Y 4. Sci 2 94-2 SS

GOVERNMENT Storage

1025-A

MENTS

DEC 7 1976

LIBRARY UNIVERSITY

[COMMITTEE PRINT]

THE ENVIRONMENTAL PROTECTION AGENCY'S RESEARCH PROGRAM

WITH PRIMARY EMPHASIS ON THE

COMMUNITY HEALTH AND ENVIRONMENTAL SURVEILLANCE SYSTEM (CHESS): AN INVESTIGATIVE REPORT

REPORT

PREPARED FOR THE

SUBCOMMITTEE ON SPECIAL STUDIES, INVESTIGATIONS AND OVERSIGHT

AND THE

SUBCOMMITTEE ON THE ENVIRONMENT AND THE ATMOSPHERE

OF THE

COMMITTEE ON SCIENCE AND TECHNOLOGY
U.S. HOUSE OF REPRESENTATIVES
NINETY-FOURTH CONGRESS

SECOND SESSION

Serial SS

NOVEMBER 1976

Printed for the use of the Committee on Science and Technology

---

U.S. GOVERNMENT PRINTING OFFICE

77-590 WASHINGTON : 1976

For sale by the Superintendent of Documents, U.S. Government Printing Office
Washington, D.C. 20402 - Price $1.35

AUTHENTICATED U.S. GOVERNMENT INFORMATION GPO

COMMITTEE ON SCIENCE AND TECHNOLOGY

OLIN E. TEAGUE, Texas, *Chairman*

KEN HECHLER, West Virginia
THOMAS N. DOWNING, Virginia
DON FUQUA, Florida
JAMES W. SYMINGTON, Missouri
WALTER FLOWERS, Alabama
ROBERT A. ROE, New Jersey
MIKE McCORMACK, Washington
GEORGE E. BROWN, JR., California
DALE MILFORD, Texas
RAY THORNTON, Arkansas
JAMES H. SCHEUER, New York
RICHARD L. OTTINGER, New York
HENRY A. WAXMAN, California
PHILIP H. HAYES, Indiana
TOM HARKIN, Iowa
JIM LLOYD, California
JEROME A. AMBRO, New York
CHRISTOPHER J. DODD, Connecticut
MICHAEL T. BLOUIN, Iowa
TIM L. HALL, Illinois
ROBERT (BOB) KRUEGER, Texas
MARILYN LLOYD, Tennessee
JAMES J. BLANCHARD, Michigan
TIMOTHY E. WIRTH, Colorado

CHARLES A. MOSHER, Ohio
ALPHONZO BELL, California
JOHN JARMAN, Oklahoma
JOHN W. WYDLER, New York
LARRY WINN, JR., Kansas
LOUIS FREY, JR., Florida
BARRY M. GOLDWATER, JR., California
MARVIN L. ESCH, Michigan
JOHN B. CONLAN, Arizona
GARY A. MYERS, Pennsylvania
DAVID F. EMERY, Maine
LARRY PRESSLER, South Dakota

SUBCOMMITTEE ON SPECIAL STUDIES, INVESTIGATIONS AND OVERSIGHT

OLIN E. TEAGUE, Texas, *Chairman*

KEN HECHLER, West Virginia
DON FUQUA, Florida
JAMES W. SYMINGTON, Missouri
MIKE McCORMACK, Washington

CHARLES A. MOSHER, Ohio
JOHN W. WYDLER, New York

SUBCOMMITTEE ON THE ENVIRONMENT AND THE ATMOSPHERE

GEORGE E. BROWN, JR., California, *Chairman*

MIKE McCORMACK, Washington
DALE MILFORD, Texas
RICHARD L. OTTINGER, New York
PHILIP H. HAYES, Indiana
JEROME A. AMBRO, New York
JAMES J. BLANCHARD, Michigan
JAMES H. SCHEUER, New York

MARVIN L. ESCH, Michigan
LARRY WINN, JR., Kansas
GARY A. MYERS, Pennsylvania
DAVID F. EMERY, Maine

CONTENTS

	Page
Letter of transmittal	VII
I. Executive summary	1
II. Introduction	5
A. Background	5
B. History of air pollution health studies leading to CHESS objectives	7
III. Findings, conclusions, and recommendations	11
A. General findings	11
B. Specific CHESS findings and recommendations	12
C. R. & D. related findings	17
D. Recapitulation of guiding questions	21
E. Closing remarks	23
IV. CHESS aerometric measurements	25
A. Introduction	25
B. Review of chemical and physical methods	27
1. The West-Gaeke method for measurement of ambient SO_2	27
2. Total suspended particulates	28
3. Suspended sulfate	28
4. Dustfall bucket, tape sampler, cascade impactor, and cyclone sampler	29
C. Findings and evaluations of measurements and data reduction	30
1. Sulfur dioxide	31
a. Spillage of reagent during shipment	32
b. Time decay of the reagent SO_2 complex	33
c. Concentration dependence of sampling method	34
d. Low flow correction	34
e. Bubbler train leakage	34
2. General assessment of CHESS SO_2 data	35
3. Total suspended particulate	37
4. Total suspended sulfate	39
D. The CHAMP air monitoring program	41
1. Introduction	41
2. System description	42
3. Findings regarding the CHAMP program	43
4. Summary	44
V. Review of CHESS air quality analysis procedures and results	45
A. Introduction	45
B. Use of estimated data	45
1. Estimated sulfur dioxide concentration, 92 $\mu g/m^3$ (Magna)	46
2. Estimated sulfur dioxide concentration, 177 $\mu g/m^3$ (Anaconda)	47
3. Estimated suspended sulfate concentration, 15 $\mu g/m^3$ (Magna)	49
4. Estimated suspended sulfate concentration, 7.2 $\mu g/m^3$ (Anaconda)	50
5. Estimates of suspended particulate, Salt Lake basin study	54
6. Estimates in the Chicago and New York studies	54
C. Use of mathematical dispersion models	54

Page

VI. An analysis of the CHESS health effects studies---------- 57
A. General problems of epidemiologic investigations of pollution effects---------- 57
B. Introduction and definitions---------- 58
C. Specific health effects studies---------- 60
1. Chronic respiratory disease (CRD) prevalence---------- 60
2. Retrospective surveys for acute lower respiratory disease (LRD) in children---------- 63
3. Acute respiratory disease (ARD) in volunteer families 65
4. Asthma panel studies---------- 68
5. Cardiopulmonary symptoms in adult panels---------- 71
6. Ventilatory function in schoolchildren---------- 73
7. Other CHESS studies---------- 75
D. Summary assessment of the population studies---------- 75
E. CHESS current status---------- 77
1. General---------- 77
2. Status of analysis---------- 78
3. Recent planning directions of Population Studies Division---------- 80
VII. Current CHESS status and future programs---------- 81
A. Retrospection---------- 81
B. Current---------- 82
C. Future---------- 83
APPENDIX A—Recapitulation of the aerometric and meteorological findings of the investigation as they relate to specific sections of the CHESS Monograph---------- 85
A. Introduction---------- 85
B. Critique---------- 85
1. Prevalence of chronic respiratory disease symptoms in adults: 1970 survey of Salt Lake Basin communities---------- 85
2. Frequency of acute lower respiratory disease in children: retrospective survey of Salt Lake Basin communities, 1967–70---------- 87
3. Aggravation of asthma by air pollutants: 1971 Salt Lake Basin studies---------- 87
4. Human exposure to air pollutants in five Rocky Mountain communities, 1940–70---------- 89
5. Prevalence of chronic respiratory disease symptoms in adults: 1970 survey of five Rocky Mountain communities---------- 90
6. Frequency of acute lower respiratory disease in children: retrospective survey of five Rocky Mountain communities, 1967–70---------- 90
7. Prevalence of chronic respiratory disease symptoms in military recruits: Chicago induction center---------- 92
8. Prospective surveys of acute respiratory disease in volunteer families: Chicago nursery school study, 1969–70---------- 93
9. Human exposure to air pollution in selected New York metropolitan communities, 1944–71---------- 93
10. Prevalence of chronic respiratory disease symptoms in adults: 1970 survey of New York communities---------- 94
11. Prospective surveys of acute respiratory disease in volunteer families: 1970–71 New York studies---------- 94
12. Aggravation of asthma by air pollutants: 1970–71 New York studies---------- 95
13. Frequency and severity of cardiopulmonary symptoms in adult panels: 1970–71 New York studies---------- 96
14. Ventilatory function in schoolchildren: 1970–71 New York studies---------- 97
15. Ventilatory function in schoolchildren: 1967–68 testing in Cincinnati neighborhoods---------- 98
C. Superficial and perfunctory treatment of meteorological information---------- 99
D. Insufficient exploration of possible relationships between meteorological conditions and asthma attack rates---------- 99
E. Failure to consider peak and episode concentrations---------- 100

Page

F. Use of a single monitoring station to determine the exposure of a community ---------- 100
G. Failure to establish similarity of exposure and stress factors between communities in the same study, excluding the exposure to specific pollutants ---------- 101
H. Impreciseness of monitoring station locations ---------- 101
I. Inexact locations of residences of individuals studied ---------- 101
APPENDIX B—Legislative history of the CHESS program ---------- 103

LETTER OF TRANSMITTAL

HOUSE OF REPRESENTATIVES,
COMMITTEE ON SCIENCE AND TECHNOLOGY,
Washington, D.C., November 19, 1976.

Hon. OLIN E. TEAGUE,
Chairman, Committee on Science and Technology, U.S. House of Representatives, Washington, D.C.

DEAR MR. CHAIRMAN: I am transmitting herewith a report entitled "The Environmental Protection Agency's Research Program with primary emphasis on the Community Health and Environmental Surveillance System (CHESS): An Investigative Report." This report provides the conclusions and recommendations resulting from an intensive evaluation of the scientific and technical adequacy of this important air pollution-health effects surveillance system and related research programs. The report indicates that there are a number of serious issues in EPA's research programs which require resolution in order to achieve more effective results.

The report was prepared by Committee staff members, Dr. R. B. Dillaway, Science Consultant and Director, EPA CHESS Investigation, and Oversight Subcommittee on Special Studies, Investigations Task Team Leader, and Dr. Radford Byerly, Jr., Science Consultant, Subcommittee on the Environment and the Atmosphere, who was detailed to assist in the investigation. Committee staff members were assisted in the preparation of the report and in the conduct of the on-site investigation by a special team of consultants. The Committee is grateful for the cooperation of the agencies which permitted their participation and expresses appreciation to the consultants for their assistance. The special consultants were: Mitchell H. Gail, M.D., Medical Statistical Researcher, Biometry Branch, National Cancer Institute; Paul A. Humphrey, Senior Scientist (Retired), Division of Meteorology, EPA/NOAA, National Oceanic and Atmospheric Administration; Dr. James M. McCullough, Senior Specialist in Science and Technology (Life Sciences), Science Policy Research Division, Congressional Research Service, Library of Congress; Cdr. John W. Poundstone, M.D., Chief, Occupational and Preventive Medicine, Navy Regional Medical Center, Great Lakes; Dr. Harry L. Rook, Chief, Activation Analysis Section, Analytical Chemistry Division, National Bureau of Standards; and Charles C. VanValin, Research Chemist, Atmospheric Physics and Chemistry Laboratory, Environmental Research Laboratories, National Oceanic and Atmospheric Administration. In addition, the Committee staff solicited the opinions and advice of a number of other scientists, statisticians, and physicians in other agencies and institutes.

In accordance with your direction, the staff in this investigation reported to me. I have nothing but praise for the professional manner with which they carried out the study. I have carefully reviewed this report and am prepared to take full responsibility for its contents. I further recommend its adoption as a Science and Technology Committee report.

In closing, however, I would like to make an additional observation that goes beyond this particular study. The need to regulate pollutants in this industralized age is likely to grow, not diminish. Regulating knowledgeably is therefore essential. This report provides several examples of the many instances in which the status of environmental research and technology is not what it should be to assure adequate and well founded protection of human health. While this investigative report has much significance with regard to research management and execution, it does not provide all the answers to the policymaker who is obliged to act in the face of less than certain proof. The policy oriented questions of "acceptable risk" and "burden of proof" can be assisted, but never fully answered by better research. I would hope that those who read this report will not use it to undermine the minimal research work that is now underway, but instead will use it as a reason to improve and enhance the on-going programs.

Sincerely,

GEORGE E. BROWN, JR., *Chairman,*
Subcommittee on the Environment and the Atmosphere.

I. EXECUTIVE SUMMARY

The Staffs of the Subcommittee on Special Studies, Investigations and Oversight and the Subcommittee on the Environment and the Atmosphere have completed an evaluation of the Environmental Protection Agency's research program related to the health effects of air pollution, with primary emphasis on the Community Health and Environmental Surveillance System (CHESS). This investigation was noted earlier in the report on joint hearings held by the House Interstate and Foreign Commerce Committee and the House Science and Technology Committee, April 9, 1976.[1] While the joint hearings, held to determine whether there had been any deliberate distortion of data during the writing of the 1974 CHESS Monograph,[2] produced general agreement that no basis existed to question the integrity or scientific honesty of the project leader for the first CHESS Monograph, the hearings did not resolvemany questions and supported the need for a comprehensive evaluation of the scientific and technical aspects of the CHESS program and related Environmental Protection Agency (EPA) research efforts.

The investigation reported herein was conducted by a team composed of members of the staff of the House Science and Technology Committee with the support of consulting specialists in aerometry and analytical chemistry, meteorology, epidemiology, medicine, and general environmental health. Conferences were held on site, in the field, with individuals in EPA and in State and local government, academic, and private institutions who had detailed knowledge of the conduct of the CHESS programs. A number of findings and recommendations are provided in the report including those highlighted in the balance of this summary.

To some the findings may appear overly critical since CHESS pioneered in a difficult research area. However, in part the report is meant retroactively to complement the CHESS Monograph which fails to fully appraise the reader of the errors, deficiencies, and other shortcomings in the information presented in the Monograph. The intent of the report is to aid in assessing the validity of the conclusions presented in the Monograph and to assist researchers performing similar studies in avoiding similar problems. The endeavor was greatly assisted by hindsight and by the splendid cooperation and self-examination by investigators both inside and outside of the Environmental Protection Agency. The report also notes where problems similar to those found in CHESS research still exist in the EPA program.

---

[1] U.S. Congress. House Interstate and Foreign Commerce and House Science and Technology Committee. 94th Congress, 2nd Session. Report on Joint Hearings on the Conduct of the Environmental Protection Agency's "Community Health and Environmental Surveillance System" (CHESS) Studies. Washington, U.S. GPO April 9, 1976.

[2] U.S. Environmental Protection Agency. Health Consequences of Sulfur Oxides: A Report from CHESS, 1970–1971. May 1974. EPA-650/1-74-004.

The report identifies the need to reexamine the wisdom of legislative and executive pressures for rapid development of information for pollution control standards, and the adverse impact of these requirements on the performance of research. Also identified were the problems which are occurring because of the Office of Management and Budget (OMB) assuming operational control over details in the EPA research programs.

The CHESS program has historical value as a first attempt at a broad based definitive study relating air quality to health effects in a precise, quantitative manner. The program is a contribution to the general field of air pollution epidemiological studies. However, many problems critical to obtaining useful quantitive results were not solved in the conduct of this research. This failure to solve procedural problems became obvious when the studies were examined. There were too many inconsistencies in the data and too many technical problems that resulted in large data uncertainties or errors associated with the aerometrics for the results of this program to provide quantitative support for policy decisions. The 25 $\mu g/m^3$ lower sensitivity limit of the method used for most SO_2 measurements coupled with the overall large error band on all measurements (possibly exceeding 100%) and the apparent bias of SO_2 concentration data on the low side, make most of the numbers presented in the CHESS Monograph unuseable. In particular, the data are too imprecise and inaccurate in the range of the SO_2 standard (80 $\mu g/m^3$) to allow the determination of any health effects threshold. For these reasons the "hockey-stick" curves have no apparent basis in the data.

Further, there are inadequate data in the Monograph for interested researchers to evaluate fully the many caveats required for interpretation of the conclusions. For this reason, it is recommended that the EPA issue a formal notice of clarification and limitations to usage of the CHESS Monograph. In addition, EPA should prepare an addendum to the Monograph pointing out, in specific detail, the limitations of the Monograph. This addendum should: (1) Contain at least the material presented in Chapters I, IV, V, and VI, and Appendix A of this report; (2) Be made a part of any future printings of the Monograph; and (3) Be made available to the public and described in the formal notice mentioned above. EPA should not use data in the 1974 CHESS Monograph as supportive justification for policy decisions without explicit qualification and should avoid unqualified reference to the document in public information statements or documents. Further, the EPA should be encouraged to point out in any references to the 1974 CHESS Monograph that serious investigators should obtain all available data on the studies before continuing investigations of the same or similar phenomena. It is critical that EPA establish with high priority an adequately supported program to correct rapidly the kinds of deficiencies in research execution which are reported in this document. Only when the proper capabilities are established should major air quality health effects programs be renewed.

To recapitulate—the overall CHESS experience, as exemplified in the Monograph, teaches two outstanding lessons:

First, CHESS corroborates the notion that elevated air pollution levels cause adverse health effects.

Second, CHESS points out the many difficult problems involved in carrying out air pollution/health effects research.

Of course, the first lesson had been learned long ago, but failures in CHESS planning and execution precluded the development of new quantitative information on the relationship of air quality to health. Thus, CHESS cannot be called a major advance, only a confirmation of previous advances. With respect to the second lesson, the value of the CHESS experience lies mainly in the use to be made of it in correcting and improving operational procedures and research tools (i.e., instrumentation, questionnaires) and in planning future health effects studies. Finally, it seems that both these lessons could have been learned at much less cost in funds, elapsed time, EPA credibility, and staff morale.

EPA has only recently been reorganized to modify the National Environmental Research Center concept. This reorganization was examined in more detail in a previous report.[3]

The results of this present investigation suggest that the changes made last year are less effective than anticipated. The issue of research management should be continuously monitored to build the Environmental Research Center, Research Triangle Park capability and reputation. At the end of the 1977 fiscal year the situation should be examined to determine further need for reorganization. Perhaps consideration should be given now to some structure which provides for a single point of review at both the Research Triangle Park and Headquarters in order to alleviate the obvious lack of coordination of multidisciplinary projects, without attempting a major reorganization.

There is considerable evidence from this investigation that some critical disciplines, such as quality control, monitoring and sampling, and statistical analysis continue to be inadequately addressed in pre-project planning of health effects studies.

In the area of research management, the investigative team's recommendations are critical of the control of research programs within EPA. There is also concern with the fact that research is being conducted by other EPA offices involving little coordination with the Office of Research and Development. This functional overlap is beginning to produce problems within the laboratories and needs to be effectively resolved before these programs continue much further. Related to this overlap and the need for the solution of the problem of poor communication between the offices and laboratories, are the needs to provide for a more effective input from the Science Advisory Board, to establish a permanent peer review system to insure coordination among laboratories, and to expedite the timely publication of research results.

Although the Administrator of EPA had indicated during the joint hearings that there was a five year plan for research on sulfates, the investigative team found that only draft plans of this program were available and that the development of this plan was being poorly coordinated. Accordingly, no specific recommendations with regard to this plan can be provided. However, the need for better information on the species of chemicals, including sulfates, which are of primary

---

[3] U.S. Congress, House Committee on Science and Technology. Organization and Management of EPA's Office of Research and Development, 94th Congress, 2nd Session. June 1976. U.S. GPO. Washington, D.C. Serial LL 40 pp.

importance in health effects is emphasized and this recommendation has direct relevance to the proposed sulfate program.

A number of other recommendations are made with regard to the need for improvements in the air monitoring and quality control programs, the need for additional strength in meteorology, and a number of specific suggestions about improvements in the type of epidemiological studies reported in the CHESS document.

It should be noted that the investigative team received friendly cooperation throughout the entire investigation and appreciation for this help is acknowledged. A number of the recommendations for change evolved from the suggestions and the very real concern of EPA personnel for the improvement of their research program. The report bears on the condition of a research area which is the result of many forces and restraints. There is no intent to impugn the integrity or ability of any individual researcher, all of whom were found to exhibit professionalism and dedication to their program.

Finally, a most important point must be made which grows from but goes beyond the study. This report documents many cases of deficiencies in EPA research both past and present. It does not discuss many of the fine research efforts ongoing. The overriding purpose of this report then is to point the way to improving and enhancing the EPA research program, not to undermine the program. EPA can regulate wisely only if its decisions are based on information generated in a sound and comprehensive research program.

II. INTRODUCTION

A. Background

The Environmental Protection Agency, in partial response to the mandate of Congress under the Clean Air Act, conducts research to study the adverse health effects of air pollution. A series of major air pollution/health effects studies were carried out between 1967 and 1975 and were considered by many experts to be the most comprehensive of their kind. This program attempted to implement a coordinated system of aerometric measurements, including measurements of sulfur dioxide (SO_2), particulates, total sulfate, carbon monoxide (CO), ozone, hydrocarbons, nitrogen dioxide (NO_2), and nitrate. In areas where these measurements were made, the EPA initiated concurrent epidemiological studies embracing acute and chronic respiratory illness, ventilatory function of children, and the aggravation of asthma symptoms. A portion of these studies eventually was given the formal title of "Community Health and Environmental Surveillance System" (CHESS).

Although data were gathered over the entire 1967–1975 period, the monograph, *Health Consequences of Sulfur Oxides, A Report from CHESS, 1970–71*, (EPA 650/1–74–004, May 1974), treats only some of these data. Some controversy exists in the professional community regarding the technical correctness of the results and conclusions as presented in that monograph. The controversy was intensified following a series of newspaper articles starting with a February 29, 1976 article in the Los Angeles *Times*. In this article, scientists both within and outside of EPA were reported to have alleged that the data and its analysis were manipulated to more strongly support certain regulatory positions taken by EPA.

Following these allegations, the House Science and Technology Committee and the House Interstate and Foreign Commerce Committee on April 9, 1976 held a joint hearing on the alleged mishandling of the CHESS program.[1]

While the hearing generally negated the idea of a planned manipulation of the data, many questions were raised as to the reliability of the data, the technical soundness of its analysis, and the subsequent validity of the conclusions reached.

Further, the EPA Administrator, Russell Train, testified that there existed a five-year program plan for sulfate air pollution and health effects R&D which would assure a sound data base for promulgating air quality standards in the future. However, the plan itself was not presented.

As a result of the questions raised by the April 9 hearings and other sources, the staff of the House Science and Technology Oversight

---

[1] U.S. Congress. House. House Interstate and Foreign Commerce Committee and the House Science and Technology Committee. 94th Congress. 2d session. Report on Joint Hearings on the Conduct of the Environmental Protection Agency's "Community Health and Environmental Surveillance System" (CHESS) Studies. Washington. U.S. Govt. Print. Off. April 9, 1976. 25 pp.

Subcommittee and the Subcommittee on the Environment and the Atmosphere undertook a technical investigation of the entire CHESS program, with particular emphasis on the work reported in the 1974 CHESS Monograph. The investigation also looked into the follow-on CHAMP aerometric program and EPA plans for research to establish a sound technical basis for air pollutant health effects relationships. As noted in the Report on Joint Hearings cited earlier:

". . . the Committee on Science and Technology is currently investigating EPA's research program. This technical investigation will focus on CHESS and EPA programs planned to develop information necessary to make a decision on a sulfate standard." [2]

This investigation was launched due to the importance of EPA research to the Nation, not because of accusations raised in newspaper articles. It is simply that recognition of the importance of environmental research as a basis for regulation has and will continue to foster reviews of the EPA R&D program. We would like to emphasize that the team conducting this review worked with the goal of helping to improve EPA's research for the benefit of the country.

In planning the investigation, the team formulated the following 13 questions as a means of setting the scope and direction to the investigation. The team did not expect necessarily to arrive at findings in a format that would explicitly answer the 13 questions in the form they were set down. The questions are as follows:

(1) Did the CHESS measurement system utilize the best available instrumentation and sound operational protocols and quality control procedures?

(2) Were the precision and accuracy of the CHESS aerometric data sufficient to quantify pollutant exposure differences between communities where health effects were studied?

(3) Were the aerometric data consistent with and accurately representative of the exposures to pollutants received by the populations used in the health effects studies?

(4) Were new knowledge, instruments, and procedures introduced into the program as they became available and proven?

(5) Were the epidemiological populations adequately selected?

(6) Were the health measurement endpoints meaningful and reliable?

(7) Were the data reduction and analysis accurately and correctly carried out?

(8) What factors have retarded the analysis and reporting of the CHESS data?

(9) Did the health and monitoring data establish unambiguously what concentrations of specific pollutants are associated with measureable adverse health effect?

(10) If so, do the quantitative dose-response estimates in the CHESS report have a firm empirical basis?

(11) Were the conclusions drawn from the CHESS program sufficiently clear and unambiguous so as to form a sound basis for future action?

(12) Does the new "CHAMP" [3] program provide a substantially improved mechanism to provide aerometric data to support future

---

[2] U.S. Congress. House Interstate and Foreign Commerce Committee and House Science and Technology Committee. Report on Joint Hearings. op. cit. p. 25.

[3] Community Health Air Monitoring Program, a systematic program for monitoring air pollution in connection with health effects studies.

epidemiological research and is an epidemiological research program planned to complement future aerometric monitorings?

(13) What additional steps, if any, should EPA take to insure that a sound technical basis will be available for future agency decisions pertaining to SO_2/sulfates?

Answers to these 13 questions were developed from the specific findings and are given at the end of Chapter III, Findings, Conclusions and Recommendations.

The investigative team consisted of at least one person with technical knowledge in each of the scientific disciplines expressly involved in CHESS, i.e., epidemiology, aerometric measurement, meteorology, data analysis, and R&D planning and management.

The first phase of the investigation involved visiting field data centers and measurement stations and interviewing the personnel who ran monitoring systems and who collected health data. The next phase involved a review of data and analysis procedures used in the CHESS monograph and intensive interviews with the key scientific and management personnel at the EPA laboratories at Research Triangle Park, North Carolina. The third phase consisted of interviews with EPA headquarters personnel who had been involved in the CHESS and CHAMP programs.

