}	{SyArSx}	{SyArSm}	{SyAgSx}	{SySxSm}		193.5	116	<.001
96	{ArAgSxSm}	{SyArSx}	{SyArSm}	{SyAgSx}	{SyAgSm}		160.4	100	<.001
97	{ArAgSxSm}	{SyArAg}	{SyAgSx}	{SyAgSm}	{SySxSm}		145.5	96	.001
98	{ArAgSxSm}	{SyArAg}	{SyArSm}	{SyAgSm}	{SySxSm}		154.6	94	<.001
99	{ArAgSxSm}	{SyArAg}	{SyArSm}	{SyAgSx}	{SySxSm}		176.4	112	<.001
100	{ArAgSxSm}	{SyArAg}	{SyArSm}	{SyAgSx}	{SyAgSm}		147.9	96	.001
101	{ArAgSxSm}	{SyArAg}	{SyArSx}	{SyAgSm}	{SySxSm}		167.5	100	<.001
102	{ArAgSxSm}	{SyArAg}	{SyArSx}	{SyAgSx}	{SySxSm}		189.7	118	<.001
103	{ArAgSxSm}	{SyArAg}	{SyArSx}	{SyAgSx}	{SyAgSm}		158.9	102	<.001
104	{ArAgSxSm}	{SyArAg}	{SyArSx}	{SyArSm}	{SySxSm}		195.2	116	<.001
105	{ArAgSxSm}	{SyArAg}	{SyArSx}	{SyArSm}	{SyAgSm}		168.4	100	<.001
106	{ArAgSxSm}	{SyArAg}	{SyArSx}	{SyArSm}	{SyAgSx}		193.5	118	<.001
107	{ArAgSxSm}	{SyArAg}	{SyArSx}	{SyArSm}	{SyAgSx}	{SySxSm}	175.0	110	<.001
108	{ArAgSxSm}	{SyAgSxSm}					140.1	80	<.001
109	{ArAgSxSm}	{SyAgSxSm}	{SyAr}				107.1	78	.016
[illegible]	{ArAgSxSm}	{SyAgSxSm}	{SyArAg}				93.0	72	.048

(Cont'd)

TABLE C-3 : SHORTNESS OF BREATH
(Cont'd)

						TEST OF MODEL		
Model	Margins Fit					L^2	df	p
11	{ArAgSxSm}	{SyAgSxSm}	{SyArSx}			106.0	76	.013
12	{ArAgSxSm}	{SyAgSxSm}	{SyArSm}			93.4	70	.032
113	{ArAgSxSm}	{SyAgSxSm}	{SyArAg}	{SyArSx}		92.1	70	.039
114*	{ArAgSxSm}	{SyAgSxSm}	{SyArAg}	{SyArSm}		80.8	64	.077
15	{ArAgSxSm}	{SyAgSxSm}	{SyArSx}	{SyArSm}		92.6	68	.025
16*	{ArAgSxSm}	{SyAgSxSm}	{SyArAg}	{SyArSx}	{SyArSm}	80.1	62	.061
117	{ArAgSxSm}	{SyArSxSm}				296.3	120	<.001
118	{ArAgSxSm}	{SyArSxSm}	{SyAg}			208.0	114	<.001
19	{ArAgSxSm}	{SyArSxSm}	{SyArAg}			190.5	108	<.001
20	{ArAgSxSm}	{SyArSxSm}	{SyAgSx}			187.3	108	<.001
21	{ArAgSxSm}	{SyArSxSm}	{SyAgSm}			161.8	90	<.001
122	{ArAgSxSm}	{SyArSxSm}	{SyArAg}	{SyAgSx}		169.7	102	<.001
23	{ArAgSxSm}	{SyArSxSm}	{SyArAg}	{SyAgSm}		148.8	84	<.001
24	{ArAgSxSm}	{SyArSxSm}	{SyAgSx}	{SyAgSm}		139.0	84	<.001
25	{ArAgSxSm}	{SyArSxSm}	{SyArAg}	{SyAgSx}	{SyAgSm}	126.5	78	.001
126	{ArAgSxSm}	{SyArAgSm}				173.3	80	<.001
127	{ArAgSxSm}	{SyArAgSm}	{SySx}			133.0	78	<.001
28	{ArAgSxSm}	{SyArAgSm}	{SyArSx}			132.7	76	<.001
29	{ArAgSxSm}	{SyArAgSm}	{SyAgSx}			112.7	72	.002
130	{ArAgSxSm}	{SyArAgSm}	{SySxSm}			120.3	70	<.001
131	{ArAgSxSm}	{SyArAgSm}	{SyArSx}	{SyAgSx}		111.7	70	.001
32	{ArAgSxSm}	{SyArAgSm}	{SyArSx}	{SySxSm}		119.9	68	<.001
33	{ArAgSxSm}	{SyArAgSm}	{SyAgSx}	{SySxSm}		99.1	64	.003
34	{ArAgSxSm}	{SyArAgSm}	{SyArSx}	{SyAgSx}	{SySxSm}	98.3	62	.003
135	{ArAgSxSm}	{SyArAgSx}				357.1	128	<.001
136	{ArAgSxSm}	{SyArAgSx}	{SySm}			198.5	120	<.001
	{ArAgSxSm}	{SyArAgSx}	{SyArSm}			184.0	112	<.001
	{ArAgSxSm}	{SyArAgSx}	{SyAgSm}			149.5	96	.001
139	{ArAgSxSm}	{SyArAgSx}	{SySxSm}			180.0	112	<.001
140	{ArAgSxSm}	{SyArAgSx}	{SyArSm}	{SyAgSm}		137.7	88	.001
41	{ArAgSxSm}	{SyArAgSx}	{SyArSm}	{SySxSm}		165.4	104	<.001
42	{ArAgSxSm}	{SyArAgSx}	{SyAgSm}	{SySxSm}		135.4	88	.001
43	{ArAgSxSm}	{SyArAgSx}	{SyArSm}	{SyAgSm}	{SySxSm}	123.7	80	.001
144	{ArAgSxSm}	{SyArAgSm}	{SyArAgSx}			102.3	64	.002
145	{ArAgSxSm}	{SyArAgSm}	{SyArAgSx}	{SySxSm}		89.4	56	.003
46	{ArAgSxSm}	{SyArSxSm}	{SyArAgSx}			158.9	96	<.001
47	{ArAgSxSm}	{SyArSxSm}	{SyArAgSx}	{SyAgSm}		116.7	72	.001
148	{ArAgSxSm}	{SyArSxSm}	{SyArAgSm}			115.0	60	<.001
149	{ArAgSxSm}	{SyArSxSm}	{SyArAgSm}	{SyAgSx}		91.3	54	.001
50*	{ArAgSxSm}	{SyAgSxSm}	{SyArAgSx}			83.3	64	.053
51*	{ArAgSxSm}	{SyAgSxSm}	{SyArAgSx}	{SyArSm}		69.9	56	.100
52*	{ArAgSxSm}	{SyAgSxSm}	{SyArAgSm}			47.1	40	.206
153*	{ArAgSxSm}	{SyAgSxSm}	{SyArAgSm}	{SyArSx}		46.3	38	.168
54	{ArAgSxSm}	{SyAgSxSm}	{SyArSxSm}			85.0	60	.018
55	{ArAgSxSm}	{SyAgSxSm}	{SyArSxSm}	{SyArAg}		74.0	54	.037
56	{ArAgSxSm}	{SyArSxSm}	{SyArAgSm}	{SyArAgSx}		81.0	48	.002
157*	{ArAgSxSm}	{SyAgSxSm}	{SyArAgSm}	{SyArAgSx}		35.0	32	.326
158*	{ArAgSxSm}	{SyAgSxSm}	{SyArSxSm}	{SyArAgSx}		63.0	48	.072
59*	{ArAgSxSm}	{SyAgSxSm}	{SyArSxSm}	{SyArAgSm}		37.7	30	.157
60*	{ArAgSxSm}	{SyAgSxSm}	{SyArSxSm}	{SyArAgSm}	{SyArAgSx}	26.7	24	.320
161*	{SyArAgSxSm}					0.0	0	---

TABLE C-4

Tests of Alternative Log-linear Models fit to 5-way Table of Symptom by Area by Age by Sex by Smoking. (Data from Boston Area Pulmonary Health Survey, 1983.)

PULMONARY TUBERCULOSIS

Model	Margins Fit					TEST OF MODEL L^2	df	p
1*	{ArAgSxSm}	{Sy}				80.9	79	.420
2*	{ArAgSxSm}	{SySm}				76.8	75	.420
3*	{ArAgSxSm}	{SySx}				80.4	78	.405
4*	{ArAgSxSm}	{SyAg}				66.1	76	>.5
5*	{ArAgSxSm}	{SyAr}				66.2	78	>.5
6*	{ArAgSxSm}	{SySx}	{SySm}			76.3	74	.405
7*	{ArAgSxSm}	{SyAg}	{SySm}			62.6	72	>.5
8*	{ArAgSxSm}	{SyAg}	{SySx}			65.7	75	>.5
9*	{ArAgSxSm}	{SyAr}	{SySm}			63.4	74	>.5
10*	{ArAgSxSm}	{SyAr}	{SySx}			65.4	77	>.5
11*	{ArAgSxSm}	{SyAr}	{SyAg}			52.3	75	>.5
12*	{ArAgSxSm}	{SyAg}	{SySx}	{SySm}		62.1	71	>.5
13*	{ArAgSxSm}	{SyAr}	{SySx}	{SySm}		62.6	73	>.5
14*	{ArAgSxSm}	{SyAr}	{SyAg}	{SySm}		49.6	71	>.5
15*	{ArAgSxSm}	{SyAr}	{SyAg}	{SySx}		51.6	74	>.5
16*	{ArAgSxSm}	{SyAr}	{SyAg}	{SySx}	{SySm}	48.9	70	>.5
17*	{ArAgSxSm}	{SyArAg}				45.0	72	>.5
18*	{ArAgSxSm}	{SyArAg}	{SySx}			44.4	71	>.5
19*	{ArAgSxSm}	{SyArAg}	{SySm}			42.3	68	>.5
20*	{ArAgSxSm}	{SyArAg}	{SySx}	{SySm}		41.5	67	>.5
21*	{ArAgSxSm}	{SyArSx}				59.6	76	>.5
22*	{ArAgSxSm}	{SyArSx}	{SyAg}			45.6	73	>.5
23*	{ArAgSxSm}	{SyArSx}	{SySm}			56.8	72	>.5
24*	{ArAgSxSm}	{SyArSx}	{SyAg}	{SySm}		42.8	69	>.5
25*	{ArAgSxSm}	{SyArSm}				58.5	70	>.5
26*	{ArAgSxSm}	{SyArSm}	{SyAg}			44.6	67	>.5
27*	{ArAgSxSm}	{SyArSm}	{SySx}			57.7	69	>.5
28*	{ArAgSxSm}	{SyArSm}	{SyAg}	{SySx}		43.7	66	>.5
29*	{ArAgSxSm}	{SyAgSx}				60.9	72	>.5
30*	{ArAgSxSm}	{SyAgSx}	{SyAr}			46.7	71	>.5
31*	{ArAgSxSm}	{SyAgSx}	{SySm}			57.3	68	>.5
32*	{ArAgSxSm}	{SyAgSx}	{SyAr}	{SySm}		44.0	67	>.5
33*	{ArAgSxSm}	{SyAgSm}				52.6	60	>.5
34*	{ArAgSxSm}	{SyAgSm}	{SyAr}			39.7	59	>.5
35*	{ArAgSxSm}	{SyAgSm}	{SySx}			52.2	59	>.5
36*	{ArAgSxSm}	{SyAgSm}	{SyAr}	{SySx}		39.0	58	>.5
37*	{ArAgSxSm}	{SySxSm}				74.4	70	.336
38*	{ArAgSxSm}	{SySxSm}	{SyAr}			60.8	69	>.5
39*	{ArAgSxSm}	{SySxSm}	{SyAg}			60.7	67	>.5
40*	{ArAgSxSm}	{SySxSm}	{SyAr}	{SyAg}		47.5	66	>.5
41*	{ArAgSxSm}	{SyAgSm}	{SySxSm}			51.1	55	>.5
42*	{ArAgSxSm}	{SyAgSm}	{SySxSm}	{SyAr}		37.7	54	>.5
43*	{ArAgSxSm}	{SyAgSx}	{SySxSm}			55.9	64	>.5
44*	{ArAgSxSm}	{SyAgSx}	{SySxSm}	{SyAr}		42.5	63	>.5
45*	{ArAgSxSm}	{SyAgSx}	{SyAgSm}			47.8	56	>.5
46*	{ArAgSxSm}	{SyAgSx}	{SyAgSm}	{SyAr}		34.3	55	>.5
47*	{ArAgSxSm}	{SyArSm}	{SySxSm}			55.9	65	>.5
48*	{ArAgSxSm}	{SyArSm}	{SySxSm}	{SyAg}		42.3	62	>.5
49*	{ArAgSxSm}	{SyArSm}	{SyAgSm}			34.6	55	>.5
50*	{ArAgSxSm}	{SyArSm}	{SyAgSm}	{SySx}		34.0	54	>.5

(Cont'd)

NOTE. Note. Variables are:
Ar: Area, 2 categories (SMSA and South Boston)
Ag: Age, 4 categories (18-34, 35-49, 50-64, 65+)
Sx: Sex, 2 Categories (Male and Female)
Sm: Smoking, 4 categories (Never Smoked Cigarettes; Ex Cigarette Smoker; 1-15 , 16-30, and 31+ Cigs. per day)
Sy: Symptom, Dependent Variable

L^2 is Likelihood-ratio chi-square statistic for fit of model

TABLE C-4 PULMONARY TUBERCULOSIS
(Cont'd)

							TEST OF MODEL		
Model	Margins Fit						L^2	df	p
51*	{ArAgSxSm}	{SyArSm}	{SyAgSx}				38.8	63	>.5
52*	{ArAgSxSm}	{SyArSx}	{SySxSm}				55.0	68	>.5
53*	{ArAgSxSm}	{SyArSx}	{SySxSm}	{SyAg}			41.3	65	>.5
54*	{ArAgSxSm}	{SyArSx}	{SyAgSm}				32.8	57	>.5
55*	{ArAgSxSm}	{SyArSx}	{SyAgSx}				41.1	70	>.5
56*	{ArAgSxSm}	{SyArSx}	{SyAgSx}	{SySm}			38.3	66	>.5
57*	{ArAgSxSm}	{SyArSx}	{SyArSm}				51.9	68	>.5
58*	{ArAgSxSm}	{SyArSx}	{SyArSm}	{SyAg}			37.7	65	>.5
59*	{ArAgSxSm}	{SyArAg}	{SySxSm}				40.1	63	>.5
60*	{ArAgSxSm}	{SyArAg}	{SyAgSm}				31.1	56	>.5
61*	{ArAgSxSm}	{SyArAg}	{SyAgSm}	{SySx}			30.6	55	>.5
62*	{ArAgSxSm}	{SyArAg}	{SyAgSx}				39.5	68	>.5
63*	{ArAgSxSm}	{SyArAg}	{SyAgSx}	{SySm}			36.7	64	>.5
64*	{ArAgSxSm}	{SyArAg}	{SyArSm}				38.1	64	>.5
65*	{ArAgSxSm}	{SyArAg}	{SyArSm}	{SySx}			37.3	63	>.5
66*	{ArAgSxSm}	{SyArAg}	{SyArSx}				38.9	70	>.5
67*	{ArAgSxSm}	{SyArAg}	{SyArSx}	{SySm}			35.9	66	>.5
68*	{ArAgSxSm}	{SyAgSx}	{SyAgSm}	{SySxSm}			46.3	52	>.5
69*	{ArAgSxSm}	{SyAgSx}	{SyAgSm}	{SySxSm}	{SyAr}		32.6	51	>.5
70*	{ArAgSxSm}	{SyArSm}	{SyAgSm}	{SySxSm}			32.6	50	>.5
71*	{ArAgSxSm}	{SyArSm}	{SyAgSx}	{SySxSm}			37.3	59	>.5
72*	{ArAgSxSm}	{SyArSm}	{SyAgSx}	{SyAgSm}			29.2	51	>.5
73*	{ArAgSxSm}	{SyArSx}	{SyAgSm}	{SySxSm}			31.4	53	>.5
74*	{ArAgSxSm}	{SyArSx}	{SyAgSx}	{SySxSm}			36.9	62	>.5
75*	{ArAgSxSm}	{SyArSx}	{SyAgSx}	{SyAgSm}			28.4	54	>.5
76*	{ArAgSxSm}	{SyArSx}	{SyArSm}	{SySxSm}			49.4	64	>.5
77*	{ArAgSxSm}	{SyArSx}	{SyArSm}	{SySxSm}	{SyAg}		35.6	61	>.5
78*	{ArAgSxSm}	{SyArSx}	{SyArSm}	{SyAgSm}			27.6	53	>.5
79*	{ArAgSxSm}	{SyArSx}	{SyArSm}	{SyAgSx}			33.0	62	>.5
80*	{ArAgSxSm}	{SyArAg}	{SyAgSm}	{SySxSm}			29.3	51	>.5
81*	{ArAgSxSm}	{SyArAg}	{SyAgSx}	{SySxSm}			35.2	60	>.5
82*	{ArAgSxSm}	{SyArAg}	{SyAgSx}	{SyAgSm}			26.3	52	>.5
83*	{ArAgSxSm}	{SyArAg}	{SyArSm}	{SySxSm}			35.9	59	>.5
84*	{ArAgSxSm}	{SyArAg}	{SyArSm}	{SyAgSm}			24.3	52	>.5
85*	{ArAgSxSm}	{SyArAg}	{SyArSm}	{SyAgSm}	{SySx}		23.8	51	>.5
86*	{ArAgSxSm}	{SyArAg}	{SyArSm}	{SyAgSx}			32.4	60	>.5
87*	{ArAgSxSm}	{SyArAg}	{SyArSx}	{SySxSm}			34.5	62	>.5
88*	{ArAgSxSm}	{SyArAg}	{SyArSx}	{SyAgSm}			25.3	54	>.5
89*	{ArAgSxSm}	{SyArAg}	{SyArSx}	{SyAgSx}			32.9	67	>.5
90*	{ArAgSxSm}	{SyArAg}	{SyArSx}	{SyAgSx}	{SySm}		30.1	63	>.5
91*	{ArAgSxSm}	{SyArAg}	{SyArSx}	{SyArSm}			31.9	62	>.5
92*	{ArAgSxSm}	{SyArSm}	{SyAgSx}	{SyAgSm}	{SySxSm}		27.5	47	>.5
93*	{ArAgSxSm}	{SyArSx}	{SyAgSx}	{SyAgSm}	{SySxSm}		26.8	50	>.5
94*	{ArAgSxSm}	{SyArSx}	{SyArSm}	{SyAgSm}	{SySxSm}		25.6	49	>.5
95*	{ArAgSxSm}	{SyArSx}	{SyArSm}	{SyAgSx}	{SySxSm}		31.3	58	>.5
96*	{ArAgSxSm}	{SyArSx}	{SyArSm}	{SyAgSx}	{SyAgSm}		22.9	50	>.5
97*	{ArAgSxSm}	{SyArAg}	{SyAgSx}	{SyAgSm}	{SySxSm}		24.7	48	>.5
98*	{ArAgSxSm}	{SyArAg}	{SyArSm}	{SyAgSm}	{SySxSm}		22.5	47	>.5
99*	{ArAgSxSm}	{SyArAg}	{SyArSm}	{SyAgSx}	{SySxSm}		30.9	56	>.5
100*	{ArAgSxSm}	{SyArAg}	{SyArSm}	{SyAgSx}	{SyAgSm}		19.3	48	>.5
101*	{ArAgSxSm}	{SyArAg}	{SyArSx}	{SyAgSm}	{SySxSm}		23.9	50	>.5
102*	{ArAgSxSm}	{SyArAg}	{SyArSx}	{SyAgSx}	{SySxSm}		28.6	59	>.5
103*	{ArAgSxSm}	{SyArAg}	{SyArSx}	{SyAgSx}	{SyAgSm}		19.6	51	>.5
104*	{ArAgSxSm}	{SyArAg}	{SyArSx}	{SyArSm}	{SySxSm}		30.0	58	>.5
105*	{ArAgSxSm}	{SyArAg}	{SyArSx}	{SyArSm}	{SyAgSm}		17.9	50	>.5
106*	{ArAgSxSm}	{SyArAg}	{SyArSx}	{SyArSm}	{SyAgSx}		25.7	59	>.5
107*	{ArAgSxSm}	{SyArAg}	{SyArSx}	{SyArSm}	{SyAgSx}	{SySxSm}	23.9	55	>.5
108*	{ArAgSxSm}	{SyAgSxSm}					39.1	40	>.5
109*	{ArAgSxSm}	{SyAgSxSm}	{SyAr}				25.2	39	>.5
110*	{ArAgSxSm}	{SyAgSxSm}	{SyArAg}				17.9	36	>.5

(Cont'd)

TABLE C-4: PULMONARY TUBERCULOSIS
(Cont'd)

Model	Margins Fit					TEST OF MODEL		
						L²	df	p
111*	{ArAgSxSm}	{SyAgSxSm}	{SyArSx}			19.8	38	>.5
112*	{ArAgSxSm}	{SyAgSxSm}	{SyArSm}			20.1	35	>.5
113*	{ArAgSxSm}	{SyAgSxSm}	{SyArAg}	{SyArSx}		11.6	35	>.5
114*	{ArAgSxSm}	{SyAgSxSm}	{SyArAg}	{SyArSm}		10.9	32	>.5
115*	{ArAgSxSm}	{SyAgSxSm}	{SyArSx}	{SyArSm}		14.3	34	>.5
116*	{ArAgSxSm}	{SyAgSxSm}	{SyArAg}	{SyArSx}	{SyArSm}	3.3	31	>.5
117*	{ArAgSxSm}	{SyArSxSm}				41.1	60	>.5
118*	{ArAgSxSm}	{SyArSxSm}	{SyAg}			27.4	57	>.5
119*	{ArAgSxSm}	{SyArSxSm}	{SyArAg}			21.9	54	>.5
120*	{ArAgSxSm}	{SyArSxSm}	{SyAgSx}			23.0	54	>.5
121*	{ArAgSxSm}	{SyArSxSm}	{SyAgSm}			17.8	45	>.5
122*	{ArAgSxSm}	{SyArSxSm}	{SyArAg}	{SyAgSx}		15.2	51	>.5
123*	{ArAgSxSm}	{SyArSxSm}	{SyArAg}	{SyAgSm}		7.9	42	>.5
124*	{ArAgSxSm}	{SyArSxSm}	{SyAgSx}	{SyAgSm}		13.0	42	>.5
125*	{ArAgSxSm}	{SyArSxSm}	{SyArAg}	{SyAgSx}	{SyAgSm}	5.1	39	>.5
126*	{ArAgSxSm}	{SyArAgSm}				23.0	40	>.5
127*	{ArAgSxSm}	{SyArAgSm}	{SySx}			22.5	39	>.5
128*	{ArAgSxSm}	{SyArAgSm}	{SyArSx}			16.6	38	>.5
129*	{ArAgSxSm}	{SyArAgSm}	{SyAgSx}			17.9	36	>.5
130*	{ArAgSxSm}	{SyArAgSm}	{SySxSm}			21.3	35	>.5
131*	{ArAgSxSm}	{SyArAgSm}	{SyArSx}	{SyAgSx}		10.8	35	>.5
132*	{ArAgSxSm}	{SyArAgSm}	{SyArSx}	{SySxSm}		14.9	34	>.5
133*	{ArAgSxSm}	{SyArAgSm}	{SyAgSx}	{SySxSm}		16.4	32	>.5
134*	{ArAgSxSm}	{SyArAgSm}	{SyArSx}	{SyAgSx}	{SySxSm}	8.3	31	>.5
135*	{ArAgSxSm}	{SyArAgSx}				32.6	64	>.5
136*	{ArAgSxSm}	{SyArAgSx}	{SySm}			29.8	60	>.5
137*	{ArAgSxSm}	{SyArAgSx}	{SyArSm}			25.4	56	>.5
138*	{ArAgSxSm}	{SyArAgSx}	{SyAgSm}			19.3	48	>.5
139*	{ArAgSxSm}	{SyArAgSx}	{SySxSm}			28.3	56	>.5
140*	{ArAgSxSm}	{SyArAgSx}	{SyArSm}	{SyAgSm}		11.8	44	>.5
141*	{ArAgSxSm}	{SyArAgSx}	{SyArSm}	{SySxSm}		23.6	52	>.5
142*	{ArAgSxSm}	{SyArAgSx}	{SyAgSm}	{SySxSm}		17.7	44	>.5
143*	{ArAgSxSm}	{SyArAgSx}	{SyArSm}	{SyAgSm}	{SySxSm}	9.1	40	>.5
144*	{ArAgSxSm}	{SyArAgSm}	{SyArAgSx}			10.5	32	>.5
145*	{ArAgSxSm}	{SyArAgSm}	{SyArAgSx}	{SySxSm}		7.6	28	>.5
146*	{ArAgSxSm}	{SyArSxSm}	{SyArAgSx}			14.9	48	>.5
147*	{ArAgSxSm}	{SyArSxSm}	{SyArAgSx}	{SyAgSm}		5.0	36	>.5
148*	{ArAgSxSm}	{SyArSxSm}	{SyArAgSm}			6.6	30	>.5
149*	{ArAgSxSm}	{SyArSxSm}	{SyArAgSm}	{SyAgSx}		4.3	27	>.5
150*	{ArAgSxSm}	{SyAgSxSm}	{SyArAgSx}			11.3	32	>.5
151*	{ArAgSxSm}	{SyAgSxSm}	{SyArAgSx}	{SyArSm}		2.3	28	>.5
152*	{ArAgSxSm}	{SyAgSxSm}	{SyArAgSm}			9.7	20	>.5
153*	{ArAgSxSm}	{SyAgSxSm}	{SyArAgSm}	{SyArSx}		1.6	19	>.5
154*	{ArAgSxSm}	{SyAgSxSm}	{SyArSxSm}			6.0	30	>.5
155*	{ArAgSxSm}	{SyAgSxSm}	{SyArSxSm}	{SyArAg}		1.1	27	>.5
156*	{ArAgSxSm}	{SyArSxSm}	{SyArAgSm}	{SyArAgSx}		4.2	24	>.5
157*	{ArAgSxSm}	{SyAgSxSm}	{SyArAgSm}	{SyArAgSx}		0.9	16	>.5
158*	{ArAgSxSm}	{SyAgSxSm}	{SyArSxSm}	{SyArAgSx}		1.0	24	>.5
159*	{ArAgSxSm}	{SyAgSxSm}	{SyArSxSm}	{SyArAgSm}		0.6	15	>.5
160*	{ArAgSxSm}	{SyAgSxSm}	{SyArSxSm}	{SyArAgSm}	{SyArAgSx}	0.4	12	>.5
161*	{SyArAgSxSm}					0.0	0	---

TABLE C-5

Tests of Alternative Log-linear Models fit to 5-way Table of Symptom by Area by Age by Sex by Smoking. (Data from Boston Area Pulmonary Health Survey, 1983.)