This report documents the investigation which was focused on the CHESS program and its published output. It was not, however, restricted to the CHESS program. The investigators looked into related programs (e.g. current epidemiology projects), and into supporting programs (e.g. development of measurement methods). As a result, the scope of the report is broader than CHESS. Since the CHESS program ended a year ago this report will have most value if it is pertinent to existing and future EPA research. It is definitely believed that many of the CHESS findings and recommendations are indeed relevant to current research programs in EPA, and that the CHESS mistakes might well be repeated if they are not documented. An Executive Summary is provided at the beginning of this Report for the convenience of the Committee in quickly getting an overall picture of the findings. The body of the report presents summarized findings and recommendations (Chapter III), followed by a detailed discussion and critique of the aerometric measurement and epidemiological parts of the CHESS program.

The field review naturally divided into air quality measurements and epidemiological studies. The first part of the field review (Chapter IV) presents the results of the evaluation of aerometric measurements, including the next generation CHAMP program. Chapter V presents an overview of the air quality analysis procedures, Chapter VI presents an analysis of the health effects studies. Appendix A presents a recapitulation of the aerometric and meteorological findings of the investigation as they relate to specific health studies. Appendix B presents a legislative history relative to the CHESS program.

B. History of Air Pollution Health Effects Studies Leading to CHESS Objectives

There has never been any question about the need for control of severe air pollution. Undisputed historical data have clearly implicated high levels of industrial pollutants, which invariably include oxides of

sulfur and particulate matter, with increased incidences of health effects—both moribidity and mortality. These historical reports clearly indicated the danger from such incidents and pointed the way toward a need for additional information on the nature and extent of the damage which such pollutants might produce.

In early December, 1930, a thick fog covered the industrial Meuse Valley in Belgium. Firket [4] recounted that several hundreds were sickened by suddenly appearing respiratory symptoms, complicated in a large number by grave circulatory failure. He notes that, "More than sixty died on the 4th and 5th of December after only a few hours' of sickness. A sizeable number of livestock had to be slaughtered." Having calculated that mortality rates were ten and one half times average during this period, Firket commented that, should a similar fog afflict London, "one would have 3,179 immediate deaths to mourn." His words were prophetic. Four thousand excess deaths were attributed to the four day London Fog of December, 1952, and Logan [5] added the following historical note to his mortality figures:

"The incident was a catastrophe of the first magnitude in which, for a few days, death rates attained a level that has been exceeded only rarely during the past hundred years—for example, at the height of the cholera epidemic of 1854 and of the influenza epidemic of 1918–19." Nor was the United States spared. Nearly half (42.7%) the inhabitants of Donora, Pennsylvania experienced symptoms during the "smog episode" of October, 1948, and twenty died during the final week of October, a period which normally occasions but two or three deaths.[6]

A statistically significant increase in mortality was also documented during a "pollution incident" in New York City in November, 1953.[7] Thus, "air pollution," or "smog," or "fog" was recognized as a lethal menace, both in the United States and abroad.

These tragic episodes raised many questions and spawned a generation of research. Although various products of sulfur combustion were considered likely noxious agents in these episodes, the commentators were duly cautious in suggesting which pollutant or combination of pollutants might prove harmful. Recent reviews [8 9 10] describe population studies, clinical experiments, and toxicological studies designed to define which pollutant(s) pose risk, and a concommitant body of research sought to develop and refine techniques of pollution measurement.

The above chronicled incidents and research logically led to the question: If high levels of pollution can kill, what can low levels do to susceptible subpopulations, such as asthmatics, and what harm might come to the general population if exposed to low levels of pollution for

---

[4] Firket, M., *Bulletin de L'Academie Royale de Medecine de Belgique* Series 5, Volume 11, Number 10 (1931), pp. 683–739.

[5] Logan, W.P.D., "Mortality in the London Fog Incident." *Lancet* Volume 264, Number 6755 (1953), pp. 336–338.

[6] Shrenk, H. H., Heimann, H., Clayton, G. D., Gafafer, W. M., Wexler, H., "Air Pollution in Donora, Pennsylvania: Epidemiology of the Unusual Smog Episode of October 1948". Public Health Bullletin Number 306 (1949).

[7] Greenberg, L., Jacongs, M. B., Drolette, B. M., Field, F., Braverman, M. M., "Report of an Air Pollution Incident in New York City, November 1953." Public Health Reports Volume 77, Number 1 (1962), pp. 7–16.

[8] Rall, D.R., "Review of the Health Effects of Sulfur Oxides." Environmental Health Perspectives Volume 8 (1974), pp. 97–121.

[9] "Air Quality and Stationary Source Emission Control." National Academy of Sciences Report for the Committee of Public Works, United States Senate Pursuant to S. Res. 135 Serial #94–4 (March 1975), pp. 58–169.

[10] "A Critical Evaluation of Current Research Regarding Health Criteria for Sulfur Oxides." Technical Report prepared for Federal Energy Administration by Tabershaw/Cooper Associates, Inc. (April 11, 1975.)

years? To answer such questions, epidemiologists began to measure sublethal effects of pollution such as eye irritation, susceptibility to chest colds, subtle changes in pulmonary function, and asthma attack frequency (see references [11] and [12].) The Community Health and Environmental Surveillance System (CHESS) represented a natural extension of efforts to gather concurrent morbidity and air pollution data. In 1970, a description of CHESS [13] stated:

> The keystone of the CHESS program is the coupling of sensitive health indicators to comprehensive environmental monitoring . . . .
> . . . EPA health research needs are practical and problem oriented. CHESS research is thus pragmatic and our goals are threefold: (1) to evaluate existing environmental standards: (2) to quantitate pollutant burdens in exposed populations; and (3) to quantitate health benefits of pollutant control.

It was perhaps this last goal which led to the rapid expansion of the CHESS system in 1968–70 and which necessitated the use of health measurements and monitoring methods which in some instances had not been adequately field-tested. Municipal regulation of industrial and power combustion, space heating, and incineration resulted in dramatic decreases in sulfur dioxide and total suspended particulates in New York City, Chicago, and elsewhere.[14] These favorable pollution trends were already evident in 1967. Thus, by 1970, it was urgent that population studies begin immediately if corresponding improvements in health indices were to be documented. CHESS data was gathered in New York City, Salt Lake City, Birmingham, Charlotte, and in Chattanooga from 1970 through 1975 and Los Angeles between 1972 and 1975. The analysis of the data and compiling of conclusions was begun aggressively as data became available. The facts relating to this program history are discussed in the following chapters.

---

[11] Nelson, D.J., Shy, C.M., English, T., Sharp, C.R., Andleman, R., Truppi, L., Van Bruggen, J., "Family Surveys of Irritation Symptoms During Acute Air Pollution Exposures." Journal of the Air Pollution Control Association Volume 23, Number 2 (1973), pp. 81–90.

[12] U.S. Environmental Protection Agency. Health Consequences of Sulfur Oxides: A Report From CHESS,1970–1971. May 1974. EPA–650/1–74–004.

[13] Riggan, W. B., Hammer, D. I., Finklea, J. F., Hasselblad, V. Sharp, C. R., Burton, R. M., Shy, C. M., "CHESS, a Community Health and Environmental Surveillance System." Proc. of the Sixth Berkeley Symposium on Mathematical Statistics and Probability (1970).

[14] U.S. EPA. CHESS. op. cit.

77–590—76——2

III. FINDINGS, CONCLUSIONS, AND RECOMMENDATIONS

A. General Findings

The findings resulting from the CHESS program review have extended into EPA's evolving task of developing a sulfate standard as well as a general perspective on EPA's entire air pollution-health effects research program.

The EPA five year plan for sulfates research cannot be commented on at this time because the plan is still in the formative stage. Thus, the recommendations deal primarily with what it is believed can be accomplished to prevent repetition of the problems produced by or related to the CHESS program.

The EPA CHESS program intended, as one of its prime goals, to provide a sound basis for relating human health effects to precise quantities of SO_2 and other oxides of sulfur in air and to provide a reliable record of the improvement of the health of the U.S. population as air pollution was controlled and reduced. The program, spanning 6 to 7 years, was to achieve these objectives by measuring and analyzing 6 health indicators together with SO_2, NO_2, suspended sulfates, total particulates, and oxidants in the air in 6 representative areas throughout the country. Only the 1969–1971 data were analyzed, correlated and published by 1974 in the CHESS Monograph. Most of the remainder of the data are in various states of analysis.

A critical review of the epidemiological and air monitoring programs with conclusions for each subprogram is presented in the body of the report. The results present a picture of a program pressured by EPA management-imposed time constraints to meet legislated mandates for promulgating new standards, hampered by inadequate mechanisms to detect and correct technical problems, and handicapped by budgetary and management restrictions placed on the program after it was well underway.

In the drive for results, the program did not adhere to standards of quality control, validation of methods, cross-checking of data, and calibration of instruments required in such research. The decision not to submit all results of on-going research for publication in professional journals deprived the CHESS program of an important source of peer review and evaluation at various stages, although the Monograph was sent to many reviewers for comments. The simultaneous publication of a large number of complex studies in a single report—the 1974 CHESS Monograph—rather than as individual on-going research reports may have been an administratively expedient method of publishing and collating results, but it undoubtedly reduced the effectiveness of peer review. The complexity of the document also impeded the public from acquiring an understanding of the results of the studies associated with policies of national importance.

As noted in Chapters II and V, the investigative team is aware that historic incidents and other clinical, toxicological, and specific epidemiological studies support the general concept that health and welfare will be improved as air pollution is reduced. However, technical errors in measurement, unresolved problems in statistical analysis, and inconsistency in data in the 1974 CHESS Monograph render it useless for determining what precise levels of specific pollutants represent a health hazard. Nonetheless, the CHESS Monograph presents some evidence of adverse health effects of air pollution, and, in particular, slightly increased prevalence of chronic respiratory disease is found in association with whatever levels of SO_2 and indeterminate particulate pollution were present in the air during the period the CHESS health panels were active. CHESS also has historical merit for undertaking the first massive air pollution health effects research program in the Western World. It is important to record and remember the lessons learned for future undertakings.

As pointed out in Chapter VII, the findings regarding CHESS epidemiological work corroborate the issues raised in previous reviews. This report collects the issues and goes into depth as to effects of errors, etc., and suggested corrective action.

In addition, this review documents serious errors and omissions in the aerometric data and analysis meteorology. Some of the more important specific findings and related recommendations related to the CHESS program are presented below, followed by RTP research lab operational issues, by suggestions as to possible improvements in health effects programs, and finally, by a review of the 13 questions set forth at the beginning of the investigation.

It should be noted that these findings do not contradict the conclusions in the report of the April 9 joint hearings.

B. Specific Findings and Recommendations

1. The data and analysis in the 1974 CHESS Monograph do not provide a reliable, quantitative basis for relating SO_2 or sulfate levels in the air to adverse health effects.

The methods and instruments used had associated measurement errors (inaccuracies) such that the threshold curves ("hockey-stick curves") presented in the Monograph are without basis in fact. The indicated thresholds occurred at levels at which the measurement methods (as used) were not reliable. Specifically in the case of SO_2 measurements the method used had a threshold of $25\mu/m^3$ and the accumulative errors found in hindsight in handling, shipping, and reading the instruments made SO_2 reading between 50 and $200\mu/m^3$ unuseable. There were similar problems with other measurements as pointed out in Chapters IV and VI.

2. The CHESS program did appear to demonstrate adverse health effects in association with local air pollution of undefined level and pollutant mix. Studies of chronic respiratory disease prevalence rates provided the most consistent results. These results are not translatable to communities other than those measured (an objective of the research) because the error uncertainties in the aerometric data made it impossible to arrive at reliable quantitative relationships.

3. The CHESS monograph has been referred to either directly or indirectly as a document supportive of regulatory and standards ac-

tions. Specific examples were found such as its use to support EPA positions on air pollution health effects in the following excerpt from the Administrator's Report to Congress for 1974:

> The CHESS studies *have demonstrated* [emphasis added] the benefits from improved air quality with respect to the chronic respiratory disease experience of subjects who have moved to communities having cleaner air. Also, the studies *have shown* [emphasis added] that children living for 3 or more years in communities having high level of air pollution have more acute respiratory disease episodes than recent immigrants to the community. . . . Data obtained from the CHESS program indicate that adverse health effects *are consistently* [emphasis added] associated with exposure to suspended sulfates, indeed, more so than to SO_2 or total suspended particulates. This information has initiated further study in the transport processes and control techniques for suspended sulfates.[1]

The CHESS study results were cited in an EPA document [2] dated May 1976. The preface of that document states "The report's objective is to review results of research by and under the sponsorship of ERC/RTP, related to criteria and hazardous pollutants since publication of the criteria and background documents." CHESS studies are explicitly mentioned in the report (e.g., page 207, page 276) with no references. Pages 303 through 308 seem to describe the CHESS studies as reported in the CHESS Monograph, although *no reference to a document is given*. Curiously, reference numbers 244 through 252 are not found in the body of the text where they would be expected (between pages 301 and 308). Inspection of the references list at the end of the pertinent chapter shows that the "missing" references, numbers 244 through 252, refer to the CHESS Monograph. Thus, it appears that CHESS results are used without attribution, and thus without caveat.

Recommendations:

(*a*) EPA should immediately publish and distribute a formal announcement to the effect that the 1974 CHESS Monograph is at best a preliminary document reporting research of varying degrees of reliability and as such should be considered only as a beginning study of larger problems. The notice should further point out: (1) That many of the problems leading to the unreliability of the research are not reported in the Monograph; (2) That no interpretations should be made based only on the data in the Monograph; And (3) that other data available at EPA, including data taken since the publication of the Monograph, indicate the need for reexamination of the conclusions in the Monograph. The notice should also announce the availability of and describe how to acquire the addendum recommended below.

(*b*) EPA should not utilize the 1974 CHESS Monograph as a source for specific quantitative data supportive of standards or regulatory decisions without explicit qualification.

(*c*) EPA should publish an addendum to the CHESS Monograph containing at least the materials in Chapters IV, V, and VI and Appendix A of this report. EPA should take the initiative to make sure this addendum is available to the public and included in all future issues of the Monograph.

4. If the EPA had drawn adequate attention to the many difficulties being encountered in the CHESS program (which was under pressure to provide data as fast as possible to support air quality standards

---

[1] U.S. Environmental Protection Agency. Report to Congress. Progress in the Prevention and Control of Air Pollution in 1974 (undated) p. 39.

[2] Scientific and Technical Data Base for Criteria and Hazardous Pollutants 1975 ERC/RTP Review. Health Effects Research Laboratory, U.S. E.P.A. (EPA-600/1-76-023) May 1976.

under conditions set by the Congress, and on which OMB subsequently placed restrictions as to the allocation of positions and funds) it is possible that the program could have been either terminated with considerable savings or restricted in scope to permit a more satisfactory research program to be completed.

Recommendations:

(*a*) Legislation affecting the research responsibilities of the EPA should be reexamined to insure that unrealistic procedures and schedules are not assigned.

(*b*) Research programs should be designed to gain information and answer questions, not to support positions.

(*c*) If public policy requires expeditious research, every effort should be made to insure that the OMB allows all necessary funding and personnel.

(*d*) In the event that budgetary restrictions evolve which will have an impact on the completion of major projects, the EPA Administrator should be required to advise the Congress of the potential for such impact. Consideration should be given to requiring the Administrator to provide an annual analysis to the Congress which estimates the probable impact of proposed budget changes on major projects such as the CHESS program.

5. At the present time, OMB has the power to control the most detailed aspects in a research program. For example, a delay in the approval by OMB of health effects questionnaires submitted by EPA prevents or significantly delays the continuation of epidemiological studies associated with the establishment of air quality standards. Thus EPA has the responsibility for its mission, but lacks the necessary authority to carry it out.

Recommendation:

The OMB should be asked to develop procedures by which questionnaires and similar documents required for research programs can be given prompt review and approved without restrictive time limits for research programs of like nature.

6. The Health Effects Research Laboratory at Research Triangle Park is currently proceeding with the processing of the remaining data from the CHESS program. If adequate resources are made available to complete both current and past data processing programs, then the CHESS data may provide some additional insight into pollution health relationships or at least identify pitfalls to avoid in future studies. The loss of some key personnel involved in the CHESS program makes unavailable important insights provided by individuals who were involved in gathering of the data.

Recommendations:

(*a*) In view of the current unsettled conditions at the Health Effects Laboratory and the severe restrictions on manpower and budget, it is recommended that the unanalyzed epidemiological and aerometric data be examined critically by year (to determine at what point improvements in measurement make the data acceptable) and that analysis be carried out only on that data that appears to have a higher degree of validity than the '69–'71 data used as a basis of the 1974 CHESS Monograph.

(*b*) Steps should be taken to attempt to publish research in traditional, referred, archival journals in a timely fashion.

(*c*) The results of large or complex projects should not be published solely in monograph form. If EPA decides to publish such results in monographs as well as journals, then EPA should take care that, within the limits of reason, all pertinent data are included.

(*d*) Projects should not be initiated unless there is full assurance that the data collected in such projects can be properly analyzed and reports prepared in a realistic time frame for policy consideration. If policy considerations dictate a time frame, then research staff must have a part in deciding what can realistically be accomplished in the allotted time.

7. While CHAMP looks more promising for providing reliable real time measurements of air quality (absolute pollutant levels), the system reliability and precision for SO_2, sulfates, and NO_2 have yet to be proven.

Recommendations:

(*a*) The aerometric and quality control programs in EPA should be further strengthened and improved.

(*b*) Every effort should be made to shorten the time between the acquisition of data and quality assurance analysis of data so that errors in instrumentation and collection may be corrected.

(*c*) The practice of employing development stage instruments in field operations before qualification testing is satisfactorily completed should be stopped.

(*d*) Field utilization of laboratory models of instruments is not acceptable until the instrument has been thoroughly checked out in the field and until all field personnel have been adequately trained in its use.

(*e*) The opening of the CHAMP operation contract to competition now, before development is complete should be reexamined by EPA policy management to see whether the merits of open bidding at this time outweigh the problems of instability.

(*f*) Health effects research personnel need to coordinate closely with air quality and monitoring personnel so that there is early understanding as to the chemical species to be monitored as well as the accuracy needed for support of health effects research.

8. Meteorological support to the epidemiological program has been passive and not fully developed.

Recommendations:

Additional meteorological support is needed in the health research-air pollution effects studies. Assuming that CHAMP stations will continue in operation, the meteorological instrumentation should be uniform and complete for all stations.

9. The characterization of species of air pollutants and combinations thereof which are of particular health or welfare significance needs to be speeded up. The useful output from future large scale epidemiological studies of general air pollution effects will be minimized if such characterization is not available. Such characterization will require a greater effort than is now possible with currently assigned resources. The present Federal effort in this area is fragmented and uncoordinated.

Recommendation:

The EPA Health Effects program as well as interagency utilization of all available Federal and extramural resources in the health effects area should be examined with the objective of significantly accelerating research in this area.

10. The overall impression left with the review group was a general awareness of many of the problems we found in the air quality health effects research area but an apathy to really drive to correct them quickly.

Recommendations:

(*a*) A truly interdisciplinary task force led by an eminent scientist should draw up a program plan for EPA to develop a solid base of knowledge and procedures in aerometric instrumentation and measurements, meteorology, field data gathering, quality control, epidemiology project design and testing and panel planning.

The plan should be thoroughly reviewed with peer groups and then funded for two years or as required at the expense of other programs so that effective tools for meeting research needs will be available for future programs.

(*b*) At the very least, CHAMP should be pushed to verify present instruments and protocols so that reliable data at pollutant concentrations within instrument capabilities can be achieved. New third generation instruments, such as the pulsed fluorescence devices should be developed and proven for field usage so that reliable values with uncertainties of less than $\pm 20\%$ can be detected for pollutants at concentrations well below current standards.

(*c*) The epidemiological questionnaires, panel selection criteria and other items criticized in this report should be worked out and approved by peer groups so that EPA will be prepared for the next round of serial or episodes investigations with reliable tools and plans. Specifically, questions such as the following must be resolved for future epidemiological studies:

(1) How do CRD* questionnaire responses change on serial administration in an area with unchanging pollution patterns?

(2) What is the sensitivity of the self-administered CRD questionnaire compared with its use in an interview?

(3) What is the nature of the statistical dependence of ARD★ attack rates, and what formal statistical methods are appropriate to the analysis of relative attack rates?

(4) What can be done to tighten the eligibility requirements for asthma and cardiopulmonary panels?

(5) How can the statistical analysis of asthma and cardiopulmonary panels be improved?

(6) What combination of CHESS health measurements is most appropriate to long-term serial surveillance?

(7) What combination of CHESS health measurements is appropriate to intensive studies of specific pollution hazards?

---

*CRD=chronic respiratory disease.
★ARD=acute respiratory disease.

(*d*) Along with correcting epidemiological program deficiencies pointed out in Chapter VI of the report, the following ideas were raised in team interviews with EPA personnel and others and are worthy of review in such a program as outlined above.

(1) Even though CHESS was stopped partially due to the cost of such massive programs, the CHESS concept of health surveillance should be reevaluated before a decision is made not to undertake similar programs in the future. Documentation of improving public health with pollution abatement; identification of new pollutants becoming threats; and baseline serial health data are important.

(2) The appropriateness of the CHESS endpoints (i.e., dependent variables) such as chronic respiratory disease (CRD), should be evaluated with respect to utilizing more objective and more easily quantified health measurements, such as mortality rates, admissions to emergency wards, or blood or pulmonary function tests as being easier to document and more reliable.

(3) Surveillance in small rural towns may yield valuable control data on serial health measurements.

(4) Epidemiological program planners should consider utilization of local physicians to help develop and screen panels and gather data in localities where there is an interest.

(5) Additional consideration should be given to the study of high risk groups such as those in nursing homes, those recently admitted to an intensive care unit for myocardial infarction, and premature infants.

(6) Resources should be available for opportunistic epidemiology to study interesting events, such as the January 1976 episode of high SO_2 in Magna, Utah. Such events as this may allow one to disentangle the acute effects of particular pollutants.

(7) Alternative populations should be considered, and cohort studies of particular occupational or medical risk groups may prove valuable.

(8) More attention might be given the need for individualized dosimetry.

(9) Under what circumstances, if any, can epidemiologic studies alone provide unambiguous dose-response data?

(10) More exotic studies might be considered. The health effects of a new event could be observed, such as the construction of a coal gasification plant. Changes in a town which is known to plan selective modifications in the regulation of a specific pollutant could be studied. Studying vacationers from small towns who spend a week in a polluted urban center might be considered. Longer term migrant studies might also be useful.

C. R&D Related Findings

During the course of this investigation the extensive interviews brought to light some issues which did not bear directly on the CHESS program above but which are important to the effectiveness of the

total EPA R. & D. program. Generally the Research Triangle Park EPA Health Effects Research Laboratory, and in fact all RTP laboratories visited, appear to be inadequately staffed and suffering from poor morale due to a number of factors including frequent changes in leadership, poorly understood organizational changes, uncertainty about responsibilities, difficulty in recruiting high quality personnel for leadership positions, poorly defined long-range task objectives, and physical separation of working units. In addition, there appear to be serious problems of coordination and communication between divisions at RTP as well as at Headquarters, Washington.

These general findings are also borne out in the report on "The Organization and Management of EPA's Office of Research and Development" prepared by the Subcommittee on the Environment and the Atmosphere of the House Committee on Science and Technology in June, 1976.[3]

There follow several specific findings and related recommendations.

11. The laboratories have not yet completely recovered from the (inevitable) disruption associated with the December, 1975 reorganization of the Office of Research and Development.

Recommendation:

Since stability is an important ingredient in any good research organization, the Center should be allowed to function in its present mode without further significant reorganization until the end of Fiscal Year 1977 with all possible help it can use from Headquarters.

12. There is evidence of a need for improved coordination between the four research laboratories at the Environmental Research Center, RTP, particularly between the Health Effects Research Laboratory and the Environmental Monitoring and Support laboratory. The exchange of data and coordination of effort among the EPA/RTP laboratories, and between EPA, contractors, and Federal laboratories involved in similar programs is wholly inadequate.

A specific example of this lack of coordination is seen in CHESS. However, even after this learning experience the Environmental Protection Agency is still not well organized to undertake and conduct multidisciplinary field projects and, in fact, certain of the recent organizational changes are believed to have exacerbated this problem. When multi-disciplinary projects are organized and run by staff having a single, particular discipline, important elements of the problem may be omitted or given insufficient attention. Objectives need to be correlated more effectively with requirements and capabilities. Statisticians and quality control experts need to play a larger role in the planning of such projects. Problems of coordination and communication seem to arise at project levels because of the Office of Research and Development organizational structure in which individual laboratories report directly to separate management offices in EPA Headquarters. This is epitomized in the draft research plan for sulfates, for example, wherein the research effort is divided into four distinct segments at RTP but there is no program management at RTP to insure coordination. As a result, the four research elements are not developed on consistent premises.

---

[3] U.S. Congress, House Committee on Science and Technology, Subcommittee on the Environment and the Atmosphere, Organization and Management of EPA's Office of Research and Development, 94th Congress, 2nd Session, June 1976. U.S. Government Printing Office, Washington, D.C. Serial LL, 40 pp.

Recommendation:

(a) In accordance with the recommendation under 11, above, improved coordination should be sought at this time without formal organizational changes. This could be achieved through establishment of an authoritative peer review panel to monitor all on-going projects and insure appropriate interface between the various disciplines. Such a review panel should represent the highest level of research skill and should assist not only in program planning and review for scientific objectivity, but in peer review of major scientific reports with policy implications. Such review should insure that future reports are released in a timely fashion after complete scientific assessment.

This need for coordination cannot be satisfactorily achieved by *ad hoc* arrangements alone. Nevertheless, the formation of *ad hoc* working groups, under the oversight of the continuing peer review panel mentioned above, could be a useful mechanism for attacking special problems, such as planning a coherent sulfate research program.

In addition, the following additional steps might be taken without affecting the stability of the laboratory:

(b) A stronger focus on management at the Environmental Research Center, RTP, would aid in interfacing with the four Deputy Assistant Administrator Offices at Washington to prevent separate tasking on essentially identical problems.

(c) Consideration should be given to creation of a systems analysis/operations research program review group at Headquarters of EPA and perhaps at each technical center. This activity could insure that proper attention is addressed to all scientific disciplines in program plans and budgets, irrespective of which scientific discipline initiates a program. This group might act as staff to the peer review panel.

(d) The Science Advisory Board's (SAB) charter should be re-examined to improve the assistance rendered to the Health Effects Research Laboratory and to Washington project directors. The SAB should have direct informal access to the research programs and opportunity to review and recommend actions relative to planning and coordination of projects involving several divisions. The SAB should have scientific review responsibility for programs, and its participation should not be limited to public evaluations after something goes wrong.

13. The EPA research programs in many cases were found to be somewhat isolated. The research staff does not seem to avail itself of a maximal technical exchange with scientific peers. It seems that to a large extent outside contracts go to the same groups again and again.

Recommendations:

(a) EPA should actively seek cooperative research programs with a variety of universities, other laboratories, other agencies. *NOTE:* This is not a recommendation that more task-limited contracts be given, but that truly cooperative programs be implemented in which the outside laboratory can contribute its own ideas.

(b) EPA should arrange for research staff to spend six to eighteen month assignments working at other labs to reduce scientific isolation and to promote career development (see below). Similarly, they should

promote reciprocal visiting research assignments for outside scientists to work at EPA laboratories.

(c) EPA should consider funding individual, worthwhile Ph. D. thesis research as a means of broadening their technical community.

(d) The Science Advisory Board should be asked to help in developing such outreach programs.

14. There are examples of increasing overlap among the Office of Research and Development (ORD), the Office of Pesticides (OP), and the Office of Toxic Substances (OTS), at Washington, in assignments to the Environmental Research Center, RTP. Instances were noted where both ORD and OTS have asked the RTP Center for research on the same substances and where it was evident that neither the divisions at the RTP Center nor the Offices at Washington had coordinated the inputs to the Center during the development and assignment of projects.

In addition it appears that research is being carried out in offices other than ORD.

Recommendation:

The Administrator should clarify the role of the Office of Research and Development in managing Agency research, and the role of the laboratories in carrying it out.

15. The various EPA groups at RTP are now scattered geographically in such a way that the peer interaction needed to insure proper coordination and exchange of ideas is physically difficult.

Recommendation:

Whatever final management concepts evolve, this separation of facilities must be resolved in an expeditious fashion if the EPA program at RTP is expected to function with any degree of cohesiveness.

16. Poor staff morale seemed largely due to the superposition of a very difficult task (research in support of EPA's regulatory mission) upon a very uncertain and fragile system for the support of research. This would seem to cause or at least contribute to anxiety among the professional research staff, especially since there appears to be no plan of career development for this staff.

Recommendation:

EPA management should develop, implement, and defend a professional career development plan for each professional employee and for the staff as a whole. Developing such a plan will contribute to better management by forcing a comparison of goals and resources. An employee secure in a peer-respected professional career will more happily tolerate the exigencies of research in a regulatory agency. Management of the ORD will sometimes have to defend its commitment to such a plan against regulatory demands.

17. As stated above, major improvements in research execution are needed. If, after the end of fiscal year 1977, the present structure is judged inadequate to achieve this, a number of alternatives and factors should be considered. The RTP laboratory directors are divided as to the type of management structure needed but there is a general impression that the National Environmental Research Center

(NERC) concept was preferred to the four-RTP-division/four Washington-manager structure which currently exists.[4] A major problem is the lack of a single manager's office at the Center to insure interdivisional review of multi-disciplinary projects and to secure input from all divisions on such projects.

If success in recruiting key leadership and improving program execution cannot be shown, even more fundamental steps may be needed.

The steps to be considered might include (1) moving the existing EPA research program into a new or existing agency; (2) establishing a new agency to provide the basic, long-term research support needed (but not being done) and effectively dedicating the present EPA laboratories to short-leash support of the agency regulatory programs; (3) assigning long-term research support responsibilities to other existing agencies.

Recommendation:

The Administrator must be prepared to answer hard and fundamental questions such as, should EPA conduct research?

D. Recapitulation of Guiding Questions

The following tabulation provides summary answers based on the findings of the investigative team to the guiding questions formulated at the beginning of the investigation.

(1) *Did the CHESS measurement system utilize the best available instrumentation and sound operational protocols and quality control procedures?*

The instrumentation was that of the state of the art at the time the program started. However, it is quite clear that neither the protocols nor quality control procedures were adequate to accomplish the objectives of the program, and that the data usefulness could have been greatly enhanced if well-known quality control procedures had been rigorously followed. For example, if the procedures described in the CHESS Monograph (but not actually used) had been implemented, the data would have been much more useful. If error bars had been associated with each data point the weakness of some of the cause-and-effect relationships might have been evident. This, in turn, would probably have prevented some of the over-interpretation of data.

(2) *Were the precision and accuracy of the CHESS aerometric data sufficient to quantify pollutant exposure differences between communities where health effects were studied?*

For many pairs of communities the answer is clearly no—both because of inadequate procedures and also because at apparent concentration levels in the lower-concentration areas the measurement methods were incapable of detecting differences or changes in those pollution concentrations.

(3) *Were the aerometric data consistent with and accurately representative of the exposures to pollutants received by the populations used in the health effects studies?*

---

[4] Under the NERC structure, the entire RTP laboratory reported to a NERC Director who in turn reported to Washington Headquarters.

This question cannot be answered, because there is no information in CHESS on the relation between individual exposure and ambient air quality. Thus, the CHESS researchers did not know this answer either. In addition, the practice of having only one aerometry station per community made it difficult to assess area-wide concentrations. This illustrates the importance of personal dosimeters.

(4) *Were new knowledge, instruments, and procedures introduced into the program as they became available and proven?*

It seems the introduction of new techniques was slow at best. For example, as stated above even well-known quality control procedures were introduced only slowly. The Committee understands the importance of consistency in a project like CHESS, but this can be assured with proper quality control techniques.

(5) *Were the epidemiological populations adequately selected?*

The findings raise doubts as to the validity of certain of the population selections.

(6) *Were the health measurement endpoints meaningful and reliable?*

Again, the findings disclosed many shortcomings in this area. As with the areometry, the health measurements could have been more useful if validated, etc.

(7) *Were the data reduction and analysis accurately and correctly carried out?*

There were instances of errors in both data reduction and analysis.

(8) *What factors have retarded the analysis and reporting of the CHESS data?*

Several factors were uncovered. The most important appear to be the forced change of computer and support contractors.

(9) *Did the health and monitoring data establish unambiguously what concentrations of specific pollutants are associated with measurable adverse health effects?*

The findings clearly say no.

(10) *If so, do the quantitative dose-response estimates in the CHESS report have a firm empirical basis?*

From (9) the answer must be no.

(11) *Were the conclusions drawn from the CHESS program sufficiently clear and unambiguous so as to form a sound basis for future action?*

Again, the answer is no. CHESS does indicate the need for a rigorous program of health effects research.

(12) *Does the new CHAMP program provide a substantially improved mechanism to provide aerometric data to support future epidemiological research and is an epidemiological research program planned to complement future aerometric monitorings?*

The CHAMP system is clearly an improvement in aerometric pollution measurement, although it is not yet fully validated. There is no clearly defined, well-targeted, coordinated epidemiological program to make use of CHAMP at this time. There are no projects in epidemiology making use of CHAMP data.

(13) *What additional steps, if any, should EPA take to insure that a sound technical basis will be available for future agency decisions pertaining to SO_2/sulfates?*

For the answer to this question, refer to all the recommendations in the sections B and C concerned with program improvements.

E. Closing Remarks

The investigative team is aware that this is a very critical report. The Environmental Protection Agency is a relatively new agency which has been confronted with the need to grow rapidly and to assume a large burden of responsibility. The enforcement of these responsibilities has increased in political significance with the growth in intensity of the overall energy problem. All EPA personnel consulted in this study were generally cooperative and hopeful of being able to provide constructive suggestions. Indeed, the recommendations presented here are, in many instances, suggestions and comments made by EPA personnel. It was not difficult to identify many deficiencies within the CHESS program or to relate these difficulties to the need for changes in organization and management within EPA. The difficult task will be to resolve these issues and construct a more effective organization to meet the very real problems of our society which have been mandated to the Environmental Protection Agency for resolution.

IV. CHESS AEROMETRIC MEASUREMENTS

A. Introduction

As pointed out in the introduction, the attainment of precise, reliable, reproducible, and real time air quality measurements in the field (e.g., SO_2 and particulates) was a critical element of the CHESS program. This chapter provides a critical review of the aerometric measurement aspects of CHESS.

However, before reporting on this review two facts about CHESS aerometry should be mentioned. First, the methods used in CHESS, especially in 1970–71, were probably as good as any available. Second, quality control procedures were slowly introduced into the CHESS program. EPA cannot be criticised, and is not criticised in this report, for using the best available methods. However, EPA can be criticized for not pursuing a vigorous program of quality control throughout CHESS. The review reported here showed that CHESS did not employ well-established quality control measures. The quality control program described in Appendix A of the Monograph was not carried out. A thorough quality control program would have discovered, for example, the temperature effects on the method used to measure SO_2 (described below). It would also have placed bounds on the validity of the data and precluded overinterpretations.

In the design and implementation of any measurement system, the single most important consideration is the end user of the data produced by that measurement system. In the simplest of all measurement processes, an individual scientist conducting his own research, both measures the parameters of interest and uses the resultant data to draw conclusions about his experiment. In such a process the individual involved has at his disposal all of the information contained in the data, especially that concerned with the limitations of the data and the constraints under which they should be used. In this type of situation, few formal qualifications of the recorded data are necessary since those qualifications are implicit in the mind of the scientist.

In larger programs however, the measurement process and the utilization process are quite often compartmentalized such that one group of scientists is responsible for the collection, quality assessment and storage of the measurement data, and a second, usually nonrelated, group of scientists is responsible for the synthesis of all pertinent information into a final set of conclusions. In this type of systems research, the determination of the fundamental quality of the measurement data and transmittance of that quality assessment are the single most important qualifier in the process of going from observation to understanding.

The CHESS program, as designed and implemented by the Environmental Protection Agency, is a classic example of the large systems approach to research. The epidemiological measurements were designed, conducted, and stored by one group of scientists; the

77-590—76——3

aerometric measurements were designed, conducted and stored by a second group of scientists. The desired end product, a correlation of health effects with atmospheric pollution was then derived from these two independent sets of data accumulated in a large data storage network. It is important to reemphasize here that in such a research program it is incumbent upon the measurement personnel to transmit to the data user all of the information contained in the resultant data, especially that relative to accuracy and precision. In order to understand the problems encountered in a large research program such as CHESS, it is necessary to understand the types of measurements that were made.

The assessment of atmospheric pollution exposure received by a defined population can be derived from one of two broad classes of measurement. The first is a measurement that yields an "index" of pollution. The second is a measurement that yields quantitative information about a specific pollutant as it is found in the atmosphere.

A pollution index is a measure of the relative level of pollution which contains little or no information as to the specific chemical or physical properties of that pollution.* These indices can be useful in assessing short-term trends of atmospheric quality in well-defined and limited geographic regions. They cannot be used to deduce information about the source or chemical nature of the material being measured. They also cannot be used to assess long-term trends of pollution burden since gradual changes in pollution sources will distort the quantitative aspect of the index. Most importantly, they cannot be used to correlate atmospheric pollutant levels among diverse geographic areas. Here again, the difference in chemical and physical makeup of the pollutants being measured distort the quantitative aspect of the index.

An example of a measurement that gives a pollution index is the dustfall observations as applied in CHESS. In this method, an open topped cylinder called a dustfall bucket is used to collect any particulate matter that falls out of the atmosphere. This collection is carried out over a long time period, usually one month; and the total dry weight of material collected is used to estimate particulate burden of the atmosphere during that time period. A detailed description of this process is given later in this Chapter. This measurement falls in the index class because all solid material, regardless of its derivation or chemical nature, is included in the final quantitative result.

The second class of pollution measurement is that which contains information both on the specific species of pollutants and on the atmospheric concentrations of those pollutants. In this type or measurement the signal that is measured is derived from a process or property which is specific to the pollutant of interest and which correlates directly with the concentration of that pollutant in the atmosphere. An example of this type of method is the West-Gaeke procedure for the measurement of atmospheric sulfur dioxide. In this procedure, air is bubbled through an absorbing solution at a known rate. The solution is specific for the absorption of SO_2 from the air. After a known duration of sampling, the quantity of SO_2 which was absorbed from the air is quantitatively determined by the formation of a

---

* N.B. This index is not the kind of "air quality index" often used popularly (in radio broadcasts, etc.) to advise citizens of the relative air quality of a city. Such popular air quality indices are usually arrived at by combining measurements of several pollutants.

colored chemical complex of SO_2. If carefully carried out, the procedure gives an accurate value for the SO_2 concentration. The procedure is described in detail later in this Chapter.

Measurements such as the West-Gaeke procedure, which are specific and quantitative, can be used to compare atmospheric pollutant burdens across diverse geographic areas and through long time periods. They can also be used to assess short-term variations in pollutant levels provided that sufficient sensitivity exists in the method to obtain a meaningful signal for the short time period used. In conducting a program such as CHESS, where an attempt is made to relate health effects to pollution burdens, only those measurements that fall in the second class, specific and quantitative, can properly be used to assess the relation between health effects and pollutant burden.

In this chapter, an attempt will be made to evaluate the methodology used to measure aerometric parameters and to assess the validity of the resultant data. The review will encompass procedures used in the field situation, the quality control exercised over the procedures, and the data storage and retrieval network. Conclusions will be drawn as to the adequacy of the measured pollution levels to assess exposures received by specific CHESS population groups.

B. Review of Chemical and Physical Methods

1. THE WEST-GAEKE METHOD FOR THE MEASUREMENT OF AMBIENT SO_2

a. Description of the Method

The West-Gaeke colorimetric procedure for SO_2 determination is the designated Reference Method (Federal Register, *36*, No. 84, 6168, April 30, 1971).* Atmospheric SO_2 is collected by bubbling air through a solution of potassium tetrachloromercurate (TCM). The product of the reaction between SO_2 and TCM is the nonvolatile dichlorosulfitomercurate that is then determined quantitatively by reaction with formaldehyde and pararosaniline hydrochloride, followed by photometric measurement of the resulting intensely colored pararosaniline methyl sulfonic acid.

b. Description of the Field Apparatus and Sample Collection

Outside air is drawn through a sample line at the rate of 200 ml min^{-1}, then through a 6-inch long glass bubbler stem (tip diameter of 0.025 in.) immersed in 35 ml (50 ml after January, 1974) of 0.1 M TCM solution contained in a 32 mm diameter by 164 mm long polypropylene sample container. The exhaust air passed through a glass wool moisture trap, then through a hypodermic needle used as a critical orifice to control the flow, through another moisture trap, and finally through a vacuum pump. A sample consisted of a 24-hour collection. Collected samples were stoppered, and mailed to EPA/RTP for analysis.

c. Validity as a Laboratory Procedure

A collaborative study by McKee et al. (H. C. McKee, R. E. Childers, and O. Saenz, Southwest Research Institute, SWRI Project 21–2811, EPA contract CPA 70–40) indicates that "the method can-

*Alternately see CFR Title 40, Part 50, Appendix A.

not detect a difference smaller than 10 percent between two observations by the same analyst in the range of 0 to 1000 $\mu g\ m^{-3}$. A difference of 20 percent or less may be detected above 300$\mu g\ m^{-3}$, and a difference of less than 50 percent may be detected above 100$\mu g\ m^{-3}$." For analyses conducted by different laboratories on the same sample, "the method cannot detect a difference of less than 20 percent between single-replicate observations of two laboratories in the range of 0 to 1000 $\mu g\ m^{-3}$. At a level of 100 $\mu g\ m^{-3}$, a difference of less than 100 percent is not detectable." The National Primary Ambient Air Quality Standard for SO_2 is: For 24 hour average, 365 $\mu g/m^3$. For annual average, 80 $\mu g/m^3$. Thus if the standard is met, most values will be around or below 80 $\mu g/m^3$, no more than one will be above 365 $\mu g/m^3$.

Regarding the lower limit of detection, the authors cited above propose a value of 25 $\mu g\ m^{-3}$ as a practical figure. "A single determination less than this value is not significantly different from zero" (Instrumentation for Environmental Monitoring, Air-SO_2, Instrumentation, Lawrence Berkeley Laboratories, March 1972).

It is therefore evident that a single analysis is of little use, considering that the expected concentrations of SO_2 will usually be less than the ambient air quality standard of 80 $\mu g\ m^{-3}$. Results should be regarded as valid only in terms of the mean of multiple determinations, and only when the analytical method has been followed rigorously by experienced analysts.

2. TOTAL SUSPENDED PARTICULATES

Total suspended particulates (TSP) were measured using the EPA Reference Method as specified in the Federal Register (*36* (84): 8191–8194, April 30, 1971‡).

Total suspended particulates (TSP) were measured by drawing air through a prewieghed 8 x 10 inch glass fiber filter for a period of 24 hours. The apparatus used for this procedure was the standard High Volume Sampler. At the end of the 24 hour time period, the filter was reweighed, and the TSP computed on the basis of total air flow. The air flow rate was approximately 60 ft^3min^{-1} at the start, and must be not less than 40 ft^3min^{-1} at the end for the measurement to be acceptable. The average air flow rate was computed on the basis of a straight-line interpolation between beginning and ending flow rates.

The National Primary Ambient Air Quality Standard for TSP is: For 24 hour average, 260 $\mu g/m^3$. For annual geometric mean, 75 $\mu g/m^3$.

3. SUSPENDED SULFATE

Suspended sulfate was analyzed, during the CHESS program, using portions of the TSP samples. From the beginning of CHESS to September 1971 the turbidimetric method of analysis was used; then the turbidimetric method was dropped in favor of the methylthymol blue method, which was used throughout the remainder of the CHESS program.

The turbidimetric method consists of the water extraction of soluble sulfates on the TSP filter, the addition of a barium chloride preparation to the extract, and measurement of the resultant turbidity (from

‡ Alternatively see CFR Title 40, Part 50, Appendix B.

the formation of insoluble barium sulfate) with a spectrophotometer or colorimeter. Accuracy of the method is affected by the kind and concentration of other ions present, as well as pH, conductance, temperature, and barium concentration in the test solution.

The methylthymol blue method also utilizes the water extraction of soluble sulfates from the TSP. The filter extract is then passed through an ion-exchange bed to remove interfering ions, and barium chloride is added under slightly acid conditions, forming barium sulfate. Then the test mixture is made alkaline and methylthymol blue is added, which forms a chelate with the excess barium. The uncomplexed methythymol blue is equivalent to the amount of sulfate present, and is measured spectrophotometrically. The methylthymol blue procedure is automated (Technicon Autoanalyzer) in all steps following water extraction of the TSP, and this part of the procedure is reproducible within a range of 2 percent. Error in the determination of sulfate occurs predominantly in the steps preceding the methylthymol blue method.

4. DUSTFALL BUCKET, TAPE SAMPLER, CASCADE IMPACTOR, AND CYCLONE SAMPLER

In addition to TSP measurements using the Hi-Vol sampler, four other means of estimating particulate concentrations were used at various times. They are the dustfall bucket, the tape sampler, the cascade impactor, and the cyclone sampler.

(*a*) The name "dustfall bucket" is adequately descriptive. It is basically an open-topped cylinder, with some protection against wind and rain loss, that is left out in the open, close to the ground or on a rooftop, for a month. At the end of that time the dry matter collected is weighed, and sometimes analyzed for trace metals. The dustfall bucket method is very crude and misses almost completely the very significant part of the aerosol, including the respirable aerosol, that does not settle rapidly. It must be considered here, however, because dustfall measurements were extrapolated to obtain estimates of suspended sulfates and sulfur dioxide in New York City during the period 1949–58 ((Table 5.2.1, CHESS Monograph), and intermittently in Chicago (Table 4.1.A.3), CHESS Monograph). Dustfall measurements were used as the basis for these extrapolations because there was no other basis for such estimates, but it must be remembered that the relationship between suspended sulfates and dustfall is unknown, and that between sulfur dioxide and dustfall is another step removed from reality.

(*b*) Coefficient of Haze (COH) is determined by the automatically operating tape sampler. It is determined by measuring the optical density of an aerosol deposited on a filter tape. The aerosol deposit is obtained by drawing air at a given flow rate through white filter paper tape for a known period of time. If one could assume that the composition and physical characteristics of the aerosol in a given location did not change with time—that only atmospheric loadings would change—then the COH would give a fairly good approximation of the variations of particulate loading and visibility.

However, this assumption is seldom justified, and even at a given location the COH only roughly approximates the true particulate loading. The COH method is worthless, or nearly so, for comparisons between areas with dissimilar aerosols. For example, the aerosols

collected at the Utah sites are primarily the light-colored aluminosilicate dust, whereas the aerosol collected within the inner core of large cities has a predominantly sooty character. For a given particulate loading the Utah aerosol will often have as little as one-tenth the optical density of the urban aerosol.

(*c*) The cascade impactor operates on the principle that particles in an air stream will tend to follow a straight line when the air stream is deflected, and thus can be impacted on a surface in their path. The cascade impactor consists of a series of parallel plates separated by precisely determined spaces. Alternate plates contain a certain number of holes of a size that is decreased as one goes through the series of plates from entrance to exit. Alternating with the plates containing the calibrated holes are plates without holes. These may be coated with a medium for the trapping of impinged particles. Air is drawn through the apparatus at a known rate, and the particles are collected in decreasing size fractions related to the decreasing size of the holes in the plates.

(*d*) The cyclone sampler is a device for the collection of the respirable size fraction of an atmospheric particulate loading. It operates on the principle that the inertia of individual particles will tend to keep the particles moving in a straight line when the air stream in which they are carried is deflected. By this means the larger size particles are removed by impaction and settling, while the respirable particles are carried along with the air stream and are subsequently collected on a filter.

C. Findings and Evaluations of Measurements and Data Reduction

It is important to preface this evaluation of the CHESS air monitoring program with a statement of the following facts. The investigative team looked backward at the program through a window in time with all of the subsequent knowledge built up during that time. More than ten years have passed since the initial planning of the CHESS program and more than six years have passed since the first data were collected. During that time there has been a vast improvement in the understanding of the methods used for pollution monitoring. Many of the procedures used in CHESS have subsequently been found to contain serious errors. These problems were often uncovered as a direct result of research and quality control programs ongoing within EPA. It would thus be unjustified to lay criticism on the principals in the CHESS program for using state of the art measurement technology.

On the other hand, some serious oversights in scientific judgement did occur. In the area of pollutant monitoring, these oversights could have been completely avoided had proper attention been paid to even rudimentary quality control procedures. Throughout the program, much more emphasis was placed on the uninterrupted collection of data than was placed on the systematic evaluation of data quality. The field investigation stage of this review identified numerous problems that resulted in the propagation of unnecessarily large errors in the aerometric data. These unevaluated errors persist even today in the data as it is stored in the CHESS computer system. They could have been avoided or easily discovered and quantified had a well-

designed quality control procedure been applied to the CHESS aerometric monitoring program. This statement is contrary to the statement of the quality control procedures in appendix A of the 1974 CHESS Monograph. Appendix A was not a manual provided to CHESS data gatherers, but was written long after the data in the 1974 Monograph were collected. However, during the field investigation of the CHESS monitoring contractors, it was found that the quality control procedures as described in Appendix A of the CHESS Monograph were routinely disregarded. In fact, for the first two years of the program, virtually no EPA-directed quality control program was implemented at any of the New York, Salt Lake City or Los Angeles CHESS monitoring sites. Problems that were found in this time period were observed and documented by contractor personnel and it was mainly through their personal professional conduct that any of the field problems were corrected. Reasons for this rather gross oversight on proper data management can only be conjecture, but it did appear that inadequate staffing of the monitoring group, coupled with the intense pressure to get the monitoring stations on line and producing data, led to the situation described.

In fairness (regarding the time perspective mentioned earlier) the problem of inadequate quality control on many large EPA programs eventually was recognized internally and in 1974 a Quality Control Branch was established in the Quality Assurance and Environmental Monitoring Laboratory. This branch was given the authority to implement proper quality control procedures on all large atmospheric monitoring programs. Since the formation of this group, there has been a significant and steady improvement in quality assurance as applied to air monitoring methods and data.

In this section, major emphasis will be placed on review and evaluation of the analytical methodology used in the CHESS program to assess population exposures to sulfur oxides and total suspended particulates. Conclusions will be general to all data taken at "official" CHESS monitoring sites, regardless of location. Where local differences in procedures or resultant data did occur, these will be described separately. Health studies, as described in the 1974 CHESS Monograph, that used aerometric data derived from non-CHESS monitoring sites will be reviewed separately.

1. SULFUR DIOXIDE

Atmospheric levels of SO_2 were determined using the EPA Reference Method, better known as the West-Gaeke or Pararosanaline method. The specific details of this method are described in the procedures section of this chapter (Part B.1.). However, a few important aspects of this method will be reiterated. This reference method is basically a laboratory method adapted for field use. It is a "wet chemical" procedure relying on a gas-liquid phase chemical reaction between SO_2 and sodium tetrachloromercurate (TCM). To accomplish this reaction, the SO_2 as a gas phase pollutant, must be quantitatively absorbed into the liquid reactant solution. This is accomplished by bubbling ambient air through the solution at a controlled flow rate, thus, its description as a "bubbler method."

In an attempt to standardize the methodology and to eliminate problems associated with interlaboratory errors, a CHESS policy was instituted whereby all air sampling equipment was assembled and tested at the central EPA research laboratory and then shipped to the contractors for field use. Also, bubbler tubes were prefilled with the appropriate absorber solution, shipped to the contractor for their daily monitoring use, and shipped back to the central laboratory for chemical analysis. It was this long distance shipment of the chemical solutions that led to the first of a series of field-use problems with the procedure. These problem areas will be summarized below with an attempt to evaluate their net effect on the resultant CHESS SO_2 data. Following this summary of individual problem areas, an assessment of the overall SO_2 data quality will be given.

a. Spillage of Reagent During Shipment

The first field data were obtained in New York City and the Salt Lake area (Utah) in November, 1970. By mid-1971, field personnel at the Utah site reported to their CHESS field engineers that severe spillage was occurring during shipment. Many bubbler tubes were arriving partially filled with reagent and some were completely empty. At the Salt Lake area an attempt was made to refill with solution from extra tubes those tubes that were low. However, due to insufficient reagent, this was only partially successful. This problem was not officially recognized until October, 1972, at which time an internal EPA/CHESS memo was written outlining the problem and suggesting corrective action. The magnitude of the problem can be best assessed by quoting from the memo. "The present reagent tubes for SO_2 and NO_2 leak during shipment. . . . The SO_2 leakage rate (was found to be) 18% of the total volume, 50% of the time. . . . It follows therefore, that the resultant pollution data *are unreliable.*" Recommendations were made in this memo as to possible corrective measures. These recommendations were not instituted until March, 1973.