CHEST COLDS

Model	Margins Fit	TEST OF MODEL L^2	df	p
1	{ArAgSxSm} {Sy}	177.9	79	<.001
2	{ArAgSxSm} {SySm}	164.6	75	<.001
3	{ArAgSxSm} {SySx}	157.6	78	<.001
4	{ArAgSxSm} {SyAg}	131.5	76	<.001
5	{ArAgSxSm} {SyAr}	177.0	78	<.001
6	{ArAgSxSm} {SySx} {SySm}	141.7	74	<.001
7	{ArAgSxSm} {SyAg} {SySm}	116.9	72	.001
8	{ArAgSxSm} {SyAg} {SySx}	107.6	75	.008
9	{ArAgSxSm} {SyAr} {SySm}	163.8	74	<.001
10	{ArAgSxSm} {SyAr} {SySx}	157.1	77	<.001
11	{ArAgSxSm} {SyAr} {SyAg}	130.6	75	<.001
12*	{ArAgSxSm} {SyAg} {SySx} {SySm}	89.7	71	.066
13	{ArAgSxSm} {SyAr} {SySx} {SySm}	141.2	73	<.001
14	{ArAgSxSm} {SyAr} {SyAg} {SySm}	116.0	71	.001
15	{ArAgSxSm} {SyAr} {SyAg} {SySx}	106.9	74	.007
16*	{ArAgSxSm} {SyAr} {SyAg} {SySx} {SySm}	89.0	70	.062
17	{ArAgSxSm} {SyArAg}	128.5	72	<.001
18	{ArAgSxSm} {SyArAg} {SySx}	105.2	71	.005
19	{ArAgSxSm} {SyArAg} {SySm}	113.7	68	.001
20	{ArAgSxSm} {SyArAg} {SySx} {SySm}	87.1	67	.050
21	{ArAgSxSm} {SyArSx}	152.4	76	<.001
22	{ArAgSxSm} {SyArSx} {SyAg}	103.1	73	.012
23	{ArAgSxSm} {SyArSx} {SySm}	136.4	72	<.001
24*	{ArAgSxSm} {SyArSx} {SyAg} {SySm}	85.0	69	.092
25	{ArAgSxSm} {SyArSm}	159.8	70	<.001
26	{ArAgSxSm} {SyArSm} {SyAg}	112.0	67	.001
27	{ArAgSxSm} {SyArSm} {SySx}	137.0	69	<.001
28*	{ArAgSxSm} {SyArSm} {SyAg} {SySx}	84.9	66	.058
29	{ArAgSxSm} {SyAgSx}	99.1	72	.019
30	{ArAgSxSm} {SyAgSx} {SyAr}	98.6	71	.017
31*	{ArAgSxSm} {SyAgSx} {SySm}	81.6	68	.124
32*	{ArAgSxSm} {SyAgSx} {SyAr} {SySm}	81.1	67	.116
33	{ArAgSxSm} {SyAgSm}	101.9	60	.001
34	{ArAgSxSm} {SyAgSm} {SyAr}	100.9	59	.001
35*	{ArAgSxSm} {SyAgSm} {SySx}	73.3	59	.099
36*	{ArAgSxSm} {SyAgSm} {SyAr} {SySx}	72.5	58	.095
37	{ArAgSxSm} {SySxSm}	119.6	70	<.001
38	{ArAgSxSm} {SySxSm} {SyAr}	119.1	69	<.001
39*	{ArAgSxSm} {SySxSm} {SyAg}	72.9	67	.289
40*	{ArAgSxSm} {SySxSm} {SyAr} {SyAg}	72.4	66	.276
41*	{ArAgSxSm} {SyAgSm} {SySxSm}	57.8	55	.372
42*	{ArAgSxSm} {SyAgSm} {SySxSm} {SyAr}	57.2	54	.357
43*	{ArAgSxSm} {SyAgSx} {SySxSm}	66.3	64	.398
44*	{ArAgSxSm} {SyAgSx} {SySxSm} {SyAr}	65.8	63	.379
45*	{ArAgSxSm} {SyAgSx} {SyAgSm}	65.0	56	.191
46*	{ArAgSxSm} {SyAgSx} {SyAgSm} {SyAr}	64.4	55	.181
47	{ArAgSxSm} {SyArSm} {SySxSm}	114.7	65	<.001
48*	{ArAgSxSm} {SyArSm} {SySxSm} {SyAg}	68.1	62	.279
49	{ArAgSxSm} {SyArSm} {SyAgSm}	96.6	55	.001
50*	{ArAgSxSm} {SyArSm} {SyAgSm} {SySx}	68.0	54	.095

(Cont'd)

NOTE. Note. Variables are:
Ar: Area, 2 categories (SMSA and South Boston)
Ag: Age, 4 categories (18-34, 35-49, 50-64, 65+)
Sx: Sex, 2 Categories (Male and Female)
Sm: Smoking, 4 categories (Never Smoked Cigarettes; Ex Cigarette Smoker; 1-15 , 16-30, and 31+ Cigs. per day)
Sy: Symptom, Dependent Variable

L^2 is Likelihood-ratio chi-square statistic for fit of model

TABLE C-5: CHEST COLDS
(Cont'd)

Model	Margins Fit						TEST OF MODEL		
							L^2	df	p
51*	{ArAgSxSm}	{SyArSm}	{SyAgSx}				76.6	63	.116
52	{ArAgSxSm}	{SyArSx}	{SySxSm}				111.4	68	.001
53*	{ArAgSxSm}	{SyArSx}	{SySxSm}	{SyAg}			66.0	65	.441
54*	{ArAgSxSm}	{SyArSx}	{SyAgSm}				68.1	57	.149
55	{ArAgSxSm}	{SyArSx}	{SyAgSx}				94.4	70	.027
56*	{ArAgSxSm}	{SyArSx}	{SyAgSx}	{SySm}			76.8	66	.171
57	{ArAgSxSm}	{SyArSx}	{SyArSm}				131.9	68	<.001
58*	{ArAgSxSm}	{SyArSx}	{SyArSm}	{SyAg}			80.6	65	.092
59*	{ArAgSxSm}	{SyArAg}	{SySxSm}				70.3	63	.246
60	{ArAgSxSm}	{SyArAg}	{SyAgSm}				98.6	56	.001
61*	{ArAgSxSm}	{SyArAg}	{SyAgSm}	{SySx}			70.7	55	.075
62	{ArAgSxSm}	{SyArAg}	{SyAgSx}				97.0	68	.012
63*	{ArAgSxSm}	{SyArAg}	{SyAgSx}	{SySm}			79.4	64	.093
64	{ArAgSxSm}	{SyArAg}	{SyArSm}				109.8	64	<.001
65	{ArAgSxSm}	{SyArAg}	{SyArSm}	{SySx}			83.1	63	.046
66	{ArAgSxSm}	{SyArAg}	{SyArSx}				101.1	70	.009
67*	{ArAgSxSm}	{SyArAg}	{SyArSx}	{SySm}			82.8	66	.079
68*	{ArAgSxSm}	{SyAgSx}	{SyAgSm}	{SySxSm}			51.2	52	>.5
69*	{ArAgSxSm}	{SyAgSx}	{SyAgSm}	{SySxSm}	{SyAr}		50.7	51	.484
70*	{ArAgSxSm}	{SyArSm}	{SyAgSm}	{SySxSm}			52.4	50	.379
71*	{ArAgSxSm}	{SyArSm}	{SyAgSx}	{SySxSm}			61.3	59	.395
72*	{ArAgSxSm}	{SyArSm}	{SyAgSx}	{SyAgSm}			59.5	51	.194
73*	{ArAgSxSm}	{SyArSx}	{SyAgSm}	{SySxSm}			50.4	53	>.5
74*	{ArAgSxSm}	{SyArSx}	{SyAgSx}	{SySxSm}			59.1	62	>.5
75*	{ArAgSxSm}	{SyArSx}	{SyAgSx}	{SyAgSm}			59.7	54	.275
76	{ArAgSxSm}	{SyArSx}	{SyArSm}	{SySxSm}			106.5	64	.001
*	{ArAgSxSm}	{SyArSx}	{SyArSm}	{SySxSm}	{SyAg}		61.3	61	.465
*	{ArAgSxSm}	{SyArSx}	{SyArSm}	{SyAgSm}			63.3	53	.157
79*	{ArAgSxSm}	{SyArSx}	{SyArSm}	{SyAgSx}			72.1	62	.179
80*	{ArAgSxSm}	{SyArAg}	{SyAgSm}	{SySxSm}			55.2	51	.318
81*	{ArAgSxSm}	{SyArAg}	{SyAgSx}	{SySxSm}			64.1	60	.336
82*	{ArAgSxSm}	{SyArAg}	{SyAgSx}	{SyAgSm}			62.8	52	.146
83*	{ArAgSxSm}	{SyArAg}	{SyArSm}	{SySxSm}			66.1	59	.245
84	{ArAgSxSm}	{SyArAg}	{SyArSm}	{SyAgSm}			94.3	52	<.001
85*	{ArAgSxSm}	{SyArAg}	{SyArSm}	{SyAgSm}	{SySx}		66.2	51	.075
86*	{ArAgSxSm}	{SyArAg}	{SyArSm}	{SyAgSx}			75.1	60	.091
87*	{ArAgSxSm}	{SyArAg}	{SyArSx}	{SySxSm}			63.5	62	.424
88*	{ArAgSxSm}	{SyArAg}	{SyArSx}	{SyAgSm}			66.0	54	.127
89	{ArAgSxSm}	{SyArAg}	{SyArSx}	{SyAgSx}			92.4	67	.021
90*	{ArAgSxSm}	{SyArAg}	{SyArSx}	{SyAgSx}	{SySm}		74.6	63	.150
91*	{ArAgSxSm}	{SyArAg}	{SyArSx}	{SyArSm}			78.4	62	.078
92*	{ArAgSxSm}	{SyArSm}	{SyAgSx}	{SyAgSm}	{SySxSm}		45.7	47	>.5
93*	{ArAgSxSm}	{SyArSx}	{SyAgSx}	{SyAgSm}	{SySxSm}		43.6	50	>.5
94*	{ArAgSxSm}	{SyArSx}	{SyArSm}	{SyAgSm}	{SySxSm}		45.3	49	>.5
95*	{ArAgSxSm}	{SyArSx}	{SyArSm}	{SyAgSx}	{SySxSm}		54.1	58	>.5
96*	{ArAgSxSm}	{SyArSx}	{SyArSm}	{SyAgSx}	{SyAgSm}		54.6	50	.305
97*	{ArAgSxSm}	{SyArAg}	{SyAgSx}	{SyAgSm}	{SySxSm}		48.9	48	.435
98*	{ArAgSxSm}	{SyArAg}	{SyArSm}	{SyAgSm}	{SySxSm}		50.4	47	.339
99*	{ArAgSxSm}	{SyArAg}	{SyArSm}	{SyAgSx}	{SySxSm}		59.6	56	.345
100*	{ArAgSxSm}	{SyArAg}	{SyArSm}	{SyAgSx}	{SyAgSm}		57.9	48	.155
101*	{ArAgSxSm}	{SyArAg}	{SyArSx}	{SyAgSm}	{SySxSm}		48.0	50	>.5
102*	{ArAgSxSm}	{SyArAg}	{SyArSx}	{SyAgSx}	{SySxSm}		56.6	59	>.5
103*	{ArAgSxSm}	{SyArAg}	{SyArSx}	{SyAgSx}	{SyAgSm}		57.7	51	.242
104*	{ArAgSxSm}	{SyArAg}	{SyArSx}	{SyArSm}	{SySxSm}		58.8	58	.445
105*	{ArAgSxSm}	{SyArAg}	{SyArSx}	{SyArSm}	{SyAgSm}		61.2	50	.134
106*	{ArAgSxSm}	{SyArAg}	{SyArSx}	{SyArSm}	{SyAgSx}		70.0	59	.155
107*	{ArAgSxSm}	{SyArAg}	{SyArSx}	{SyArSm}	{SyAgSx}	{SySxSm}	51.8	55	>.5
108*	{ArAgSxSm}	{SyAgSxSm}					44.9	40	.276
109*	{ArAgSxSm}	{SyAgSxSm}	{SyAr}				44.4	39	.256
[illegible]*	{ArAgSxSm}	{SyAgSxSm}	{SyArAg}				42.5	36	.211

(Cont'd)

TABLE C-5: CHEST COLDS
Cont'd)

Model	Margins Fit					TEST OF MODEL L^2	df	p
111*	{ArAgSxSm}	{SyAgSxSm}	{SyArSx}			37.3	38	>.5
112*	{ArAgSxSm}	{SyAgSxSm}	{SyArSm}			39.6	35	.274
113*	{ArAgSxSm}	{SyAgSxSm}	{SyArAg}	{SyArSx}		34.9	35	.474
114*	{ArAgSxSm}	{SyAgSxSm}	{SyArAg}	{SyArSm}		37.7	32	.223
115*	{ArAgSxSm}	{SyAgSxSm}	{SyArSx}	{SyArSm}		32.2	34	>.5
116*	{ArAgSxSm}	{SyAgSxSm}	{SyArAg}	{SyArSx}	{SyArSm}	29.8	31	>.5
117	{ArAgSxSm}	{SyArSxSm}				104.1	60	.001
118*	{ArAgSxSm}	{SyArSxSm}	{SyAg}			59.4	57	.388
119*	{ArAgSxSm}	{SyArSxSm}	{SyArAg}			57.0	54	.365
120*	{ArAgSxSm}	{SyArSxSm}	{SyAgSx}			52.3	54	>.5
121*	{ArAgSxSm}	{SyArSxSm}	{SyAgSm}			43.4	45	>.5
122*	{ArAgSxSm}	{SyArSxSm}	{SyArAg}	{SyAgSx}		49.9	51	>.5
123*	{ArAgSxSm}	{SyArSxSm}	{SyArAg}	{SyAgSm}		41.2	42	>.5
124*	{ArAgSxSm}	{SyArSxSm}	{SyAgSx}	{SyAgSm}		36.4	42	>.5
125*	{ArAgSxSm}	{SyArSxSm}	{SyArAg}	{SyAgSx}	{SyAgSm}	34.2	39	>.5
126	{ArAgSxSm}	{SyArAgSm}				84.3	40	<.001
127	{ArAgSxSm}	{SyArAgSm}	{SySx}			56.0	39	.038
128*	{ArAgSxSm}	{SyArAgSm}	{SyArSx}			51.0	38	.077
129*	{ArAgSxSm}	{SyArAgSm}	{SyAgSx}			48.3	36	.083
130*	{ArAgSxSm}	{SyArAgSm}	{SySxSm}			39.8	35	.264
131*	{ArAgSxSm}	{SyArAgSm}	{SyArSx}	{SyAgSx}		43.0	35	.167
132*	{ArAgSxSm}	{SyArAgSm}	{SyArSx}	{SySxSm}		31.9	34	>.5
133*	{ArAgSxSm}	{SyArAgSm}	{SyAgSx}	{SySxSm}		33.9	32	.378
134*	{ArAgSxSm}	{SyArAgSm}	{SyArSx}	{SyAgSx}	{SySxSm}	25.5	31	>.5
135	{ArAgSxSm}	{SyArAgSx}				88.4	64	.023
136*	{ArAgSxSm}	{SyArAgSx}	{SySm}			71.2	60	.152
	{ArAgSxSm}	{SyArAgSx}	{SyArSm}			66.4	56	.160
	{ArAgSxSm}	{SyArAgSx}	{SyAgSm}			54.8	48	.233
139*	{ArAgSxSm}	{SyArAgSx}	{SySxSm}			53.1	56	>.5
140*	{ArAgSxSm}	{SyArAgSx}	{SyArSm}	{SyAgSm}		49.4	44	.266
141*	{ArAgSxSm}	{SyArAgSx}	{SyArSm}	{SySxSm}		48.1	52	>.5
142*	{ArAgSxSm}	{SyArAgSx}	{SyAgSm}	{SySxSm}		38.4	44	>.5
143*	{ArAgSxSm}	{SyArAgSx}	{SyArSm}	{SyAgSm}	{SySxSm}	32.7	40	>.5
144*	{ArAgSxSm}	{SyArAgSm}	{SyArAgSx}			39.6	32	.166
145*	{ArAgSxSm}	{SyArAgSm}	{SyArAgSx}	{SySxSm}		22.0	28	>.5
146*	{ArAgSxSm}	{SyArSxSm}	{SyArAgSx}			46.3	48	>.5
147*	{ArAgSxSm}	{SyArSxSm}	{SyArAgSx}	{SyAgSm}		31.1	36	>.5
148*	{ArAgSxSm}	{SyArSxSm}	{SyArAgSm}			30.7	30	.432
149*	{ArAgSxSm}	{SyArSxSm}	{SyArAgSm}	{SyAgSx}		24.3	27	>.5
150*	{ArAgSxSm}	{SyAgSxSm}	{SyArAgSx}			31.7	32	.482
151*	{ArAgSxSm}	{SyAgSxSm}	{SyArAgSx}	{SyArSm}		26.3	28	>.5
152*	{ArAgSxSm}	{SyAgSxSm}	{SyArAgSm}			27.4	20	.124
153*	{ArAgSxSm}	{SyAgSxSm}	{SyArAgSm}	{SyArSx}		19.6	19	.419
154*	{ArAgSxSm}	{SyAgSxSm}	{SyArSxSm}			30.5	30	.439
155*	{ArAgSxSm}	{SyAgSxSm}	{SyArSxSm}	{SyArAg}		28.2	27	.399
156*	{ArAgSxSm}	{SyArSxSm}	{SyArAgSm}	{SyArAgSx}		20.9	24	>.5
157*	{ArAgSxSm}	{SyAgSxSm}	{SyArAgSm}	{SyArAgSx}		15.5	16	.486
158*	{ArAgSxSm}	{SyAgSxSm}	{SyArSxSm}	{SyArAgSx}		24.9	24	.414
159*	{ArAgSxSm}	{SyAgSxSm}	{SyArSxSm}	{SyArAgSm}		18.5	15	.237
160*	{ArAgSxSm}	{SyAgSxSm}	{SyArSxSm}	{SyArAgSm}	{SyArAgSx}	14.6	12	.266
161*	{SyArAgSxSm}					0.0	0	---

TABLE C-6

Tests of Alternative Log-linear Models fit to 5-way Table of Symptom by Area by Age by Sex by Smoking. (Data from Boston Area Pulmonary Health Survey, 1983.)

WHEEZE

Model	Margins Fit					L^2	df	p
						TEST OF MODEL		
1	{ArAgSxSm}	{Sy}				169.0	79	<.001
2	{ArAgSxSm}	{SySm}				101.8	75	.021
3	{ArAgSxSm}	{SySx}				168.5	78	<.001
4	{ArAgSxSm}	{SyAg}				165.9	76	<.001
5	{ArAgSxSm}	{SyAr}				151.7	73	<.001
6	{ArAgSxSm}	{SySx}	{SySm}			101.1	74	.020
7	{ArAgSxSm}	{SyAg}	{SySm}			101.5	72	.013
8	{ArAgSxSm}	{SyAg}	{SySx}			165.2	75	<.001
9*	{ArAgSxSm}	{SyAr}	{SySm}			92.6	74	.071
10	{ArAgSxSm}	{SyAr}	{SySx}			151.3	77	<.001
11	{ArAgSxSm}	{SyAr}	{SyAg}			148.1	75	<.001
12	{ArAgSxSm}	{SyAg}	{SySx}	{SySm}		100.8	71	.011
13*	{ArAgSxSm}	{SyAr}	{SySx}	{SySm}		92.2	73	.064
14	{ArAgSxSm}	{SyAr}	{SyAg}	{SySm}		92.2	71	.046
15	{ArAgSxSm}	{SyAr}	{SyAg}	{SySx}		147.7	74	<.001
16	{ArAgSxSm}	{SyAr}	{SyAg}	{SySx}	{SySm}	91.8	70	.041
17	{ArAgSxSm}	{SyArAg}				133.8	72	<.001
18	{ArAgSxSm}	{SyArAg}	{SySx}			133.4	71	<.001
19*	{ArAgSxSm}	{SyArAg}	{SySm}			80.6	68	.141
20*	{ArAgSxSm}	{SyArAg}	{SySx}	{SySm}		80.1	67	.131
21	{ArAgSxSm}	{SyArSx}				151.1	76	<.001
22	{ArAgSxSm}	{SyArSx}	{SyAg}			147.5	73	<.001
23*	{ArAgSxSm}	{SyArSx}	{SySm}			92.1	72	.055
[illegible]	{ArAgSxSm}	{SyArSx}	{SyAg}	{SySm}		91.7	69	.035
[illegible]*	{ArAgSxSm}	{SyArSm}				90.1	70	.053
26	{ArAgSxSm}	{SyArSm}	{SyAg}			89.8	67	.033
27	{ArAgSxSm}	{SyArSm}	{SySx}			89.8	69	.047
28	{ArAgSxSm}	{SyArSm}	{SyAg}	{SySx}		89.5	66	.029
29	{ArAgSxSm}	{SyAgSx}				157.2	72	<.001
30	{ArAgSxSm}	{SyAgSx}	{SyAr}			139.4	71	<.001
31	{ArAgSxSm}	{SyAgSx}	{SySm}			94.8	68	.017
32*	{ArAgSxSm}	{SyAgSx}	{SyAr}	{SySm}		85.9	67	.060
33	{ArAgSxSm}	{SyAgSm}				83.4	60	.025
34*	{ArAgSxSm}	{SyAgSm}	{SyAr}			75.0	59	.078
35	{ArAgSxSm}	{SyAgSm}	{SySx}			82.9	59	.022
36*	{ArAgSxSm}	{SyAgSm}	{SyAr}	{SySx}		74.7	58	.069
37	{ArAgSxSm}	{SySxSm}				97.4	70	.017
38*	{ArAgSxSm}	{SySxSm}	{SyAr}			88.4	69	.058
39	{ArAgSxSm}	{SySxSm}	{SyAg}			97.2	67	.009
40	{ArAgSxSm}	{SySxSm}	{SyAr}	{SyAg}		88.2	66	.035
41	{ArAgSxSm}	{SyAgSm}	{SySxSm}			79.2	55	.018
42*	{ArAgSxSm}	{SyAgSm}	{SySxSm}	{SyAr}		70.8	54	.062
43	{ArAgSxSm}	{SyAgSx}	{SySxSm}			91.5	64	.014
44	{ArAgSxSm}	{SyAgSx}	{SySxSm}	{SyAr}		82.6	63	.049
45	{ArAgSxSm}	{SyAgSx}	{SyAgSm}			75.9	56	.039
46*	{ArAgSxSm}	{SyAgSx}	{SyAgSm}	{SyAr}		67.4	55	.123
47	{ArAgSxSm}	{SyArSm}	{SySxSm}			86.0	65	.041
48	{ArAgSxSm}	{SyArSm}	{SySxSm}	{SyAg}		85.8	62	.024
49*	{ArAgSxSm}	{SyArSm}	{SyAgSm}			71.9	55	.062
50*	{ArAgSxSm}	{SyArSm}	{SyAgSm}	{SySx}		71.7	54	.054

(Cont'd)

NOTE. Note. Variables are:
Ar: Area, 2 categories (SMSA and South Boston)
Ag: Age, 4 categories (18-34, 35-49, 50-64, 65+)
Sx: Sex, 2 Categories (Male and Female)
Sm: Smoking, 4 categories (Never Smoked Cigarettes; Ex Cigarette Smoker; 1-15 , 16-30, and 31+ Cigs. per day)
Sy: Symptom, Dependent Variable

L^2 is Likelihood-ratio chi-square statistic for fit of model

TABLE C-6 : WHEEZE
(Cont'd)

Model	Margins Fit						TEST OF MODEL L^2	df	p
51	{ArAgSxSm}	{SyArSm}	{SyAgSx}				83.7	63	.042
52	{ArAgSxSm}	{SyArSx}	{SySxSm}				88.3	68	.049
53	{ArAgSxSm}	{SyArSx}	{SySxSm}	{SyAg}			88.1	65	.030
54*	{ArAgSxSm}	{SyArSx}	{SyAgSm}				74.6	57	.059
55	{ArAgSxSm}	{SyArSx}	{SyAgSx}				139.4	70	<.001
56*	{ArAgSxSm}	{SyArSx}	{SyAgSx}	{SySm}			85.8	66	.051
57	{ArAgSxSm}	{SyArSx}	{SyArSm}				89.6	68	.041
58	{ArAgSxSm}	{SyArSx}	{SyArSm}	{SyAg}			89.2	65	.025
59*	{ArAgSxSm}	{SyArAg}	{SySxSm}				76.5	63	.119
60*	{ArAgSxSm}	{SyArAg}	{SyAgSm}				64.8	56	.197
61*	{ArAgSxSm}	{SyArAg}	{SyAgSm}	{SySx}			64.5	55	.179
62	{ArAgSxSm}	{SyArAg}	{SyAgSx}				124.8	68	<.001
63*	{ArAgSxSm}	{SyArAg}	{SyAgSx}	{SySm}			73.8	64	.188
64*	{ArAgSxSm}	{SyArAg}	{SyArSm}				79.0	64	.098
65*	{ArAgSxSm}	{SyArAg}	{SyArSm}	{SySx}			78.6	63	.088
66	{ArAgSxSm}	{SyArAg}	{SyArSx}				133.2	70	<.001
67*	{ArAgSxSm}	{SyArAg}	{SyArSx}	{SySm}			80.0	66	.116
68*	{ArAgSxSm}	{SyAgSx}	{SyAgSm}	{SySxSm}			69.5	52	.053
69*	{ArAgSxSm}	{SyAgSx}	{SyAgSm}	{SySxSm}	{SyAr}		60.7	51	.166
70	{ArAgSxSm}	{SyArSm}	{SyAgSm}	{SySxSm}			67.8	50	.047
71	{ArAgSxSm}	{SyArSm}	{SyAgSx}	{SySxSm}			80.4	59	.033
72*	{ArAgSxSm}	{SyArSm}	{SyAgSx}	{SyAgSm}			64.5	51	.097
73*	{ArAgSxSm}	{SyArSx}	{SyAgSm}	{SySxSm}			70.8	53	.052
74	{ArAgSxSm}	{SyArSx}	{SyAgSx}	{SySxSm}			82.5	62	.042
75*	{ArAgSxSm}	{SyArSx}	{SyAgSx}	{SyAgSm}			67.3	54	.106
76	{ArAgSxSm}	{SyArSx}	{SyArSm}	{SySxSm}			85.8	64	.036
[illegible]	{ArAgSxSm}	{SyArSx}	{SyArSm}	{SySxSm}	{SyAg}		85.6	61	.020
[illegible]	{ArAgSxSm}	{SyArSx}	{SyArSm}	{SyAgSm}			71.5	53	.046
[illegible]	{ArAgSxSm}	{SyArSx}	{SyArSm}	{SyAgSx}			83.5	62	.036
80*	{ArAgSxSm}	{SyArAg}	{SyAgSm}	{SySxSm}			60.7	51	.167
81*	{ArAgSxSm}	{SyArAg}	{SyAgSx}	{SySxSm}			70.4	60	.169
82*	{ArAgSxSm}	{SyArAg}	{SyAgSx}	{SyAgSm}			56.7	52	.304
83*	{ArAgSxSm}	{SyArAg}	{SyArSm}	{SySxSm}			74.9	59	.079
84*	{ArAgSxSm}	{SyArAg}	{SyArSm}	{SyAgSm}			63.4	52	.133
85*	{ArAgSxSm}	{SyArAg}	{SyArSm}	{SyAgSm}	{SySx}		63.2	51	.118
86*	{ArAgSxSm}	{SyArAg}	{SyArSm}	{SyAgSx}			72.5	60	.129
87*	{ArAgSxSm}	{SyArAg}	{SyArSx}	{SySxSm}			76.3	62	.105
88*	{ArAgSxSm}	{SyArAg}	{SyArSx}	{SyAgSm}			64.3	54	.159
89	{ArAgSxSm}	{SyArAg}	{SyArSx}	{SyAgSx}			122.7	67	<.001
90*	{ArAgSxSm}	{SyArAg}	{SyArSx}	{SyAgSx}	{SySm}		72.1	63	.203
91*	{ArAgSxSm}	{SyArAg}	{SyArSx}	{SyArSm}			78.3	62	.079
92*	{ArAgSxSm}	{SyArSm}	{SyAgSx}	{SyAgSm}	{SySxSm}		57.9	47	.133
93*	{ArAgSxSm}	{SyArSx}	{SyAgSx}	{SyAgSm}	{SySxSm}		60.6	50	.144
94	{ArAgSxSm}	{SyArSx}	{SyArSm}	{SyAgSm}	{SySxSm}		67.7	49	.039
95	{ArAgSxSm}	{SyArSx}	{SyArSm}	{SyAgSx}	{SySxSm}		80.3	58	.023
96*	{ArAgSxSm}	{SyArSx}	{SyArSm}	{SyAgSx}	{SyAgSm}		64.3	50	.084
97*	{ArAgSxSm}	{SyArAg}	{SyAgSx}	{SyAgSm}	{SySxSm}		50.1	48	.392
98*	{ArAgSxSm}	{SyArAg}	{SyArSm}	{SyAgSm}	{SySxSm}		59.4	47	.107
99*	{ArAgSxSm}	{SyArAg}	{SyArSm}	{SyAgSx}	{SySxSm}		68.9	56	.115
100*	{ArAgSxSm}	{SyArAg}	{SyArSm}	{SyAgSx}	{SyAgSm}		55.5	48	.213
101*	{ArAgSxSm}	{SyArAg}	{SyArSx}	{SyAgSm}	{SySxSm}		60.6	50	.146
102*	{ArAgSxSm}	{SyArAg}	{SyArSx}	{SyAgSx}	{SySxSm}		68.7	59	.181
103*	{ArAgSxSm}	{SyArAg}	{SyArSx}	{SyAgSx}	{SyAgSm}		55.2	51	.320
104*	{ArAgSxSm}	{SyArAg}	{SyArSx}	{SyArSm}	{SySxSm}		74.7	58	.069
105*	{ArAgSxSm}	{SyArAg}	{SyArSx}	{SyArSm}	{SyAgSm}		63.0	50	.103
106*	{ArAgSxSm}	{SyArAg}	{SyArSx}	{SyArSm}	{SyAgSx}		70.4	59	.146
107*	{ArAgSxSm}	{SyArAg}	{SyArSx}	{SyArSm}	{SyAgSx}	{SySxSm}	67.3	55	.124
108*	{ArAgSxSm}	{SyAgSxSm}					52.4	40	.091
109*	{ArAgSxSm}	{SyAgSxSm}	{SyAr}				44.1	39	.264
110*	{ArAgSxSm}	{SyAgSxSm}	{SyArAg}				32.9	36	>.5

(Cont'd)

TABLE C-6 : WHEEZE
(Cont'd)

Model	Margins Fit					TEST OF MODEL L^2	df	p
111*	{ArAgSxSm}	{SyAgSxSm}	{SyArSx}			44.0	38	.231
112*	{ArAgSxSm}	{SyAgSxSm}	{SyArSm}			41.6	35	.205
113*	{ArAgSxSm}	{SyAgSxSm}	{SyArAg}	{SyArSx}		31.5	35	>.5
114*	{ArAgSxSm}	{SyAgSxSm}	{SyArAg}	{SyArSm}		32.0	32	.468
115*	{ArAgSxSm}	{SyAgSxSm}	{SyArSx}	{SyArSm}		41.5	34	.177
116*	{ArAgSxSm}	{SyAgSxSm}	{SyArAg}	{SyArSx}	{SyArSm}	30.8	31	.478
117	{ArAgSxSm}	{SyArSxSm}				80.2	60	.042
118	{ArAgSxSm}	{SyArSxSm}	{SyAg}			80.0	57	.024
119*	{ArAgSxSm}	{SyArSxSm}	{SyArAg}			68.8	54	.084
120	{ArAgSxSm}	{SyArSxSm}	{SyAgSx}			74.7	54	.033
121	{ArAgSxSm}	{SyArSxSm}	{SyAgSm}			62.2	45	.045
122*	{ArAgSxSm}	{SyArSxSm}	{SyArAg}	{SyAgSx}		61.4	51	.150
123*	{ArAgSxSm}	{SyArSxSm}	{SyArAg}	{SyAgSm}		53.7	42	.107
124*	{ArAgSxSm}	{SyArSxSm}	{SyAgSx}	{SyAgSm}		52.2	42	.134
125*	{ArAgSxSm}	{SyArSxSm}	{SyArAg}	{SyAgSx}	{SyAgSm}	42.1	39	.338
126*	{ArAgSxSm}	{SyArAgSm}				45.9	40	.242
127*	{ArAgSxSm}	{SyArAgSm}	{SySx}			45.7	39	.215
128*	{ArAgSxSm}	{SyArAgSm}	{SyArSx}			45.5	38	.189
129*	{ArAgSxSm}	{SyArAgSm}	{SyAgSx}			38.1	36	.372
130*	{ArAgSxSm}	{SyArAgSm}	{SySxSm}			42.5	35	.179
131*	{ArAgSxSm}	{SyArAgSm}	{SyArSx}	{SyAgSx}		36.6	35	.397
132*	{ArAgSxSm}	{SyArAgSm}	{SyArSx}	{SySxSm}		42.3	34	.156
133*	{ArAgSxSm}	{SyArAgSm}	{SyAgSx}	{SySxSm}		32.4	32	.447
134*	{ArAgSxSm}	{SyArAgSm}	{SyArSx}	{SyAgSx}	{SySxSm}	30.7	31	.479
135	{ArAgSxSm}	{SyArAgSx}				111.2	64	<.001
136*	{ArAgSxSm}	{SyArAgSx}	{SySm}			59.8	60	.483
[illegible]*	{ArAgSxSm}	{SyArAgSx}	{SyArSm}			58.2	56	.393
[illegible]	{ArAgSxSm}	{SyArAgSx}	{SyAgSm}			43.0	48	>.5
139*	{ArAgSxSm}	{SyArAgSx}	{SySxSm}			56.5	56	.455
140*	{ArAgSxSm}	{SyArAgSx}	{SyArSm}	{SyAgSm}		41.9	44	>.5
141*	{ArAgSxSm}	{SyArAgSx}	{SyArSm}	{SySxSm}		55.1	52	.358
142*	{ArAgSxSm}	{SyArAgSx}	{SyAgSm}	{SySxSm}		36.8	44	>.5
143*	{ArAgSxSm}	{SyArAgSx}	{SyArSm}	{SyAgSm}	{SySxSm}	36.0	40	>.5
144*	{ArAgSxSm}	{SyArAgSm}	{SyArAgSx}			24.6	32	>.5
145*	{ArAgSxSm}	{SyArAgSm}	{SyArAgSx}	{SySxSm}		22.7	28	>.5
146*	{ArAgSxSm}	{SyArSxSm}	{SyArAgSx}			49.8	48	.401
147*	{ArAgSxSm}	{SyArSxSm}	{SyArAgSx}	{SyAgSm}		31.9	36	>.5
148*	{ArAgSxSm}	{SyArSxSm}	{SyArAgSm}			37.4	30	.166
149*	{ArAgSxSm}	{SyArSxSm}	{SyArAgSm}	{SyAgSx}		21.8	27	>.5
150*	{ArAgSxSm}	{SyAgSxSm}	{SyArAgSx}			19.8	32	>.5
151*	{ArAgSxSm}	{SyAgSxSm}	{SyArAgSx}	{SyArSm}		17.8	28	>.5
152*	{ArAgSxSm}	{SyAgSxSm}	{SyArAgSm}			17.0	20	>.5
153*	{ArAgSxSm}	{SyAgSxSm}	{SyArAgSm}	{SyArSx}		16.8	19	>.5
154*	{ArAgSxSm}	{SyAgSxSm}	{SyArSxSm}			35.6	30	.223
155*	{ArAgSxSm}	{SyAgSxSm}	{SyArSxSm}	{SyArAg}		25.8	27	>.5
156*	{ArAgSxSm}	{SyArSxSm}	{SyArAgSm}	{SyArAgSx}		15.2	24	>.5
157*	{ArAgSxSm}	{SyAgSxSm}	{SyArAgSm}	{SyArAgSx}		10.9	16	>.5
158*	{ArAgSxSm}	{SyAgSxSm}	{SyArSxSm}	{SyArAgSx}		15.8	24	>.5
159*	{ArAgSxSm}	{SyAgSxSm}	{SyArSxSm}	{SyArAgSm}		8.9	15	>.5
160*	{ArAgSxSm}	{SyAgSxSm}	{SyArSxSm}	{SyArAgSm}	{SyArAgSx}	2.8	12	>.5
161*	{SyArAgSxSm}					0.0	0	---

TABLE C-7

Tests of Alternative Log-linear Models fit to 5-way Table of Symptom by Area by Age by Sex by Smoking. (Data from Boston Area Pulmonary Health Survey, 1983.)

EV. CHEST COLDS

Model	Margins Fit					TEST OF MODEL L^2	df	p
1	{ArAgSxSm}	{Sy}				144.6	79	<.001
2	{ArAgSxSm}	{SySm}				139.0	75	<.001
3	{ArAgSxSm}	{SySx}				128.0	78	<.001
4	{ArAgSxSm}	{SyAg}				121.6	76	.001
5	{ArAgSxSm}	{SyAr}				144.5	78	<.001
6	{ArAgSxSm}	{SySx}	{SySm}			121.9	74	.001
7	{ArAgSxSm}	{SyAg}	{SySm}			115.8	72	.001
8	{ArAgSxSm}	{SyAg}	{SySx}			103.0	75	.018
9	{ArAgSxSm}	{SyAr}	{SySm}			138.9	74	<.001
10	{ArAgSxSm}	{SyAr}	{SySx}			127.8	77	<.001
11	{ArAgSxSm}	{SyAr}	{SyAg}			121.6	75	.001
12	{ArAgSxSm}	{SyAg}	{SySx}	{SySm}		96.7	71	.023
13	{ArAgSxSm}	{SyAr}	{SySx}	{SySm}		121.6	73	<.001
14	{ArAgSxSm}	{SyAr}	{SyAg}	{SySm}		115.6	71	.001
15	{ArAgSxSm}	{SyAr}	{SyAg}	{SySx}		102.9	74	.015
16	{ArAgSxSm}	{SyAr}	{SyAg}	{SySx}	{SySm}	96.4	70	.020
17	{ArAgSxSm}	{SyArAg}				114.9	72	.001
18	{ArAgSxSm}	{SyArAg}	{SySx}			96.4	71	.024
19	{ArAgSxSm}	{SyArAg}	{SySm}			109.3	68	.001
20	{ArAgSxSm}	{SyArAg}	{SySx}	{SySm}		90.4	67	.030
21	{ArAgSxSm}	{SyArSx}				125.6	76	<.001
22	{ArAgSxSm}	{SyArSx}	{SyAg}			101.1	73	.016
	{ArAgSxSm}	{SyArSx}	{SySm}			119.3	72	.001
	{ArAgSxSm}	{SyArSx}	{SyAg}	{SySm}		94.5	69	.022
25	{ArAgSxSm}	{SyArSm}				131.9	70	<.001
26	{ArAgSxSm}	{SyArSm}	{SyAg}			108.8	67	.001
27	{ArAgSxSm}	{SyArSm}	{SySx}			114.4	69	.001
28	{ArAgSxSm}	{SyArSm}	{SyAg}	{SySx}		89.2	66	.030
29*	{ArAgSxSm}	{SyAgSx}				89.8	72	.076
30*	{ArAgSxSm}	{SyAgSx}	{SyAr}			89.6	71	.067
31*	{ArAgSxSm}	{SyAgSx}	{SySm}			82.9	68	.106
32*	{ArAgSxSm}	{SyAgSx}	{SyAr}	{SySm}		82.4	67	.098
33	{ArAgSxSm}	{SyAgSm}				110.5	60	<.001
34	{ArAgSxSm}	{SyAgSm}	{SyAr}			110.3	59	<.001
35	{ArAgSxSm}	{SyAgSm}	{SySx}			90.3	59	.006
36	{ArAgSxSm}	{SyAgSm}	{SyAr}	{SySx}		89.9	58	.005
37	{ArAgSxSm}	{SySxSm}				111.4	70	.001
38	{ArAgSxSm}	{SySxSm}	{SyAr}			111.1	69	.001
39	{ArAgSxSm}	{SySxSm}	{SyAg}			88.6	67	.040
40	{ArAgSxSm}	{SySxSm}	{SyAr}	{SyAg}		88.2	66	.035
41	{ArAgSxSm}	{SyAgSm}	{SySxSm}			82.8	55	.009
42	{ArAgSxSm}	{SyAgSm}	{SySxSm}	{SyAr}		82.4	54	.008
43*	{ArAgSxSm}	{SyAgSx}	{SySxSm}			75.5	64	.155
44*	{ArAgSxSm}	{SyAgSx}	{SySxSm}	{SyAr}		74.9	63	.145
45	{ArAgSxSm}	{SyAgSx}	{SyAgSm}			76.5	56	.036
46	{ArAgSxSm}	{SyAgSx}	{SyAgSm}	{SyAr}		75.9	55	.032
47	{ArAgSxSm}	{SyArSm}	{SySxSm}			103.5	65	.002
48*	{ArAgSxSm}	{SyArSm}	{SySxSm}	{SyAg}		80.8	62	.055
49	{ArAgSxSm}	{SyArSm}	{SyAgSm}			103.7	55	<.001
50	{ArAgSxSm}	{SyArSm}	{SyAgSm}	{SySx}		83.1	54	.007

(Cont'd)

OTE. Note. Variables are:
Ar: Area, 2 categories (SMSA and South Boston)
Ag: Age, 4 categories (18-34, 35-49, 50-64, 65+)
Sx: Sex, 2 Categories (Male and Female)
Sm: Smoking, 4 categories (Never Smoked Cigarettes; Ex Cigarette Smoker; 1-15 , 16-30, and 31+ Cigs. per day)
Sy: Symptom, Dependent Variable

^{2} is Likelihood-ratio chi-square statistic for fit of model

TABLE C-7: SEV. CHEST COLDS
(Cont'd)

Model	Margins Fit						TEST OF MODEL L^2	df	p
51*	{ArAgSxSm}	{SyArSm}	{SyAgSx}				75.1	63	.141
52	{ArAgSxSm}	{SyArSx}	{SySxSm}				107.5	68	.002
53	{ArAgSxSm}	{SyArSx}	{SySxSm}	{SyAg}			85.3	65	.046
54	{ArAgSxSm}	{SyArSx}	{SyAgSm}				87.9	57	.006
55*	{ArAgSxSm}	{SyArSx}	{SyAgSx}				87.9	70	.073
56*	{ArAgSxSm}	{SyArSx}	{SyAgSx}	{SySm}			80.6	66	.107
57	{ArAgSxSm}	{SyArSx}	{SyArSm}				111.3	68	.001
58	{ArAgSxSm}	{SyArSx}	{SyArSm}	{SyAg}			86.5	65	.038
59*	{ArAgSxSm}	{SyArAg}	{SySxSm}				82.2	63	.053
60	{ArAgSxSm}	{SyArAg}	{SyAgSm}				104.2	56	<.001
61	{ArAgSxSm}	{SyArAg}	{SyAgSm}	{SySx}			84.4	55	.007
62*	{ArAgSxSm}	{SyArAg}	{SyAgSx}				83.5	68	.098
63*	{ArAgSxSm}	{SyArAg}	{SyAgSx}	{SySm}			76.8	64	.131
64	{ArAgSxSm}	{SyArAg}	{SyArSm}				102.3	64	.002
65	{ArAgSxSm}	{SyArAg}	{SyArSm}	{SySx}			83.0	63	.046
66	{ArAgSxSm}	{SyArAg}	{SyArSx}				94.5	70	.027
67	{ArAgSxSm}	{SyArAg}	{SyArSx}	{SySm}			88.4	66	.034
68	{ArAgSxSm}	{SyAgSx}	{SyAgSm}	{SySxSm}			70.2	52	.047
69	{ArAgSxSm}	{SyAgSx}	{SyAgSm}	{SySxSm}	{SyAr}		69.6	51	.043
70	{ArAgSxSm}	{SyArSm}	{SyAgSm}	{SySxSm}			75.2	50	.012
71*	{ArAgSxSm}	{SyArSm}	{SyAgSx}	{SySxSm}			67.5	59	.210
72	{ArAgSxSm}	{SyArSm}	{SyAgSx}	{SyAgSm}			68.9	51	.048
73	{ArAgSxSm}	{SyArSx}	{SyAgSm}	{SySxSm}			79.3	53	.011
74*	{ArAgSxSm}	{SyArSx}	{SyAgSx}	{SySxSm}			72.2	62	.177
75	{ArAgSxSm}	{SyArSx}	{SyAgSx}	{SyAgSm}			73.9	54	.037
76	{ArAgSxSm}	{SyArSx}	{SyArSm}	{SySxSm}			98.9	64	.004
77*	{ArAgSxSm}	{SyArSx}	{SyArSm}	{SySxSm}	{SyAg}		76.8	61	.083
[illegible]	{ArAgSxSm}	{SyArSx}	{SyArSm}	{SyAgSm}			80.2	53	.009
79*	{ArAgSxSm}	{SyArSx}	{SyArSm}	{SyAgSx}			72.6	62	.169
80	{ArAgSxSm}	{SyArAg}	{SyAgSm}	{SySxSm}			76.7	51	.011
81*	{ArAgSxSm}	{SyArAg}	{SyAgSx}	{SySxSm}			69.4	60	.191
82	{ArAgSxSm}	{SyArAg}	{SyAgSx}	{SyAgSm}			70.6	52	.044
83*	{ArAgSxSm}	{SyArAg}	{SyArSm}	{SySxSm}			74.4	59	.085
84	{ArAgSxSm}	{SyArAg}	{SyArSm}	{SyAgSm}			97.8	52	<.001
85	{ArAgSxSm}	{SyArAg}	{SyArSm}	{SyAgSm}	{SySx}		77.5	51	.010
86*	{ArAgSxSm}	{SyArAg}	{SyArSm}	{SyAgSx}			69.5	60	.189
87*	{ArAgSxSm}	{SyArAg}	{SyArSx}	{SySxSm}			79.0	62	.072
88	{ArAgSxSm}	{SyArAg}	{SyArSx}	{SyAgSm}			82.2	54	.008
89*	{ArAgSxSm}	{SyArAg}	{SyArSx}	{SyAgSx}			81.2	67	.114
90*	{ArAgSxSm}	{SyArAg}	{SyArSx}	{SyAgSx}	{SySm}		74.4	63	.154
91*	{ArAgSxSm}	{SyArAg}	{SyArSx}	{SyArSm}			80.1	62	.061
92*	{ArAgSxSm}	{SyArSm}	{SyAgSx}	{SyAgSm}	{SySxSm}		62.4	47	.066
93*	{ArAgSxSm}	{SyArSx}	{SyAgSx}	{SyAgSm}	{SySxSm}		66.6	50	.058
94	{ArAgSxSm}	{SyArSx}	{SyArSm}	{SyAgSm}	{SySxSm}		71.2	49	.021
95*	{ArAgSxSm}	{SyArSx}	{SyArSm}	{SyAgSx}	{SySxSm}		63.8	58	.281
96*	{ArAgSxSm}	{SyArSx}	{SyArSm}	{SyAgSx}	{SyAgSm}		66.2	50	.062
97*	{ArAgSxSm}	{SyArAg}	{SyAgSx}	{SyAgSm}	{SySxSm}		64.2	48	.059
98	{ArAgSxSm}	{SyArAg}	{SyArSm}	{SyAgSm}	{SySxSm}		69.5	47	.018
99*	{ArAgSxSm}	{SyArAg}	{SyArSm}	{SyAgSx}	{SySxSm}		61.9	56	.275
100*	{ArAgSxSm}	{SyArAg}	{SyArSm}	{SyAgSx}	{SyAgSm}		63.7	48	.064
101	{ArAgSxSm}	{SyArAg}	{SyArSx}	{SyAgSm}	{SySxSm}		73.4	50	.017
102*	{ArAgSxSm}	{SyArAg}	{SyArSx}	{SyAgSx}	{SySxSm}		65.9	59	.252
103*	{ArAgSxSm}	{SyArAg}	{SyArSx}	{SyAgSx}	{SyAgSm}		68.1	51	.055
104*	{ArAgSxSm}	{SyArAg}	{SyArSx}	{SyArSm}	{SySxSm}		70.2	58	.130
105	{ArAgSxSm}	{SyArAg}	{SyArSx}	{SyArSm}	{SyAgSm}		74.5	50	.014
106*	{ArAgSxSm}	{SyArAg}	{SyArSx}	{SyArSm}	{SyAgSx}		66.2	59	.243
107*	{ArAgSxSm}	{SyArAg}	{SyArSx}	{SyArSm}	{SyAgSx}	{SySxSm}	57.3	55	.391
108	{ArAgSxSm}	{SyAgSxSm}					61.3	40	.017
109	{ArAgSxSm}	{SyAgSxSm}	{SyAr}				60.7	39	.015
110	{ArAgSxSm}	{SyAgSxSm}	{SyArAg}				55.9	36	.018