During the subsequent years, many attempts were made to correct this leakage problem. However, none were wholly successful and as late as January 1975, another EPA memo described losses of solution in SO_2 bubblers during shipment and suggesting appropriate corrective action.

The effects of the reagent spillage problem on the SO_2 data can be only grossly estimated. Certainly, many samples were totally lost. These lost samples were not the major problem. Of more significance was the undetermined amount of daily SO_2 data that were in error due to the loss of sample by spillage and yet included in the network system.

If the reagent was partially lost during shipment to the sampling site and used as received, an increased concentration of TCM-SO_2 complex would occur relative to normal sampling. This potential positive bias would be corrected by for the analytical procedure used (Page A–6 CHESS monograph—Analysis Procedure). "At the laboratory, the sample is brought back to its original volume by the addition of distilled water to compensate for water loss during sampling." If however, the reagent spillage occurred after sampling, the required addition of water would result in data that were biased low in proportion to the amount spilled relative to the total volume of solution.

According to the EPA Memo of October, 1972, one half of all SO_2 data taken between November, 1970 and March, 1973 are likely to have been biased low by an average of 17%. This problem was corrected after April, 1973.

b. Time Delay of the Reagent—SO_2 Complex

The Reference Method as originally described in the Federal Register, was to be conducted at 20° C. There was a known error in the method associated with time delay between sampling and analysis which was dependent on temperatures. This error was derived from the spontaneous decomposition over time of the TCM–SO_2 complex as a function of temperature. The magnitude of the error and its exact dependence on temperature was not known but a brief study was conducted to determine its magnitude by scientists of the CHESS monitoring group in November, 1971. As a result of this study, a correction factor of +1.5% per day was arithmetically applied to all CHESS SO_2 data to compensate for the time delay between sampling and analysis.

A more recent and comprehensive study has been carried out within the Quality Control Branch, Environmental Monitoring Laboratory at EPA on the effect of temperature on "The Stability of SO_2 Samples Collected by the Federal Reference Method." This study indicated a much more severe problem than was estimated by the original CHESS study. The evaluation was carried out over the range of 35 to 278 $\mu g/m^3$ SO_2 concentration. The following findings were presented in the report:

> Over a normal range of temperature, the rate of decay of the TMC–SO_2 complex increases five-fold for every 10°C increase in temperature, respectively.
>
> The rate of decay is independent of SO_2 concentration.
>
> At 20, 30, 40, and 50° C the following SO_2 losses were observed: 0.9, 5, 25, and 74% loss per day, respectively.

This study makes abundantly clear a second and even more severe error associated with the SO_2 measurements conducted by CHESS. During the summer months, when the SO_2 absorber solutions were subjected to high and unknown temperatures between field sampling and laboratory analysis, significant degradation of the samples did occur. Estimates of time delay between sampling and analysis range from 7 to 14 days. Estimates of summer temperature exposures range from 25 to 40° C being most severe for the Utah CHESS sites. Thus, CHESS SO_2 data can be estimated to be negatively biased, mainly during the summer months. It would normally be difficult or impossible to estimate the magnitude of the bias except to say that it is probably large. However, simultaneous SO_2 measurements were taken by the New York City Department of Air Resources and by the Utah State Division of Health. These results were obtained by an independent method not susceptable to the temperature related error. A consistent pattern emerged when side by side data are compared. From May to October, the CHESS SO_2 data were low with the largest error occurring in the middle three summer months. The magnitude of the error varied from month to month and year to year, but the CHESS data were consistently low and represented only a portion of the true ambient SO_2 concentration.

c. Concentration Dependence of Sampling Method

The SO_2 reference method was subjected to a collaborative study program in 1973. Four participating laboratories tested the 24-hour version of the Federal Reference Method. A previously unknown source of error was documented that applies to the CHESS SO_2 data. It was found that the 24-hour sampling method does have a concentration dependent bias which becomes significant at the high concentration levels (200$\mu g/m^3$). Observed values tend to be lower than the expected (known) SO_2 concentration levels. This error source will yield a negative bias on the daily CHESS SO_2 data when they exceed 200 $\mu g/m^3$ and on all monthly and yearly average data.

d. Low flow correction

The determination of atmospheric SO_2 concentration was dependent on, among other factors, the accurate measurement of air that passed through the TCM solution. This flow was controlled by a critical flow orifice in the form of a standard hypodermic needle. In practice, the air flow through the sampling system was measured at the start and end of each 24-hour sampling period. This was done to detect low flow due to needle blockage. The Federal Register Method (Reference Method) calls for an air flow of 200 ± 20 ml/min. In field operation, the CHESS procedure substantially broadened these tolerances. Replacement needles were installed if the initial air flow was greater than 220 ml/min which is consistent with the Reference Method; however, needles were not replaced nor were samples voided until the measured flow dropped below 100 ml/min. Integrated flows were calculated by assuming a linear decrease in flow between the start and end of the 24-hour sampling period. If, however, the needle was partially blocked near either the beginning or the end of the sampling period, the linear flow correction would be in error. Using the Reference Method flow tolerance, only small errors would be introduced by this correction (less than 10%). Using the CHESS procedure, however, errors as large as 50% could be introduced and not detected. These errors would be random (either positive or negative) depending on when during the sampling period the needle blockage occurred. Thus a large random error component was added to the SO_2 daily data but this component was somewhat damped statistically in the monthly or yearly averages.

The modification of flow tolerance by the CHESS aerometric group is a procedure that would not have withstood the critical review of a competent quality assurance program.

e. Bubbler train leakage

The West-Gaeke method, as described in the Federal Register, employs a vacuum bubbler train. That is, the sampled air is drawn through the bubbler train by a vacuum pump rather than being pushed through by a positive pressure pump. There are many advantages to the vacuum procedure, most important is that the air does not come in contact with any internal pump mechanism. However, there is a modest pressure differential between the atmosphere and the internal bubbler; thus all fittings and joints must be gas tight. The bubbler train used in the CHESS program had two points where frequent air leak problems were encountered. One was around the rubber stoppers for the bubbler tube and moisture trap and the other was the

rubber tubing used to hold the glass assembly pieces together. Field operators reported consistent problems with leakage in the routine field use of the bubbler train. In a severe leak situation, the samples were voided due to out of tolerance (Low) flow rates. There were many cases however, where small leaks occurred but the final flow was within specifications so the sample was included as valid. In cases where the leaks formed around the rubber stoppers, no significant error would be introduced except due to the linear flow correction as applied to instantaneously developing leaks. This error is similar in nature to that discussed in the flow section. In the case of leaks upstream of the bubbler train, room air instead of outside air is drawn through reagent. In normal situations, it has been observed that room air is significantly less polluted than outside air. (See page 6–6, CHESS Monograph—comparison of school air to outside air). This effect may not be as large for the small buildings used to house CHESS stations, but a somewhat decreased pollutant level would undoubtedly be sampled. The absolute magnitude of this error cannot be adequately assessed but it can be stated that the error would be in a negative direction, that is, again to underestimate SO_2 levels.

2. GENERAL ASSESSMENT OF CHESS SO_2 DATA

The SO_2 data, accumulated at "official" CHESS sites, followed a remarkably uniform trend as the program progressed. The method used was the EPA Reference Method which is specific for the chemical species, SO_2. Thus, regional changes in pollutant mix, i.e., the proportion of other pollutant species relative to SO_2, had minimal effect on the SO_2 data. However, the sum effect of the errors detailed in this section did have a profound effect on both the accuracy and the precision of the data.

Under normal circumstances, a retrospective evaluation of a monitoring effort that occurred a number of years in the past and which had been terminated, could yield only the broadest of estimates of data quality. Fortunately for this review, two geographically different locations with six different monitoring sites were involved in the collection of simultaneous SO_2 data. Further, the groups responsible for the two data sets were managed independently and the methodology used was also independent. This fortunate circumstance enabled the reviewers to acquire a quantitative understanding of absolute differences among data sets as well as correlations with respect to time.

The locations where side by side data existed were the New York City sites at Bronx and Queens and the Salt Lake Basin sites at Ogden, Salt Lake City, Kearns, and Magna. In these locations, the local environmental monitoring agencies had sites located within 50 meters of the CHESS sites and at similar elevations. At these sites, the local agencies collected daily SO_2 and TSP data for the entire life of the CHESS program. The SO_2 methodologies used by both State agencies were variations of the peroxide bubbler method in which twenty-four 1-hour samples were integrated to form a single 24-hour SO_2 measurement. In New York the samples were measured acidimetricly and in Salt Lake City they were quantified conductiometrically. Neither method is as specific for SO_2 as is the Reference Method, that is,

pollutants that are in a significant concentration, relative to SO_2 and that also oxidize to form an acidic compound will be interpreted as SO_2. For this reason, when the NYC Department of Air Resources initially brought to the attention of the CHESS Aerometric team the large discrepancy between their respective data, the discrepancy was dismissed as method bias on the part of the New York method. An EPA memo dated November 3, 1971 described a limited study into the Reference Method. The conclusion reached was "On the basis of (this study) . . . I feel there is no sound basis for discrediting the EES (Environmental Exposure System) methodology."

No further attempt was made to uncover the cause of the discrepancy in SO_2 data. Had the CHESS EES team obtained and compared the Salt Lake Basin data, especially that from Magna site, a disturbing similarity would have been immediately apparent. This data confirmed in detail the discrepancies observed in New York. It is important that the Magna site data were confirmatory since it was in a region of single source pollution, that from the nearby copper smelter. In this site very low levels of other pollutants existed relative to SO_2, thus the peroxide method was capable of giving reasonably reliable estimates of the SO_2 concentration. Of equal importance the general pollutant mix was very different between this rural smelter site and the urban area of New York City. Despite these differences the comparison of side by side Federal-State data indicate the same discrepancies in both trends and absolute concentrations. The following conclusions as to SO_2 data validity can thus be reasonably drawn from the review of methodological errors and the comparison of existing side by side data.

From November 1970 until December 1971 the SO_2 data generated from CHESS sites using the modified Reference method were biased low by 50 to 100 percent in the High Exposure sites when compared with existing State SO_2 data. Thus, the 1971 annual average SO_2 exposure estimates of 60$\mu g/m^3$ as reported for Magna in the CHESS monograph (page 2–24) are more likely in the vicinity of 100 $\mu g/m^3$. Also, the same phenomenon occurred in New York and the reported values are also in similar error.

A confirming fact is that during cool months *after* 1971 SO_2 data correlated well both in trends and absolute concentrations between State and Federal analyses. It thus seems likely that the State data were reasonably accurate throughout that time period. However, one consideration must be applied here: namely, *that due to the difference between the independent methods an error bar of at least one hundred percent must be applied to the data and explicitly correct data cannot be drawn from these observations.* In other words, where two or more independent observations are in disagreement by a significant amount it cannot be said by inference alone that one data set is more correct than the other. It is reasonable to assume, however, from our review of all State and Federal data in the time period of 1970 through 1971, that the Federal SO_2 data as collected in the CHESS program were substantially low and went through an abrupt upward transition in concentration in December 1971 at all CHESS sites and Federal data taken before that time may reasonably be expected to have a large, unknown negative bias.

In November 1971, the CHESS monthly mean SO_2 data underwent an abrupt change in the positive direction. The cause of this change is not apparent. However, the result was profound. From that time until the conclusion of the CHESS program in July of 1975, the fall-winter data were in very good agreement with other existing data and very likely gave reliable estimates of SO_2 exposures.

Throughout the entire program, the CHESS SO_2 data had an associated negative bias during the summer months, becoming most severe during the hottest periods of July and August. This error usually reached a maximum of 60 to 80 percent underestimation of exposures and was variable. As a result, even though wintertime monthly SO_2 averages appear valid from 1972–1975, annual averages of the same data are biased low due to the inclusion of the summer errors. The best estimate of error in the annual average data 1972–1975 is approximately minus 15–20 percent relative.

The individual daily SO_2 levels, when compared to city or State data or to replicate CHESS measurements taken after 1973 had so large a random error component that they are not useful to assess daily SO_2 exposure (as attempted in the asthma panels). The random errors associated with the daily values were much larger than the differences observed over time.

Due to inherent methodological errors, the following may be considered as minimum differences between High and Low SO_2 exposures which may be considered "real." These are based on EPA's collaborative study of the reference method and used a 95 percent confidence interval.

Below 100 $\mu g/m^3$ SO_2, a difference of at least 50 $\mu g/m^3$ is necessary to be statistically significant.

Between 100 and 300 $\mu g/m^3$ SO_2, a difference of at least 60 $\mu g/m^3$ is necessary to be significant.

Below 25 $\mu g/m^3$, a single determination is not significantly different from zero.

3. TOTAL SUSPENDED PARTICULATE

The Reference Method for the determination of total suspended particulate matter (TSP) is probably the simplest and most reliable method used by CHESS. It has been well studied and most error sources are known. However, it is a method that measures an arbitrary and poorly defined portion of the total atmospheric particulate burden and the portion measured has unknown relevance to the human respirable portion. The size fraction measured is somewhat dependent on the design of the shelter used for Hi-Volume sampler. The design and dimensions of the Reference Method shelter are specified in the Federal Register, thus the portion of TSP that is collected by the method is generally uniform. Best estimates of particle size range included in the Reference Method are from 0.05 to 60 μm diameter. Above 60 μm diameter, the particle fall velocity is too great to navigate the bend around the roof of the shelter. Below 0.05 μm the collection efficiency of the glass fiber filter used in the method diminishes.

A collaborative study was conducted on the Reference Method using 12 different groups sampling ambient air at a common location. The results of this study indicate the method is capable of reproducible

measurements with less than 5 percent error at the 95 percent confidence level. Also, the minimum detectable amount of TSP is approximately 2 $\mu g/m^3$ for a 24-hour sampling period. This sensitivity is more than sufficient for most 24-hour TSP measurements.

The TSP measurement method, as used in CHESS, had one notable difference from the laboratory procedure which was collaboratively studied. The weighing procedure to determine TSP was performed at EPA/RTP laboratory not by the CHESS contractors on site. This necessitated the shipment of individual filter samples through the mail and the subsequent storages of the samples at EPA. During laboratory reorganizations at RTP, periods as long as 6 months elapsed between actual field sampling and laboratory analysis.

The following is a summary of individual errors and an assessment of overall TSP data quality.

Loss of particulate matter before weighing

In the TSP methodology there were field-related procedures that resulted in partial loss of particulate matter from the Hi-Volume filter samples. Due to the exposed location of the Hi-Vol TSP samplers, wind and cold sometimes made it very difficult to remove the filter paper from the apparatus without losing part of the sample. No estimate has been made of loss due to this problem; it would, of course bias the reported results only in the direction of lower-than-actual atmospheric loadings. This was not a constant problem among CHESS sites. It was noted by field operators as being a particularly severe problem in the Salt Lake City area during the winter months.

Two other error sources have been identified in the determination of TSP, both of which would also produce a low-side bias: (1) the shaking-off of particles from the filter during transit from the field site to EPA/RTP, and (2) the evaporation of organic substances. In an attempt to quantify the mass loss during transit, David Hinton, EPA/RTP, made a comparison of filters collected in Utah, before and after mailing from Salt Lake City to RTP (22). He found that there was a average 4 percent loss. Carl Broadhead, of the Utah Division of Health, conducted a similar comparison; however, he noted an apparent loss of approximately 25%. This difference may, in part, be due to the time of year the studies were conducted, During the dry summer months in the Salt Lake City area, much of the TSP loading is due to windborn crustal material (sand). This material is much more easily lost in sample handling that is the finer anthropogenic particulate material.

A final error source, one more difficult to assess, derives from wind velocity versus collection efficiency. On days with relatively high wind (>15 mph), the Hi-Vol sampler is more susceptible to the inclusion of large diameter particulate material. To compound this problem, the design of the shelter makes the magnitude of the error dependent on the wind direction relative to the orientation of the shelter. The main result of this problem is that two side by side Hi-Vol samplers, oriented 90 degrees relative to each other, will produce dissimilar measurements with the discrepancy increasing as the daily wind velocity increases.

The overall effect of the summed errors with the Hi-Vol TSP measurement is a slight negative bias. This bias may be as small as 10% or may be as large as 30%. Side by side data from New York

and Salt Lake indicate that this assessment is reasonable. These data also indicate that the TSP data were by far the best quality data taken in the CHESS monitoring program. Differences measured between High and Low sites are probably reasonable estimates of the differences of TSP exposures as received by populations within these areas. Some local source variations undoubtedly did occur, but average annual exposures were reasonable.

In any overall assessment of the CHESS TSP data it should be noted that all of the sources of errors mentioned previously related almost exclusively to the loss of large particulate matter and most likely that matter is associated with crustal weathering. This material is outside of the normal human respirable size fraction and by composition, it would be unlikely to be associated with aggravated health. Thus, loss of that portion of the total material may not have diminished the quality of data for health effects studies. It may in fact have rendered that data a closer estimate of the respirable TSP exposure to which the CHESS population groups were subjected.

It has been suggested by some environmental scientists that whenever Hi-Vol measurements are made for health related studies, the filter pads should be "shaken out" much like a housewife does when shaking crumbs from a used tablecloth. The resultant TSP exposure estimates derived from such a procedure would then more closely relate to the human respirable size fraction of the total atmospheric particulate burden. Although never actually implemented, this suggestion indicates the general level of dissatisfaction with the TSP Hi-Vol measurement method.

4. TOTAL SUSPENDED SULFATE

The determination of atmospheric sulfate concentrations, as carried out in the CHESS program, was a methodological extension of the Hi-Vol TSP method. Thus, all errors associated with the TSP method also affect the sulfate method. Subsamples were cut from the exposed Hi-Vol filters and were analyzed for total water soluble sulfate. Methods available for sulfate analysis at the time of CHESS determined all water-soluble sulfates as a class rather than distinguishing them by chemical species. Two different methods were available for total sulfate and both were used in CHESS. From November 1970 until September 1971, the manual turbidimetric method was employed. From September 1971 until July 1975, the methylthymol blue (MTB) method was used. The methods are somewhat similar and are described in detail above.

The turbidimetric method is subject to interferences, many of them being other common pollutants. In areas like the Salt Lake Basin where the pollutants are dominated by a single source, the procedure may be adequate. However, in urban areas like Cincinnati or New York City, where the pollutant mix is derived from many independent sources and is variable even within the city, the method is capable of only the crudest estimates of sulfate levels. It should not be thought of as an accurate measurement of atmospheric sulfate. Especially, small differences between High and Low exposure communities, such as were reported in the Cincinnati Study in the CHESS Monograph (page 6–5) cannot be identified as real differences. When a realistic error estimate is applied to the reported sulfate concentrations, the differ-

ence becomes statistically insignificant. Any correlation of CHESS health effects with sulfate levels where the sulfate data were obtained using the turbidimetric method must be carefully qualified.

The MTB method is basically a better measurement method because most of the aerometric interferences have been eliminated by its revised methodology. The two remaining interferents, phosphate and barium, are not normally found in atmospheric concentrations high enough to cause inordinate problems. However, problems associated with the sampling aspect of the method have been documented and do impact on the general CHESS sulfate data quality.

First, problems associated with sulfate blanks (the level of sulfate on the filter pad as manufactured) were reported to be high and variable. In the 1971–1973 time period, problems of variable blanks within the EPA NASN program were documented. The general blank level was equivalent to an atmospheric sulfate concentration of 1–2$\mu g/m^3$. However, the major problem was variability of the blank among manufactured lots of the filters. The blank level often varied by more than 100 percent among lots so that routine and continuous blank assessment should have been mandatory.

No evidence of routine sulfate blank determination was found in the CHESS monitoring program until 1974. From that time period on, adequate blank assessment and correction were applied to the data. From 1971 until 1974 however, the blank contribution to the CHESS sulfate data was not adequately assessed and consequently a positive and highly variable bias of unknown magnitude was included in the data.

Second, adsorption of atmospheric SO_2 onto the fiberglass filter material followed by spontaneous oxidation of the SO_2 to sulfate had been well documented. A 1966 publication by R. E. Lee and J. Wagman provided results of their investigation of the problem. The conversion was clearly documented with severe effects demonstrated on four-hour samples. The conversion did appear to be an active-site catalytic conversion that decreased in magnitude after an initial saturation of sites. Thus, 24-hour samples were much less affected by this problem than were those taken for shorter time intervals. Even so, the paper by Lee and Wagman, presented data in which routinely 0.5 to 1 $\mu g/m^3$ of the measured sulfate was derived from SO_2 conversion products. The maximum conversion presented was 2.1$\mu g/m^3$ derived from SO_2; this constituted a 10 percent positive bias of the sulfate data. A more realistic average bias is likely in the 5 percent range. However, there is clear evidence that in regions of high levels of SO_2, relative to sulfate, the positive measurement bias becomes much more severe. This is probably the case in the Salt Lake Basin area.

The third and most devastating problem associated with the CHESS sulfate data occurred when the laboratory analysis of sulfates was contracted to an outside firm. During this time period (October 1972–June 1974) the reported sulfate data underwent a sudden and sustained decrease in apparent atmospheric sulfate level. Upon investigation it was determined that the laboratory analysis of all sulfate data from all CHESS sites were biased low by approximately 50 percent. The reason for this negative bias was and still is not completely clear, but the continued dissemination of poor data was clearly due to inadequate quality controls. An interim EPA report

on a retrospective quality assurance evaluation of CHESS Sulfate Data states:

> A quality control protocol was designed for CHESS chemical analysis but has not been implemented as per the contract . . . . The quality control protocol should be implemented immediately.

In a series of following studies the magnitude of the affected data and of the error were documented and an attempt was made to correct and therefore recover the data. This type of procedure is difficult at best and impossible in most cases. The validity of this data correction was again assessed by the EPA Quality Assurance Branch. Their finding was:

> The basic question . . . is—How does one make bad data good? Whatever is tried will be attacked for a multitude of (justifiable) reasons. Using the existing data set for relative pollution level assessment will be acceptable, but statements concerning absolute levels will not be. It would not be wise to submit these data to the NADB,[1] but rather answer all requests for these data internally.

Their statement gives a reasonable assessment of the CHESS sulfate data between 1972 and 1974. The assessment of other year CHESS sulfate data is more difficult. No comparative sulfate data exists from the local agencies as it did for SO_2 and TSP. Based on the intrinsic capabilities of the methods, and the error assessment of the field use procedures, it can generally be stated that:

1. From 1970 to September 1971 the sulfate data were obtained using the turbidimetric method. It should be used only as a sulfate level indicator. Due to interferences, there will be severe problems if an attempt is made to correlate sulfate levels in one part of the country with sulfate levels in another.
2. From October 1971 until October 1972, the data are subject to the following considerations:
   a. The data are likely biased in the positive direction from 1–2 $\mu g/m^3$. This bias may be more severe in areas of high SO_2 concentration relative to sulfate.
   b. The random error component of the measurement is probably in the order of $\pm 25\%$ at an atmospheric concentration of 10 $\mu g/m^3$.
3. From October 1972 until June 1974, all CHESS sulfate data were biased negatively by approximately 50% on an annual average basis due to improper laboratory analysis by the contractor. These data should be used only on an adjusted annual average basis to establish local trends within site locations. The unknown cause of the bias prohibits use of the data in shorter time structure (i.e., day, week, month) increments.
4. From July 1974 until July 1975, CHESS sulfate data underwent a marked improvement and was somewhat better than that collected in the 1971–1972 era. The positive bias of the data is probably similar to that of the earlier period but the random error component was improved due to improved sulfate blanks on the TSP filters.

D. The CHAMP Air Monitoring Program

1. INTRODUCTION

Early in the execution of the CHESS program in 1969, a number of staff members in the air quality measurements organization of EPA

[1] National Aerometric Data Bank.

decided it was desirable, indeed imperative, to improve the efficiency and accuracy of short-term air quality data monitoring coverage. EPA coined the term CHAMP (Community Health Air Monitoring Program) for this concept of a second generation automatic system of air monitoring stations. Seven prototype stations were operated in California from January, 1972 to February 1974. The manpower ceiling placed on EPA resulted in a decision to contract for the development, installation, and operation of the CHAMP system. A contract for the development of the CHAMP system was awarded in February, 1973. The developmental monitoring system was to contain the newest technology in monitoring instrumentation. Accurate measurement of all critical air and liquid flows in the system was incorporated to enhance the accuracy of the system. The development continued to mid 1974 when the first station systems were installed in the Los Angeles area for field evaluation.

2. SYSTEM DESCRIPTION

The CHAMP air quality measurement system assembles the available discrete pollutant measurement devices and associated meteorological instruments into a complete system in an air-conditioned portable building. EPA specified the pollutants to be measured and selected the instruments with the advice of the CHAMP contractor. All data are recorded digitally in a mini-computer integral to each system. The data are checked and stored on tape at each CHAMP site for transmittal to the EPA/RTP Laboratory at Durham, North Carolina. SO_2 and NO_2, and TSP measurements are also taken periodically using older CHESS-type bubblers and Hi-Vol sampler instruments described previously for backup and validation of the CHAMP instruments. These bubbler and filter samples are sent to the contractor's chemical laboratory in California for analysis.

All the CHAMP systems measure ozone, total gaseous sulphur NO/NO_2, TSP,/RSP combinations, temperature, wind direction and velocity, and humidity. Selected systems also incorporate CO and hydrocarbon sampling. The CHAMP system while automatic in principle, requires periodic calibration and servicing by an operator to maintain a high duty factor and an acceptable quality of data (less than 15% error band). The operator repairs and adjusts instruments as required, checks for failures, and does periodic calibrations and data verifications. A quality assurance specialist continually spot-monitors the CHAMP sites carrying-out calibration and quality checks.

It should be noted that the instrumentation of the CHAMP stations is not completely uniform. Some stations do not have wind and pressure instruments; not all have CO and hydrocarbon instruments.

The manner in which meteorological data from the CHAMP stations is being analyzed and used has not been investigated. This is a subject of interest depending on the future of the CHAMP program.

CHAMP stations were visited in Thousand Oaks, California, and Salt Lake City, Magna, and Kearns, Utah. The kind of meteorological instruments in use appeared to be appropriate and they appear to be well-located and properly maintained. Problems have occurred with new dew-point measuring equipment that is now being replaced (this has to do with humidity measurement. Except for occasional failures of the sensing element of the dew-point apparatus, collecting meteorological data from the CHAMP stations should be routine.