(Cont'd)

TABLE C-7 : SEV. CHEST COLDS
(Cont'd)

Model	Margins Fit					TEST OF MODEL L^2	df	p
111	{ArAgSxSm}	{SyAgSxSm}	{SyArSx}			57.6	38	.021
112	{ArAgSxSm}	{SyAgSxSm}	{SyArSm}			53.9	35	.022
113	{ArAgSxSm}	{SyAgSxSm}	{SyArAg}	{SyArSx}		52.1	35	.031
114	{ArAgSxSm}	{SyAgSxSm}	{SyArAg}	{SyArSm}		49.0	32	.028
115	{ArAgSxSm}	{SyAgSxSm}	{SyArSx}	{SyArSm}		49.9	34	.038
116*	{ArAgSxSm}	{SyAgSxSm}	{SyArAg}	{SyArSx}	{SyArSm}	44.3	31	.058
117	{ArAgSxSm}	{SyArSxSm}				95.6	60	.003
118*	{ArAgSxSm}	{SyArSxSm}	{SyAg}			73.9	57	.066
119*	{ArAgSxSm}	{SyArSxSm}	{SyArAg}			67.7	54	.100
120*	{ArAgSxSm}	{SyArSxSm}	{SyAgSx}			60.9	54	.243
121	{ArAgSxSm}	{SyArSxSm}	{SyAgSm}			68.0	45	.015
122*	{ArAgSxSm}	{SyArSxSm}	{SyArAg}	{SyAgSx}		54.8	51	.333
123	{ArAgSxSm}	{SyArSxSm}	{SyArAg}	{SyAgSm}		62.5	42	.021
124*	{ArAgSxSm}	{SyArSxSm}	{SyAgSx}	{SyAgSm}		55.5	42	.079
125*	{ArAgSxSm}	{SyArSxSm}	{SyArAg}	{SyAgSx}	{SyAgSm}	49.8	39	.115
126	{ArAgSxSm}	{SyArAgSm}				83.9	40	<.001
127	{ArAgSxSm}	{SyArAgSm}	{SySx}			62.5	39	.010
128	{ArAgSxSm}	{SyArAgSm}	{SyArSx}			59.3	38	.015
129*	{ArAgSxSm}	{SyArAgSm}	{SyAgSx}			48.6	36	.078
130	{ArAgSxSm}	{SyArAgSm}	{SySxSm}			54.1	35	.020
131*	{ArAgSxSm}	{SyArAgSm}	{SyArSx}	{SyAgSx}		45.0	35	.119
132	{ArAgSxSm}	{SyArAgSm}	{SyArSx}	{SySxSm}		49.4	34	.042
133*	{ArAgSxSm}	{SyArAgSm}	{SyAgSx}	{SySxSm}		41.6	32	.119
134*	{ArAgSxSm}	{SyArAgSm}	{SyArSx}	{SyAgSx}	{SySxSm}	36.5	31	.228
135*	{ArAgSxSm}	{SyArAgSx}				80.2	64	.083
136*	{ArAgSxSm}	{SyArAgSx}	{SySm}			73.4	60	.115
[illegible]*	{ArAgSxSm}	{SyArAgSx}	{SyArSm}			65.4	56	.183
[illegible]	{ArAgSxSm}	{SyArAgSx}	{SyAgSm}			67.0	48	.036
139*	{ArAgSxSm}	{SyArAgSx}	{SySxSm}			64.9	56	.195
140*	{ArAgSxSm}	{SyArAgSx}	{SyArSm}	{SyAgSm}		59.5	44	.059
141*	{ArAgSxSm}	{SyArAgSx}	{SyArSm}	{SySxSm}		56.5	52	.309
142*	{ArAgSxSm}	{SyArAgSx}	{SyAgSm}	{SySxSm}		59.6	44	.059
143*	{ArAgSxSm}	{SyArAgSx}	{SyArSm}	{SyAgSm}	{SySxSm}	51.7	40	.101
144*	{ArAgSxSm}	{SyArAgSm}	{SyArAgSx}			44.1	32	.076
145*	{ArAgSxSm}	{SyArAgSm}	{SyArAgSx}	{SySxSm}		35.4	28	.160
146*	{ArAgSxSm}	{SyArSxSm}	{SyArAgSx}			54.1	48	.254
147*	{ArAgSxSm}	{SyArSxSm}	{SyArAgSx}	{SyAgSm}		49.2	36	.070
148	{ArAgSxSm}	{SyArSxSm}	{SyArAgSm}			46.6	30	.027
149*	{ArAgSxSm}	{SyArSxSm}	{SyArAgSm}	{SyAgSx}		33.4	27	.183
150	{ArAgSxSm}	{SyAgSxSm}	{SyArAgSx}			51.0	32	.018
151	{ArAgSxSm}	{SyAgSxSm}	{SyArAgSx}	{SyArSm}		43.5	28	.031
152	{ArAgSxSm}	{SyAgSxSm}	{SyArAgSm}			32.8	20	.035
153*	{ArAgSxSm}	{SyAgSxSm}	{SyArAgSm}	{SyArSx}		28.2	19	.080
154	{ArAgSxSm}	{SyAgSxSm}	{SyArSxSm}			46.8	30	.026
155	{ArAgSxSm}	{SyAgSxSm}	{SyArSxSm}	{SyArAg}		41.7	27	.035
156*	{ArAgSxSm}	{SyArSxSm}	{SyArAgSm}	{SyArAgSx}		32.5	24	.116
157	{ArAgSxSm}	{SyAgSxSm}	{SyArAgSm}	{SyArAgSx}		27.0	16	.042
158	{ArAgSxSm}	{SyAgSxSm}	{SyArSxSm}	{SyArAgSx}		41.1	24	.016
159*	{ArAgSxSm}	{SyAgSxSm}	{SyArSxSm}	{SyArAgSm}		24.7	15	.054
160	{ArAgSxSm}	{SyAgSxSm}	{SyArSxSm}	{SyArAgSm}	{SyArAgSx}	23.5	12	.024
161*	{SyArAgSxSm}					0.0	0	---

TABLE C-8

Tests of Alternative Log-linear Models fit to 5-way Table of Symptom by Area by Age by Sex by Smoking. (Data from Boston Area Pulmonary Health Survey, 1983.)

ASTHMA

Model	Margins Fit				TEST OF MODEL L^2	df	p
1*	{ArAgSxSm}	{Sy}			96.8	79	.085
2*	{ArAgSxSm}	{SySm}			94.0	75	.068
3*	{ArAgSxSm}	{SySx}			96.8	78	.074
4*	{ArAgSxSm}	{SyAg}			87.7	76	.169
5*	{ArAgSxSm}	{SyAr}			96.5	78	.076
6*	{ArAgSxSm}	{SySx}	{SySm}		93.9	74	.059
7*	{ArAgSxSm}	{SyAg}	{SySm}		84.3	72	.152
8*	{ArAgSxSm}	{SyAg}	{SySx}		87.7	75	.150
9*	{ArAgSxSm}	{SyAr}	{SySm}		93.6	74	.062
10*	{ArAgSxSm}	{SyAr}	{SySx}		96.4	77	.066
11*	{ArAgSxSm}	{SyAr}	{SyAg}		87.3	75	.157
12*	{ArAgSxSm}	{SyAg}	{SySx}	{SySm}	84.3	71	.134
13*	{ArAgSxSm}	{SyAr}	{SySx}	{SySm}	93.4	73	.054
14*	{ArAgSxSm}	{SyAr}	{SyAg}	{SySm}	83.7	71	.144
15*	{ArAgSxSm}	{SyAr}	{SyAg}	{SySx}	87.3	74	.139
16*	{ArAgSxSm}	{SyAr}	{SyAg}	{SySx} {SySm}	83.7	70	.127
17*	{ArAgSxSm}	{SyArAg}			85.1	72	.140
18*	{ArAgSxSm}	{SyArAg}	{SySx}		85.1	71	.122
19*	{ArAgSxSm}	{SyArAg}	{SySm}		81.3	68	.129
20*	{ArAgSxSm}	{SyArAg}	{SySx}	{SySm}	81.3	67	.113
21*	{ArAgSxSm}	{SyArSx}			95.3	76	.067
22*	{ArAgSxSm}	{SyArSx}	{SyAg}		86.2	73	.139
23*	{ArAgSxSm}	{SyArSx}	{SySm}		92.3	72	.054
[illegible]*	{ArAgSxSm}	{SyArSx}	{SyAg}	{SySm}	82.7	69	.125
[illegible]*	{ArAgSxSm}	{SyArSm}			90.1	70	.053
26*	{ArAgSxSm}	{SyArSm}	{SyAg}		80.8	67	.120
27	{ArAgSxSm}	{SyArSm}	{SySx}		89.9	69	.046
28*	{ArAgSxSm}	{SyArSm}	{SyAg}	{SySx}	80.8	66	.105
29*	{ArAgSxSm}	{SyAgSx}			84.4	72	.150
30*	{ArAgSxSm}	{SyAgSx}	{SyAr}		84.1	71	.138
31*	{ArAgSxSm}	{SyAgSx}	{SySm}		80.9	68	.137
32*	{ArAgSxSm}	{SyAgSx}	{SyAr}	{SySm}	80.3	67	.128
33*	{ArAgSxSm}	{SyAgSm}			70.6	60	.166
34*	{ArAgSxSm}	{SyAgSm}	{SyAr}		69.8	59	.159
35*	{ArAgSxSm}	{SyAgSm}	{SySx}		70.6	59	.144
36*	{ArAgSxSm}	{SyAgSm}	{SyAr}	{SySx}	69.8	58	.139
37*	{ArAgSxSm}	{SySxSm}			89.6	70	.058
38*	{ArAgSxSm}	{SySxSm}	{SyAr}		89.2	69	.052
39*	{ArAgSxSm}	{SySxSm}	{SyAg}		80.2	67	.130
40*	{ArAgSxSm}	{SySxSm}	{SyAr}	{SyAg}	79.6	66	.122
41*	{ArAgSxSm}	{SyAgSm}	{SySxSm}		66.4	55	.139
42*	{ArAgSxSm}	{SyAgSm}	{SySxSm}	{SyAr}	65.7	54	.132
43*	{ArAgSxSm}	{SyAgSx}	{SySxSm}		76.7	64	.133
44*	{ArAgSxSm}	{SyAgSx}	{SySxSm}	{SyAr}	76.2	63	.123
45*	{ArAgSxSm}	{SyAgSx}	{SyAgSm}		67.6	56	.137
46*	{ArAgSxSm}	{SyAgSx}	{SyAgSm}	{SyAr}	66.9	55	.130
47	{ArAgSxSm}	{SyArSm}	{SySxSm}		85.6	65	.044
48*	{ArAgSxSm}	{SyArSm}	{SySxSm}	{SyAg}	76.5	62	.102
49*	{ArAgSxSm}	{SyArSm}	{SyAgSm}		66.4	55	.139
50*	{ArAgSxSm}	{SyArSm}	{SyAgSm}	{SySx}	66.4	54	.120

(Cont'd)

NOTE. Note. Variables are:
Ar: Area, 2 categories (SMSA and South Boston)
Ag: Age, 4 categories (18-34, 35-49, 50-64, 65+)
Sx: Sex, 2 Categories (Male and Female)
Sm: Smoking, 4 categories (Never Smoked Cigarettes; Ex Cigarette Smoker; 1-15 , 16-30, and 31+ Cigs. per day)
Sy: Symptom, Dependent Variable

L^2 is Likelihood-ratio chi-square statistic for fit of model

TABLE C-8: ASTHMA
(Cont'd)

Model	Margins Fit						TEST OF MODEL L^2	df	p
51*	{ArAgSxSm}	{SyArSm}	{SyAgSx}				77.3	63	.106
52*	{ArAgSxSm}	{SyArSx}	{SySxSm}				87.8	68	.053
53*	{ArAgSxSm}	{SyArSx}	{SySxSm}	{SyAg}			78.4	65	.122
54*	{ArAgSxSm}	{SyArSx}	{SyAgSm}				68.5	57	.142
55*	{ArAgSxSm}	{SyArSx}	{SyAgSx}				83.1	70	.136
56*	{ArAgSxSm}	{SyArSx}	{SyAgSx}	{SySm}			79.4	66	.125
57	{ArAgSxSm}	{SyArSx}	{SyArSm}				89.1	68	.044
58*	{ArAgSxSm}	{SyArSx}	{SyArSm}	{SyAg}			80.0	65	.100
59*	{ArAgSxSm}	{SyArAg}	{SySxSm}				77.2	63	.108
60*	{ArAgSxSm}	{SyArAg}	{SyAgSm}				66.3	56	.163
61*	{ArAgSxSm}	{SyArAg}	{SyAgSm}	{SySx}			66.4	55	.141
62*	{ArAgSxSm}	{SyArAg}	{SyAgSx}				81.6	68	.125
63*	{ArAgSxSm}	{SyArAg}	{SyAgSx}	{SySm}			77.7	64	.117
64*	{ArAgSxSm}	{SyArAg}	{SyArSm}				78.3	64	.108
65*	{ArAgSxSm}	{SyArAg}	{SyArSm}	{SySx}			78.3	63	.093
66*	{ArAgSxSm}	{SyArAg}	{SyArSx}				84.2	70	.119
67*	{ArAgSxSm}	{SyArAg}	{SyArSx}	{SySm}			80.5	66	.109
68*	{ArAgSxSm}	{SyAgSx}	{SyAgSm}	{SySxSm}			63.5	52	.132
69*	{ArAgSxSm}	{SyAgSx}	{SyAgSm}	{SySxSm}	{SyAr}		62.8	51	.124
70*	{ArAgSxSm}	{SyArSm}	{SyAgSm}	{SySxSm}			62.1	50	.117
71*	{ArAgSxSm}	{SyArSm}	{SyAgSx}	{SySxSm}			73.0	59	.104
72*	{ArAgSxSm}	{SyArSm}	{SyAgSx}	{SyAgSm}			63.5	51	.112
73*	{ArAgSxSm}	{SyArSx}	{SyAgSm}	{SySxSm}			64.3	53	.138
74*	{ArAgSxSm}	{SyArSx}	{SyAgSx}	{SySxSm}			75.1	62	.122
75*	{ArAgSxSm}	{SyArSx}	{SyAgSx}	{SyAgSm}			65.7	54	.132
76	{ArAgSxSm}	{SyArSx}	{SyArSm}	{SySxSm}			84.3	64	.045
77*	{ArAgSxSm}	{SyArSx}	{SyArSm}	{SySxSm}	{SyAg}		75.5	61	.101
[illegible]*	{ArAgSxSm}	{SyArSx}	{SyArSm}	{SyAgSm}			65.4	53	.118
79*	{ArAgSxSm}	{SyArSx}	{SyArSm}	{SyAgSx}			76.6	62	.100
80*	{ArAgSxSm}	{SyArAg}	{SyAgSm}	{SySxSm}			62.3	51	.134
81*	{ArAgSxSm}	{SyArAg}	{SyAgSx}	{SySxSm}			73.5	60	.113
82*	{ArAgSxSm}	{SyArAg}	{SyAgSx}	{SyAgSm}			63.3	52	.136
83*	{ArAgSxSm}	{SyArAg}	{SyArSm}	{SySxSm}			74.1	59	.090
84*	{ArAgSxSm}	{SyArAg}	{SyArSm}	{SyAgSm}			62.6	52	.149
85*	{ArAgSxSm}	{SyArAg}	{SyArSm}	{SyAgSm}	{SySx}		62.6	51	.128
86*	{ArAgSxSm}	{SyArAg}	{SyArSm}	{SyAgSx}			74.7	60	.096
87*	{ArAgSxSm}	{SyArAg}	{SyArSx}	{SySxSm}			76.3	62	.105
88*	{ArAgSxSm}	{SyArAg}	{SyArSx}	{SyAgSm}			65.3	54	.140
89*	{ArAgSxSm}	{SyArAg}	{SyArSx}	{SyAgSx}			81.0	67	.117
90*	{ArAgSxSm}	{SyArAg}	{SyArSx}	{SyAgSx}	{SySm}		77.1	63	.108
91*	{ArAgSxSm}	{SyArAg}	{SyArSx}	{SyArSm}			77.7	62	.086
92*	{ArAgSxSm}	{SyArSm}	{SyAgSx}	{SyAgSm}	{SySxSm}		59.1	47	.111
93*	{ArAgSxSm}	{SyArSx}	{SyAgSx}	{SyAgSm}	{SySxSm}		61.6	50	.126
94*	{ArAgSxSm}	{SyArSx}	{SyArSm}	{SyAgSm}	{SySxSm}		60.9	49	.119
95*	{ArAgSxSm}	{SyArSx}	{SyArSm}	{SyAgSx}	{SySxSm}		72.1	58	.102
96*	{ArAgSxSm}	{SyArSx}	{SyArSm}	{SyAgSx}	{SyAgSm}		62.6	50	.110
97*	{ArAgSxSm}	{SyArAg}	{SyAgSx}	{SyAgSm}	{SySxSm}		59.1	48	.130
98*	{ArAgSxSm}	{SyArAg}	{SyArSm}	{SyAgSm}	{SySxSm}		58.3	47	.124
99*	{ArAgSxSm}	{SyArAg}	{SyArSm}	{SyAgSx}	{SySxSm}		70.3	56	.095
100*	{ArAgSxSm}	{SyArAg}	{SyArSm}	{SyAgSx}	{SyAgSm}		59.4	48	.125
101*	{ArAgSxSm}	{SyArAg}	{SyArSx}	{SyAgSm}	{SySxSm}		61.2	50	.134
102*	{ArAgSxSm}	{SyArAg}	{SyArSx}	{SyAgSx}	{SySxSm}		72.9	59	.105
103*	{ArAgSxSm}	{SyArAg}	{SyArSx}	{SyAgSx}	{SyAgSm}		62.5	51	.129
104*	{ArAgSxSm}	{SyArAg}	{SyArSx}	{SyArSm}	{SySxSm}		73.2	58	.087
105*	{ArAgSxSm}	{SyArAg}	{SyArSx}	{SyArSm}	{SyAgSm}		61.8	50	.123
106*	{ArAgSxSm}	{SyArAg}	{SyArSx}	{SyArSm}	{SyAgSx}		74.3	59	.087
107*	{ArAgSxSm}	{SyArAg}	{SyArSx}	{SyArSm}	{SyAgSx}	{SySxSm}	69.8	55	.087
108*	{ArAgSxSm}	{SyAgSxSm}					46.6	40	.221
109*	{ArAgSxSm}	{SyAgSxSm}	{SyAr}				45.7	39	.213
110*	{ArAgSxSm}	{SyAgSxSm}	{SyArAg}				42.8	36	.203

(Cont'd)

TABLE C-8 : ASTHMA
(Cont'd)

Model	Margins Fit					L^2	df	p
						TEST OF MODEL		
111*	{ArAgSxSm}	{SyAgSxSm}	{SyArSx}			44.4	38	.221
112*	{ArAgSxSm}	{SyAgSxSm}	{SyArSm}			42.0	35	.193
113*	{ArAgSxSm}	{SyAgSxSm}	{SyArAg}	{SyArSx}		41.9	35	.198
114*	{ArAgSxSm}	{SyAgSxSm}	{SyArAg}	{SyArSm}		38.7	32	.194
115*	{ArAgSxSm}	{SyAgSxSm}	{SyArSx}	{SyArSm}		40.8	34	.197
116*	{ArAgSxSm}	{SyAgSxSm}	{SyArAg}	{SyArSx}	{SyArSm}	37.9	31	.184
117*	{ArAgSxSm}	{SyArSxSm}				72.8	60	.125
118*	{ArAgSxSm}	{SyArSxSm}	{SyAg}			63.8	57	.249
119*	{ArAgSxSm}	{SyArSxSm}	{SyArAg}			61.9	54	.214
120*	{ArAgSxSm}	{SyArSxSm}	{SyAgSx}			60.6	54	.251
121*	{ArAgSxSm}	{SyArSxSm}	{SyAgSm}			50.0	45	.281
122*	{ArAgSxSm}	{SyArSxSm}	{SyArAg}	{SyAgSx}		58.7	51	.214
123*	{ArAgSxSm}	{SyArSxSm}	{SyArAg}	{SyAgSm}		46.9	42	.279
124*	{ArAgSxSm}	{SyArSxSm}	{SyAgSx}	{SyAgSm}		47.2	42	.268
125*	{ArAgSxSm}	{SyArSxSm}	{SyArAg}	{SyAgSx}	{SyAgSm}	44.0	39	.269
126*	{ArAgSxSm}	{SyArAgSm}				50.1	40	.133
127*	{ArAgSxSm}	{SyArAgSm}	{SySx}			50.1	39	.110
128*	{ArAgSxSm}	{SyArAgSm}	{SyArSx}			49.1	38	.107
129*	{ArAgSxSm}	{SyArAgSm}	{SyAgSx}			47.0	36	.105
130*	{ArAgSxSm}	{SyArAgSm}	{SySxSm}			45.7	35	.106
131*	{ArAgSxSm}	{SyArAgSm}	{SyArSx}	{SyAgSx}		46.3	35	.096
132*	{ArAgSxSm}	{SyArAgSm}	{SyArSx}	{SySxSm}		44.4	34	.109
133*	{ArAgSxSm}	{SyArAgSm}	{SyAgSx}	{SySxSm}		42.5	32	.102
134*	{ArAgSxSm}	{SyArAgSm}	{SyArSx}	{SyAgSx}	{SySxSm}	41.7	31	.095
135*	{ArAgSxSm}	{SyArAgSx}				77.2	64	.124
136*	{ArAgSxSm}	{SyArAgSx}	{SySm}			73.4	60	.115
137*	{ArAgSxSm}	{SyArAgSx}	{SyArSm}			70.5	56	.092
[illegible]*	{ArAgSxSm}	{SyArAgSx}	{SyAgSm}			58.9	48	.134
139*	{ArAgSxSm}	{SyArAgSx}	{SySxSm}			69.2	56	.111
140*	{ArAgSxSm}	{SyArAgSx}	{SyArSm}	{SyAgSm}		55.2	44	.120
141*	{ArAgSxSm}	{SyArAgSx}	{SyArSm}	{SySxSm}		65.9	52	.092
142*	{ArAgSxSm}	{SyArAgSx}	{SyAgSm}	{SySxSm}		54.9	44	.126
143*	{ArAgSxSm}	{SyArAgSx}	{SyArSm}	{SyAgSm}	{SySxSm}	50.8	40	.118
144*	{ArAgSxSm}	{SyArAgSm}	{SyArAgSx}			42.0	32	.111
145*	{ArAgSxSm}	{SyArAgSm}	{SyArAgSx}	{SySxSm}		37.1	28	.116
146*	{ArAgSxSm}	{SyArSxSm}	{SyArAgSx}			55.3	48	.219
147*	{ArAgSxSm}	{SyArSxSm}	{SyArAgSx}	{SyAgSm}		41.3	36	.249
148*	{ArAgSxSm}	{SyArSxSm}	{SyArAgSm}			33.9	30	.285
149*	{ArAgSxSm}	{SyArSxSm}	{SyArAgSm}	{SyAgSx}		30.8	27	.280
150*	{ArAgSxSm}	{SyAgSxSm}	{SyArAgSx}			37.5	32	.232
151*	{ArAgSxSm}	{SyAgSxSm}	{SyArAgSx}	{SyArSm}		32.5	28	.256
152*	{ArAgSxSm}	{SyAgSxSm}	{SyArAgSm}			26.8	20	.141
153*	{ArAgSxSm}	{SyAgSxSm}	{SyArAgSm}	{SyArSx}		25.4	19	.148
154*	{ArAgSxSm}	{SyAgSxSm}	{SyArSxSm}			29.9	30	.468
155*	{ArAgSxSm}	{SyAgSxSm}	{SyArSxSm}	{SyArAg}		28.7	27	.375
156*	{ArAgSxSm}	{SyArSxSm}	{SyArAgSm}	{SyArAgSx}		27.5	24	.282
157*	{ArAgSxSm}	{SyAgSxSm}	{SyArAgSm}	{SyArAgSx}		17.1	16	.382
158*	{ArAgSxSm}	{SyAgSxSm}	{SyArSxSm}	{SyArAgSx}		25.3	24	.388
159*	{ArAgSxSm}	{SyAgSxSm}	{SyArSxSm}	{SyArAgSm}		18.9	15	.217
160*	{ArAgSxSm}	{SyAgSxSm}	{SyArSxSm}	{SyArAgSm}	{SyArAgSx}	13.0	12	.369
161*	{SyArAgSxSm}					0.0	0	---

TABLE C-9

Tests of Alternative Log-linear Models fit to 5-way Table of Symptom by Area by Age by Sex by Smoking. (Data from Boston Area Pulmonary Health Survey, 1983.)

POSITE OF COUGH, PHLEGM AND WHEEZE

Model	Margins Fit	TEST OF MODEL L^2	df	p
1	{ArAgSxSm} {Sy}	958.0	553	<.001
2*	{ArAgSxSm} {SySm}	537.4	525	.344
3	{ArAgSxSm} {SySx}	937.9	546	<.001
4	{ArAgSxSm} {SyAg}	920.1	532	<.001
5	{ArAgSxSm} {SyAr}	901.0	546	<.001
6*	{ArAgSxSm} {SySx} {SySm}	517.2	518	>.5
7*	{ArAgSxSm} {SyAg} {SySm}	507.1	504	.454
8	{ArAgSxSm} {SyAg} {SySx}	899.7	525	<.001
9*	{ArAgSxSm} {SyAr} {SySm}	505.8	518	>.5
10	{ArAgSxSm} {SyAr} {SySx}	879.9	539	<.001
11	{ArAgSxSm} {SyAr} {SyAg}	863.1	525	<.001
12*	{ArAgSxSm} {SyAg} {SySx} {SySm}	486.7	497	>.5
13*	{ArAgSxSm} {SyAr} {SySx} {SySm}	484.4	511	>.5
14*	{ArAgSxSm} {SyAr} {SyAg} {SySm}	476.0	497	>.5
15	{ArAgSxSm} {SyAr} {SyAg} {SySx}	841.7	518	<.001
16*	{ArAgSxSm} {SyAr} {SyAg} {SySx} {SySm}	454.4	490	>.5
17	{ArAgSxSm} {SyArAg}	833.0	504	<.001
18	{ArAgSxSm} {SyArAg} {SySx}	811.5	497	<.001
19*	{ArAgSxSm} {SyArAg} {SySm}	448.2	476	>.5
20*	{ArAgSxSm} {SyArAg} {SySx} {SySm}	426.3	469	>.5
21	{ArAgSxSm} {SyArSx}	873.6	532	<.001
22	{ArAgSxSm} {SyArSx} {SyAg}	835.4	511	<.001
23*	{ArAgSxSm} {SyArSx} {SySm}	478.1	504	>.5
24*	{ArAgSxSm} {SyArSx} {SyAg} {SySm}	448.1	483	>.5
25*	{ArAgSxSm} {SyArSm}	467.0	490	>.5
26*	{ArAgSxSm} {SyArSm} {SyAg}	437.1	469	>.5
27*	{ArAgSxSm} {SyArSm} {SySx}	445.3	483	>.5
28*	{ArAgSxSm} {SyArSm} {SyAg} {SySx}	415.2	462	>.5
29	{ArAgSxSm} {SyAgSx}	879.7	504	<.001
30	{ArAgSxSm} {SyAgSx} {SyAr}	820.9	497	<.001
31*	{ArAgSxSm} {SyAgSx} {SySm}	464.9	476	>.5
32*	{ArAgSxSm} {SyAgSx} {SyAr} {SySm}	432.6	469	>.5
33*	{ArAgSxSm} {SyAgSm}	405.0	420	>.5
34*	{ArAgSxSm} {SyAgSm} {SyAr}	374.2	413	>.5
35*	{ArAgSxSm} {SyAgSm} {SySx}	385.6	413	>.5
36*	{ArAgSxSm} {SyAgSm} {SyAr} {SySx}	354.9	406	>.5
37*	{ArAgSxSm} {SySxSm}	488.1	490	>.5
38*	{ArAgSxSm} {SySxSm} {SyAr}	455.4	483	>.5
39*	{ArAgSxSm} {SySxSm} {SyAg}	457.4	469	>.5
40*	{ArAgSxSm} {SySxSm} {SyAr} {SyAg}	425.4	462	>.5
41*	{ArAgSxSm} {SyAgSm} {SySxSm}	357.1	385	>.5
42*	{ArAgSxSm} {SyAgSm} {SySxSm} {SyAr}	325.3	378	>.5
43*	{ArAgSxSm} {SyAgSx} {SySxSm}	434.9	448	>.5
44*	{ArAgSxSm} {SyAgSx} {SySxSm} {SyAr}	402.8	441	>.5
45*	{ArAgSxSm} {SyAgSx} {SyAgSm}	366.5	392	>.5
46*	{ArAgSxSm} {SyAgSx} {SyAgSm} {SyAr}	334.3	385	>.5
47*	{ArAgSxSm} {SyArSm} {SySxSm}	416.4	455	>.5
48*	{ArAgSxSm} {SyArSm} {SySxSm} {SyAg}	386.1	434	>.5
49*	{ArAgSxSm} {SyArSm} {SyAgSm}	331.4	385	>.5
50*	{ArAgSxSm} {SyArSm} {SyAgSm} {SySx}	312.0	378	>.5

(Cont'd)

NOTE. Note. Variables are:
Ar: Area, 2 categories (SMSA and South Boston)
Ag: Age, 4 categories (18-34, 35-49, 50-64, 65+)
Sx: Sex, 2 Categories (Male and Female)
Sm: Smoking, 4 categories (Never Smoked Cigarettes; Ex Cigarette Smoker; 1-15 , 16-30, and 31+ Cigs. per day)
Sy: Symptom, Dependent Variable

L^2 is Likelihood-ratio chi-square statistic for fit of model
of freedom for test statistics.

TABLE C-9: COMPOSITE OF SYMPTOMS
(Cont'd)

							TEST OF MODEL		
Model	Margins Fit						L^2	df	p
51*	{ArAgSxSm}	{SyArSm}	{SyAgSx}				392.8	441	>.5
52*	{ArAgSxSm}	{SyArSx}	{SySxSm}				449.4	476	>.5
53*	{ArAgSxSm}	{SyArSx}	{SySxSm}	{SyAg}			419.3	455	>.5
54*	{ArAgSxSm}	{SyArSx}	{SyAgSm}				348.7	399	>.5
55	{ArAgSxSm}	{SyArSx}	{SyAgSx}				814.6	490	<.001
56*	{ArAgSxSm}	{SyArSx}	{SyAgSx}	{SySm}			426.3	462	>.5
57*	{ArAgSxSm}	{SyArSx}	{SyArSm}				439.5	476	>.5
58*	{ArAgSxSm}	{SyArSx}	{SyArSm}	{SyAg}			409.3	455	>.5
59*	{ArAgSxSm}	{SyArAg}	{SySxSm}				396.9	441	>.5
60*	{ArAgSxSm}	{SyArAg}	{SyAgSm}				344.8	392	>.5
61*	{ArAgSxSm}	{SyArAg}	{SyAgSm}	{SySx}			325.7	385	>.5
62	{ArAgSxSm}	{SyArAg}	{SyAgSx}				790.9	476	<.001
63*	{ArAgSxSm}	{SyArAg}	{SyAgSx}	{SySm}			403.8	448	>.5
64*	{ArAgSxSm}	{SyArAg}	{SyArSm}				410.0	448	>.5
65*	{ArAgSxSm}	{SyArAg}	{SyArSm}	{SySx}			388.1	441	>.5
66	{ArAgSxSm}	{SyArAg}	{SyArSx}				805.1	490	<.001
67*	{ArAgSxSm}	{SyArAg}	{SyArSx}	{SySm}			420.0	462	>.5
68*	{ArAgSxSm}	{SyAgSx}	{SyAgSm}	{SySxSm}			334.9	364	>.5
69*	{ArAgSxSm}	{SyAgSx}	{SyAgSm}	{SySxSm}	{SyAr}		302.8	357	>.5
70*	{ArAgSxSm}	{SyArSm}	{SyAgSm}	{SySxSm}			283.1	350	>.5
71*	{ArAgSxSm}	{SyArSm}	{SyAgSx}	{SySxSm}			363.1	413	>.5
72*	{ArAgSxSm}	{SyArSm}	{SyAgSx}	{SyAgSm}			290.8	357	>.5
73*	{ArAgSxSm}	{SyArSx}	{SyAgSm}	{SySxSm}			319.6	371	>.5
74*	{ArAgSxSm}	{SyArSx}	{SyAgSx}	{SySxSm}			396.4	434	>.5
75*	{ArAgSxSm}	{SyArSx}	{SyAgSx}	{SyAgSm}			328.4	378	>.5
76*	{ArAgSxSm}	{SyArSx}	{SyArSm}	{SySxSm}			411.0	448	>.5
77*	{ArAgSxSm}	{SyArSx}	{SyArSm}	{SySxSm}	{SyAg}		380.5	427	>.5
78*	{ArAgSxSm}	{SyArSx}	{SyArSm}	{SyAgSm}			306.1	371	>.5
79*	{ArAgSxSm}	{SyArSx}	{SyArSm}	{SyAgSx}			387.0	434	>.5
80*	{ArAgSxSm}	{SyArAg}	{SyAgSm}	{SySxSm}			295.7	357	>.5
81*	{ArAgSxSm}	{SyArAg}	{SyAgSx}	{SySxSm}			373.5	420	>.5
82*	{ArAgSxSm}	{SyArAg}	{SyAgSx}	{SyAgSm}			303.7	364	>.5
83*	{ArAgSxSm}	{SyArAg}	{SyArSm}	{SySxSm}			359.1	413	>.5
84*	{ArAgSxSm}	{SyArAg}	{SyArSm}	{SyAgSm}			301.6	364	>.5
85*	{ArAgSxSm}	{SyArAg}	{SyArSm}	{SyAgSm}	{SySx}		282.8	357	>.5
86*	{ArAgSxSm}	{SyArAg}	{SyArSm}	{SyAgSx}			365.2	420	>.5
87*	{ArAgSxSm}	{SyArAg}	{SyArSx}	{SySxSm}			391.0	434	>.5
88*	{ArAgSxSm}	{SyArAg}	{SyArSx}	{SyAgSm}			319.6	378	>.5
89	{ArAgSxSm}	{SyArAg}	{SyArSx}	{SyAgSx}			781.7	469	<.001
90*	{ArAgSxSm}	{SyArAg}	{SyArSx}	{SyAgSx}	{SySm}		394.9	441	>.5
91*	{ArAgSxSm}	{SyArAg}	{SyArSx}	{SyArSm}			382.2	434	>.5
92*	{ArAgSxSm}	{SyArSm}	{SyAgSx}	{SyAgSm}	{SySxSm}		259.8	329	>.5
93*	{ArAgSxSm}	{SyArSx}	{SyAgSx}	{SyAgSm}	{SySxSm}		296.9	350	>.5
94*	{ArAgSxSm}	{SyArSx}	{SyArSm}	{SyAgSm}	{SySxSm}		277.4	343	>.5
95*	{ArAgSxSm}	{SyArSx}	{SyArSm}	{SyAgSx}	{SySxSm}		357.3	406	>.5
96*	{ArAgSxSm}	{SyArSx}	{SyArSm}	{SyAgSx}	{SyAgSm}		284.8	350	>.5
97*	{ArAgSxSm}	{SyArAg}	{SyAgSx}	{SyAgSm}	{SySxSm}		272.0	336	>.5
98*	{ArAgSxSm}	{SyArAg}	{SyArSm}	{SyAgSm}	{SySxSm}		253.7	329	>.5
99*	{ArAgSxSm}	{SyArAg}	{SyArSm}	{SyAgSx}	{SySxSm}		335.1	392	>.5
100*	{ArAgSxSm}	{SyArAg}	{SyArSm}	{SyAgSx}	{SyAgSm}		259.8	336	>.5
101*	{ArAgSxSm}	{SyArAg}	{SyArSx}	{SyAgSm}	{SySxSm}		289.2	350	>.5
102*	{ArAgSxSm}	{SyArAg}	{SyArSx}	{SyAgSx}	{SySxSm}		364.6	413	>.5
103*	{ArAgSxSm}	{SyArAg}	{SyArSx}	{SyAgSx}	{SyAgSm}		294.9	357	>.5
104*	{ArAgSxSm}	{SyArAg}	{SyArSx}	{SyArSm}	{SySxSm}		353.8	406	>.5
105*	{ArAgSxSm}	{SyArAg}	{SyArSx}	{SyArSm}	{SyAgSm}		277.0	350	>.5
106*	{ArAgSxSm}	{SyArAg}	{SyArSx}	{SyArSm}	{SyAgSx}		356.5	413	>.5
107*	{ArAgSxSm}	{SyArAg}	{SyArSx}	{SyArSm}	{SyAgSx}	{SySxSm}	327.1	385	>.5
108*	{ArAgSxSm}	{SyAgSxSm}					274.0	280	>.5
109*	{ArAgSxSm}	{SyAgSxSm}	{SyAr}				242.5	273	>.5
110*	{ArAgSxSm}	{SyAgSxSm}	{SyArAg}				212.3	252	>.5

(Cont'd)

BLE C9 A: COMPOSITE OF SYMPTOMS
(cont d)

Model	Margins Fit					TEST OF MODEL L^2	df	p
111*	{ArAgSxSm}	{SyAgSxSm}	{SyArSx}			236.6	266	>.5
112*	{ArAgSxSm}	{SyAgSxSm}	{SyArSm}			200.2	245	>.5
113*	{ArAgSxSm}	{SyAgSxSm}	{SyArAg}	{SyArSx}		204.1	245	>.5
114*	{ArAgSxSm}	{SyAgSxSm}	{SyArAg}	{SyArSm}		170.5	224	>.5
115*	{ArAgSxSm}	{SyAgSxSm}	{SyArSx}	{SyArSm}		194.3	238	>.5
116*	{ArAgSxSm}	{SyAgSxSm}	{SyArAg}	{SyArSx}	{SyArSm}	163.3	217	>.5
117*	{ArAgSxSm}	{SyArSxSm}				381.2	420	>.5
118*	{ArAgSxSm}	{SyArSxSm}	{SyAg}			350.8	399	>.5
119*	{ArAgSxSm}	{SyArSxSm}	{SyArAg}			323.9	378	>.5
120*	{ArAgSxSm}	{SyArSxSm}	{SyAgSx}			328.2	378	>.5
121*	{ArAgSxSm}	{SyArSxSm}	{SyAgSm}			248.2	315	>.5
122*	{ArAgSxSm}	{SyArSxSm}	{SyArAg}	{SyAgSx}		298.7	357	>.5
123*	{ArAgSxSm}	{SyArSxSm}	{SyArAg}	{SyAgSm}		219.5	294	>.5
124*	{ArAgSxSm}	{SyArSxSm}	{SyAgSx}	{SyAgSm}		225.1	294	>.5
125*	{ArAgSxSm}	{SyArSxSm}	{SyArAg}	{SyAgSx}	{SyAgSm}	194.6	273	>.5
126*	{ArAgSxSm}	{SyArAgSm}				225.4	280	>.5
127*	{ArAgSxSm}	{SyArAgSm}	{SySx}			205.5	273	>.5
128*	{ArAgSxSm}	{SyArAgSm}	{SyArSx}			199.8	266	>.5
129*	{ArAgSxSm}	{SyArAgSm}	{SyAgSx}			183.2	252	>.5
130*	{ArAgSxSm}	{SyArAgSm}	{SySxSm}			176.6	245	>.5
131*	{ArAgSxSm}	{SyArAgSm}	{SyArSx}	{SyAgSx}		174.6	245	>.5
132*	{ArAgSxSm}	{SyArAgSm}	{SyArSx}	{SySxSm}		170.9	238	>.5
133*	{ArAgSxSm}	{SyArAgSm}	{SyAgSx}	{SySxSm}		153.2	224	>.5
134*	{ArAgSxSm}	{SyArAgSm}	{SyArSx}	{SyAgSx}	{SySxSm}	145.5	217	>.5
135	{ArAgSxSm}	{SyArAgSx}				752.8	448	<.001
136*	{ArAgSxSm}	{SyArAgSx}	{SySm}			364.5	420	>.5
137*	{ArAgSxSm}	{SyArAgSx}	{SyArSm}			326.4	392	>.5
138*	{ArAgSxSm}	{SyArAgSx}	{SyAgSm}			264.6	336	>.5
139*	{ArAgSxSm}	{SyArAgSx}	{SySxSm}			334.3	392	>.5
140*	{ArAgSxSm}	{SyArAgSx}	{SyArSm}	{SyAgSm}		222.6	308	>.5
141*	{ArAgSxSm}	{SyArAgSx}	{SyArSm}	{SySxSm}		296.6	364	>.5
142*	{ArAgSxSm}	{SyArAgSx}	{SyAgSm}	{SySxSm}		233.4	308	>.5
143*	{ArAgSxSm}	{SyArAgSx}	{SyArSm}	{SyAgSm}	{SySxSm}	190.4	280	>.5
144*	{ArAgSxSm}	{SyArAgSm}	{SyArAgSx}			146.7	224	>.5
145*	{ArAgSxSm}	{SyArAgSm}	{SyArAgSx}	{SySxSm}		114.3	196	>.5
146*	{ArAgSxSm}	{SyArSxSm}	{SyArAgSx}			271.2	336	>.5
147*	{ArAgSxSm}	{SyArSxSm}	{SyArAgSx}	{SyAgSm}		171.4	252	>.5
148*	{ArAgSxSm}	{SyArSxSm}	{SyArAgSm}			143.1	210	>.5
149*	{ArAgSxSm}	{SyArSxSm}	{SyArAgSm}	{SyAgSx}		112.1	189	>.5
150*	{ArAgSxSm}	{SyAgSxSm}	{SyArAgSx}			173.0	224	>.5
151*	{ArAgSxSm}	{SyAgSxSm}	{SyArAgSx}	{SyArSm}		126.8	196	>.5
152*	{ArAgSxSm}	{SyAgSxSm}	{SyArAgSm}			97.1	140	>.5
153*	{ArAgSxSm}	{SyAgSxSm}	{SyArAgSm}	{SyArSx}		87.4	133	>.5
154*	{ArAgSxSm}	{SyAgSxSm}	{SyArSxSm}			164.9	210	>.5
155*	{ArAgSxSm}	{SyAgSxSm}	{SyArSxSm}	{SyArAg}		126.0	189	>.5
156*	{ArAgSxSm}	{SyArSxSm}	{SyArAgSm}	{SyArAgSx}		98.3	168	>.5
157*	{ArAgSxSm}	{SyAgSxSm}	{SyArAgSm}	{SyArAgSx}		67.3	112	>.5
158*	{ArAgSxSm}	{SyAgSxSm}	{SyArSxSm}	{SyArAgSx}		107.6	168	>.5
159*	{ArAgSxSm}	{SyAgSxSm}	{SyArSxSm}	{SyArAgSm}		63.9	105	>.5
160*	{ArAgSxSm}	{SyAgSxSm}	{SyArSxSm}	{SyArAgSm}	{SyArAgSx}	51.0	84	>.5
161*	{SyArAgSxSm}					0.0	0	---

APPENDIX D
ESTIMATES OF SAMPLE DESIGN AND INTERVIEWER EFFECTS FOR MORE COMPLETE SET OF SURVEY MEASUREMENTS.

Table D-1

Interviewer Effects on Variance of Estimates for Metric and Dichotomous Variables

Variable	R_i	deft	deff	N_r	N_i	$\overline{N}_i$	F	p
119 Usually Cough	0.007	1.185	1.404	2663	45	59.2	1.414	0.04
121 Usually Cough - a.m.	0.007	1.186	1.407	2662	45	59.2	1.417	0.04
122 Usually Cough - Other	0.027	1.599	2.556	2657	45	59.0	2.626	<0.01
127 Usually Pleghm	0.017	1.416	2.004	2661	45	59.1	2.039	<0.01
129 Usually Phlegm - a.m	0.029	1.640	2.690	2660	45	59.1	2.771	<0.01
130 Usually Phlegm - Other	0.013	1.332	1.774	2652	45	58.9	1.798	<0.01
135 Cough & Phlegm - 3 wks	0.006	1.165	1.358	2659	45	59.1	1.366	0.06
138 Ever Wheeze	0.001	1.037	1.075	2648	45	58.8	1.076	0.34
143 Chest Wheeze - Cold	0.007	1.185	1.404	2655	45	59.0	1.414	0.04
144 Chest Wheeze - No Cold	0.006	1.158	1.342	2657	45	59.0	1.350	0.06
148 Cold Usually in Chest	0.011	1.279	1.637	2636	45	58.6	1.655	0.01
49 Chest Cold	0.004	1.120	1.254	2658	45	59.1	1.259	0.12
154 Able to Walk	0.018	1.434	2.056	2664	45	59.2	2.094	<0.01
157 Breathless - on hill	0.001	1.030	1.061	2648	45	58.8	1.062	0.36
208 Bronchitis	0.015	1.370	1.877	2665	45	59.0	1.906	<0.01
212 Pneumonia	0.002	1.043	1.088	2661	45	59.1	1.090	0.32
216 Hay Fever	0.013	1.334	1.780	2658	45	59.1	1.804	<0.01
220 Sinus	0.000	1.002	1.005	2661	45	59.1	1.005	0.46
224 Pulmonary T.B.	-0.004	0.870	0.757	2660	45	59.1	0.754	0.88
228 Emphysema	0.015	1.360	1.850	2659	45	59.1	1.877	<0.01
233 Chronic Bronchitis	0.005	1.144	1.308	2654	45	59.0	1.315	0.08
238 Asthma	0.006	1.164	1.356	2663	45	59.2	1.364	0.06
246 Hypertension	0.013	1.329	1.765	2661	45	59.1	1.789	<0.01
248 Heart Trouble	0.005	1.139	1.296	2660	45	59.1	1.303	0.09
250 Headaches	0.009	1.237	1.530	2659	45	59.1	1.544	0.01
251 Lung Trouble <16	0.009	1.242	1.542	2654	45	59.0	1.557	0.01

R_i: intra-interviewer response correlation; deff: Kish design effect coefficient; deft: square root of deff; N_r: number of respondents; N_i:number of interviewers; F: F ratio ith $d_{f=N_c/N_r}$.