There are at present 18 CHAMP stations on line at locations selected by EPA; six in the Los Angeles Basin, three in Birmingham, Alabama, four in New York City, four in the Salt Lake Valley, and one at the EPA Health Effects Research Laboratory at Research Triangle Park, North Carolina.

3. FINDINGS REGARDING THE CHAMP PROGRAM

As in the CHESS program, all the instruments incorporated in the CHAMP station were developed by the manufacturer for laboratory use. In fact, some non-commercial instruments were selected by EPA to try to use the most advanced technology. The CHESS experience has demonstrated the need for validation in field use and the contractor appears to be attempting to do this.

There was apparently some attempt to standardize on one instrument manufacturer for ease of maintenance, etc. Bendix ozone and NO_x instruments were employed. Flame photometric measurement was selected for SO_2 EPA apparently was interested in a pulsed flourescence device but the equipment cost was too high for the budget. The present instrument actually measures total gaseous sulphur and it is assumed that this is SO_2. (The only other likely gaseous sulfur compound H_2S, does not seem to be widely present.) The rest of the measurements appear to be well-validated. The backup measurement with bubbler methods have validated NO_2, to the extent possible. The TSP/Hi-Vol measurements were apparently validated at the beginning of the CHAMP program. However, because of the non-linear calibration character of the flame photometric instrument in the low concentration ranges of interest from 0 to 50 $\mu g/M^3$, calibration and range setting by the operator still results in 5% to 15% range of uncertainty in the total sulphur readings. Further, while the West-Gaeke bubblers used to check CHAMP SO_2 are stored at 70° F at the sites, they are shipped to the contractor's facilities for analysis without temperature control and are subject to the unpredictable temperature dependent decay of solutions prior to analysis. Thus, the SO_2 validation in the CHAMP system may be in greater error than EPA expects.

The execution of the CHAMP program has yielded validation and quality control of field measurements better than CHESS. However, there are clearly numerous unresolved problems with the operation which have led to delays in validating the data bank and which require high level attention for resolution before reliable quantitative aerometric data can be obtained.

The data processing was 2,900 data-days behind at the time of this investigation and no date agreed on for total backlog elimination. Drift of zero setting and data span of instruments have invalidated part of the earlier analyses. The data are only about 60 percent machine validated. Field operator problems have arisen possibly due, in part, to a lack of standardized operating procedures. Successful operation of the CHAMP system requires well-trained instrument technicians, and people of this high level of skill have not been employed in the past. Because of such circumstances, the SO_2 data obtained through 1975 have been lost and apparently are not recoverable.

Some months ago EPA found that significant data were lost in transmitting over leased lines to the RTP laboratory. Thus, the primary data source is the data tapes from the CHAMP site computer which are mailed to RTP.

The CHAMP contract is up for renewal in November 1976 and the bids are being solicited competitively. It is believed that at this time competitive bidding could be a destabilizing step in this program and could delay the achievement of reliable routine data gathering another year. On the other hand there are obvious advantages to open competitive bidding. When system development is more nearly complete, it would certainly be appropriate for competitive bidding to be adopted. The competition should include quality control considerations. Unfortunately, the EPA quality assurance group was not consulted on the renewal request for proposal, although that group did participate in evaluating proposals received.

4. SUMMARY

CHAMP appears to be an improvement in real time field measurement of air pollutants in comparison with CHESS. However, the system is still not completely validated and may not be ready for routine use for 6 to 12 months. Data should not be stored in an accessible data bank until it is validated.

The present best estimate of expected accuracy is ± 15 to 20% on the CHAMP measurements. However, this will be a significant improvement over previous CHESS aerometric network mesurement systems when and if it is realized.

V. REVIEW OF CHESS AIR QUALITY ANALYSIS PROCEDURES AND RESULTS

A. Introduction

This chapter presents the results of the investigative team's critical review of the utilization of aerometric data in the analysis and data modeling presented in the CHESS Monograph. The citations to pages, figures and paragraph numbers are to the 1974 CHESS Monograph. The findings are highlighted in terms of examples wherein it appears that estimates have been extended beyond the range of credibility, models have been misused, or miscellaneous errors of various types have occurred which lead to misinterpretation or over-interpretation of data or results of analyses.

B. Use of Estimated Data

A serious weakness in the CHESS study was acknowledged in the last paragraph on page 7–9, which refers to the Salt Lake Basin study and the Rocky Mountain study. It is in part:

> Several factors should be remembered when interpreting the results of the lower respiratory disease studies . . . a majority of the pollution exposure data in both studies were estimated from emissions data.

This statement applies to one of the most important and controversial paragraphs in the CHESS report, also on page 7–9, which follows:

> It is interesting to note that larger increases in total lower respiratory disease and two of its components were observed in the High pollution community of the Salt Lake Basin study than in the corresponding communities in the Rocky Mountain study. Also, the mean annual suspended sulfate concentration was higher in the High pollution community in the Salt Lake Basin study than in the Rocky Mountain study; the opposite was true for sulfur dioxide. This suggests that increases in lower respiratory disease frequency are probably associated with suspended sulfates rather than sulfur dioxide.

The paragraph summarizes the argument that exposure to suspended sulfates over a period of years produces significant adverse health effects.

Analysis of the background material leading to the conclusion shows that it is derived from an interpretation of the relationship of four numbers all of which are estimated values. The sulfur dioxide values are estimated from smelter emissions and the sulfate values are estimated from estimates of sulfur dioxide in one case and estimates of suspended particulate based on smelter emissions in the other, assuming no difference in the ratio of sulfate to suspended particulate in the communities, Kellogg, Idaho; Helena-East Helena and Anaconda, Montana; and Magna, Utah.

The "High pollution community of the Salt Lake Basin" is Magna, Utah. It is less clear what is meant by the words "than in the Rocky

Mountain study". However, this paragraph refers to the preceding paragraph of the CHESS report, which speaks of concentrations, "as low as 7.2 $\mu g/m^3$ in the Rocky Mountain Study".

From this it can be concluded that reference is being made to concentrations of sulfates in Anaconda, Montana.

A comparison is being made, therefore, between average sulfur dioxide concentrations and average sulfate concentrations in Magna and Anaconda. The period of the records being compared covers the years 1968–1970.

From the preceding paragraph the values being compared may be obtained. They are as follows:

[The concentration values are given in micrograms per cubic meter, written as $\mu g/m^3$]

	Sulfur dioxide	Sulfates
Magna	92	15.0
Anaconda	177	7.2

Because of the methods used for making estimates, the absolute values of these concentrations are questionable. The next four sections discuss these estimates.

1. ESTIMATED SULFUR DIOXIDE CONCENTRATION, 92 $\mu G/M^3$ (MAGNA)

The concentration value 92 $\mu g/m^3$ for Magna can be obtained from Table 2.1.A.14 or Table 2.1.A.16. It is based on the following *estimated* values for three years:

Year:	$\mu g/m^3$
1970	84
1969	103
1968	90
Average	92

These estimates of annual sulfur dioxide exposures were derived by multiplying the yearly smelter emission for sulfur dioxide by the ratio of the 1971 measured annual average sulfur dioxide concentration (61.8 $\mu g/m^3$) to the same year's sulfur dioxide emission rate (193 tons/day). The last chapter established that these data could be off by 100 percent, probably on the low side.

$$61.8/193 = .320 \ (\mu g/m^3)/(\text{tons/day})$$

The emission rates used were as follows (page 2–37):[1]

Year:	*Tons/day (SO_2)*
1970	261
1969	322
1968	281

In order to obtain the estimated sulfur dioxide concentrations, it must be first assumed that the meteorological conditions for each of the years 1968, 1969 and 1970, were identical to those conditions in

[1] These rates of emission are off by a factor of two. Tons of sulfur, not tons of sulfur dioxide, are listed. These values corrected should be 522, 644 and 562 tons/day. However, this does not change the estimates of sulfur dioxide concentrations, which depend on a ratio between measured 1971 concentrations and 1971 emissions, whatever they might be. Doubling the emission rate also doubles concentations estimated by the application of a mathematical diffusion model (Page 2–23).

1971. There was no presentation in the Monograph of the use of climatological data to show that 1971 was similar to the other years, an average year, or a generally representative year. Even if the meteorological conditions for all four years had been identical, there is still a problem because the year 1971, on which the estimates are based is not a normal year for smelter operations. Emissions were zero, or practically zero for two weeks during July, and nearly zero for six weeks in July and August. Therefore, the emission/concentration ratio is deficient in showing the effects of the summer season, when wind direction frequencies from the smelter to Magna might have been less than during the remainder of the year. This suggests that the average concentration of sulfur dioxide in Magna is likely to have been slightly over-estimated, but it supports rather than changes the conclusion that average concentrations of sulfur dioxide are less in Magna than in Anaconda. Primarily this estimate is criticized because it is not supported by climatological information.

Also it should be realized that the method used for estimating the annual average concentration can result in an incorrect estimate if there is a significant background of sulfur dioxide from a source or sources other than the smelter. Multiplying the emission rate of the smelter by a factor assumes that all individual observational values that make up the annual average can be multiplied by this same factor, when actually only those values totally resulting from the smelter emissions would be effected. The Salt Lake City airport wind rose (Figure 2.1.2) is probably not representative for estimating the percentage of time that Magna is downwind from the smelter because the smelter stack is at the base of the Oquirrh mountain range. However, the frequency of west northwest and northwest winds at the airport suggest that Magna is only downwind about 5% of the time. Allowing for the effect of calm and variable winds, it seems unlikely that Magna would be under the influence of the smelter more than 10% of the time. It follows then, sulfur dioxide values for only these hours would be affected. On the other hand, if the smelter is the only significant source of sulfur dioxide, as may be the case, then multiplying individual observation values of zero concentration would yield only zero, and the procedure for estimating yields a true result, assuming no change in meteorological or emission conditions. Since the sulfur dioxide background in Magna is not known, the error that could be produced by background concentrations cannot be determined. Probably most of the sulfur dioxide does come from the smelter, so this source of error is not significant.

2. ESTIMATED SULFUR DIOXIDE CONCENTRATION, 177 $\mu g/m^3$ (ANACONDA)

A paragraph in the right hand column of page 3–12 explains how the average concentration of 177 $\mu g/m^3$ for sulfur dioxide was estimated for Anaconda for the period 1968–70 using sulfation plate data and emission rates. However, the explanation is incomplete, because it requires the 1971 emission rate of the smelter, which has been omitted from the Monograph. Thus, the validity of the entire procedure is impossible to verify. Table 3.1.2., which lists the emission rates by year begins with the year 1970. The ratio of $0.343 \pm .253$ $(\mu g/m^3)/(ton/day)$ was obtained by a very dubious procedure. To begin with, sulfation plate data are of somewhat uncertain nature. The document "Air

Quality Criteria for Sulfur Oxides", U.S. Department of Health, Education and Welfare, Public Health Service, National Air Pollution Control Administration, Washington, D.C., January, 1969, pp 24–25 says that sulfation "candles" (and plates) give only "an empirical estimate of the average concentration". It also says "results are influenced by wind movement and humidity" and that "the lead peroxide candle provides intelligence on the oxidizable sulfur compounds in the atmosphere which seldom can be directly related to sulfur dioxide".

The CHESS Monograph paragraphs refer to sulfation plate data for 1965. The sulfation plate is a variation of the lead peroxide candle. Developmental work on the plate was reported in the following reference: Huey, N.A. "The Lead Peroxide Estimation of Sulfur Dioxide Pollution" J. Air Pollution Control Association, Vol. 18, pp 610–611, Sept. 1968. Consequently it is unlikely that sulfation plates were in use in Anaconda in 1965.*

In order to determine sulfur dioxide from a lead peroxide candle or plate an empirical relationship must be used. For example, in the Helena Valley, Montana, Area Environmental Study, (EPA, Office of Air Programs, Research Triangle Park, North Carolina, January 1972) the sulfation values were converted to sulfur dioxide values by means of the relationship: 1 mg SO_3 per 100 cm^2 per day is equivalent to 0.035 ppm SO_2. In the history of the use of lead peroxide devices, there has not been general agreement as to what ratio should be used, and a belief prevails that sulfation candle or plate data are conservative, i.e., that sulfur dioxide concentrations are sometimes higher than indicated. Further, more information is needed concerning the location of the station, or stations, in the Anaconda area, where the sulfation data were obtained. In order to validate the Anaconda sulfur dioxide data further work needs to be done.

In 1965 the annual average concentration of sulfur dioxide was reported to be 80 $\mu g/m^3$ with an emission rate of 609 tons/day. Since the 1971 emission rate is omitted from the report it cannot be compared with the corresponding concentration of 286 $\mu g/m^3$. Assuming that the 1971 emission rate is also on the order of 600–700 tons/day, then there seems to be too great a difference between the 80 $\mu g/m^3$ concentration and the 286 $\mu g/m^3$ concentration. (Center paragraph, right hand side, page 3–12.)‡

The ratio $0.343 \pm .253$ has a large error factor. The range is from .090 to .597. If the low value is multiplied by the emissions for the years 1968–1970, the following concentrations are obtained:

[Tons per day]

Year	Table 3.1.2 (SO_2)	New value Montana SDES
1971	Omitted	636
1970	635	856
1969	545	824
1968	367	662
1967	346	459

NOTE.—The omission of the 1971 emission rates makes it impossible to check the effect of using the new value for 1971 on the estimated emission rates.

*The chemical reaction for "candles" and "plates" is the same.

‡According to information recently received from the Montana State Department of Health and Environmental Sciences, the emission rates listed for the Anaconda smelter are low.

Year	SO_2 emissions (tons per day)	Estimated average concentrations ($\mu g/m^3$)
1970	635	57
1969	545	49
1968	367	33

The average of these values is 46 $\mu g/m^3$. This concentration is considerably less than the 92$\mu g/m^3$ value at Magna. However, information received from a representative of the Montana State Department of Health and Environmental Services, suggests that the 80 $\mu g/m^3$ value and the 286 $\mu g/m^3$ value were measured in two different locations in Anaconda, and that the 80 $\mu g/m^3$ value is too low. This indicates that the estimated values of sulfur dioxide in the table comparing Anaconda and Magna values are somewhat too low.

The estimates are further weakened by the fact that an assumption is made that meteorological conditions during all of the year is identical for all years. No supporting climatological information is presented.

Also, note that the Table 3.1.7 lists a sulfur dioxide concentration of 177 $\mu g/m^3$ for 1971 instead of the 286 $\mu g/m^3$ value obtained from the Montana State Department of Health.

The procedure for estimating sulfur dioxide concentrations in Anaconda seems unnecessarily crude, making the average concentration value for the years 1968–1970 uncertain. However, since the reported 1971 values for Anaconda and Magna are 286 $\mu g/m^3$ and 61.8 $\mu g/m^3$, and these values are the basis for estimates, it would appear that it was fairly certain that there was more sulfur dioxide present in Anaconda, than at Magna during the '68 to '70 period of the CHESS studies.

3. ESTIMATED SUSPENDED SULFATE CONCENTRATION, 15 $\mu g/m^3$ (MAGNA)

The 15 $\mu g/m^3$ estimate is a *double estimate* since the sulfur dioxide concentration data on which it is based is also *estimated.* The sulfate value seems to be an average for the years 1968–1970. It is obtained by using the following regression equation, which is found on page 2–39.

$$\text{Magna}-SS=0.09(SO_2)+6.66$$

This equation is based on 1971 conditions.

It is of interest to note that with a zero concentration of sulfur dioxide there would still be 6.66 $\mu g/m^3$ of sulfate, or approximately half the average annual value reported on 1971, which was 12.4 $\mu g/m^3$ Further, 44% of the 15 $\mu g/m^3$ of interest for the years 1968–1970 is unrelated to sulfur dioxide concentrations. The Figures 2.4.2 and 2.4.4 suggest some lack of complete correlation between sulfur dioxide and sulfate concentrations.

During the strike with zero sulfur dioxide concentrations, there still is an appreciable amount of suspended sulfate. Also, a peak value of sulfate occurred during the third week that does not correspond with sulfur dioxide value behavior during the same period. Similarly, the very large rise in sulfur dioxide that peaked in the ninth week hardly shows in the sulfate values. Consequently, the regression equation can be questioned because the reason for the

sulfate values is not understood. What is the physical source of the sulfates?

Since the sulfur dioxide concentrations used in the regression equation are themselves estimated, uncertainties in the sulfur dioxide estimates are compounded in the sulfate estimates. Further, since the source of a considerable amount of the sulfate seems to be not associated with the sulfur dioxide, it is not clear what effect the strike period has on the estimates.

The CHESS report lists the suspended sulfate concentration as 12.4 $\mu g/m^3$ in 1971 and this is the basis for the estimate of 15 $\mu g/m^3$ for the 1968–1970 period. Observations of sulfate in Magna area subsequent to 1971 support the argument that average annual concentrations are in the neighborhood of 15 $\mu g/m^3$, or that they are significantly higher than reported for Anaconda.

On page 2–79, in Table 2.4.1, it may be noted that suspended sulfate values for the High community do not follow the sulfur dioxide concentrations, particularly for the Spring and Summer. This raises a question about using sulfur dioxide as an indicator of sulfate, as was done with the regression equation on page 2–39. (Median values for the High community are: Sulfur dioxide, Spring 64, Summer 9, whereas for suspended sulfate they are 8 and 7, respectively.)

Wind blowing from the smelter stack to Magna would generally cross a portion of the Great Salt Lake and, therefore, might carry more moisture, thereby facilitating the conversion of sulfur dioxide to sulfate. Perhaps this mechanism helps to account for the high sulfate concentrations observed in Magna.

4. ESTIMATED SUSPENDED SULFATE CONCENTRATION, 7.2 $\mu g/m^3$ (ANACONDA)

The 7.2 $\mu g/m^3$ suspended sulfate value can be obtained from Table 3.1.7, page 3–12, by taking an average of sulfate values for three years, as follows:

Year:	$\mu g/m^3$
1970	8.9
1969	7.6
1968	5.1
Average	7.2

These sulfate values are estimates, based on estimates of total suspended particulate and an estimate of the ratio of suspended sulfate concentration to total suspended particulate concentration, based on results from East Helena and Helena, Montana, and Magna, Utah. The same procedure was used for Kellogg, Idaho.

On page 3–11, in an attempt to explain how the suspended sulfate estimates were made for Kellogg, it is stated that "Data observed for Magna during the period January 1971–June 1972 indicated an average ratio of suspended sulfate concentration to total suspended particulate of 0.159." Following this is the reference number "22," referring to National Air Pollution Control Administration Publication No. AP–61, "Characteristics of Particulate Patterns 1957–1966." This publication presents graphs of suspended particulate concentrations for various cities over a ten year period. In it, suspended sulfates are not mentioned, the time period is wrong, and there are no data

for Magna; therefore, it must be concluded that the reference is an error.

An obvious reference for this paper would have been the paper by Marvin B. Hertz, et al., "Human Exposure to Air Pollution in Salt Lake Communities, 1940–1971," however, it is not referenced. Perhaps this was the reference intended. Even so, the ratio 0.159 cannot be obtained from the Hertz paper.

In the Hertz paper, page 2–11, Table 2.1.2, which gives CHESS 1971 Annual Averages for Magna, the suspended sulfate concentration is 9.6 $\mu g/m^3$ and the total suspended particulate concentration is 53.9, which gives a ratio of 0.178. In Tables 2.1.5 and 2.1.A.16, the following concentrations are given: TSP, 66 $\mu g/m^3$, SS, 12.4 $\mu g/m^3$. Here the ratio is 0.188. Other ratios can be determined for various time periods from Tables 2.1.A.4 and 2.1.A.5, but none of these is 0.159.

Note (page 3–11) that the unexplained ratio 0.159 for Magna is used with the 0.063 ratio for East Helena to obtain the ratio 0.111 plus or minus 0.057 that is used to estimate suspended sulfate concentrations for Kellogg, and the 0.11 plus or minus 0.06 ratio for Anaconda (page 3–13).

(Pages 3–8 and 3–9) Particulate emissions for East Helena are given in two tables on pages that face each other. The headings of the second column in Table 3.1.4 should be "Emissions, Tons/year," not "Emissions, Tons/day."

On page 3–7 it is stated that estimates of stack emissions for both particulate and sulfur dioxide for East Helena for the years 1941–1970 were provided by Asarco. Presumably the data in Table 3.1.3 are Asarco data. The source of the data in Table 3.1.4 is not stated.

The Office of Air Programs Publication No. AP–91, Helena Valley, Montana, Area Environmental Pollution Study, gives more information about the industrial complex at East Helena. This study was conducted during the period June 1969 through June 1970. The table below is from this study.

EMISSIONS FROM EAST HELENA INDUSTRIAL COMPLEX

[Tons per day]

Company and operation	Emissions					
	SO_2 production			Particulates production		
	Reduced	Normal	Maximum	Reduced	Normal	Maximum
Asarco:						
Sintering	184.6	315.6	355.1	0.8	0.5	0.5+
Smelting	8.4	14.6	23.2	(1)	(1)	(1)
Miscellaneous	(1)	(1)	(1)	(2)	(2)	(2)
Subtotal	193.0	330.2	378.3	.3	.5	.5+
Anaconda:						
Fuming	13.0	13.0	13.0	(3)	(3)	(3)
Miscellaneous	(4)	(4)	(4)	1.0	1.0	1.0
Subtotal	13.0	13.0	13.0	1.0	1.0	1.0
American Chemet: Pigment production	(1)	(1)	(1)	(5)	(5)	(5)
Total	206.0	343.2	391.3	1.3	1.5	1.5+

1 Negligible.
2 The outside storage of concentrates contributes a significant but undetermined amount of particulates.
3 Emissions also occur during the slag charging and the coal mill, but no estimates have been made.
4 Emissions occur when slag is dumped, but no estimate of their quantity has been made.
5 Emissions are controlled by cyclones and bag filters with high collection efficiencies.

It may be noted that ASARCO is only one of several particulate sources for the East Helena area. Fuming and other slag processing activities of the Anaconda Co. are estimated to produce 1.0 tons per day of particulates, resulting in a normal total of 1.5 tons per day, not a rate in the neighborhood of 0.3 tons per day as Table 3.1.3 suggests. Further, the total normal sulfur dioxide emission rate in the preceding table is 343.2 tons per day, a considerably higher rate than is given in Table 3.1.2. (i.e., 1969: 221 tons/day; 1970: 239 tons/day).

On page 3–7, right hand side, is given an explanation of how the data in Table 3.1.4 were used to obtain a ratio of total suspended particulate concentration to tons of particulate emitted per day for East Helena. However, after giving this explanation, the estimates of TSP in Table 3.1.5, that were used to make the suspended sulfate estimates were *not* obtained by means of this ratio. They seem to have been obtained from the particulate emission data in Table 3.1.3, using the factor 383.22 ($\mu g/m^3$)/(tons/day). The derivation of this factor is not explained. The ratio that is explained never seems to have been used. The suspended sulfate estimates are obtained by multiplying the total suspended particulate concentrations by the factor 0.063, which is explained on page 3–8.

Both observed and estimated suspended particulate concentrations are given in Table 3.1.4 and 3.1.5. It may be noted that the estimated TSP values are used to estimate the suspended sulfate concentrations and not the observed values for the years 1966 through 1969. In 1966, the observed value was 87 $\mu g/m^3$, whereas the estimated value is 114.2 $\mu g/m^3$. No explanation is given for rejecting the observed values.

Data for Magna during the period January 1971-June 1972 indicated an average ratio of suspended sulfate concentration to total suspended particulate of 0.159. The available data for East Helena indicated a suspended sulfate to total suspended particulate ratio of 0.063 ± 0.022 $\mu g/m^3$. For Kellogg, the assumption has been made that the ratio of suspended sulfate to total suspended particulate is the average of these values, or 0.111 ± 0.057. For Anaconda, this value was rounded to 0.11 ± 0.06. It is multiplied by the estimated concentrations of total suspended particulate listed in Table 3.1.7, to obtain the suspended sulfate values for each year.

The following table has been prepared from the Helena Valley study, June through October 1969.

Station	Location [1]		Suspended particulate	Particulate sulfate	Ratio
	Degrees	Miles			
1	34	0.8	108	3.5	0.032
2	105	2.5	74	3.7	.05
3	112	.4	59	4.4	.069
4	274	4.5	62	2.9	.047
Average			76	3.6	.050

[1] With respect to the smelter stack.

The data from stations 1 and 3, the stations nearest the stack, were used to obtain a ratio range (0.037, pages 3–8), but for some curious reason the available ratios from the Helena Valley study were not used. The average ratio for stations 1 and 3 is 0.051.

The ratio chosen for East Helena, 0.063 plus or minus 0.022 ($\mu g/m^3$)/($\mu g/m^3$), is not significantly different from that which might have been obtained had more use been made of the Helena Valley study, but there is no basis for the assumption that the ratio of suspended sulfate to suspended particulate is similar in Magna, East Helena, Helena, and Anaconda.

The dubious nature of using suspended particulate concentrations to estimate suspended sulfate can be seen by comparing Figures 2.4.3 and 2.4.4. In the Low Exposure Community, the sulfate level remains low and nearly constant while the suspended particulate concentrations fluctuate.

In the High Exposure Community, the highest concentration of suspended particulate occurred on the fourth week whereas the peak sulfate value occurred on the third week. On the fourth week, sulfate levels dropped. A corresponding drop in the sulfate levels does not occur until the fifth week. Only during the last seven or eight weeks do suspended particulate and suspended sulfate concentrations fluctuate together. There may be some situations where suspended particulate and suspended sulfate concentrations are well correlated. Justification for assuming correlation in the Salt Lake Basin and the Rocky Mountain communities is inadequately supported by scientific evidence presented in the CHESS Monograph.

Further, the 7.2 $\mu g/m^3$ suspended sulfate estimate for Anaconda is based on an estimate that comes from another estimate of suspended particulate values based on rates of emission from the smelter. During the period 1961–1962, the annual total suspended particulate concentration was found to be 84.5 $\mu g/m^3$. In 1971, the average suspended particulate level was observed to be 52 $\mu g/m^3$. By comparing the observed total suspended particulate concentration with the particulate emitted from the Anaconda plant, a ratio of 9.1 ± 2.3 ($\mu g/m^3$)/(ton/day) was determined. This ratio was multiplied by the particulate emission for Anaconda shown in Table 3.1.3 to estimate the total suspended particulate concentrations for the years 1940–1970. This ratio cannot be actually obtained from the data presented in the report because particulate emissions for the year 1971 are not given, i.e., they are not listed in Table 3.1.3.