Table D-1 cont'd

Variable	R_i	deft	deff	N_r	N_i	$\overline{N}_i$	F	p
252 Food Allergy	0.013	1.320	1.741	2650	45	58.9	1.764	<0.01
254 Pollen Allergy	0.007	1.182	1.397	2653	45	59.0	1.407	0.04
256 Skin Allergy	0.004	1.121	1.256	2657	45	59.0	1.262	0.12
259 Working Full Time	0.029	1.644	2.704	2660	45	59.1	2.786	<0.01
260 Ever Worked	0.054	1.474	2.172	1030	45	22.9	2.295	<0.01
269 Usual Job?	0.004	1.099	1.207	2483	45	55.2	1.212	0.16
276 Gas/Chemical Fumes	0.010	1.240	1.537	2494	45	55.4	1.552	0.01
308 Dusty Job	0.010	1.253	1.569	2496	45	55.5	1.586	0.01
312 Gas/Fumes-Not at Work	0.001	1.017	1.034	2661	45	59.1	1.035	0.49
324 Ever Smoked Cigarettes	-0.002	0.937	0.879	2664	45	59.2	0.877	0.70
325 Smoked Cigarettes at all	0.045	1.414	2.001	1045	45	23.2	2.095	<0.01
329 Smoke Cigarettes Now	0.007	1.126	1.267	1662	45	36.9	1.277	0.11
342 Ever Smoke Cigar or Pipe	0.036	1.490	2.219	1566	45	34.8	2.302	<0.01
343 Ever Smoke Pipe	0.025	1.278	1.633	1139	44	25.9	1.676	<0.01
355 Ever Smoke Cigar	0.050	1.494	2.231	1129	44	25.7	2.348	<0.01
369 Ever Smoke non-Tobacco	0.013	1.319	1.740	2658	45	59.1	1.762	<0.01
370 Ever Live w/ Smoker	0.007	1.180	1.392	2662	45	59.2	1.401	0.04
373 Live w/ Smoker Now	0.010	1.211	1.466	2194	45	48.8	1.480	0.02
375 Usually Exposed Smokers	0.006	1.170	1.368	2656	45	59.0	1.377	0.05
449 Ever Live South Boston	-0.002	0.967	0.935	1373	44	31.2	0.933	0.60
467 Gas Log	0.007	1.132	1.281	1762	45	39.2	1.291	0.10
469 Heat with Stove	0.009	1.192	1.421	2127	45	47.3	1.434	0.03
471 Air Condition-All/Part	0.003	1.041	1.083	1290	45	28.7	1.086	0.33
472 A.C. Living Room	0.005	1.055	1.114	1094	45	24.3	1.119	0.28
473 A.C. Bedroom	0.010	1.109	1.229	1096	45	24.4	1.241	0.14
474 Sick from Water	0.003	1.081	1.169	2642	45	58.7	1.172	0.20
475 Sick from Air	0.000	1.004	1.008	2652	45	58.9	1.008	0.46
146 # Yrs Wheezy Chest	0.033	1.325	1.757	1064	44	24.2	1.816	<0.01
327 Age Start Smoking	-0.002	0.960	0.922	1660	45	36.9	0.920	0.62

Table D-2

Interviewer Effects on Variance of Estimates for Non-Metric Variables

Variable	R_i	deft	deff	N_r	N_i	$\overline{N}_i$	F	p
377:1 Intensity Smoke Expos:1	0.006	1.160	1.346	2665	45	59.2	1.354	0.06
377:2 Intensity Smoke Expos:2	0.008	1.218	1.485	2665	45	59.2	1.497	0.02
377:3 Intensity Smoke Expos:3	0.008	1.210	1.465	2665	45	59.2	1.477	0.02
376:1 Freq Smoke Expos:1	0.006	1.173	1.376	2665	45	59.2	1.385	0.05
376:2 Freq Smoke Expos:2	0.007	1.195	1.428	2665	45	59.2	1.439	0.03
376:3 Freq Smoke Expos:3	0.018	1.422	2.022	2665	45	59.2	2.058	<0.01
374:1 Intensity Home Smoke:1	0.005	1.130	1.277	2665	45	59.2	1.283	0.10
374:2 Intensity Home Smoke:2	0.008	1.202	1.446	2665	45	59.2	1.457	0.03
374:3 Intensity Home Smoke:3	-0.005	0.856	0.732	2665	45	59.2	0.729	0.91
336:1 Inhale Cigarette:deep	-0.004	0.860	0.740	2665	45	59.2	0.737	0.90
336:2 Inhale Cigarette:mod	0.002	1.069	1.142	2665	45	59.2	1.145	0.24
`36:3 Inhale Cigarette:mild	0.009	1.237	1.531	2665	45	59.2	1.545	0.01
336:4 Inhale Cigarette:none	0.003	1.074	1.153	2665	45	59.2	1.156	0.22
339:1 Inhaled:deep	-0.000	0.987	0.974	2665	45	59.2	0.974	0.52
339:2 Inhaled:moderate	0.007	1.177	1.386	2665	45	59.2	1.395	0.04
339:3 Inhaled:mild	0.009	1.237	1.531	2665	45	59.2	1.545	0.01
339:4 Inhaled:none	0.016	1.384	1.916	2665	45	59.2	1.947	<0.01
311:1 Dust:mild	-0.002	0.952	0.906	2665	45	59.2	0.905	0.65
311:2 Dust:moderate	0.005	1.134	1.286	2665	45	59.2	1.292	0.09
311:3 Dust:severe	0.010	1.256	1.577	2665	45	59.2	1.593	0.01
280:1 Fumes at work:mild	0.002	1.065	1.133	2665	45	59.2	1.136	0.25
280:2 Fumes at work:moderate	0.015	1.360	1.849	2665	45	59.2	1.876	<0.01
280:3 Fumes at work:severe	0.002	1.055	1.113	2665	45	59.2	1.115	0.28
279:1 Fumes at work:often	0.010	1.250	1.563	2665	45	59.2	1.578	0.01
279:2 Fumes at work:some	-0.001	0.978	0.956	2665	45	59.2	0.955	0.56
279:3 Fumes at work:rarely	0.006	1.152	1.327	2665	45	59.2	1.334	0.07
118:1 Compare Health:better	0.016	1.398	1.954	2665	45	59.2	1.987	<0.01

R_i: intra-interviewer response correlation; deff: Kish design effect coefficient; deft: square root of deff; N_r: number of respondents; N_i:number of interviewers; F: F ratio with $d_{f=N_i/N_r}$.

Variable	R_i	deft	deff	N_r	N_i	$\overline{N}_i$	F	p
118:2 Compare Health:same	0.014	1.340	1.796	2665	45	59.2	1.821	<0.01
118:3 Compare Health:worse	-0.001	0.956	0.914	2665	45	59.2	0.913	0.64
422:1 Married	0.002	1.058	1.120	2665	45	59.2	1.122	0.27
422:2 Widowed	0.008	1.218	1.483	2665	45	59.2	1.495	0.02
422:34 Divorced/Sep.	-0.000	0.991	0.982	2665	45	59.2	0.982	0.51
422:5 Never Married	0.003	1.072	1.149	2665	45	59.2	1.152	0.23
423:1 White	0.009	1.224	1.497	2665	45	59.2	1.510	0.02
423:2+ Non-White	0.009	1.241	1.540	2665	45	59.2	1.554	0.01
463:1 Heat:gas	0.004	1.121	1.256	2665	45	59.2	1.262	0.12
463:2 Heat:oil	0.006	1.157	1.338	2665	45	59.2	1.346	0.06
463:3+ Heat:other	0.003	1.096	1.201	2665	45	59.2	1.205	0.17
464:1 Heat:water	0.017	1.408	1.983	2665	45	59.2	2.017	<0.01
464:2 Heat:air	0.012	1.301	1.694	2665	45	59.2	1.714	<0.01
464:3+ Heat:electric	0.018	1.436	2.061	2665	45	59.2	2.099	<0.01
466:1 Cook:gas	0.013	1.317	1.735	2665	45	59.2	1.757	<0.01
466:2 Cook:electric	0.012	1.311	1.718	2665	45	59.2	1.739	<0.01
265:1 Employed by Self	0.004	1.113	1.240	2665	45	59.2	1.245	0.13
265:2 Employed by Other	0.015	1.361	1.852	2665	45	59.2	1.880	<0.01
266:1 Work:home	0.027	1.604	2.572	2665	45	59.2	2.643	<0.01
266:2 Work:away	0.029	1.647	2.714	2665	45	59.2	2.796	<0.01
340:1 Filters:all	-0.000	0.996	0.991	2665	45	59.2	0.991	0.49
340:2 Filters: >1/2	-0.003	0.894	0.800	2665	45	59.2	0.797	0.83
340:3 Filters: <1/2	-0.003	0.923	0.851	2665	45	59.2	0.849	0.75
340:4 Filters: 1/2	-0.002	0.945	0.893	2665	45	59.2	0.891	0.68
340:5 Filters: none	0.002	1.052	1.107	2665	45	59.2	1.109	0.29
117:1 General Health	0.010	1.268	1.608	2665	45	59.2	1.625	0.01
117:2 General Health	0.005	1.127	1.269	2665	45	59.2	1.275	0.11
117:3 General Health	0.006	1.152	1.327	2665	45	59.2	1.335	0.07
117:4 General Health	0.001	1.032	1.066	2665	45	59.2	1.067	0.35

Table D-3

Sample Design Effects on Variance of Estimates
for Metric and Dichotomous Variables (SMSA sample)

Variable	R_i	deft	deff	N_r	N_c	$\overline{N}_c$	F	p
119 Usually Cough	0.052	1.090	1.187	1398	303	4.6	1.252	0.01
121 Usually Cough - a.m.	0.060	1.103	1.216	1396	303	4.6	1.293	<0.01
122 Usually Cough - Other	0.021	1.038	1.077	1395	303	4.6	1.101	0.14
127 Usually Phlegm	0.067	1.114	1.241	1395	303	4.6	1.330	<0.01
129 Usually Phlegm - a.m.	0.047	1.081	1.169	1394	303	4.6	1.226	0.01
130 Usually Phlegm - Other	0.077	1.130	1.277	1394	303	4.6	1.384	<0.01
135 Cough & Phlegm - 3 wks	0.057	1.099	1.207	1396	303	4.6	1.281	<0.01
138 Ever Wheeze	0.038	1.065	1.135	1390	303	4.6	1.179	0.03
143 Chest Wheeze - Cold	0.023	1.041	1.083	1393	303	4.6	1.109	0.12
144 Chest Wheeze - No Cold	0.042	1.074	1.153	1395	303	4.6	1.204	0.02
148 Cold Usually in Chest	0.027	1.047	1.096	1380	303	4.6	1.126	0.09
49 Chest Cold	0.075	1.127	1.269	1396	303	4.6	1.372	<0.01
154 Able to Walk	0.147	1.238	1.533	1398	303	4.6	1.798	<0.01
157 Breathless - Hill	0.060	1.103	1.216	1393	302	4.6	1.293	<0.01
208 Bronchitis	-0.013	0.976	0.953	1392	303	4.6	0.941	0.74
212 Pneumonia	0.047	1.081	1.168	1395	303	4.6	1.225	0.01
216 Hay Fever	0.038	1.067	1.138	1395	303	4.6	1.184	0.03
220 Sinus	0.050	1.087	1.181	1396	303	4.6	1.244	0.01
224 Pulmonary T.B.	-0.029	0.945	0.894	1396	303	4.6	0.868	0.93
228 Emphysema	-0.016	0.971	0.943	1396	303	4.6	0.928	0.78
233 Chronic Bronchitis	0.038	1.066	1.137	1392	303	4.6	1.182	0.03
238 Asthma	0.048	1.083	1.173	1398	303	4.6	1.232	0.01
246 Hypertension	0.066	1.113	1.238	1396	303	4.6	1.326	<0.01
248 Heart Trouble	0.051	1.088	1.184	1396	303	4.6	1.247	0.01
250 Headaches	0.070	1.120	1.254	1397	303	4.6	1.349	<0.01
251 Lung Trouble <16	0.033	1.058	1.120	1394	303	4.6	1.158	0.05

R_i: intra-cluster response correlation; deff: Kish design effect coefficient; deft: square root of deff; N_r: number of respondents; N_c:number of clusters. F: F ratio with $d_{f=N_c/N_r}$.

Table D-3 (cont'd)

Variable	R_i	deft	deff	N_r	N_c	$\overline{N}_c$	F	p
252 Food Allergy	0.021	1.038	1.077	1394	303	4.6	1.100	0.14
254 Pollen Allergy	0.038	1.066	1.136	1393	303	4.6	1.180	0.03
256 Skin Allergy	0.029	1.051	1.106	1396	303	4.6	1.139	0.07
259 Working Full Time	0.034	1.059	1.121	1394	303	4.6	1.160	0.05
269 Usual Job?	0.043	1.070	1.146	1316	302	4.4	1.198	0.02
276 Gas/Chemical Fumes	0.064	1.102	1.215	1320	302	4.4	1.298	<0.01
308 Dusty Job Ever	0.049	1.079	1.165	1320	302	4.4	1.225	0.01
312 Gas/Fumes non-work	0.055	1.094	1.198	1396	303	4.6	1.267	<0.01
324 Ever Smoke	0.040	1.070	1.144	1398	303	4.6	1.192	0.02
369 Ever Smoke Non-tobacco	0.087	1.147	1.315	1395	303	4.6	1.441	<0.01
370 Ever Live w/ Smoker	0.066	1.113	1.238	1395	303	4.6	1.326	<0.01
373 Live w/ Smoker Now	0.082	1.108	1.228	1142	301	3.8	1.337	<0.01
375 Exposed to Smoker	0.052	1.089	1.187	1391	303	4.6	1.252	0.01
319 Weight	0.065	1.109	1.231	1384	303	4.6	1.316	<0.01
332 Age Quit Smoking	0.060	1.103	1.216	1393	303	4.6	1.293	<0.01
449 Ever Live S. Boston	0.031	1.053	1.109	1379	303	4.6	1.144	0.07
454 Length Residence	0.178	1.274	1.624	1367	303	4.5	1.974	<0.01
469 Stove for Heat	0.090	1.137	1.293	1293	303	4.3	1.420	<0.01
461 # Moves	0.226	1.346	1.812	1389	302	4.6	2.340	<0.01
470 Air Conditioning	0.065	1.111	1.234	1397	303	4.6	1.320	<0.01
474 Sick from Water	0.058	1.099	1.207	1384	303	4.6	1.281	<0.01
475 Sick from Air	0.110	1.182	1.396	1390	302	4.6	1.569	<0.01
476 How Cooperative	0.086	1.141	1.302	1364	303	4.5	1.425	<0.01

Table D-4

Sample Design Effects on Variance of Estimates
for Non-Metric Variables (SMSA only)

Variable	R_i	deft	deff	N_r	N_c	$\bar{N}_c$	F	p
377:1 Intensity Smoke Expos:1	0.006	1.011	1.022	1398	303	4.6	1.028	0.38
377:2 Intensity Smoke Expos:2	0.002	1.043	1.088	1398	303	4.6	1.115	0.11
377:3 Intensity Smoke Expos:3	0.046	1.079	1.165	1398	303	4.6	1.221	0.01
376:1 Freq Smoke Expos:1	0.057	1.098	1.206	1398	303	4.6	1.279	<0.01
376:2 Freq Smoke Expos:2	0.064	1.109	1.230	1398	303	4.6	1.314	<0.01
376:3 Freq Smoke Expos:3	0.058	1.099	1.209	1398	303	4.6	1.283	<0.01
374:1 Intensity Home Smoke:1	0.031	1.054	1.111	1398	303	4.6	1.146	0.07
374:2 Intensity Home Smoke:2	0.072	1.122	1.259	1398	303	4.6	1.356	<0.01
374:3 Intensity Home Smoke:3	0.066	1.113	1.238	1398	303	4.6	1.325	<0.01
336:1 Inhale Cigarette:deep	0.006	1.012	1.023	1398	303	4.6	1.030	0.37
336:2 Inhale Cigarette:mod	0.024	1.043	1.088	1398	303	4.6	1.115	0.11
336:3 Inhale Cigarette:mild	0.052	1.090	1.189	1398	303	4.6	1.254	0.01
336:4 Inhale Cigarette:none	0.042	1.073	1.151	1398	303	4.6	1.201	0.02
339:1 Inhaled:deep	0.010	1.018	1.036	1398	303	4.6	1.046	0.31
339:2 Inhaled:moderate	0.024	1.043	1.087	1398	303	4.6	1.114	0.12
339:3 Inhaled:mild	-0.002	0.996	0.991	1398	303	4.6	0.989	0.54
339:4 Inhaled:none	0.076	1.129	1.275	1398	303	4.6	1.380	<0.01
311:1 Dust:mild	0.027	1.047	1.097	1398	303	4.6	1.127	0.09
311:2 Dust:moderate	0.031	1.055	1.114	1398	303	4.6	1.150	0.06
311:3 Dust:severe	0.039	1.067	1.139	1398	303	4.6	1.185	0.03
280:1 Fumes at work:mild	0.047	1.082	1.170	1398	303	4.6	1.228	0.01
280:2 Fumes at work:moderate	0.056	1.097	1.203	1398	303	4.6	1.275	<0.01
280:3 Fumes at work:severe	0.007	1.013	1.026	1398	303	4.6	1.033	0.36
279:1 Fumes at work:often	0.016	1.029	1.059	1398	303	4.6	1.077	0.20
279:2 Fumes at work:some	0.082	1.138	1.295	1398	303	4.6	1.410	<0.01
279:3 Fumes at work:rarely	0.023	1.041	1.084	1398	303	4.6	1.110	0.12
118:1 Compare Health:better	0.059	1.102	1.214	1398	303	4.6	1.291	<0.01

R_i: intra-cluster response correlation; deff: Kish design effect coefficient; deft: square root of deff; N_r: number of respondents; N_c:number of clusters. F: F ratio with df=N_c/N

Table D-4 cont'd

Variable	R_i	deft	deff	N_r	N_c	$\bar{N}_c$	F	p
118:2 Compare Health:same	0.050	1.086	1.179	1398	303	4.6	1.241	0.01
118:3 Compare Health:worse	0.038	1.066	1.136	1398	303	4.6	1.180	0.03
422:1 Married	0.096	1.160	1.346	1398	303	4.6	1.489	<0.01
422:2 Widowed	-0.005	0.991	0.983	1398	303	4.6	0.978	0.59
422:3+ Divorced/Sep.	-0.008	0.985	0.971	1398	303	4.6	0.963	0.65
422:5 Never Married	0.119	1.196	1.430	1398	303	4.6	1.623	<0.01
423:1 White	0.322	1.471	2.164	1398	303	4.6	3.192	<0.01
423:2+ Non-White	0.326	1.476	2.179	1398	303	4.6	3.234	<0.01
463:1 Heat:gas	0.099	1.165	1.358	1398	303	4.6	1.507	<0.01
463:2 Heat:oil	0.069	1.117	1.249	1398	303	4.6	1.341	<0.01
463:3+ Heat:other	0.039	1.068	1.141	1398	303	4.6	1.187	0.03
464:1 Heat:water	0.057	1.099	1.207	1398	303	4.6	1.281	<0.01
464:2 Heat:air	0.084	1.142	1.303	1398	303	4.6	1.423	<0.01
464:3+ Heat:electric	0.052	1.090	1.188	1398	303	4.6	1.253	0.01
466:1 Cook:gas	0.165	1.264	1.598	1398	303	4.6	1.914	<0.01
466:2 Cook:electric	0.173	1.275	1.626	1398	303	4.6	1.966	<0.01
265:1 Employed by Self	0.016	1.029	1.059	1398	303	4.6	1.076	0.21
265:2 Employed by Other	0.006	1.011	1.023	1398	303	4.6	1.029	0.37
340:1 Filters:all	0.069	1.117	1.248	1398	303	4.6	1.340	<0.01
340:2 Filters: >1/2	0.083	1.140	1.300	1398	303	4.6	1.417	<0.01
340:3 Filters: <1/2	0.005	1.009	1.019	1398	303	4.6	1.024	0.39
340:4 Filters: 1/2	0.076	1.129	1.275	1398	303	4.6	1.380	<0.01
340:5 Filters: none	0.000	1.000	1.000	1398	303	4.6	1.000	1.50
266:1 Work:home	0.038	1.066	1.137	1398	303	4.6	1.182	0.03
266:2 Work:away	0.047	1.082	1.170	1398	303	4.6	1.228	0.01
117:1 General Health	0.093	1.155	1.355	1398	303	4.6	1.471	<0.01
117:2 General Health	0.035	1.062	1.127	1398	303	4.6	1.168	0.04
117:3 General Health	0.048	1.083	1.174	1398	303	4.6	1.233	0.01
117:4 General Health	0.115	1.190	1.416	1398	303	4.6	1.601	<0.01

APPENDIX E: SURVEY QUESTIONNAIRE

:1-2 C173

ID# ______
FOR OFFICE USE ONLY

:3 - 6

1
:7

INTERVIEWER NAME: ______

:8-9 INTERVIEWER ID NUMBER: ______

:10-13 INTERVIEW DATE: ______ (MONTH) ______ (DAY)

INTERVIEW LENGTH: ______

:14-15 SUB-SAMPLE NUMBER: ______

:16 FORM NUMBER B: 2

BOSTON AREA HEALTH QUESTIONNAIRE

Center for Survey Research

A Facility of the University of Massachusetts/Boston
and the Joint Center of Urban Studies
of Harvard University and M.I.T.

July/August, 1983

Confidential: No information shall be presented or published in any way that would identify any household or individual.

INTRODUCTION: a. MENTION CSR

b. PURPOSE: To learn more about the health of people living in the Boston area -- and about things in their lives that may affect their health.

Before we begin, there are a few points I need to cover.

1. I want to assure you that all information you give will be completely confidential and that none of it will be released in any way that would permit identification of you or your family.

2. Your participation in this study is, of course, voluntary.

3. If there is any question you would prefer not to answer, just tell me and we will go on to the next question.

TIME NOW: __________

I would like to begin with some general questions about your health.

1:17 A1. NOT IN THIS FORM. GO TO A2.

:18 A2. Compared to most other people your age, is your health better than most, about the same as most, or worse than most.

1 [ ] BETTER

2 [ ] ABOUT THE SAME

3 [ ] WORSE

These next questions are about coughing.

By "cough" I mean almost anything that you consider to be a cough. A cough includes coughing when you first go out of doors or when you first smoke a cigarette. A cough does not include just clearing your throat.

1:19 A3. Do you usually have a cough?

1 [ ] YES

2 [ ] NO (SKIP TO A5)

:20 A4. Do you usually cough as much as 4 times a day at least 4 days a week?

1 [ ] YES

2 [ ] NO

:21 A5. Do you usually cough at all when you get up or first thing in the morning?

1 [ ] YES

2 [ ] NO

:22 A6. What about the rest of the day or night -- do you usually cough then?

1 [ ] YES

2 [ ] NO

:23 A7. INTERVIEWER CHECK

1 [ ] YES IN ANY OF THE ABOVE (GO TO A8)

2 [ ] NO IN ALL OF THE ABOVE (SKIP TO A10)

:24 A8. Do you usually cough on most days for at least three months in a row during the year?

1 [ ] YES

2 [ ] NO

:25-26 A9. For how many years have you had this cough?

_____ NUMBER OF YEARS

Now some questions about phlegm and other chest problems.

1:27 A10. Do you usually bring up phlegm from your chest?

1 [ ] YES

2 [ ] NO (SKIP TO A12)

:28 A11. Do you usually bring up phlegm as much as twice a day at least 4 days a week?

1 [ ] YES

2 [ ] NO

:29 A12. Do you usually bring up phlegm <u>at all</u> when you get up or first thing in the morning?

1 [ ] YES

2 [ ] NO

:30 A13. What about the rest of the day or night -- do you usually bring up phlegm then?

1 [ ] YES

2 [ ] NO

:31 A14. INTERVIEWER CHECK

1 [ ] YES TO ANY OF THE ABOVE (GO TO A15)

2 [ ] NO TO ALL OF THE ABOVE (SKIP TO A17)

:32 A15. Have you brought up phlegm on most days for as much as three months in a row?

1 [ ] YES

2 [ ] NO

:33-34 A16. For how many years have you had trouble with phlegm?

_____ NUMBER OF YEARS

1:35 A17. Have you had periods of both coughing and bringing up phlegm that have lasted for at least 3 weeks?

1 [ ] YES

2 [ ] NO (SKIP TO A18)

:36-37 A17a. For how many years have you had at least one such episode per year?

___________ NUMBER OF YEARS

:38 A18. Have you ever had an attack of wheezing that has made you feel short of breath?

1 [ ] YES

2 [ ] NO (SKIP TO A19)

:39-40 A18a. How old were you when this first happened to you?

_____ AGE IN YEARS

:41 A18b. Has this happened more than once?

1 [ ] YES

2 [ ] NO

:42 A18c. Did you ever get medicine or medical treatment for the(se) attack(s)?

1 [ ] YES

2 [ ] NO

1:43 A19. Does your chest ever sound wheezy or whistling when you have a cold?

1 [ ] YES

2 [ ] NO

:44 A20. Does your chest ever sound wheezy or whistling when you don't have a cold?

1 [ ] YES

2 [ ] NO (SKIP TO A21)

:45 A20a. Would you say this happens only occasionally or most days or nights?