The basis for this ratio is unfounded since there are sources for the suspended particulate other than the smelter emissions.

Although there are no actual sulfate observations from the Anaconda area included in the CHESS report there are some actual observations of suspended sulfate versus total suspended particulate available for the year 1971, that were obtained from the Montana State Department of Health and Environmental Sciences. These suggest that annual average suspended sulfate levels in Anaconda are in the neighborhood of 4 or 5 $\mu g/m^3$, even less than the estimated value (7.2 $\mu g/m^3$).

There are also pronounced seasonal effects, with much higher values in winter than in summer. The months of February and April had values of 7 and 9 $\mu g/m^3$ whereas the months of July and August have values of less than 1 ug/m^3. Local heating emissions and relative humidity may be significant factors determining the measured concentration as well as the smelter emissions.

5. ESTIMATES OF SUSPENDED PARTICULATE, SALT LAKE BASIN STUDY

On page 2-23 it is stated that "the number of sulfric acid plants utilizing sulfur recovered from emissions have increased from one in 1940 to seven in 1971, and that air pollution control devices in the form of baghouses, scrubbers, cyclones, and mist eliminators have been installed. Such changes in the smelter operations would greatly effect the ratio of suspended particulate to tons of copper produced. Therefore, aside from the fact that there would be differences from year to year because of meteorology, the procedure described in the first paragraph, right hand column, page 2–24, for estimating suspended particulate from copper production in tons for 1971, is highly questionable.

6. ESTIMATES IN THE CHIGACO AND NEW YORK STUDIES

In the Chicago and New York studies suspended sulfate concentrations were estimated from suspended particulate concentrations. In Chicago, the estimates were used to fill in data for some years when no data were available. In the New York study measured values for suspended sulfates for 1956–1970 were available from the Manhattan 121st Street station, and these values were used as citywide values. The observed annual ratios of suspended sulfate to dustfall for New York City were used to estimate the suspended sulfate levels in Queens and Bronx. In Table 5.3.1 suspended sulfate levels for the Low Community (Riverhead) are listed as about 10 $\mu g/m^3$ for the years 1961 through 1970. The basis for this estimate is not given, although it was probably determined from the 1971 concentration, which was 10.2 $\mu g/m^3$.

In summary, it appears that some values, on which are based important conclusions that sulfates may be harmful to health, are estimated values.

C. Use of Mathematical Dispersion Models

The dispersion model shown in Figure 2.1.16 is incorrectly applied. It was used in the Salt Lake Basin study to determine sulfur dioxide contours around the smelter source and to show that annual exposure estimates obtained from the ratio of 1971 observed air quality to 1971 emissions were not unreasonably high or low. First, the contours are incorrect because the model used does not take into account the elevation of the terrain and the wind direction frequencies for the Salt Lake City airport, which were used are different from those affecting the smelter plume, which originates at the base of the Oquirrh Mountains. Second, a dispersion model is based on numerous assumptions and applied in this way might be off by a factor of two, or more. It does not make sense to use a model to check observations.

The usual application is to apply observational data to calibrate, or verify, a model. A model such as the one used might have been applied to show some sort of relative distribution of concentrations across the Salt Lake Valley, however, it should not have been used to justify estimates of concentrations over the period 1940–1970. (See Tables 2.1.A.14 and 2.1.A.16). Further, during this review of the CHESS report it was discovered that smelter emissions used for the model estimates were tons of sulfur, not tons of sulfur dioxide. Therefore, the model estimate is only half what it should have been. Doubling the emission rate and reducing the wind direction frequency somewhat with respect to Magna might result in an estimated concentration near that measured, which was 61 $\mu g/m^3$.

Apparently the dispersion model was run only once and then the ratio between the emission at the smelter for 1971 and the calculated concentration was applied to emission values for the other years in order to obtain the other listed concentrations in the column headed "Diffusion Model". No account is taken of the fact that meteorological conditions, or perhaps stack conditions, were not the same for all years. More information should have been included in this report on exactly what meteorological data were used in the model. The model requires the use of the STAR program, which is obtained from the National Climate Center. Frequently the results of running this program are based on data for the year 1964, which is the only year when wind directions were punched on data cards to the nearest 10 degrees each hour rather than each 3-hours. Therefore, the model is likely to have incorporated meteorological data for some year other than 1971, the year of the emission data. No attempts is made to show that the year (or period) of the meteorological data is average, good or bad. Similarly there is no attempt to show that 1971 was an average year, yet all of the estimates are based on this assumption.

Considering how the model estimates for the years 1940–1970 were obtained it is misleading to include them in the table, and they serve little purpose since the ratio for the year 1970 is repeated throughout.

On page 2–43, bottom of right hand column, the following statements appear: "Estimates of sulfur dioxide, total suspended particulates, and suspended sulfate concentrations in the High exposure community for 1940–1970 and the Intermediate II exposure community for 1950–1970 were obtained by a mathematical dispersion model, which utilized emissions from the industrial source and extensive local meteorological data, and by observed relationships among pollutants. Observed suspended particulate, suspended sulfate, and sulfur dioxide concentrations for 1970–1971 were used to calibrate the models used to estimate exposure levels for previous years." This is an overstatement. The estimates were obtained from simple ratios and the application of a regression equation. See page 2–39. The model was only applied once to demonstrate that annual exposure estimates obtained from a ratio were not unreasonably high or low.

In the Chicago study, another attempt was made to apply a dispersion model (Figure 4.1.10). This model gives a false picture of pollution conditions that prevailed in the study area because it is based only on pollution sources within the city limits of Chicago, omitting effects of adjoining large industrial sources in Indiana and of some suburban communities to the southwest of the Loop area, which have considerable air pollution.

Maps recently published by the Chicago Department of Environment Control, for the years 1970 and 1975 clearly show that pollution concentrations are not simply concentric around the urban core as the model indicates.

On page 4–8, it is stated Measured data from the City network, from which the exposure estimates were made, were best supported by the Mitre model. It is not clear why a greater use was not made of the available actual measurements instead of the model estimates. Also, it is not sufficiently clear why the model happens to be for the year 1968.

VI. AN ANALYSIS OF THE CHESS HEALTH EFFECTS STUDIES

A. General Problems of Epidemiologic Investigations of Pollution Effects

Before discussing health effects problems specific to CHESS, some discussion of general difficulties inherent to pollution epidemiology may be helpful.

Exposure to suspect pollutants is not controlled in population studies. Indeed with current technologies, it is not possible to be sure that the correct pollutant is even being measured. Combinations of pollutants may be more harmful than any single pollutant, and the number of studies needed to investigate such synergisms (interactions) increases rapidly with the number of pollutants under consideration. The analysis of synergisms is often impractical since sites with the needed configurations of pollutants are seldom at hand.

Not only is exposure uncontrolled, it is often difficult to measure. Even when aerometric measurements are valid, special meteorologic conditions or personal habits may cause a given subject to experience pollution levels very different from those measured at a nearby fixed monitoring station. These problems are exacerbated in long term studies during which the quality of aerometric data has been variable and individuals have changed jobs and residences. Aerometric methods for measuring hourly or daily pollution levels are often less reliable than required for studies associating pollution levels with short term health effects.

The health measurements are often subjective responses to a questionnaire or interview. An individual may give one answer on a self-administered questionnaire and another to a friendly interviewer. Other factors, such as the public announcement of a pollution alert, can also influence subjective health measurements. Some health measurements, such as pulmonary function tests or blood analyses, are less influenced by poorly defined conditions surrounding the meaurements and are said to be objective. However, even objective endpoints respond to uncontrolled events like an undetected influenza epidemic or high pollen count.

Whether the health measurement is subjective or objective, the response is often affected by factors (covariates) associated with the subject studied and unrelated to pollutant exposure. Whether the individual smokes or is subjected to cigarette smoke at home or work is a covariate of dominant importance in pollution studies. Educational attainment may affect responses to questions about phlegm or pneumonia. Occupation, age, sex, race, immunity to influenza, allergy, access to air-conditioning and countless other covariates complicate the interpretation of epidemiologic data. Epidemiologists treat covariates in two ways. They try to choose study populations which

77-590—76——5

have similar covariate characteristics so that health differences between such populations can be ascribed to pollution effects. Alternatively, they make mathematical adjustments to nullify the effects of covariate imbalances. Both strategies have weaknesses, and neither works if the investigator is unaware of an important covariate or has failed to measure it.

The epidemiologist has little control over the subjects studied. He cannot assign them at random to reside in polluted communities of interest. Thus, a clean town may contain many asthmatics because asthmatics have wisely chosen to live there rather than in a more polluted community. This fundamental problem of self-selection must qualify any conclusions obtained from non-randomized population studies: it may be possible to demonstrate temporal or spatial associations between health and pollution measurements, but a causal relationship cannot be inferred on the basis of a single epidemiologic study.

Students of pollution counter these weaknesses in several ways. One strategy is to replicate an epidemiologic study in a variety of circumstances and serially in time. If a consistent association between pollution and health measurement is observed, it is held to be reliable since covariate imbalances and problems of self-selection are unlikely to affect all sites and to persist over time. Clinical studies, in which healthy volunteers are subjected to controlled pollution exposures, and toxicological studies, in which animals are subjected to various combinations and doses of pollutants, complement information obtained from epidemiologic studies. This body of information from clinical and toxicological studies and from several epidemiologic studies may substantiate an interesting association suggested by the health and pollution measurements of a single epidemiologic study.

In addition to these general issues, several questions directly pertinent to the CHESS health measurements were examined, namely:

(1) Was the health measurement a reliable and meaningful indicator of public health?

(2) Was the statistical analysis sound and impartial?

(3) Were the methods used to ascribe specific health effects to specific pollutants and to establish dose-response relationships logically compelling?

The sources of information used in this assessment include:

(1) The CHESS Monograph cited previously which contains data gathered in 1970–1971 as well as some earlier studies.

(2) Preliminary internal EPA drafts of 1971–1972 studies.

(3) External peer review documents.

(4) Interviews and progress reports of contractors and EPA personnel who gathered health and aerometric data.

(5) Interviews with EPA personnel who gathered and analyzed the data.

Time limitations prohibited a complete reanalysis of the primary health and aerometric data.

B. Introduction and Definitions

What follows is a detailed discussion of each type of study found in the CHESS Monograph. Each health measurement is considered together with special problems associated with particular CHESS sites. Comparisons across CHESS sites and comparisons with sub-

sequent CHESS studies are made (when possible). A relatively nontechnical summary follows the discussion of each health measurement. Unfortunately, much of the terminology is specialized. A short glossary of statistical terms is given below for those who wish to read the detailed assessments. This glossary is intended to convey the purpose and use of certain statistical terms rather than to give precise definitions, which are available from standard statistics texts. The definitions are particularized to pollution applications.

Adjustment Procedure.—This is a mathematical method, based on an assumed model, to compare populations exposed to various pollution levels when other factors which might affect health measurements, such as age, sex, or race also differ among communities.

Autocorrelation.—This term describes the way in which the asthma attack rate on a particular day depends on the attack rate on other days. Pollution levels, which are also time series, also have autocorrelation structure, since a pollutant which is elevated above usual levels on one day is likely to be elevated on the next day also.

Correlation.—Correlation is a measure of the strength of the linear relationship between two quantities. Positively correlated quantities tend to rise and fall together, whereas negatively correlated quantities tend to rise and fall out of phase.

Goodness-of-Fit Statistic.—This quantity is used to gauge how well a mathematical model fits a set of data.

Least Squares Method.—This is a method of estimating the parameters of a mathematical model so that the final estimates bring the model into closest possible congruence with the data.

Linear Model for Categorical Data.—Linear models are models in which the average value of the observation (dependent variable) is assumed to be a linear function of independent variables, such as pollution category, age or sex. If the expected frequency of a cell in a contingency table is assumed to be a linear function of such independent variables, one has a linear model for categorical data. The value of such a model, if it fits the data, is that it enables one to estimate the relative contributions of the various independent variables.

Multiple Regression Analysis.—Multiple regression analysis is a method for relating an observation (dependent variable) such as asthma attack rate to several independent variables, such as suspended sulfate levels, sulfur dioxide and perhaps other pollutants. Usual applications require that the expected value of the observation is a linear function of the independent variables, and that observations are statistically independent with constant variability about the expected value. A multiple regression model can be used to describe a given set of data and to predict the outcome of future observations for known values of the independent variables. However, the fact that a multiple regression model fits a given data set well or yields good predictions of future observations need not imply that the independent variables determine the observations or bear any causal relationship to the observations.

Pooled.—Pooled data is data obtained by combining data from two or more samples and then ignoring the fact that the combined data came from various sources.

Residuals.—Residuals are differences between observed data and values expected under a mathematical model. Residuals are helpful in assessing the appropriateness of the model.

Ridge Regression.—If independent variables are highly correlated, estimates of the effect of a particular independent variable on the dependent variable can be greatly influenced by what other independent variables are included in the multiple regression model. Ridge regression is a method of determining the importance of a particular independent variable in the presence of other highly correlated independent variables.

Statistical Significance.—A statistical hypothesis is tested by computing how probable a set of data is under that hypothesis. This level of probability is called the significance level and a level $p<.05$ means that the chances are less than 5 in 100 that the observed outcome could have occurred if the hypothesis were true. The significance level is often taken as a measure of the strength of evidence against the hypothesis, with smaller significance levels providing more evidence against the hypothesis. Suppose, for example, the hypothesis being tested is that two populations have the same risk of chronic respiratory disease. This is called a "null" hypothesis—that is, there is no difference. If, in view of the data, this hypothesis is found to be highly improbable (that is, to have low probability of occurrence) one rejects the hypothesis (of no difference in risk of disease between populations) and assumes instead that there *is* a difference in risk of chronic respiratory disease between the two populations.

C. Specific Health Effects Studies

1. chronic respiratory disease (crd) prevalence

CHESS estimates adult chronic respiratory disease prevalence by means of a self-administered questionnaire which inquires whether the subject coughs and produces phlegm for at least three months of the previous year. These studies compare CRD prevalences in adjacent communities with different pollution levels, often designated "High," "Intermediate" or "Low." The Monograph reports on data gathered in the New York City and Salt Lake City areas in 1970–1971 and also includes studies from the Chicago area and five Rocky Mountain communities. CHESS monitoring stations were never present in Chicago or the Rocky Mountain communities, and exposure estimates were based on local data sources and theoretical extrapolations in these areas (see Chapters IV and V). CHESS monitoring stations only became operational in Salt Lake City in December, 1970. Attention in this investigation was confined to the four surveys, which were reviewed by Chapman et al,[1] and to New York City CHESS follow-up data [2] from 1971–1972.

The CHESS CRD questionnaire was adapted for self-administration from an interview-administered questionnaire used by the British Medical Research Council.[3] Although a similar questionnaire had been validated for self-administration in a 1971 Japanese study,[4] the

---

[1] Chapman, R. S., Shy, C. M., Finklea, J. F. House, D. E., Goldberg, H. E., Hayes, C. G., "Chronic Respiratory Disease in Military Inductees and Parents of Schoolchildren." *Archives of Environmental Health*, Volume 27, Number 3 (1973), pp. 138–142.

[2] Galke, W. A., House, D., "Prevalence of Chronic Respiratory Disease Symptoms in New York Area Adults, 1972." *EPA In-House Technical Report* (June 7, 1976).

[3] "Standards for Epidemiologic Surveys in Chronic Respiratory Disease." National Tuberculosis and Respiratory Disease Association. (1969).

[4] Tsunetoshi, Y., Shimiza, T., Takahashi, H., Ichinosawa, A., Ueda, M., Nakayama, N., Yagamata, Y., "Epidemiological Study of Chronic Bronchitis with Special Reference to Effect of Air Pollution." Int Arch Arbeitsmed. Volume 29 (1971), pp. 1–27.

Monograph does not contain validating data for U.S. surveys. An internal EPA report [5] dated 1973 showed that the self-administered questionnaire detected only three of nine cases of chronic bronchitis found through interview. Survey supervisors noted that the word "phlegm" was not understood by some subjects, suggesting that educational attainment affects CRD response. However, even if the CHESS questionnaire underestimates CRD prevalence in a given community, it may give a reliable indication of the difference in CRD prevalence between two communities of similar educational attainment.

The Bronx, Queens and Riverhead, Long Island represented two Intermediate and one Low level of urban pollution in the New York City area. The Intermediate regions exceeded Riverhead in total suspended particulate (TSP), sulfur dioxide (SO_2), suspended sulfate (SS) and suspended nitrate (SN). Generally, questionnaires were given to school children who gave them to their parents to fill out. Thus, the sample is not representative of adults in general but only of parents of school children. Only 73% of Riverhead parents responded to the questionnaire which probably reflects the fact that a large number of Riverhead questionnaires were mailed to the parents rather than distributed through the children. Only white respondents who provided sufficient questionnaire data and who had stable residence histories were included in the analysis. Queens was shown to have the highest income and educational level and unpublished data showed Queens to have a larger proportion of Jews than Bronx. These facts may account for the higher CRD prevalence rates reported for the Queens than Bronx. No formal adjustments were applied for these socio-economic covariates. While smokers and non-smokers were treated separately, no covariate adjustments were made for exogenous exposure to cigarette smoke at home or work or for occupational exposure (which affected less than 1% of respondents). The statistical methods were sound. Sex- and smoking-specific prevalence rates were consistently lower in Riverhead than in Queens or Bronx, the difference being 5% typically. For comparison, smokers had prevalence rates nearly 10% above non-smokers in the same community. A linear model for categorical data [6] was used to make smoking-, age-, and sex-adjusted tests for community differences in prevalence rates, and the differences were statistically significant. A parallel analysis using severity scores for CRD confirmed the analysis of prevalence. While the statistical methods were appropriate, additional information would be helpful. In particular, confidence intervals on the prevalence rates in Table 5.2.7 (unless otherwise indicated, figure, page and Table citations refer to the CHESS Monograph) and observed and expected rates under the linear model of Table 5.2.8. would allow the reader to make specific prevalence comparisons and to verify that the linear model held for all subcategories.

Communities in the Salt Lake region provided an opportunity to study the effects of sulfur oxide (SS, SO_2) smelter emissions. Other urban pollutants, including SN and TSP, were moderate or low. Magna was exposed to intermittent smelter fumigations of SO_2 and

---

[5] House, D., "Reliability of the CHESS School and Family Health Questionnaire." EPA Human Studies Laboratory Internal Report (April 12, 1973).

[6] Grizzle, J.E., Starmer, C.F., Koch, G.G., "Analysis of Categorical Data by Linear Models." *Biometrics* Volume 25, Number 3 (September 1969), pp. 489–504.

had the highest average SO_2 and SS levels of these pollutants, and Ogden was designated Low. Aerometric copper data was not available. The questionnaire was distributed through elementary school children and mailed to parents of high school students. Response rates of 85% and 35% were found for child-carried and mailed questionnaires respectively. Although the 65% nonresponse rate to mailed questionnaires admits the possibility of serious reporting bias, the authors assure the reader that inter-community CRD differences (presumably similar) were observed for both sets of parents. Respondents were excluded for incomplete questionnaires, for a residential change within the previous two years, and for occupational exposure to irritants such as coal dust, cutting oils, asbestos, mine dust, smelter fumes, cotton dust and foundry dust. Subsequent analysis showed that the occupational exclusion gives a conservative estimate of effects attributable to pollution. All races were included, but the proportion of black respondents was trivial. No covariate measurements were made to assess religion, exposure to exogenous cigarette smoke at home or work, or racial composition, although Salt Lake City has proportionately fewer Mormons, and Magna more Spanish Americans. Educational attainment was comparable in the four communities. CRD prevalence rates reflected pollution levels faithfully, and rates in Magna (High) exceeded those in Ogden (Low) by 2–7%, depending on sex and smoking status. These differences were found to be statistically significant using the sex, smoking and age adjusted linear model.[7] For comparison, differences in CRD prevalence attributable to smoking and to occupational exposure were 10–20% and 2–8% respectively. Thus, air pollution and occupational exposure were associated with comparable increases in CRD prevalence, and personal smoking habits seemed to be a more important determinant.

The exposure data from the Rocky Mountain communities were least adequate. Two smelter communities, Kellogg and Anaconda, had relatively high levels of SO_2. East Helena and two non-smelter communities, Bozeman and Helena, were classified as Low. Ambient zinc, copper and lead were not measured. Nearly 85% of the questionnaires distributed by elementary school children were returned. Respondents were excluded for occupational exposure, but no residence duration requirement is mentioned. Over 97% of respondents were white. Low communities were better educated than High communities, and most occupational exclusions were from the latter, but these factors tend to reduce the apparent pollution effect. Conversely, the increased residential crowding in High communities might increase apparent pollution effects. Sex-, education-, smoking- and age-adjusted prevalence rates were statistically significantly higher in High than Low communities (Table 3.2.6). High communities had CRD prevalence rates 1–3% higher than found in Low communities (Table 3.2.5). Corresponding differences associated with smoking and occupational exposure were 10–16% and 0.1–3.6% respectively.

The Chicago studies compared urban (High), suburban (Intermediate) and clean surrounding areas of Illinois and Indiana (Low). Local SO_2, SS and TSP measurements were used, and other urban pollutants were not measured. Gary, Indiana was included in urban

---

[7] Grizzle N. E., et al., op. cit.

and may represent a special industrial pollution hazard. A CRD questionnaire designed by Julius Goldberg was administered to recruits reporting to the Chicago induction center. This population obviously differs from the previous three and contains only young male adults. Some with serious respiratory problems may have been exempted prior to induction. To be analyzed, the questionnaire had to be adequately filled out and the respondent had to have lived at his present address for at least three years. Statistically significant age-, race-, smoking-, and education-adjusted community differences in prevalence rates were found for black smokers and white non-smokers (Table 4.2.8). However, the significance levels for black smokers and the adjusted rates in Table 4.2.7 are suspect since the linear model does not fit the data well for black and white smokers. (The "adjusted" rates in Table 4.2.7 are really expected rates under the linear model, which is a useage unfamiliar to epidemiologists.) Race and smoking specific prevalence rates tend to be 0.4–3.6 percent higher among urban than Low communities, while smoking effects are 0.3–3.5 percent for blacks and 12.4–15.1 percent for whites.

To summarize, these four studies demonstrate that higher adult CRD prevalence rates are associated with pollution in two urban sites and in two smelter-exposed sites with high sulfur oxide emissions. The magnitude of these differences is comparable to that associated with occupational exposure but smaller than that associated with smoking. Such differences were not seen in a 1970 CHESS study in Chattanooga, comparing communities with different nitrogen dioxide levels.[8] New York CHESS studies in 1971–1972 showed no significant differences in CRD prevalence among Riverhead, Queens Bronx and Sheepshead Bay,[9] and prevalence rates were dramatically reduced compared to 1970–1971 values. Perhaps subjects tire of the CRD questionnaire and report less symptomatology on repeated sampling. This would limit the usefulness of the CRD questionnaire as a health surveillance measurement. An alternate explanation is that decreasing CRD prevalence reflects decreasing pollution in New York.[10]

These studies demonstrate that small CRD prevalence increases are associated with pollution levels which are moderate by historic standards in urban and smelter sites. The Monograph contains no formal methods for associating these effects with a specific pollutant or for establishing a dose-response relationship. Any such inferences are tenuous, especially since the aerometric data is of dubious quality and completeness (See Chapters IV and V). The authors are to be complimented for applying newly developed and appropriate statistical methods.

2. RETROSPECTIVE SURVEYS FOR ACUTE LOWER RESPIRATORY DISEASE (LRD) IN CHILDREN

The Monograph contains data from questionnaires asking mothers to recall how many times each child under age twelve had had pneumonia, croup, or bronchitis during the previous three years. Data was also gathered on related hospitalizations and doctor visits.

---

[8] Chapman, R.S., et. al., Prevalence of Chronic Respiratory Disease In Chattagnooga: Effect of Community Exposure to Nitrogen Oxides. EPA Internal Draft. July, 1973.

[9] Galke, W.A. and D. House. Prevalence of Chronic Respiratory Symptoms in New York Area Adults-1972. EPA Internal Draft. June 7, 1976.

[10] Galke, W.A. and D. House. op. cit.

These questionnaires were distributed through school children together with the 1970 Salt Lake and Rocky Mountain CRD questionnaires. Children less than one year old, those with incomplete questionnaires, and asthmatics were excluded from analysis. This last exclusion tended to minimize estimates of pollution affects (See Table 2.3.4). The sample is not a random selection from all children aged 1–12, but represents only schoolchildren and their siblings Because the data is retrospective, validation is crucial. Nurses examined medical records to document LRD (any croup, bronchitis or pneumonia) detected by questionnaire. In Ogden (Low), 70% of questionnaire-detected cases were confirmed, whereas 78% were documented in Kearns (Intermediate) and Magna (High). These differences would tend to reduce estimates of pollution effects. The corresponding figures were 88% and 75% for pooled Low and pooled High Rocky Mountain communities, which would tend to increase estimates of pollution effects.

Parents smoked less in Magna than in other Salt Lake communities (Table 2.3.3), which would minimize pollution effects. The Monograph presents a higher percentage of parental smoking in pooled-High Rocky Mountain communities (Table 3.3.2), but argues no adjustment is needed.

The monograph does not specify how age-, sex-, and socioeconomic-"adjusted" rates are computed. Since the epidemiologic interpretation is predicated on these adjusted rates, some clarification, obtained from EPA interviews, is given.

A linear model [11] containing community, age, sex, and socioeconomic main effects and no interactions was fitted separately for the three residence designations in Table 2.3.12. The "adjusted" rates in each column of this table are simply the sum of the estimated overall mean and community effect. Differences of rates within each such column are estimated community effects under the linear model. These comments may explain the anomaly that adjusted rates exceed unadjusted rates in every community, both for Salt Lake and Rocky Mountain studies. A more reasonable "adjusted" rate would be obtained by applying the total estimated linear model, including all main effects, to the actual covariate composition of each community. The "adjusted" rates given in the Monograph still provide reasonable estimates of community effects, provided the linear model used fits the data well. The Salt Lake studies give no goodness-of-fit statistic or analysis of residuals, and the model is not even specified by Table 2.3.9. Thus the Salt Lake "adjusted" rates and significance levels must be regarded as provisional. Although Rocky Mountain studies give no analysis of residuals, the goodness-of-fit statistics suggest the linear model is valid.