1 [ ] OCCASIONALLY → (SKIP TO A22)

2 [ ] MOST DAYS OR NIGHTS → (SKIP TO A22)

A21. INTERVIEWER CHECK

1 [ ] YES IN ANY OF THE ABOVE (GO TO A22)

2 [ ] NO IN BOTH OF THE ABOVE (SKIP TO A23)

:46-47 A22. How many years ago did you first have a wheezy or whistling chest?

______ NUMBER OF YEARS

:48 A23. If you get a cold does it <u>usually</u> go to your chest?

1 [ ] YES

2 [ ] NO

1:49 A24. During the past 3 years, have you had any colds or chest illnesses that have kept you indoors at home, or in bed?

1 [ ] YES

2 [ ] NO (SKIP TO A25)

:50 A24a. Did you produce phlegm with any of these chest illnesses?

1 [ ] YES

2 [ ] NO (SKIP TO A25)

:51 A24b. In the last three years did any of these illnesses with phlegm last at least one week?

1 [ ] YES

2 [ ] NO (SKIP TO A25)

:52-53 A24c. How many illnesses?

______ NUMBER OF ILLNESSES

Now for some different questions.

:54 A25. Are you able to walk?

1 [ ] YES (SKIP TO A26)

2 [ ] NO

:55 A25a. Are you completely unable to walk?

1 [ ] YES

2 [ ] NO (SKIP TO A26)

:56 A25b. Are you unable to walk because of heart disease or lung disease?

1 [ ] YES → (SKIP TO B1)

2 [ ] NO → (SKIP TO B1)

1:57 A26. Are you troubled by shortness of breath when hurrying on level ground or walking up a slight hill?

1 [ ] YES

2 [ ] NO (SKIP TO B1)

:58 A27. When walking on level ground do you have to walk more slowly than people your own age because of breathlessness?

1 [ ] YES

2 [ ] NO (SKIP TO A31)

:59 A28. When you are walking on level ground at your usual pace do you ever have to stop for breath?

1 [ ] YES

2 [ ] NO (SKIP TO A31)

:60 A29. Do you ever have to stop for breath after walking a few minutes on level ground?

1 [ ] YES

2 [ ] NO (SKIP TO A31)

:61 A30. Are you too breathless to leave the house or breathless on dressing or undressing?

1 [ ] YES

2 [ ] NO

:62-63 A31. For how many years have you had these symptoms of shortness of breath?

_____ NUMBER OF YEARS

:1-6 Duplicate
2
:7

B. SYMPTOMS

Now some more questions about illnesses.

		B1. Have you ever had (READ "A")? NO	YES	B2. How old were you when you first had (READ "A")? YEARS	B3. Did a doctor ever say you had (READ "A")? YES	NO
___ ___ ___ ___ :8-11	a. Attacks of bronchitis	[2] ↓	[1] →	________	[1]	[2]
___ ___ ___ ___ :12-15	b. Pneumonia	[2] ↓	[1] →	________	[1]	[2]
___ ___ ___ ___ :16-19	c. Hay fever	[2] ↓	[1] →	________	[1]	[2]
___ ___ ___ ___ :20-23	d. Sinus trouble	[2] ↓	[1] →	________	[1]	[2]
___ ___ ___ ___ :24-27	e. Pulmonary tuberculosis	[2]	[1] →	________	[1]	[2]

		B1. Have you ever had (READ "F")? NO	YES	B2. How old were you when you first had (READ "F")? YEARS	B3. Did a doctor ever say you had (READ "F")? YES	NO	B4. Do you still hav (READ "F YES	
___ ___ ___ ___ ___ :28-32	f. Emphysema	[2] ↓	[1] →	________	[1]	[2]	[1]	[
___ ___ ___ ___ ___ :33-37	g. Chronic bronchitis	[2]	[1] →	________	[1]	[2]	[1]	[

:38 B5. Have you ever had asthma?

1 [ ] YES

2 [ ] NO (SKIP TO B6)

:39-40 B5a. How old were you when you first had asthma?

________ AGE IN YEARS

2:49 B7a. In the past 10 years, have you taken any medicine or had any treatment for heart trouble?

1 [ ] YES

2 [ ] NO

:50 B8. Are you troubled by frequent headaches?

1 [ ] YES

2 [ ] NO

:51 B9. Did you have any lung trouble before you were 16?

1 [ ] YES

2 [ ] NO

B10. Have you ever had an allergic reaction to (READ "A")?

B11. Was your allergy confirmed by a doctor?

		B10 NO	B10 YES		B11 YES	B11 NO
:52-53	a. Food or medicine or anything else you ate or drank?	[2] ↓	[1]	⟶	[1]	[2]
:54-55	b. Pollen, dust or anything you inhaled?	[2] ↓	[1]	⟶	[1]	[2]
:56-57	c. Detergents, metals, or any-anything you touched?	[2]	[1]	⟶	[1]	[2]

:58 B12. Would you say your health is excellent, very good, good, fair, or poor?

1 [ ] EXCELLENT

2 [ ] VERY GOOD

3 [ ] GOOD

4 [ ] FAIR

5 [ ] POOR

C. OCCUPATIONAL HISTORY

These next questions are about your work.

2:59 C1. Are you presently working at least 30 hours a week for pay?

1 [ ] YES (SKIP TO C3)

2 [ ] NO

:60 C2. Have you ever worked for pay full time or almost full time for at least 6 months in a row?

1 [ ] YES (READ C2a)

2 [ ] NO (SKIP TO C8)

C2a. I'm going to ask some questions about the last job you had where you worked full time for 6 months or more [GO TO C3, ASK ABOUT THAT JOB]

:61-62 C3. What sort of work (do you do on your job/did you do on that job)? (What are/were some of your main activities or duties?)

____

____

C3a. Tell me a little more about what you (do/did) on your job?

____

____

:63-64 C3b. What kind of business or industry (is/was) that in? (What do they make or do?)

____

____

:65 C3c. (Are/Were) you self employed or (do/did) you work for someone else?

1 [ ] SELF EMPLOYED

2 [ ] SOMEONE ELSE

2:41 B5b. Do you still have it?

1 [ ] YES (SKIP TO B5d)

2 [ ] NO

:42-43 B5c. How old were you when it stopped?

______ AGE IN YEARS

:44 B5d. Did a doctor ever say that you had asthma?

1 [ ] YES

2 [ ] NO

:45 B5e. Do you presently take any medication or get any treatment for asthma?

1 [ ] YES

2 [ ] NO

:46 B6. Has a doctor ever told you that you have high blood pressure or hypertension?

1 [ ] YES

2 [ ] NO (SKIP TO B7)

:47 B6a. In the past 10 years have you taken any medicine or had any treatment for it?

1 [ ] YES

2 [ ] NO

:48 B7. Has a doctor ever told you that you had heart trouble?

1 [ ] YES

2 [ ] NO (SKIP TO B8)

2:66 C3d. [DON'T ASK IF CLEAR, BUT MARK ANSWER] Where (do/did) you do this work - at home or away from home?

1 [ ] AT HOME

2 [ ] AWAY FROM HOME

:67-68 C4. How long (have you worked at your present job/did you work at that job)?

_______________ YEARS WORKED

:69 C5. Would you say that the job you just described is the same kind of work you have done for most of the time you have ever been employed in your life?

1 [ ] YES (SKIP TO C6)

2 [ ] NO

(NOTE: IF R HAS TROUBLE HERE, ASK ABOUT THE JOB R WORKED AT THE LONGEST.)

:70-71 C5a. What sort of work have you done most of the time in your life? (What were some of your main activities and duties?)

____

____

C5b. Tell me a little more about what you did on your usual job?

____

____

:72-73 C5c. What business or industry was that in? (What do they make or do?)

____

____

:74-75 C5d. For how many years did you work in a job like that?

______ YEARS WORKED

2:76 C6. Have you ever been exposed to gas or chemical fumes in your work?

1 [ ] YES

2 [ ] NO (SKIP TO C7)

:77-78 C6a. For how many years total did you work in a job where you were exposed to gas or chemical fumes?

_____________ YEARS

:79 C6b. Would you say that you were exposed often, sometimes, or rarely?

1 [ ] OFTEN

2 [ ] SOMETIMES

3 [ ] RARELY

:80 C6c. Would you say that the fumes were mild, moderate, or severe?

1 [ ] MILD

2 [ ] MODERATE

3 [ ] SEVERE

C6d. What kind of fumes were you exposed to?

____

____

:1-6 Duplicate

3
:7

:8 C7. Have you ever worked for at least a year in a dusty job?

1 [ ] YES

2 [ ] NO (SKIP TO C8)

:9-10 C7a. For how many years total did you work in a dusty job?

_____________ YEARS

:11 C7b. Would you say that the dust was mild, moderate, or severe?

1 [ ] MILD

2 [ ] MODERATE

3 [ ] SEVERE

C7c. What kind of dust was it?

____

____

3:12 C8. Have you ever been affected by gas or fumes anywhere when you were not at work?

1 [ ] YES

2 [ ] NO (SKIP TO C9)

:13 C8a. Has this happened more than once?

1 [ ] YES

2 [ ] NO

If YES:

:14 C8b. Did you have to see a doctor any of the times it happened?

1 [ ] YES

2 [ ] NO (SKIP TO C9)

:15 C8c. Was that one of the times or more than one of the times?

1 [ ] ONE

2 [ ] MORE THAN ONE

:16 C8d. Were you (ever) hospitalized for a day or more because of fumes or gas?

1 [ ] YES

2 [ ] NO (SKIP TO C9)

:17-18 C8e. How many times were you hospitalized?

______ NUMBER OF TIMES

If NO:

C8b. Did you have to see a doctor?

1 [ ] YES

2 [ ] NO (SKIP TO C9)

C8d. Were you hospitalized for a day or more because of fumes or gas?

1 [ ] YES

2 [ ] NO

:19-21 C9. How much do you weigh now?

______________ POUNDS

:22-23 C10. How tall are you?

_________ FEET __________ INCHES

D. SMOKING HISTORY

These next questions are about smoking.

3:24 D1. Have you ever smoked cigarettes regularly?

1 [ ] YES (SKIP TO D2)

2 [ ] NO

:25 D1a. Have you ever smoked cigarettes at all?

1 [ ] YES

2 [ ] NO (SKIP TO D9)

:26 D1b. Would you say you have smoked as many as 400 cigarettes in your lifetime?

1 [ ] YES

2 [ ] NO (SKIP TO D9)

:27-28 D2. At what age did you start smoking cigarettes regularly?

______ AGE

:29 D3. Do you smoke cigarettes now?

1 [ ] YES

2 [ ] NO (SKIP TO D5)

:30-31 D4. How many cigarettes do you smoke per day now?

______ PER DAY (SKIP TO D7)

:32-33 D5. At what age did you stop smoking?

______ AGE

3:34-35 D6. About how many cigarettes did you smoke a day during most of the time while you were a smoker?

_____ PER DAY

:36 D6a. Would you say you inhaled the smoke deeply, moderately, mildly, or not at all?

1 [ ] DEEPLY
2 [ ] MODERATELY
3 [ ] MILDLY
4 [ ] NOT AT ALL
→ (SKIP TO D8)

:37-38 D7. On average, how many cigarettes have you smoked a day during most of the time since you started to smoke?

_____ PER DAY

:39 D7a. Would you say you inhaled the smoke deeply, moderately, mildly, or not at all?

1 [ ] DEEPLY
2 [ ] MODERATELY
3 [ ] MILDLY
4 [ ] NOT AT ALL

:40 D8. During all the time you (have) smoked cigarettes, would you say you smoked filter tips all the time, more than half the time, about half the time, less than half the time, or not at all?

1 [ ] ALL THE TIME
2 [ ] MORE THAN HALF THE TIME
3 [ ] HALF THE TIME
4 [ ] LESS THAN HALF THE TIME
5. [ ] NOT AT ALL

:41 D9. INTERVIEWER OBSERVATION: R'S SEX

1 [ ] MALE (SKIP TO D11)

2 [ ] FEMALE

:42 D10. Have you ever smoked other tobacco products such as cigars or pipes?

1 [ ] YES (GO TO D11)

2 [ ] NO (SKIP TO D21)

:43 D11. Have you ever smoked a pipe regularly?

1 [ ] YES

2 [ ] NO (SKIP TO D14)

:44-45 D11a. How old were you when you started smoking a pipe regularly?

_____ AGE IN YEARS

:46 D11b. Do you still smoke a pipe?

1 [ ] YES (SKIP TO D12)

2 [ ] NO

:47-48 D11c. How old were you when you stopped smoking a pipe?

_________ AGE IN YEARS

D12. On the average, over the entire time you (have) smoked a pipe, how much pipe tobacco (did you/have you) smoked per week?

:49-51 _____ OUNCES PER WEEK

OR

:52-53 _____ POUCHES PER WEEK

3:54 D13. (Did you/Do you) inhale the pipe smoke deeply, moderately, mildly, or not at all?

1 [ ] DEEPLY

2 [ ] MODERATELY

3 [ ] MILDLY

4 [ ] NOT AT ALL

:55 D14. Have you ever smoked cigars?

1 [ ] YES

2 [ ] NO (SKIP TO D21)

:56 D14a. Did you ever smoke at least one cigar a week for at least one year?

1 [ ] YES

2 [ ] NO (SKIP TO D21)

:57-58 D15. How old were you when you started smoking cigars regularly?

_____ AGE IN YEARS

:59 D16. Do you smoke them now?

1 [ ] YES (SKIP TO D18)

2 [ ] NO

:60-61 D17. How old were you when you stopped smoking cigars?

_____ AGE [SKIP TO D19]

:62-63 D18. On average, how many cigars do you smoke per week now?

_____ CIGARS PER WEEK

:64-65 D19. On average, over the entire time you (have) smoked cigars, how many cigars (did you/have you) smoked per week?

_____ CIGARS PER WEEK

3:66 D20. (Do you/Did you) inhale the cigar smoke <u>deeply</u>, <u>moderately</u>, <u>mildly</u>, or <u>not at all</u>?

1 [ ] DEEPLY

2 [ ] MODERATELY

3 [ ] MILDLY

4 [ ] NOT AT ALL

:67 D21. Have you ever chewed tobacco or used snuff regularly?

1 [ ] YES

2 [ ] NO (SKIP TO D22)

:68 D21a. Which did you do?

1 [ ] CHEWED TOBACCO

2 [ ] USED SNUFF

3 [ ] BOTH

:69 D22. Have you ever smoked non-tobacco products regularly?

1 [ ] YES

2 [ ] NO

:70 D23. Since you were born, have you ever lived with someone who smoked cigarettes, cigars or pipes regularly?

1 [ ] YES

2 [ ] NO (SKIP TO D24)

3:71-72 D23a. For how many years total did you live in a household with (other) smokers?

_____________ TOTAL YEARS

:73 D23b. Does anyone who lives with you in your household now smoke?

1 [ ] YES

2 [ ] NO (SKIP TO D24)

:74 D23c. Would you say your exposure to cigarette smoke in your house is light, moderate, or heavy?

1 [ ] LIGHT

2 [ ] MODERATE

3 [ ] HEAVY

:75 D24. Are you usually exposed to (other) smokers away from home?

1 [ ] YES

2 [ ] NO (SKIP TO D25)

:76 D24a. Is that almost every day, a couple of times a week, or once a week or less?

1 [ ] ALMOST EVERY DAY

2 [ ] A COUPLE OF TIMES A WEEK

3 [ ] ONCE A WEEK OR LESS

:77 D24b. Would you say the smoke is light, moderate or heavy?

1 [ ] LIGHT

2 [ ] MODERATE

3 [ ] HEAVY

:1-6 Duplicate

4
:7

Now, a few questions about your parents' health.

D25. Were either of your natural parents told that they had a chronic lung condition such as (READ "A")

D26. Was it your father, your mother, or both parents?

		NO	YES	FATHER	MOTHER	BOTH
:8-9	a. Chronic Bronchitis	[2] ↓	[1] →	[1]	[2]	[3]
:10-11	b. Emphysema	[2] ↓	[1] →	[1]	[2]	[3]
:12-13	c. Asthma	[2] ↓	[1] →	[1]	[2]	[3]
:14-15	d. Lung cancer	[2] ↓	[1] →	[1]	[2]	[3]
:16-17	e. Another lung or breathing condition	[2]	[1] →	[1]	[2]	[3]

E. BACKGROUND INFORMATION

Now, we would like some background information on you.

4:18-21 E1. What is the month and year of your birth?

__________ MONTH __________ YEAR

:22 E2. Are you married, widowed, separated, divorced, or have you never been married?

1 [ ] MARRIED

2 [ ] WIDOWED

3 [ ] SEPARATED

4 [ ] DIVORCED

5 [ ] NEVER BEEN MARRIED

:23 E3. Which of these best describes your background - White, Black, Hispanic, American Indian, or Asian?

1 [ ] WHITE

2 [ ] BLACK

3 [ ] HISPANIC

4 [ ] AMERICAN INDIAN

5 [ ] ASIAN

:24 E4. What is the highest grade of school you completed? [IF HIGH SCHOOL OR COLLEGE: Did you graduate?]

1 [ ] 8TH GRADE OR LESS

2 [ ] 1-3 YEARS HIGH SCHOOL

3 [ ] HIGH SCHOOL GRADUATE

4 [ ] 1-3 YEARS COLLEGE

5 [ ] COLLEGE GRADUATE

6 [ ] GRADUATE EDUCATION

4:25 E5. In 1982, was your total family income before taxes, more than $15,000?

[ ] NO → Was it more than $10,000?

- [ ] NO → Was it more than $5,000?
  - [ ] NO
  - [ ] YES
- [ ] YES (SKIP TO E6)

[ ] YES → Was it more than $25,000?

- [ ] NO → More than $20,000?
  - [ ] NO
  - [ ] YES
- [ ] YES → More than $30,000?
  - [ ] NO
  - [ ] YES

:26-28 E6. What city or town do you live in now? (NOTE: IF R SAYS BOSTON: ASK "What part of Boston?")

1 [ ] SOUTH BOSTON

2 [ ] ______________________ (SKIP TO E9)
CITY OR TOWN

E7. What city or town were you living in ....

		CITY OR TOWN
:29-31	a. 5 years ago?	______________________
:32-34	b. 10 years ago?	______________________
:35-37	c. 20 years ago?	______________________

:38-39 E8. For how many years total have you lived in South Boston?

______________ NUMBER OF YEARS (SKIP TO E14)

E9. What city or town were you living in......

4:40-42 a. 5 years ago? [ ] SOUTH BOSTON

______________________ CITY OR TOWN

:43-45 b. 10 years ago? [ ] SOUTH BOSTON

______________________ CITY OR TOWN

:46-48 c. 20 years ago? [ ] SOUTH BOSTON

______________________ CITY OR TOWN

E10. INTERVIEWER CHECK:

1 [ ] SOUTH BOSTON MENTIONED IN E9 (SKIP TO E12)

2 [ ] SOUTH BOSTON NOT MENTIONED IN E9 (GO TO E11)

:49 E11. Have you ever lived in South Boston?

1 [ ] YES

2 [ ] NO (SKIP TO E13)

:50-51 E12. What was the last year you lived in South Boston?

______________ YEAR

:52-53 E12a. For how many years total did you live in South Boston?

______________ NUMBER OF YEARS

:54-55 E13. For how many years total have you lived in (CITY/TOWN R LIVES IN NOW)?

______________ NUMBER OF YEARS

[ ] ALL OF R'S LIFE

4:56-60 E14. What is your zip code?

_____________ ZIP CODE

:61-62 E15. How many times have you moved in the past 10 years?

______ NUMBER OF MOVES

:63 E16. What kind of fuel is used most for heating your home? Is it gas, oil, electricity or what?

1 [ ] GAS

2 [ ] OIL

3 [ ] ELECTRIC

4 [ ] WOOD

5 [ ] COAL OR COKE

6 [ ] OTHER (SPECIFY)___________________________________

____

:64 E17. Is your house heated with hot water or steam, forced hot air, electric units or what?

1 [ ] HOT WATER OR STEAM

2 [ ] FORCED HOT AIR

3 [ ] ELECTRIC UNITS

4 [ ] OTHER (SPECIFY)___________________________________

____

:65 E18. During the winter, do you ever use a kerosene space heater to heat rooms in your house?

1 [ ] YES

2 [ ] NO

4:66 E19. What kind of fuel is used most for cooking in your home--gas, electricity, coal, wood, kerosene, or something else?

1 [ ] GAS

2 [ ] ELECTRIC

3 [ ] COAL OR COKE

4 [ ] WOOD

5 [ ] KEROSENE

6. [ ] OTHER (SPECIFY): ______________________

(2–6 → SKIP TO E21)

:67 E20. (ASK ONLY IF R COOKS WITH GAS) Does your gas stove have what is called a gas log -- that is, a heater built into the stove?

1 [ ] YES

2 [ ] NO (SKIP TO E21)

:68 E20a. Do you ever use it to heat your house?

1 [ ] YES

2 [ ] NO

(1–2 → SKIP TO E22)

:69 E21. Have you ever used your kitchen stove for heating your house?

1 [ ] YES

2 [ ] NO

:70 E22. Is there any air conditioning in your house?

1 [ ] YES

2 [ ] NO (SKIP TO E23)

:71 E22a. Is it all air conditioned or just part of it?

1 [ ] ALL (SKIP TO E23)

2 [ ] PART

4:72 E22b. Is the living room air conditioned?

1 [ ] YES

2 [ ] NO

:73 E22c. Is your bedroom air conditioned?

1 [ ] YES

2 [ ] NO

These last questions are about the air and water in your neighborhood.

:74 E23. Have you ever felt ill or had a health problem you believe was caused by the drinking water in your neighborhood?

1 [ ] YES

2 [ ] NO

:75 E24. Have you ever felt ill or had a health problem that you believe was caused by the air in your neighborhood?

1 [ ] YES

2 [ ] NO

Thank you very much. You have been very helpful. We may be calling you again soon to find out a little more about your health.

TIME NOW: ______________

INTERVIEWER OBSERVATIONS

4:76 F1. HOW COOPERATIVE WAS R?

1 [ ] COOPERATIVE

2 [ ] NEUTRAL

3 [ ] ANTAGONISTIC

:77 F2. OVERALL, HOW WAS R'S INTEREST IN THE INTERVIEW?

1 [ ] VERY INTERESTED

2 [ ] INTERESTED

3 [ ] INDIFFERENT

4 [ ] NEGATIVE, DEFENSIVE, UPSET

:78 F3. LANGUAGE INTERVIEW WAS TAKEN IN?

1 [ ] ENGLISH

2 [ ] PORTUGUESE

3 [ ] SPANISH

4 [ ] OTHER (SPECIFY): ______________________________

APPENDIX F: DEFINITION OF MAJOR DEPENDENT AND CONTROL VARIABLES USED IN ANALYSIS

APPENDIX F

DEPENDENT VARIABLES

The major dependent variables are defined as follows: pulmonary disroders (chronic cough, asthma, etc.) take the value of zero unless either (A) the condition specified below is met, in which instance the value of the variable would be set to 1, or (B) the relevant question was not answered, in which case the variable is coded as missing, and the case is excluded from analyses involving that symptom. The survey questions defining the symptom conditions are:

1. CHRONIC COUGH. "YES" (1) to question A8: "Do you usually cough on most days for at least three months in a row during the year?"

2. CHRONIC PHLEGM. "YES" (1) to question A15: "Have you brought up phlegm on most days for as much as three months in a row?"

3. CHRONIC WHEEZE. "MOST DAYS OR NIGHTS" (2) to question A20a: "Would you say this [chest sound wheezy or whistling] happens only occasionally or most days or nights?"

4. BREATHLESSNESS
   Level 0: "NO" to question A26: "Are you troubled by shortness of breath when hurrying on level ground or walking up a slight hill?"

   Level 1: "NO" to question A26: "Are you troubled by shortness of breath when hurrying on level ground or walking up a slight hill?"

   Level 2: "YES" (1) to question A27: "When walking on level ground do you have to walk more slowly than people your own age because of breathlessness?" [NOTE: For analyses using 2-category Breathlessness variable, levels 0 and 1 are combined].

5. RECENT CHEST ILLNESS. "YES" (1) to question A24: "During the past 3 years, have you had any colds or chest illnesses that have kept you indoors at home, or in bed?"

6. ASTHMA. "YES" (1) to question B5a: "Did a doctor ever say that you had asthma?"

7. RECENT SEVERE CHEST ILLNESS. "YES" (1) to question A24: "During the past 3 years, have you had any colds or chest illnesses that have kept you indoors at home, or in bed?" and "YES" (1) to question A24b: "In the last three years did any of these illnesses with phlegm last at least one week?"

8. PULMONARY TUBERCULOSIS. "YES" (1) to question B1.e: "Have you ever had Pulmonary Tuberculosis?" and YES (1) to question B3.e "Did a doctor ever say you had Pulmonary Tuberculosis?"

MAJOR CONTROL VARIABLES

The major control variables used in the analyses were defined as follows:

1. AGE. 4 categories from responses to question E1: "What is the month and year of your birth?" (1) 18-34 (2) 35-49 (3) 50-64 (4) 65+.

2. SEX. 2 categories: (1) MALE (2) FEMALE [Interviewer observation]

3. CIGARETTE SMOKING. 5 categories:
   1. NEVER SMOKED. "NO" (2) to question D1a: "Have you ever smoked cigarettes at all?" or "NO" (2) to question D1b: "Would you say that you have smoked as many as 400 cigarettes in your lifetime?" plus "NO" (2) to question D10. "Have you ever smoked other tobacco products such as cigars or pipes?" or "NO" (2) to both question D11: "Have you ever smoked a pipe regularly?" and question D14: "Have you ever smoke cigars?" or question D14a: "Did you ever smoke at least one cigar a week for at least one year?"

2. EX-SMOKER. "NO" (2) to question D3: "Do you smoke cigarettes now?" and response to question D5: "At what age did you stop smoking? at least 1 year less than R°s present age."

   3. 1 to 15 CIGARETTES PER DAY. Response to question D4: "How many cigarettes to you smoke per day now?"
   4. 16-30 CIGARETTES PER DAY. Response to question D4: "How many cigarettes do you smoke per day now?"
   5. MORE THAN 30 CIGARETTES PER DAY. Response to question D4. "How many cigarettes do you smoke per day now?"

APPENDIX G: ESTIMATES OF 5-WAY INTERACTION EFFECTS IN CHRONIC COUGH ANALYSIS

Table G-1

5-Way Interaction Effects
For Chronic Cough by Age by Smoking
by Sex by Area

ESTIMATED LAMBDA EFFECTS, (and Standard Errors)

Age	Never Smoked	Ex- Smoker	0-15 per day	16-30 per day	31+ per day
18-34	0.315	-0.290	0.011	0.089	-0.125
	(0.240)	(0.643)	(0.546)	(0.234)	(0.423)
	1.311	-0.451	0.020	0.383	-0.296
35-49	0.124	-0.353	0.442	-0.065	-0.148
	(0.393)	(1.201)	(0.930)	(0.388)	(0.519)
50-64	-0.231	1.129	-1.094	0.059	0.137
	(0.397)	(1.201)	(0.930)	(0.388)	(0.519)
65+	-0.207	-0.485	0.641	-0.084	0.136
	(0.505)	(1.239)	(1.194)	(0.503)	(1.148)

ESTIMATED TAU PARAMETERS

18-34	1.370	0.748	1.011	1.094	0.882
35-49	1.132	0.702	1.555	0.937	0.863
50-64	0.794	3.091	0.335	1.061	1.147
65+	0.813	0.616	1.898	0.919	1.145

Note: Lambda >0 and Tau<1, for "excess" of healthy, males in SMSA for indicated category of age-and-smoking. (Estimates are from saturated model of 5-way table, hence "excess" means net of all 4-way and lower-order effects.) (This table, unlike others in this section, uses Goodman and Fay's convention of representing "contrasts" as difference between specified category and"average" of all categories.)

APPENDIX H. PERCENTAGE DISTRIBUTIONS FOR SELECTED ADDITIONAL VARIABLES

TABLE H - 1

Percentage Distributions for Additional Demographic Variables, Environmental Risk Factors, and Other Health and Symptom Reports in South Boston Boston SMSA Samples

Variable	SMSA	South Boston	X^2	df
MAJOR CONTROL VARIABLES				
Smoking				
Never smoked	38.0%	31.8%	56.8	5
Ex. Smoker	22.1	20.0		
Cigar or Pipe Smoker	7.7	4.6		
Cigarette Smoker: 0-15/day	10.7	11.4		
Cigarette Smoker: 16-30/day	15.7	21.4		
Cigarette Smoker: 31+	5.8	10.8		
Age				
18-34	41.7	44.6		
35-49	26.6	20.8	15.1	3
50-64	18.4	21.3		
65+	13.3	13.4		
Sex:				
Male	46.8	43.3	3.8	1
Female	53.2	56.7	3.8	1
DETAILED HEALTH & SYMPTOM REPORTS				
Self-assessed Health (A)				
Excellent	33.7	30.2	21.7	4
Very good	34.4	28.8		
Good	23.1	24.9		
Fair	6.5	11.5		
Poor	2.2	4.6		
Self-Assessed Health Compared to Others				
Better	41.0	35.3	11.3	2
About the Same	53.4	57.5		
Worse	5.6	7.2		
Usually Cough: Yes	12.2	18.8	25.2	1
Cough in Morning: Yes	12.4	17.7	16.8	1
Cough Rest of Day: Yes	10.0	16.0	23.5	1
Usually Phlegm: Yes	14.4	25.1	54.1	1
Phlegm in Morning: Yes	14.4	20.4	18.4	1
Phlegm Rest of Day	8.0	15.3	39.1	1

NOTE:
Chi-square statistics do not account for non-SRS sample design; hence, common inferences, e.g., p = .05 at 3.84 with 1 d.f., 5.99 with 2 d.f, 7.81 with 3 d.f., 9.49 with 4 d.f., will not be precisely accurate. Estimates of sample design and interviewer effects upon the variance of these measurements are presented in Section VI.

Variable	Boston SMSA	South Boston	X^2	df
Cough & Phlegm for 3 weeks: Yes	15.9%	16.0%	0.0	1
Attacks of Wheezing: Yes	18.6	21.0	2.7	1
Chest Wheeze during Colds: Yes	38.7	44.0	8.9	1
Chest Wheeze at other Times: Yes	10.6	12.3	2.1	1
Colds Usually go to Chest: Yes	35.3	40.1	7.3	1
Colds/Chest Ill. Kept you at Home: Yes	34.4	35.2	0.2	1
Shortness of Breath (Highest Level)				
0: None	81.7	71.9	53.8	5
I: Br. when hurrying on hill	13.5	18.8		
II: Walk slowly due to Br.	2.8	4.3		
III: Br. walking at own pace	0.6	2.1		
IV: Br. after few minutes on level	1.2	1.5		
V: Too Br. to go out or dressing	0.2	1.3		
Ever had Bronchitis: Yes	21.7	17.8	7.0	1
Ever had Pneumonia: Yes	22.8	22.0	0.2	1
Ever had Hay Fever: Yes	23.6	18.2	12.8	1
Ever had Sinus Trouble: Yes	31.9	25.7	14.0	1
Ever had Pulmonary TB: Yes	0.7	1.9	7.8	1
Ever had Emphysema: Yes	1.1	1.8	2.8	1
Ever had Chronic Bronchitis: Yes	3.6	4.7	2.6	1
Ever had Asthma: Yes	6.8	6.2	0.4	1
Doctor Says have Hypertension: Yes	22.8	24.8	1.7	1
Doctor Says have Heart Trouble: Yes	8.1	11.4	9.1	1
Had Lung Trouble before age 16: Yes	4.3	5.0	0.9	1
Allergic to Food/Medicine: Yes	30.7	25.6	9.9	1
Allergic to Pollen/Dust: Yes	28.3	22.6	12.4	1
Allergic to Metals/Detergent: Yes	13.4	13.1	0.1	1
Ever Ill from Drinking Water: Yes	2.3	3.1	1.6	1
Ever Ill from Air: Yes	5.4	14.8	74.1	1
Have Frequent Headaches: Yes	12.0	14.7	4.8	1

(cont'd)

Variable	Boston SMSA	South Boston	X^2	df
DEMOGRAPHIC VARIABLES				
Work 30 Hours a Week: Yes	62.9%	56.6%	12.3	1
Self-employed: Yes	10.7	5.1	29.5	1
Marital Status				
Married	56.4	43.1	60.0	4
Widowed	7.9	11.5		
Separated	2.0	3.8		
Divorced	6.6	9.8		
Never Married	27.0	31.8		
Race/Ethnicity				
White	93.1	98.5	72.2	4
Black	4.8	0.1		
Hispanic	0.7	0.2		
Amer. Indian	0.4	0.5		
Asian	1.0	0.3		
Education				
0-11 Years	10.4	23.9	218.6	2
12 Years	30.8	43.2		
13+ Years	58.7	32.9		
Family Income				
0 - 4,999	4.7	9.5	184.6	6
5,000 - 9,999	10.3	17.7		
10,000 - 14,999	13.7	20.1		
15,000 - 19,999	11.2	14.4		
20,000 - 24,999	12.5	14.0		
25,000 - 29,999	10.6	8.8		
30,000 or more	36.9	15.6		
Ever Lived in South Boston: Yes	2.4	na		
ENVIRONMENTAL AND RISK FACTORS				
Exposed to Gas/Chemical Fumes	24.6	24.2	0.1	1
Ever Work on Dusty Job for 1 Yr: Yes	17.9	26.6	30.7	1
Exposed to Gas/Fumes outside Work: Yes	10.1	9.5	0.3	1
Live with Smoker Now: Yes	42.0	48.4	10.3	1
Parents had Chronic Bronchitis: Yes	8.0	8.4	0.1	1
Parents had Emphesyma: Yes	8.1	8.3	0.0	1
Parents had Asthma: Yes	6.5	7.3	0.7	1
Parents had Lung Cancer: Yes	4.9	6.3	2.7	1
Parents had Other Lung Disorder: Yes	6.0	5.7	0.1	1

(cont'd)

Variable	Boston SMSA	South Boston	X^2	df
Fuel Used for Heating				
Gas	37.4	59.0	157.5	5
Oil	51.4	36.6		
Electric	8.2	4.1		
Wood	1.8	0.2		
Coal or Coke	0.6	0.0		
Other	0.6	0.1		
Type of Heating System				
Hot Water or Steam	61.3	60.4	111.5	3
Forced Hot Air	27.7	18.9		
Electric Units	6.6	5.6		
Other	4.4	15.1		
Ever Use Kerosene Heater: Yes	2.0	1.5	1.1	1
Cooking Fuel				
Gas	49.9	87.5	485.7	4
Electric	49.8	12.4		
Coal or Coke	0.0	0.0		
Wood	0.2	0.0		
Other	0.1	0.1		
Have a Gas Log: Yes	14.0	39.8	147.3	1
(Use for Heat as % of those having one)	(45.9)	(77.5)	(44.4)	(1)
Ever Use Kitchen Stove for Heat: Yes	10.5	20.2	44.0	1
Have Air Conditioning: Yes	51.0	44.9	10.9	1
INTERVIEWER ASSESSMENTS				
How Cooperative was Respondent				
Cooperative	85.4	83.6	(NOT APPLICABLE)	
Neutral	11.9	12.7		
Antagonistic	2.6	3.7		
Respondent's Interest in Interview				
Very Interested	17.0	21.4	(NOT APPLICABLE)	
Interested	60.9	54.3		
Indifferent	19.8	21.8		
Negative/Defensive	2.4	2.5		
Sample Sizes	1547	1444		

APPENDIX I: REPORTING BIASES AND INTERGROUP COMPARISONS OF SELF-REPORTED PULMONARY DISORDER

APPENDIX I

REPORTING BIASES AND INTERGROUP COMPARISONS OF SELF-REPORTED PULMONARY DISORDER

All self-report measurements are vulnerable to bias. The hallmark of scientific measurement, in survey research as elsewhere, is the incorporation of appropriate validation procedures as part of the measurement program, the aim being to estimate the magnitude and direction of biases which affect one's measurements.[1]

If the intent of a measurement program is to contrast groups, it is of critical importance to determine if any biases that may exist are equivalent across the groups which are to be contrasted. Without such a determination, the self-report data collected in a survey can be dangerously misleading since intergroup differences found in the self-report data may reflect either: (1) true differences in the state of the population, or (2) differences in the magnitude or direction of the systematic errors affecting measurements in each group.

----

1. The Center for Survey Research proposed as part of its survey proposal that such validation be commissioned. (See proposal addendum No. 2 entitled, "Differential Validity Analysis.") At present, the requisite validation data for subsamples of survey respondents in the SMSA and South Boston are not available.

It is for this reason that scientific users of subjective survey data have come to recognize the need to place subjective measurements in context with appropriate auxilliary data which provide the necessary framework for establishing the (differential) validity of one's measurements. A study group of the Committee on National Statistics (of the National Academy of Sciences)[2] has recently observed that:

> [Our] review indicates that the survey enterprise is quite substantial in size, and it appears that the data it produces can sometimes have notable (and not well understood) effects on the public. The collection and reporting of survey data . . . are not, however, of uniformly high quality, and there exist numerous causes for concern both about common practices and also about the foundations of basic knowledge upon which contemporary practice rests. (p. 41)

This study group strongly recommended that all ongoing programs of survey data collection incorporate appropriate asssements of the random and systematic components of the errors which affect the survey measurement process (i.e., the variance components which reflect the effects of factors that, in the language of psychometrics, generate measurement unreliability and invalidity).

In the present instance it might wrongly be assumed that the use of a "previously validated" measure of pulmonary health relieves one of the burden of worrying about bias (i.e., systematic error). However, as the National Academy of Sciences' panel pointed out, this would not be appropriate:

----

2. C. F. Turner and E. Martin (Eds.), Surveys of Subjective Phenomena: Summary Report. Washington D.C.: National Academy of Sciences Press, 1981.

> [While] we do not question the need for preliminary and auxilliary studies to develop new questions or procedures . . . we believe that such studies will be most useful when they are part of the process of designing serious substantive surveys. Even more important, they cannot be relied upon by themselves to provide the needed estimates of sampling and nonsampling variation. Those estimates must come from the main survey itself.

In the present instance a concurrent study of the differential validity (and sensitivity) of the self-report data is not only proper scientific practice, but it also provides a meaningful basis for responding to claims that differences between South Boston vs. the control area might reflect only differences in "health consciousness," [or "environmental concerns," or sophistication, etc.] and not true differences in the pulmonary health of residents of South Boston (vs. residents of the control area).

Thus, someone with a stake in a finding of "no difference" between symptom rates in South Boston and the control area may argue that, for example, higher self-reports of Shortness of Breaths in South Boston, "only indicate a greater concern of the South Boston population with this issue, and thus a better memory for events impinging on their health and a greater readiness to report such symptoms to the survey interviewers who telephone them." Alternatively, someone with a different stake in the outcome of the research might note that the survey instrument asked about things like "asthma", and for physicians' confirmation of this disease. It might alternatively be suggested that the relatively poorer population of South Boston underreports such symptoms because of its more limited access to the medical care required to identify and confirm the existence of such pulmonary disorders.

A Scientific Prophylaxis. Systematic error inevitably haunts all measurement programs. The scientific prophylaxis is external validation. Such external validation not only allows the identification of systematic error components in one's measurements, but in a study such as the present one, it would permit recalibration of one's data to take account of any non-equivalence in the systematic errors affecting data from each subpopulation.

Briefly stated, what is required is a small-scale program to obtain objective data on the pulmonary health of subsamples of the subjects who responded to the survey. Given such objective measurements, the predictive validity of the self-report measures could be assessed within each group [South Boston and the control area(s)]. Estimates of the relationship between self-report measurements and physical measurements within each area would allow us to determine whether equivalent levels of self-reported pulmonary dysfunction were generated by equivalent physical symptoms in each area. If residents of one area were differentially sensitive to pulmonary health matters (or for any other reason were prone to higher levels of self-reported dysfunction), this would readily appear when, for example, the physical data were plotted against the self-report data for each subpopulation. Some hypothetical examples of differential reporting of health symptoms are shown in Figures I-1a and I-1b.

Recalibration of measurements to correct for such problems can be carried out if, and only if, objective measures of pulmonary health status are gathered concurrently with the survey. Given such data, differences of the sort shown in Figure 2, could be taken into account in the survey analysis, and meaningful inferences could thus be drawn about differences in the objective health status of the populations of each area.

In the absence of a concurrent subsample validation of the self-report data, conjectures can be offered (and defended) to interpret the subjective self-report data to suit anyone's purpose.[3] This is possible because systematic error is well-known to affect all self-report measures, and because the

---

3. If for example, self-report data for South Boston were found to show the same prevalence of reported symptoms as in the control area, it could be argued that South Boston residents were healthier because they were "predisposed by publicity, general urban awareness of pollution-related health problems (or whatever other mechanisms the fertile immagination of critics might generate) to over-report symptoms."

   Alternatively, the same results might be interpreted as "proving" that South Boston residents were less healthy (than residents of the control area) since "their lower levels of income, and education, caused them to under-report medical problems of all sorts, and thus "equivalent" survey results would indicate that a "true" difference exists."

scientific burden of calibrating those errors would not have been met.[4]

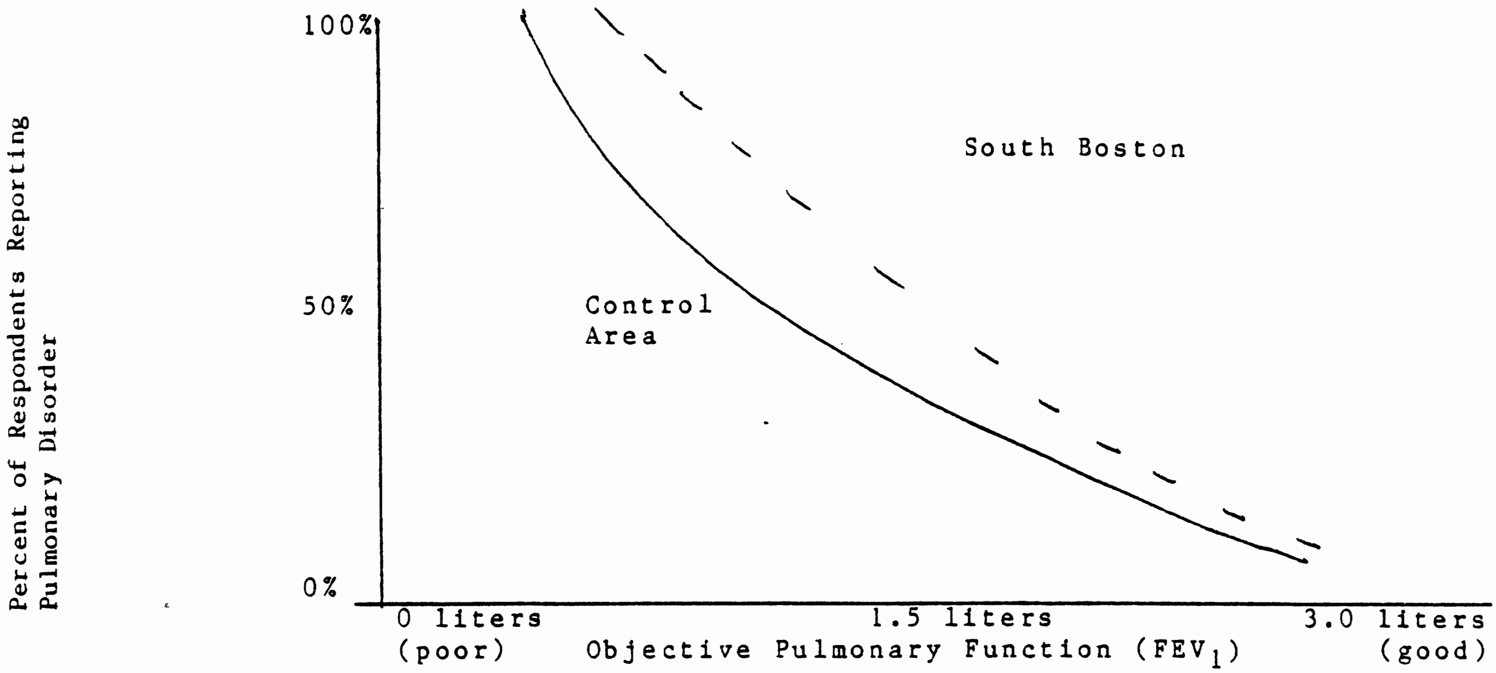
At present, since we lack the necessary validational data for survey respondents in both South Boston and the Boston SMSA, we have no scientific basis for inferring either the magnitude or the direction of the any biases which may affect symptom reporting in South Boston and the control area.

---

4. This is analogous to the situation which would occur if we were trying to assess differences in the weight of two children, each of whom mounted our scales with suitcase in hand. The weight of the suitcases would add an unknown component of bias to the two measurements. The conclusions to be drawn from our scale readings would lead contentious (and scientifically-minded) persons to wonder if the suitcases weighed the same, or if, perhaps, one contained feathers and the other lead. The prudent scientist, of course, would choose to either part the child from his suitcase before weighing him, or failing that, to weigh the suitcase separately, so that a proper inference might be drawn.

   In the present case, we cannot part our respondents from their biases; they will answer our questions with whatever biases they bring. Our obligation is to carefully weigh them and to take them into account when we use their self-report data to infer differences in the health status of the two populations.

I-1A Hypothesis: South Boston Residents are "sensitized" to reporting pulmonary disorders.



I-1B: Hypothesis: South Boston Residents more prone to "overlook" pulmonary (and other medical) disorders.

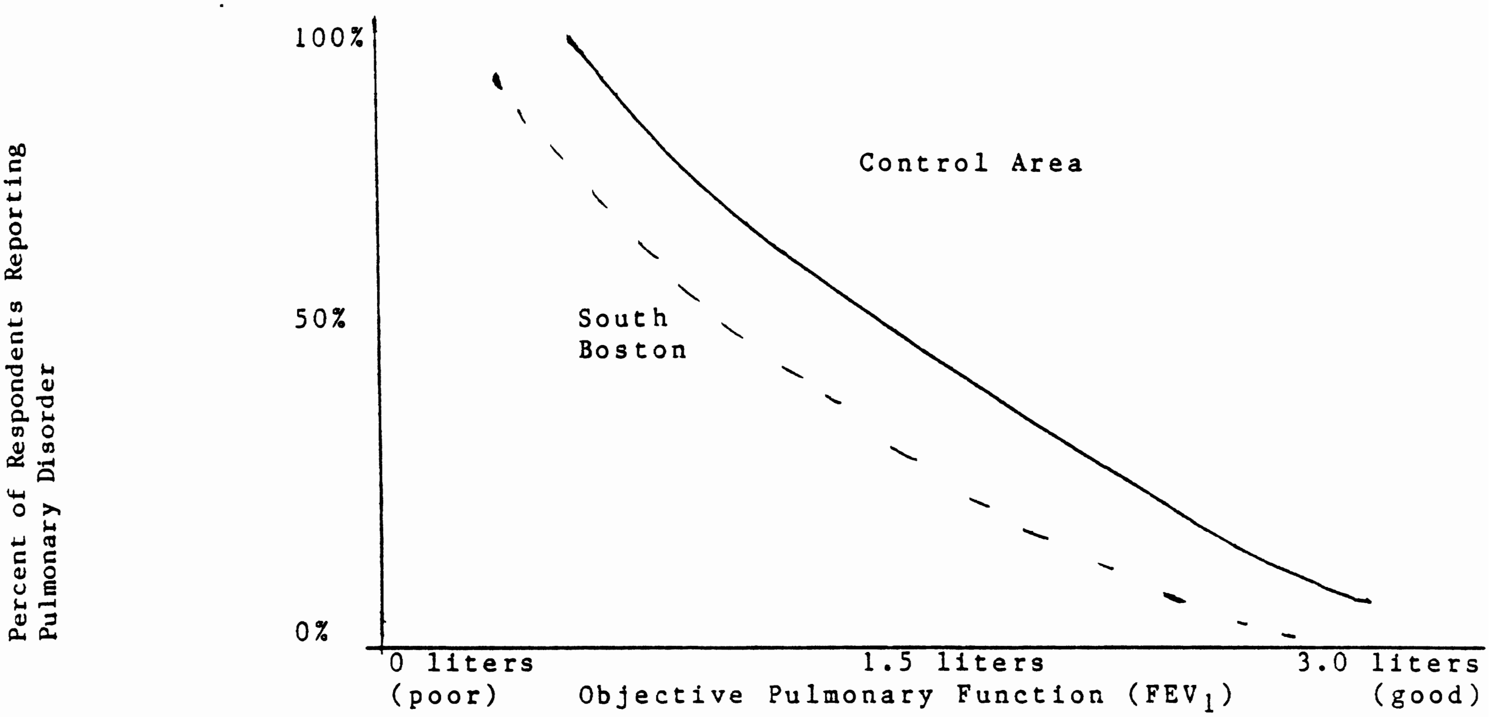


FIGURE I-1: Relationship between Self-Reported Pulmonary Disorder and an Objective Measure of Pulmonary Functioning, using different assumptions about nature of reporting biases. (Hypothetical data; population results would be standardized to some height/weight norm).