If one assumes the validity of the linear model used, the Salt Lake studies demonstrated statistically significant community differences in three-year croup, bronchitis, and LRD incidence rates after adjustment for sex, age, and socioeconomic status (actually educational attainment of the mother). Magna residents of three years duration had LRD, croup and bronchitis rates which exceed corresponding rates in Ogden by 10.9 percent, 9.5 percent, and 7.1 percent respecively. However, these rates did not increase in strict accordance with

---

[11] Grizzle, J. E. et al., op. cit.

pollution levels and Magna residents of fewer than three years actually experienced the lowest croup, bronchitis, and LRD rates. No significant community differences were found for pneumonia or hospitalization rates.

Pooled High Rocky Mountain communities had significantly higher adjusted croup rates than pooled Low communities. No significant differences (at $p<.05$) were found for LRD, bronchitis, pneumonia, or hospitalization. Those residents in pooled Low communities for fewer than three years experienced higher adjusted croup, bronchitis, and LRD rates than corresponding inhabitants of pooled High communities.

The basic observation that children of families resident in smelter-exposed communities for three years or more experience higher croup rates than those in nearby communities may be valid. However, these rates do not increase consistently with increasing pollution (Tables 3.3.A.2, 2.3.8, and 2.3.A.2), and recent migrants to High areas have lower croup rates than migrants to Low areas (Tables 2.3.8, 2.3.A.2, 3.3.A.2, and 3.3.8). These inconsistencies deserve further elucidation. The documented insensitivity of the LRD questionaire reminds us that this endpoint is no better than a mother's ability to recall a three-year illness history for each of her children and to remember specific diagnoses, such as croup and bronchitis. Finally, the conclusion in the Monograph (p. 7–9) that "increases in lower respiratory disease frequency are probably associated with suspended sulfates rather than sulfur dioxide" is tenuous at best, since no formal methods of relating health effects to a specific pollutant are described and since suspended sulfate measurements are not available in Salt Lake and Rocky Mountain communities for the years 1967–1970 surveyed by these questionnaires.

3. ACUTE RESPIRATORY DISEASE (ARD) IN VOLUNTEER FAMILIES

The Monograph contains New York ARD studies from 1970–1971. Telephone interviewers made biweekly calls to mothers of families enrolled in the study to inquire whether any family member had developed upper or lower respiratory illness in the preceding two weeks, and, if so, whether a doctor had been consulted and how many days of restricted activity had eventuated. If an individual was reported to have both upper and lower respiratory symptoms, his illness was classified as lower respiratory disease. Thus, the least ambiguous diagnostic category is "all respiratory illness." A 10 percent subsample of families who had cooperated on the CRD questionnaire was recruited for the ARD study. An eligible family had to have resided at least the past year within 1.5 miles of a CHESS monitoring site, to be white, to have one or more children age 12 or less, and to have a working phone. Priority was given to families with many preschool siblings, and in a subsequent protocol, such "priorities" were specified in writing. These priority schedules introduce ambiguity into the eligibility requirements. The major response variables were the number of respiratory illnesses per hundred person-weeks exposure (the attack rate) and the severity score, which reflected physician visits, fever, and restricted activity. The severity score scale has been criticized as arbitrary, and inversions for scores II and III in the upper

respiratory column of Table 5.3.9 support this view. Table 5.3.4 (which has obvious misprints for Bronx data) shows that only 750 of 1,136 families in Riverhead participated for the full 32 weeks of study, while 887 of 1,153 families did so for Queens. Selected interviews were repeated a few days after the initial call and concordant results were obtained in over 90 percent of those previously reporting ARD and 98 percent of those previously denying ARD. This validation shows the interviews were reproducible, but it does not verify the medical content.

The authors point out that the pooled Bronx-Queens community has a higher proportion of adult smokers and of children exposed to smoke at home than Riverhead. More cooking gas was also used in Bronx-Queens. But no formal adjustments were made for these covariates. The percent of crowded families was 15 percent in Riverhead and 17 percent in Bronx-Queens, a difference the authors judged to be inconsequential. (Table 5.3.14).

The analysis of these panel studies is complicated by the fact that biweekly attack rates are statistically dependent, since the same people are repeatedly surveyed. The same comment applies to severity scores. Significance levels or confidence limits cannot be estimated without an analysis of such dependence, which the Monograph does not provide. In particular the Monograph does not specify the methods used for computing significance levels in Tables 5.3.5, 5.3.7, and 5.3.15, and the quoted significance levels are suspect. The formal statistical properties of "attack rates" and "relative risks" are insufficiently defined to allow us to calculate confidence intervals and perform significance tests, even though these quantities may be useful descriptive statistics.

Figure 5.3.1 shows that Riverhead attack rates are usually lower than either Queens or Bronx rates. It is inconsistent with the pollution hypothesis that Queens has higher rates than Bronx for 13 of the 15 surveys shown. Perhaps this reflects socioeconomic differences mentioned in the CRD discussion.

Table 5.3.6 contains ratios of attack rates with the Low community rate as denominator.

Bronx and Queens rates usually exceed Riverhead rates for lower respiratory illness and all respiratory illness, but the relation is inconstant for upper respiratory illness. This difference may be an artifact of classifying all those with upper and lower respiratory symptoms as lower. For the unambiguous category of "all acute respiratory disease," Table 5.3.6 shows the pooled Queens-Bronx rates to exceed Riverhead rates for fathers, mothers, school children, and preschool children. Table 5.3.13 divides the populations shown in Table 5.3.6 into those who have lived five or more years at their present residence (stable) and those who have moved (mobile). It is surprising that stable fathers and mothers actually experience lower attack rates in pooled Queens-Bronx than Riverhead, and this inconsistency warrants special consideration. Likewise confusing is the result (Table 5.3.14) that crowded families have lower attack rates than less crowded families.

The 1969–1970 Chicago ARD study compared Intermediate, High, and Highest Chicago neighborhoods. Twenty monitoring sites run by the City of Chicago were ranked in order of increasing 1960 TSP

values. The lowest 10 were designated as Intermediate, the next five as High, and the five with highest TSP values as Highest. Such designations did not correspond to SO_2 ranking (Table 4.3.1) and SS was not measured for these sites. Families with 2–5 year old children in day care centers located near monitoring stations were recruited. The family was assigned the exposure classification of a monitoring station in the same census tract, if any, but the Monograph is unclear about exposure designations for families in unmonitored census tracts. Over 500 of the families initially enrolled participated throughout the study. Biweekly telephone interviews were used to gather data for attack rates and severity scores as in New York. Review of doctors' records documented 95.3% of telephone-survey-detected ARD.

The Highest exposure category contains more smokers and more professionals than the two other categories. The Monograph gives smoking-adjusted relative attack rates (Table 4.3.15) which show this covariate does not alter conclusions, but the method of adjustment is not presented. Although families of professionals appear at higher risk than those of non-professionals (Table 4.3.13), adjustment for this covariate was said to produce little change. The racial composition of the exposure categories is not given.

Figure 4.3.1 shows that Intermediate families experienced the lowest attack rates during 20 of 25 biweekly surveys, but the largest differences were evident during the three month influenza season. One wonders why the first three surveys were omitted in Chicago but not New York. Severity score rankings show inversions in days of restricted activity (Table 4.3.7) as in the New York study, which argues against this scoring system. The relative risk data in Table 4.3.5 shows higher ARD attack rates in Highest than in Intermediate exposure families for fathers, mothers, older siblings, nursery school children, and younger siblings. This result holds both for mobile and stable (three year residency) families (Table 4.3.11), in contrast to the New York study. The Monograph does not define Table 4.3.12 adequately. If the analyses were performed on attack rates, the significance levels quoted are untrustworthy. If by "frequency" the authors mean the proportion of families ever experiencing ARD during the study period, then use of the linear model is justified, but more explication is needed.

These two ARD studies have the advantage that health data were gathered prospectively by telephone interview, so the recall of the mother was not taxed as in the LRD studies. The quality of the health data are probably adequate, and descriptive statistics, such as "attack rates" and "relative risks" are no doubt useful guides. However, some mathematical statistical investigations to compute confidence intervals and significance tests would be a valuable addition to this methodology, especially since these problems are non-trivial. It is likely true that families exposed to high levels of urban pollution experience higher ARD attack rates than those less exposed. However, two surprising findings require further study, namely: (1) Why do residentially stable fathers and mothers cxperience lower attack rates in Bronx-Queens than in Riverhead? (2) Why do crowded families in New York and less educated families in Chicago appear at lower risk?

Summary statements linking specific SO_2, SS and TSP levels to excesses in ARD attack rates must be viewed with customary caution.

4. ASTHMA PANEL STUDIES

The 1970–1971 Salt Lake asthma panel studies attempt to correlate the daily proportion of panelists suffering an asthmatic attack with daily pollutant levels. The studies were complicated by practical and theoretical problems. Many relevant factors, including medication (steroids), humidity, exercise, daily temperature changes, nitrogen dioxide levels, copper levels, and exposure to smokers at home or work were not evaluated, and the study focused on TSP, SS, SO_2, SN (suspended nitrates), and temperature. Daily SO_2 measurements may be too unreliable for meaningful assessment (Chapter 3), and while midnight to midnight daily asthma rates were used, air monitoring samples were obtained from 10:00 A.M. one day to 10:00 A.M. the next, introducing a lag.

Potential panelists, recruited from families who had responded to the CRD questionnaire, or suggested by doctors or known asthmatics, were interviewed to obtain covariate and eligibility data. To be eligible, the asthmatic had to have been previously diagnosed by a physician and had to give a history of two or more episodes of wheezing and shortness-of-breath during the previous year. In addition he had to live within two miles of a monitoring site. This residency requirement was violated from time to time, and in 1975, 38 of 222 asthma panelists resided outside a 2.5 mile radius.[12]

Since a physician was not available to make the diagnosis of asthma for prospective candidates, certain individuals with unlikely asthma histories (such as elderly smokers with a sudden onset of "asthma") were eliminated. While this second level of panelist selection may have reduced the number of non-asthmatics, it introduces uncertainty as to what the exact eligibility requirements were. The monograph states only: "Highest priority for selection was given to non-smokers over sixteen years of age." In 1974, EPA staff suggested that an improved initial interview questionnaire was needed [13] for separating asthma from other cardiorespiratory disease, but these suggestions were not implemented. Asthma panelists were mailed weekly questionnaires and asked to record day by day whether no attack, one attack, or more than one attack had occurred. Some failed to return diaries, and some filled them out for the wrong week. Those who consistently failed to return diaries were eliminated from the panel, as were those who never reported an attack and those who reported an attack every day.

The 1970–1971 Utah CHESS Progress Report [14] gives summary asthma panel data indicating that mailed diaries are inconvenient. During week eight, 183 diaries were mailed out. Of these, 126 were returned on time, and 30 additional diaries were returned after telephone prompting. Ten calls were required to obtain or correct particular entries. Of the 183 diaries sent out, 156 provided "useable" data for this typical week. The mailed diary was discontinued in the 1973–1974 protocol and replaced by a weekly telephone interview to gather diary data kept by panelists at home.

---

[12] Beck, M. M. and P. Carl, Bureau of Environmental Epidemiology Technical Report Number 3. Utah State Department of Social Services, Division of Health. 1975.

[13] Calafiore, D. Memorandum to CHESS Field Epidemiology Director. HSL., and Attachments, (January 31, 1974).

[14] Utah CHESS Progress Report. October 22, 1970 to December 31, 1971.

The Monograph deemphasizes covariates since no intercommunity comparisons are made. Some community differences in age, sex, education, and smoking habits are noted (Table 2.4.2).

Dropout rates varied by community. Magna, Kearns, Salt Lake City, and Ogden had respective dropout rates of 13/48, 24/55, 20/46 and 14/46 during 43 weeks of study.[15]

The measured response was the daily proportion of panelists reporting an asthma attack. The daily composition of the panels (age, sex, severity or illness) changes as panelists drop-out and are replaced. Since some panelists continue throughout the study, daily asthma attack rates are statistically dependent, which complicated the analysis.

Figures 2.4.1–2.4.4 relate mean weekly attack rates for each panel to weekly minimum temperature, weekly TSP, weekly SO_2, and weekly SS. By superimposing these graphs one can see that Ogden (Low) often has higher attack rates than Magna (High). The Monograph makes no such comparison. Conspicuously absent are corresponding graphs for Salt Lake City and Kearns.

Sample correlations for daily attack rate with daily pollutant levels are shown in Table 2.4.3. Sulfur dioxide has negligible positive correlations with daily attack rate. Correlations for TSP are positive in all sites and negative for SN and minimum temperature in all sites. Suspended sulfates are negatively correlated with daily attack rate in three of four sites. These correlations indicate that SS, SN, and minimum temperature are usually low when daily attack rates are high. Quite the opposite was observed in New York (Table 5.4.3.) where SO_2 shows negative correlations in two of three sites, SN is positively correlated in all sites, and SS is positively correlated in two of three sites. While these conflicting correlations may have descriptive value, the significance levels in Tables 2.4.3. and 5.4.3. are untrustworthy since daily attack rates are statistically dependent.

This remark applies to the multiple regression analysis (Tables 2.4.4. and 5.4.4.). Researchers at EPA and collaborators have recognized that both the daily attack rates and pollution levels are time series with autocorrelation structure, and they have begun to explore alternative analyses. Stebbings [16] noted that autocorrelations affect multiple regression parameter estimates and mentioned the need for methods which could detect lagged relationships between pollutant and attack rate.

Bloomfield [17] responded by suggesting an analysis of spectral coherence to relate attack rate and pollutant level time series Hasselblad [18] recently suggested methods based on first order Markov structure, and French et. al. [19] recently used an informal but sensible analysis to relate pollutant combinations to attack rate.

Even if daily attack rates were statistically independent, so that the significance levels in Table 2.4.4 had meaning, one would need to exercise caution when inferring that particular pollutants are associated with daily attack rate. This is because daily pollutant

---

[15] Utah CHESS Progress Report, op. cit.

[16] Stebbings, J. H. Some Problems In the Design and Analysis of Panel Studies In Epidemiology. Seminar at Princeton's Department of Statistics. October 16, 1975.

[17] Bloomfield, P. "Spectrum Analysis of Epidemiological Data." Updated reprint from the Department of Statistics, Princeton University.

[18] Hasselblad, V., "Analysis of Panel Studies." *EPA Internal Draft (June 1976)*.

[19] French, J. G., Hasselblad, V., Sharp, G., Truppi, L., "A Study of Asthma in the Los Angeles Basin: 1972–1973." *EPA Internal Draft (June 10, 1976)*.

levels are correlated (Table 2.4.3.). Ridge regression methods [20] applied to the Salt Lake data to investigate this "multicollinearity" of pollutants show that temperature adjusted SS effects are stable. This result tends to support the hypothesis that SS may be important, but formal statements of statistical significance are not justified.

To compute the relative risks in Table 2.4.5. (and 5.4.5.), the data from all days with minimum temperature between 30–50° F and integrated SO_2 less than or equal to 60 $\mu g/M^3$ were pooled, and the total number of attacks was divided by the total number of persons-days to obtain the rate 16.9 (Table 2.4.5.). Ratios of such numbers were termed "relative risks." The statistical properties of these relative risks require further study, since the days contribute statistically dependent data. Also no consideration is given special combinations of pollutants since the same days are used (except for days with missing pollutant data) for each pollutant. The only large increases in relative risk occurs on days with high SS and high minimum temperatures, but these relative risks are based on numerators with few person-days exposure, and the corresponding increases in New York (Table 5.4.5.) are unimpressive.

The "hockey stick" model [21] used to determine threshold is a three parameter non-linear model (attack rate intercept, threshold, and suprathreshold slope) which the authors fit by least squares. Confidence limits are not given for the estimated slopes and thresholds; nor are these appropriate since the daily attack rates are dependent and possibly heteroscedastic (i.e., with different variances). The Monograph gives no analysis of residuals, nor is the original data plotted on the fitted graphs (Figs. 2.4.5.–2.4.7.). Thus, the reader cannot assess whether the "hockey Stick" model fits the data or whether the data suggest that a threshold even exists. One such scatterplot of SO_2 data does not support the concept of a threshold [22] and it is likely that uncritical use of the hockey stick function has led to estimates of non-existent thresholds on occasion.

The New York asthma studies had similar practical and theoretical difficulties as documented in the 1970 Biannual Report.[23] Figure 12 of that Report shows 74 of 80 mailed diaries were useable, of which 14 required telephone prompting or interviewing. Eleven of 43 Bronx panelists dropped out during the study whereas the rates for Riverhead (8/38) and Queens (11/52) were smaller (Figure 10 of the Biannual Report [24]). Superimposition of the Figures 5.4.1.–5.4.4. shows that Riverhead often has lower weekly attack rates than Queens, but such data is not given for Bronx. Preliminary analysis of 1971–1972 New York asthma panels using the same statistical methods as in the Monograph suggest that SN is a more important risk factor than SS. The Monograph gives no relative risk calculations for SN, and SN was discounted in correlation and multiple regression analyses in 1970–1971 Salt Lake and New York studies. Relative risk tables from the 1971–1972 New York study confirm the 1970–1971 associa-

---

[20] Smith, K. J. C., "A Ridge Regression Analysis of the CHESS Studies; Salt Lake Asthma." Undated EPA Internal Draft.

[21] Hassalblad, V., Creason, J. P., Nelson, W. C., "Regression Using Hockey Stick Functions." EPA-600/1-76-024 (June 1976).

[22] Demill, K. M., Wride, W. P., "Technical Memorandum 76-2." Kennecott Copper Corporation, Operations Research, Salt Lake City.

[23] Pravda, M., "Community Health Effects Surveillance Studies: New York City." Internal Biannual Report to EPA (December 1970).

[24] Ibid.

tion of SS with daily attack rate, but the effects of SN cannot be disentangled since SN is highly correlated with TSP, RSP (respirable suspended particulates), SS, and SO_2 in the 1971–1972 study. Comparison of the conclusions of these two New York studies highlights the difficulty of associating a specific pollutant with asthma attack rate, but both studies show attack rate is associated with daily urban "pollution."

These panel studies suffered from several practical and theoretical deficiencies. Panelists were not examined by a physician and eligibility requirements were somewhat obscured by a "priority" system of secondary selection. Numerous factors known to precipitate asthma attacks were ignored. The analyses failed to consider autocorrelation in the asthma attack rate and daily pollution levels, and the authors did not present spectral or lag analyses. The significance levels presented are untrustworthy, since the attack rates are dependent. The statistics used have descriptive value, but the patterns of correlation between pollutants and attack rate are strikingly discordant in Salt Lake and New York. It is disturbing that Ogden (Low) often has higher attack rates than Magna (High), and this observation requires further investigation. Although SS may well precipitate asthma attacks, the hockey stick model used to estimate thresholds was not shown to fit the data. Tentative findings of the 1971–1972 New York asthma study highlight the difficulty of relating asthma attack rate to a specific pollutant.

This commentary emphasizes limitations of the analytical methodology and problems in data collection because these topics are scarcely mentioned in the Monograph. The panel strategy is nonetheless an appropriate tool for studying the effects of pollution on selected high risk populations, and the statistical methods used provided considerable insight in exploring possible relationships between pollutants and asthma attack rates. The fact that time series methods were not used or that theory could not provide formal inference (without further research) for statistics such as the "relative risk" should not obscure the descriptive and exploratory merits of the methods used.

5. CARDIOPULMONARY SYMPTOMS IN ADULT PANELS

These 1970–1971 New York studies attempted to associate symptom aggravation in high risk elderly panels with daily pollutant levels. Since these studies were similar to the asthma panel studies, they shared similar practical and methodological problems. Figure 14 of the Biannual Report [25] shows that recruitment varied by community. Riverhead contacts came from trailer parks and from recreation and golden age clubs whereas Queens and Bronx participants were contacted on park benches or through the New York Housing Authority. It is hardly surprising, therefore, that Riverhead (Low) participants were generally younger and healthier than those from Queens and Bronx (Table 5.5.A.1.) Candidates were interviewed and questionnaire information was used to classify them as well, heart, lung, or heart-lung panelists. No previous or concurrent physician diagnosis was specified in the 1971–1975 protocols (although the Monograph summary on page 5–104 mentions physician diagnosis). Eligible candidates had to reside within 1.5 miles of a monitoring site and to

---

[25] Ibid.

be 60 years of age or more. In later studies, younger panelists were accepted and some candidates were found outside the 1.5 mile monitoring radius, especially in Riverhead.

Substantial numbers of enrollees dropped out during the study, many because they were too sick, had difficulty filling out the weekly diary, or just found the study bothersome (Figure 16 of Biannual Report [26]).

The Monograph states that dropout rates exceeded 50% in Bronx and Queens and approached 30% in Riverhead.

Many panelists failed to return weekly diaries. Figure 15 of the Biannual Report [27] shows that only 49 of 86 Riverhead panelists provided useable diaries for the week November 22–28, and corresponding fractions for Bronx and Queens were 38/99, and 9/59. The fraction of useable diaries for the heart, lung and heart-lung panels was usually even smaller. Thus Bronx, Queens and Riverhead yielded respectively 24/47, 18/40 and 10/13 heart-lung diaries, 14/31, 6/21, and 20/26 lung diaries and 26/59, 19/52, and 27/39 heart diaries. Mailed diaries were replaced by telephone interview in the 1973–1974 protocol.

The high dropout rates and poor weekly diary response rates limit the generality of these studies since the data was generated by a self-selected population. This fact also complicates community comparison.

Panelists were asked to specify daily whether symptoms were "worse," "much worse," "the same," "better," or "never present" and daily aggravation rates were computed as the sum of "worse" plus "much worse" divided by all participants that day. (The "much worse" category was eliminated in subsequent protocols.) Table 5.5.2 shows that Bronx and Queens aggravation rates are usually higher than Riverhead rates, but the authors do not ascribe these differences to pollution since Riverhead is known to contain younger, healthier panelists.

Figures 5.5.3–5.5.6 relate composite weekly symptoms aggravation rates to weekly SO_2, TSP, SS, and SN for the Riverhead and Queens heart-lung panels. The aggravation rate seems to rise and fall with SS levels. No such graphs are given for Bronx nor for well, heart and lung panels.

Table 5.5.4 shows that SS is positively correlated with daily symptom aggravation in the well and heart-lung panels, but the stated significance levels are untrustworthy since daily aggravation rates and pollution levels are each autocorrelated. Multiple correlation analysis confirms that temperature adjusted SS is associated with symptom aggravation, especially in the heart-lung panel (Table 5.5.6), but these significance levels too can only be taken as qualitative indicators. Because the daily pollutant levels are highly correlated (Table 5.5.5), specification of a particular pollutant for particular concern is speculative. However, it is impressive that other pollutants have little explanatory power after temperature and SS adjustment whereas SS still explains substantial variation in daily aggravation rate after adjustment for temperature, SN, TSP, and SO_2. Table 5.5.3 shows that daily SS was strongly positively correlated with shortness of breath aggravation in six panels and strongly negatively correlated in none. These figures are derived from Table 5.5.A.3

---

[26] Ibid.
[27] Ibid.

which shows that the correlation was strongly positive in all three heart-lung panels, in two of three heart panels and in one lung panel (Queens). Temperature specific relative risk calculations (Tables 5.5.7–5.5.9) also suggest that SS is associated with increased symptom aggravation rates, especially in the heart-lung panel on warm days. Such a table is not presented for SN. All these methods suggest an association between SS and symptom aggravation rate in heart-lung panels. It would be interesting to know if the SS data indeed exhibited a threshold effect and fit the hockey stick model in Figure 5.5.8.

The severe dropout rates and low weekly diary returns (especially for the heart-lung, heart and lung panels) raise important questions as to the type of respondent and the meaning of the response "worse." The statistical methods have the limitations mentioned for asthma panels. Moreover, a subsequent New York 1971–1972 study [28] failed to confirm the dominant importance of SS but did find that SS, SN, SO_2 and RSP were associated with symptom aggravation of the well panels. (This study did not use the same analysis as the Monograph and, in particular, did not include temperature adjusted multiple regression or temperature specific relative risk calculations.) Nonetheless, the 1970–1971 data suggest an association between symptom aggravation rates and various pollutants, especially suspended sulfates.

6. VENTILATORY FUNCTION IN SCHOOL CHILDREN

The Cincinnati study of ventilatory function in school children preceded CHESS. Children from schools in an industrial valley of Cincinnati were compared with children from schools in a non-industrial river valley on the east side of the metropolitan area. Two upper-middle white, lower-middle white and lower-middle black schools were selected from each valley. Air monitoring stations within three blocks of the schools showed that seven month average TSP values were from 18 to 32 $\mu g/m^3$ higher in the industrial valley than in the non-industrial valley, but corresponding differences for SS, SN and SO_2 ranged from 0.1 to 1.1, 0.1 to 0.8, and 0.6 to 10.4 respectively (Figure 6.1.1). Thus, the industrial valley had more TSP than the non-industrial valley, but its levels of SS, SN, and SO_2 exceeded those in the non-industrial valley by very small margins. Ventilatory function was measured as the forced expiratory volume at three quarters second (FEV .75), and height, sex, and race were noted to make adjusted FEV .75 comparisons.

The study was confined to 394 second graders who participated in weekly measurements during November 1967, February 1968 and May 1968. These students represented 93 percent of second graders in the classrooms selected. Mothers were interviewed to obtain socio-economic data. The educational attainment of fathers was similar for corresponding schools in the industrial and non-industrial valleys (Figure 6.1.2).

Figure 6.1.3 shows that average height adjusted FEV .75 in "clean" schools exceeded that in "polluted" schools in all three months for lower-middle whites and in two of three months for upper-middle

---

[28] Stebbings, J. H., Hayes, C. G., "Panel Studies of Acute Health Effects of Air Pollution I. Cardiopulmonary Symptoms in Adults." Environmental Research Volume 11 (1976), pp. 89–111.

whites. Blacks had consistently lower FEV .75 values, and a "pollution" effect was only seen among blacks during one of three months. The absolute differences in average FEV .75 were roughly 40–120 milliliters (or less than 10%) in most cases. Table 6.1.5 presents a multivariate analysis of variance which allows one to test for community effects adjusted for a possible month effect and for the covariates height, sex, race, and social class. The dependent variable for each child was his vector of three monthly average FEV .75 values. No test for a community effect (industrial valley versus non-industrial valley) was performed, but the significance levels given suggest that community differences exist. The conclusion that "suspended sulfates exerted the strongest influence" (pages 6–7) is unwarranted in view of the high correlations commonly found among these pollutants. Nonetheless, these studies support the notion that FEV .75 is 3–10 percent less among white second graders in the industrial valley than among those in the non-industrial valley.

The 1970–1971 New York ventilatory study included children ages 5 to 13. Schools were situated within 1.5 miles of CHESS air monitors, and Riverhead, Bronx, and Queens were represented by three schools each. Only white children were eligible for analysis.

A new electronic spirometer, the National Gas Cylinder (NGC), exhibited serious drift (perhaps 350 milliliters). This problem was detected by field teams who had difficulty reproducing measurements and who made several trips to the manufacturer for adjustments. The June 1971 Progress Report [29] gives an account of difficulties with the spirometer, and the Assistant Commissioner of the New York City Department of Health expressed serious reservations about the reliability of those measurements in a January 18, 1972 letter to the Acting Chief of the EPA Epidemiology Section.[30] It is possible that drift of this magnitude biased the results, since one community may have been systematically subjected to a spirometer with extreme drift (especially if drift varied in phase with the rotation of spirometers through communities). Even if rotation of the spirometers among communities distributed drift equally, the variability of the observations is greatly increased by random distribution of drift, since the community effects (60 milliliters or less) are much smaller than the drift. The NGC spirometer was replaced in late 1971 by a bellows-type spirometer described in the May 1972 CHESS protocol. However, the NGC spirometer was used throughout the 1970–1971 study reported in the Monograph.

The 1970–1971 protocol instructs investigators to "deduct one inch for stacked heels for those girls wearing such shoes." In the 1972–1973 protocol this height adjustment was reduced to ½ inch.

No data were presented to compare families of children studied in Riverhead, Queens, and Bronx, but the age distribution and parental smoking habits were said to be similar. Income and educational attainment were said to decrease in the order Queens, Riverhead, Bronx. The Monograph contains no comparison of children's smoking habits in these communities; this covariate may be relevant since statistically significant differences were found only for older children.

---

[29] "New York City CHESS Progress Report." (June 1971).

[30] Bergner, Dr. L. (Assistant Commissioner for the Department of Health, New York City), Letter to Dr. Douglas L. Hammer, Acting Chief, Epidemiology Section, EPA (January 18, 1972).

Figure 5.6.1 shows that male height and age adjusted FEV .75 values from Riverhead were intermediate between Queens and Bronx values for three of four test periods. For females, the Riverhead values exceeded Bronx and Queens values in each test period, but the differences were usually less than 50 milliliters. Table 5.6.5 show that Riverhead height adjusted FEV .75 values were largest during one of four test periods for young males, three of four for older males, one of four for young females, and three of four for older females. Thus these differences were inconstant, and even for older females, the average difference between Riverhead and Queens was only 42 milliliters (about 2.4%). Table 5.6.6 is puzzling since it includes a degree of freedom for "ethnic differences" not described in the text. Perhaps blacks were included in this analysis. The multivariate analysis of variance used in Table 6.1.5 is more appropriate than averaging over all test periods as in Table 5.6.6. However, the analyses for individual test periods in Table 5.6.6 show statistically significant differences for older males and females.

The Cincinnati study demonstrated small (less than 10%) differences in FEV .75 between white children in schools located in a non-industrial valley and those in schools in an industrial valley. For New York, small, inconstant differences were found between Riverhead and Queens-Bronx only for older (9–13 years old) children. Preliminary analysis of 1971–1972 New York CHESS winter and spring data obtained with a bellows-type spirometer shows that older Riverhead children had statistically significantly higher FEV .75 values than Bronx children in winter and statistically significantly lower values in spring.[31] In view of these findings, the problem of spirometer drift, and the possibility that older children in Bronx and Queens smoke more than in Riverhead, one must be cautious in linking New York pollution effects with ventilatory differences.

7. OTHER CHESS STUDIES

CHESS has conducted two types of studies not found in the Monograph. The "Episode Study" measures rates of discomfort (cough, shortness of breath, restricted activity, and eye, throat and chest discomfort) during control periods and during air pollution episodes. A 1973 report[32] shows higher rates of cough, chest discomfort and restricted activity during episodes of high air pollution. "Pollutant Burden" studies are designed to quantitate the levels of metals and other pollutants found in human tissues and to relate these levels to exposure history. Scalp hair was collected from families participating in ARD studies, maternal hair, blood, cord blood, and placental samples were obtained from willing mothers, and various tissues were obtained at autopsy from pathology laboratories.

D. Summary Assessment of the Population Studies

This assessment is made from the point of view of an epidemiologist. More global judgments, made in consideration of the quality of both epidemiologic and aerometric data, are found in Chapter III.

---

[31] Chapman, R. S., Hasselblad, V., Burton, R., Williams, J., "Air Pollution and Ventilatory Function." *EPA Draft* (April 8, 1975).

[32]†Nelson, C. J., Shy, C. M., English, T., Sharp, C. R., Andleman, R., Truppi, L., Van Bruggen, J., "Family Surveys of Irritation Symptoms During Acute Air Pollution Exposures." Journal of the Air Pollution Control Association, Volume 23, Number 2 (February 1973), pp. 81–90.

No formal methods are used to link specific pollutants with specific health effects in the CRD, LRD, ventilation, and ARD studies. If a demonstrated health difference between communities cannot be explained in terms of imbalances in known covariates, it is generally ascribed to pollution. The CRD studies convincingly demonstrated small increases in the prevalence of productive cough in polluted urban and smelter sites. The retrospective LRD studies may have demonstrated higher croup and bronchitis rates in children in polluted smelter communities, and the prospective ARD are rather more convincing in demonstrating an increase in acute respiratory disease in polluted urban settings. The Cincinnati ventilation study demonstrated small differences in FEV .75 between white second graders attending schools in industrial and non-industrial urban valleys. If the aerometric measurements had been entirely accurate and reproducible during these studies, it might be safe to conclude that health effects had been demonstrated in association with known patterns of pollution. Unfortunately, this was not the case. It is not possible to know which specific pollutants, if any, or what concentrations of any suspect pollutants, were responsible for the health effects. The health effects data provide at most a rough guide for making general judgments about probable health effects in other communities with similar pollutant sources, meteorology and population composition. However, even though the aerometric data are of poor quality, the studies appear to demonstrate some health effects of air pollution at concentrations which might possibly be moderate (lower) compared with some high urban levels of the early 1960's. In this light, the data in the CHESS study provide no support for those who would significantly relax current standards; on the other hand, neither do they support a need for more stringent standards. Thus, the CHESS studies have proved to be valueless for assessing or supporting setting of quantitative regulatory standards which was a goal of the program.

The methodology used in the panel studies (asthma and cardiopulmonary) attempts to disentangle the effects of the several pollutants. The multiple regression and relative risk calculations are interpreted as implicating suspended sulfates, particularly in the Salt Lake asthma and New York cardiopulmonary studies. While this interpretation appears reasonable, these formal methods do not provide logically compelling evidence that SS, or indeed any of the measured pollutants is of dominant importance. Despite the many practical and methodological difficulties of these studies, they support the belief that asthma attacks and cardiovascular symptoms occur with greater frequency when the air is polluted.

These findings may appear overly critical especially since CHESS pioneered many research techniques. However, this report complements the CHESS Monograph, which often fails to apprise the reader of practical or theoretical problems encountered in this research. These remarks are meant to aid in the assessment of the validity of the conclusions presented in the Monograph and to assist researchers performing similar studies and encountering similar difficulties. This endeavor was greatly assisted by hindsight and by the splendid cooperation and self-examination of investigators both inside and outside the Environmental Protection Agency.

E. CHESS Current Status

1. GENERAL

In the process of interviewing EPA personnel concerning CHESS background, a good deal of information was obtained as to present and future plans for the program. Many other program and management ideas were also offered by the EPA personnel.

The CHESS program no longer officially exists within EPA except as a historical reference point. Administrator Train on April 9th committed to have all data from the program analyzed by mid-1977. However, the effort within the Health Effects Research Laboratory directed at processing the several years of data remaining (1972–75) unanalyzed from the CHESS studies is progressing slowly. A forced change of computers (see below) caused major delay. In addition, the investigative team sensed ambivalence as to how much effort is justified. Discussions with EPA Headquarters personnel indicated that, in keeping with the Administrator's commitment, the analysis is of sufficient importance in its own right as well as to continuing programs to warrant early completion. Yet, from discussions at Research Triangle Park, it can only be concluded that many would rather spend resources initiating new projects or refining CHESS methodology. Perhaps this is just as well, since there is serious doubt that the analysis even when completed will ever be sufficiently credible to support the stated objectives of the program. Any results from the final CHESS analysis may be examined with suspicion since the doubt concerning the reliability of aerometric measurements and the accuracy and acceptability of health measurements which evolved following the first CHESS report will affect any further reports from this same program. The past loss of experienced personnel has impaired the long term research effort required by CHESS objectives and has hurt morale among the remaining researchers.

The investigative team found that the disorganized state of affairs described above is compounded by the fact that the Health Effects Research Laboratory still lacks key personnel in the health research field. The Laboratory Director has been actively seeking persons to fill several key positions, but without success.

Several personnel problems continue to frustrate members of the Population Studies Division, of the Health Effects Research Laboratory. An impending retirement leaves the Division without leadership in a period when long range plans are to be reformulated. Members of the Division wonder, in view of recent investigations including this one, whether yet another reorganization is soon to come, and many regard this organizational instability as an impediment to sound work and career development. Most believe that Population Studies staffing is inadequate in relation to program funds. This disproportion requires that most real work be done by contractors. Not only are RTP staff deprived of the incentive of planning and executing their own studies, but they lack time and travel funds to monitor the contracts satisfactorily. The Population Studies Division responds to special regulatory needs (e.g. sulfate studies) and to unforseen events such as the demand for assessment of Kepone effects in Hopewell, Virginia. While these situations and demands might form the basis of a unified program in the future, at present they are regarded as "fire-fighting" distractions.

The team found strong evidence of a continuing lack of coordination between field and Headquarters and insufficient coordination among regional programs, Environmental Research Center Laboratories, contractors, other Federal agencies studying health effects of air pollutants, and Headquarters, not only in total program objectives but also in program coordination of quality control, sample collection and analysis, and standardization of health effects data collection tools.

2. STATUS OF ANALYSIS

Schedules [33] for analysis of CHESS data show that CHESS health data were collected from 1970 through 1975 for asthma panels, cardiopulmonary panels, and ventilatory function panels. The last surveys for acute respiratory disease (ARD), and chronic respiratory disease (CRD), and acute episodes took place in 1973, 1974, and 1975 respectively. CHESS surveys were completed for 25 asthma panels, 12 cardiopulmonary panels, 28 ventilatory function panels, 15 ARD, 12 CRD, 8 episodes and 8 acute lower respiratory disease panels during this period. The Monograph reported on only eight of these panels conducted in the 1970–1971 New York and Salt Lake studies. (The Rocky Mountain, Chicago, and Cincinnati studies contained in the Monograph were not CHESS studies.) Thus, CHESS generated only a portion of the data analyzed and reported in the 1974 Monograph, and the Monograph reports on only a small fraction of CHESS data.

The present analysis of the studies proceeds in steps. First, the survey data is edited for computer processing and subjected to preliminary statistical analyses. Then an epidemiologist writes a draft on the basis of the preliminary statistical analyses of health and aerometric data. Other analyses may be done if judged necessary by the epidemiologist or statistician. EPA staff have tabulated the fraction of surveys subjected to preliminary statistical analysis and the fraction for which a preliminary draft has been written.[34] Drafts exist for 6 of 8 episode and 7 of 8 LRD panel studies. Drafts have been written for 7 of 15 ARD and 8 of 12 CRD studies, but no CRD drafts have been written for survey years 1973 or 1974. By comparison, asthma, cardiopulmonary and ventilatory function analyses are far behind schedule. Of 28 ventilatory surveys, 10 have undergone preliminary analysis, and 8 drafts have been written (as of June 28, 1976). Of 25 asthma surveys, nine have undergone preliminary analysis and nine drafts have been written, and of 12 cardiopulmonary panels, two have been analyzed and two drafts written. These figures substantiate the degree to which the CHESS program has fallen far behind in analyzing and writing up survey data (61 data sets are yet to be written up).

The figures quoted above indicate that a major bottleneck in the analysis of survey data has been the preliminary editing and statistical processing. It should be noted that EPA had little or no control over two disruptive events which retarded preliminary analyses by a year or more. GSA ordered a change of computer from IBM to Univac in August 1974 as part of a general government policy related to competitive bids. Both hardware and software conversions were difficult, and most programs had to be completely rewritten. Compounding

---

[33] Unpublished tables giving dates of CHESS data collection prepared by Dr. Dorothy Calafiore, EPA/RTP.

[34] Unpublished tables showing chronology of survey data analysis and draft writing prepared by Drs. Carl Hayes and/or Kathryn McClain (June 28, 1976), EPA/RTP.

this problem was the need to rely increasingly on outside data processing contractors, since increases in EPA staff needed to meet the increased data load never materialized even though verbal commitments had been given to the laboratory directors. GSA contracting procedures required two changes in data processing contractors between 1972 and 1974. Other problems impeded the preliminary statistical analyses. The loss of key personnel imposed new scientific and administrative burdens on remaining statisticians and epidemiologists. Moreover, responsibilities of the Statistics and Data Management Office were broadened during this period to include projects with the Toxicology and Clinical Studies Divisions as well as special tasks, such as preparing data tapes in reponse to outside requests for EPA data. CHESS analyses (Population Studies Division) now account for about half the Statistics and Data Management Office work load. However, the most important delays in the preliminary analysis are attributable to the forced computer change and to problems with outside data processing contractors.

The epidemiologists charged with writing a draft monograph must await receipt of aerometry computer output and biometry output. Epidemiologists interviewed attributed delays to the failure of these preliminary analyses to arrive and to subsequent changes in aerometric values used in the manuscript. Other problems retard the epidemiologists once preliminary analyses are in hand. To begin with, the Population Studies Division has lost many of the important early authors or researchers within CHESS. It is hard for an epidemiologist, not familiar with the background and history of the data collection and faced with two computer outputs, to assess the quality of the data and to obtain specific scientific guidance.

Morale is low, and this may be due in part to the loss of key personnel, to investigations of the CHESS program, and to organizational instability affecting Population Studies. Furthermore, during the past year, epidemiologists have been encouraged to pursue competing scientific interests, and they have had to shoulder administrative burdens associated with increased reliance on external contractors and to meet special tasking such as the Kepone Task Force.

It was interesting to find that several epidemiologists interviewed complained that the current team structure for analyzing survey data does not function well without a strong authority to resolve scientific and scheduling conflicts between epidemiologists and statisticians. While these factors and lack of staff have not seriously impeded the drafting of CHESS reports until now, they may become pressing problems as the backlog of preliminary analyses clears the new computer system in the near future.

The continued delays in analyzing CHESS data have had several adverse effects. Contractors gathering survey data and panelists have been embittered and demoralized by lack of data and feedback from EPA. The failure to provide contractors with timely aerometric and health data summaries stifled local initiative in analyzing the data. Thus, the CHESS program did not benefit fully from local field level insights as to local pecularities of the survey and populations or from comparisons with independent local air pollution measurements and data analysis.

3. RECENT PLANNING DIRECTIONS OF THE POPULATION STUDIES DIVISION

The following comments are largely speculations and impressions based on interviews with EPA staff.

CHESS data acquisition ended in 1974–1975 when OMB did not approve questionnaires needed for the study. Had OMB approved the forms, however, it is likely that 6 asthma, 6 ventilatory function, 2 LRD and 2 CRD surveys would have taken place in 1975–1976.[35] Instead the Populations Studies Division worked on old CHESS data and made several proposals for non-CHESS studies to be run on contract. One of these is a cohort study of ventilatory function in adolescents which uses a CHESS endpoint (measured variable) but does not use a cross-sectional CHESS replicate population. Another uses the CHESS ARD endpoint to study the effects of peak NO_2 values. One study entitled "A study to Determine the Health Effects Associated with Emissions from Coal Combustion and Coal Gasification" uses the CHESS ARD and FEV .75 endpoints to study the impact of coal combustion and gasification in local populations (Hair, blood, and urine samples are also obtained to assay for products of combustion and gasification). Other proposed studies do not use CHESS endpoints. These include: (1) a study of mortality patterns in Montana, where excessive cancer mortality has been noted, (2) a general household health survey given to members of a community located near a new sewage treatment plant, (3) a study of trace metal burdens found in the hair, blood, and urine of people living near non-ferrous smelters, and (4) an evaluation of hyperactivity and cognitive function in children with asymptomatic lead burdens. These proposals build on CHESS experience, but they do not embody any continuation of CHESS surveillance. However, only one study plan and questionnaire was approved by OMB at the time of this investigation. The continued resistance by OMB to approve new study forms and questionnaires is a great frustration to the researchers. There was a strong feeling evident that this resistance stems from a misguided assessment of scientific merit as much as from a Presidential concern for the rights of privacy.

The Scientific Advisory Board (SAB) has fallen short of EPA staff expectations. The staff members had hoped to work closely with SAB to formulate and refine new proposals. Instead, the SAB has provided only formal criticism, expressed in public meetings attended by press and industry. This is a time when the EPA staff needs new leadership from a Director of Population Studies and from close informal working relations with the SAB.

---

[35] Table of proposed CHESS studies for 1975–1976 prepared by Dr. Carl Hayes, EPA/RTP.

VII. CURRENT CHESS STATUS AND FUTURE PROGRAMS

A. Retrospection

It is always easier to look at the past and detect mistakes than it is to look forward and avoid them. There are lessons which can be learned from the research which produced the CHESS Monograph as well as from the efforts to process the balance of the CHESS data which remain unanalyzed and unpublished.

This investigative report supports much of the criticism of the CHESS Monograph as developed by other earlier and independent examinations. For example, a report prepared for the Federal Energy Administration in 1975 [1] describes the use of the unvalidated version of the British Medical Research Council questionnaire as limiting the reliability of the health effects analysis in programs utilizing this health indicator. The Subcommittee's investigation noted this same defect. Questions regarding the analysis of the data on asthma attacks indicated serious discrepancies. In the report to FEA, however, the significance of the errors in the aerometric data which are detailed in the Subcommittee's investigation was not fully recognized. In their review of the CHESS Monograph, the EPA Science Advisory Board (SAB) identified many of the same problems noted by the investigative team, again primarily in the epidemiological areas. In the SAB, or so called Whittenberger, Report [2] other deficiencies included: limited population sampling control; shortcomings in the methods used for assessment of past air pollution levels; variations in sample populations which limit comparison of communities; high drop-out rates in asthma panels; inadequate consideration of temperature as an important variable in consideration of asthma attacks; and serious concern about the problem of gaining information on individual exposure to pollutant levels. All of these points and others are confirmed by the Subcommittee's investigation.

As noted in the legislative history of this report (Appendix B), other independent investigators also had offered criticism of the CHESS Monograph during the brief period of peer review which occurred prior to publication and this criticism also appears to be well founded for the most part. In separate talks with researchers who had either participated in some phase of CHESS or had reason to follow the progress of the CHESS investigations and review of data samples, the Subcommittee investigators had occasion to gain separate confirmation of many of these same concerns about the validity of the instruments, questionnaires, and methods used for collecting data as well as the methods of analysis.

---

[1] Preliminary Report. A Critical Evaluation of Current Research Regarding Health Criteria for Sulfur Oxides. Technical Report Prepared for the Federal Energy Administration. Tabershaw/Cooper Asso., Inc. April 11, 1975.

[2] Review of the CHESS Program. A Report of a Review Panel of the Science Advisory Board—Executive Committee, March 14, 1975.

77-590—76——7

Thus, the conclusions contained within Chapter III of this report are not exceptional in their general tenor. However, the serious errors in the aerometric data and the verified significant effect of the poor quality control during the early stages of the CHESS air monitoring, measurements and analyses efforts are identified more precisely for the first time by an outside investigation.

B. Current

Moving from the historical perspective, an appropriate question to be asked is whether the errors suggested in other critiques and confirmed in greater detail by the investigative team are being adequately addressed in the current and planned Environmental Protection Agency's air pollution/health effects programs.

It was of interest, first, to determine the approximate status of the CHESS data which remain unanalyzed and to ascertain whether these data will suffer from the same or other discrepancies as the data presented in the 1974 CHESS Monograph. As is documented at the end of the last chapter (VI), in discussions with EPA personnel at Research Triangle Park, the investigative team determined that there still remains a considerable amount of analysis to be completed. The impressions gained from these interviews was that it would be at least until the summer of 1977 before the final assembly of data and examination for potential analysis and publication would be possible.

However, as noted earlier in the report (Chapter VI), a number of draft studies were completed at the time of the visits to RTP. For example, draft reports have been prepared for 6 of the 8 remaining episode type health studies and for seven of the 8 remaining lower respiratory disease studies. Seven of the remaining acute respiratory disease studies have been prepared and 8 of the remaining 12 chronic respiratory disease studies are in hand. No draft of the data on chronic respiratory disease studies appears to be ready for the survey years 1973 or 1974. Analyses of a number of other health effects studies have not been initiated.

The general impression is that there is still considerable effort required for the basic transfer of field data to computer tapes for analysis. Unfortunately, serious reservations are retained by the investigative team about some of these data since some of the same criticisms about the techniques used to collect the health effects information in the CHESS Monograph continue to apply to some of the data accumulated after 1971.

The establishment of a more vigorous quality control system appears to have improved the collection of aerometric data after 1974. The aerometric review (Chapter IV) points out that CHESS instrumentation, as used in the CHESS studies, did not consistently generate valid, reliable data in the range of the National Ambient Air Quality Standards. Thus, it is likely that no amount of effort on analysis of CHESS data would achieve the desired objectives. There is doubt that adequate meteorological support is being provided. Further, as noted in other sections of this report, there does not yet appear to be an adequate systematic planning effort to insure the full participation of all research components.

C. Future

Since sulfates were identified within the CHESS monograph as being a pollutant requiring an immediate increased priority of investigation, the investigative team attempted to determine the status of work on this pollutant. Mr. Train had indicated during testimony before the joint hearing on April 9, 1976, that a five year plan for study of sulfates was available. Actually, it was found during the investigation in the summer of 1976 that the sulfate plan was still in draft form although passing through a second iteration. There was some indication that even now there is not a full appreciation of the need for planning input with regard to quality control, air monitoring, instrument use planning, structuring of health effects data collection, processing of data, analysis and publication scheduling. The required level of cooperative preplanning does not seem to be present. In a letter to the chairman of the Subcommittee on the Environment and the Atmosphere (undated) in May, 1976, Mr. Train summarized EPA's projected research program on sulfates.

In measurement and monitoring, the emphasis in the draft plan is placed upon the development of reliable analytical and sampling techniques and instrumentation to support laboratory and field measurements of sulfates and sulfur compounds; the development of quality assurance programs; data audits, and providing technical support. The ongoing research and near-term research is aimed at increased data collection on sulfates with emphasis on measurement methodology, instrument development and field tests. Emphasis is to be placed also on the development of reliable models to predict sulfate concentrations and precursor emissions over long distances. The plan indicates a recognition of the need for instrument and field tests.

A continuation of the concept that more data on the health effects of sulfates are needed is reflected in an intention to expand emphasis on the toxicological, epidemiological and clinical studies on the effects of sulfates. The plan includes a recognition by EPA of the difficulty of conducting epidemiological studies, particularly in correlating ambient pollutant (sulfates) concentrations with health effects.

With the exception of the particular emphasis on sulfates, these objectives are essentially the same as the objectives cited as necessary upon the initiation of the original CHESS studies. If these objectives are to be achieved in the near future, it will be necessary for EPA to avoid the errors of CHESS. It is not clear either from the proposed programs or from discussion with individuals who will be implementing these programs that some of the basic errors committed during the CHESS studies will be corrected before such field, air monitoring, and epidemiological studies are continued with different pollutants. Certainly, the manpower problems have not been resolved.

In the aerometric area, the CHAMP aerometric measurement and analysis is entering a new phase of development. It is essential that the aerometric measurement and analysis errors detected during the CHESS time period be eliminated as the CHAMP effort is expanded and greater emphasis is placed upon this program. The CHAMP instruments, when verified, may be capable of identifying reliable differences in absolute air pollutant levels, but the instruments will

be pushed to their limits of sensitivity and will require constant attention by skilled technicians. New instruments just developed may be able to meet the ultimate requirements of the air pollution health effects program.

A sound program of development and testing (verifying) questionnaires and other approaches to epidemiological research also must be mounted and completed before another CHESS-type program is initiated. Completion of both phases of such a program to provide the tools and techniques for a sound health effects program might require up to 2 years and several million dollars. If this work is not carried out, progress in the area is doubtful.

APPENDIX A

A Recapitulation of the Aerometric and Meteorological Findings of the Investigation as They Relate to Specific Sections of the CHESS Monograph and the Health Findings

A. INTRODUCTION

This section contains citations of errors and omissions found in a careful review of the CHESS Monograph which show that the use of aerometric and meteorological data in correlation with health effects end point measurements can easily mislead the reader of the CHESS document into inferences which are not wholly or even partially supported by the data in the report. Page, paragraph, and figure references are to the 1974 CHESS Monograph.

Since an important application of the aerometric data is to determine correlations with health effects, any errors or overusage of aerometric data based upon estimates or improper measurements will obviously reduce or negate the value of any health effects correlations which are attempted. This misusage or overusage of aerometric data will be particularly damaging as the extension of the conclusions is made in an attempt to discover possible threshold effects.

B. CRITIQUE

1. Prevalence of Chronic Respiratory Disease Symptoms in Adults: 1970 Survey of Salt Lake Basin Communities

Observed concentrations for only one year have been used to crudely estimate concentrations of sulfur dioxide and suspended sulfates relating to a 4–7 year exposure. The 1971 observed annual average concentration of sulfur dioxide was used with the 1971 emission rate from the smelter to obtain a ratio that was then multiplied by emission rates for other years to estimate concentrations for the other years. The estimated sulfur dioxide concentrations were then used in a regression equation based on a 1971 relationship to estimate suspended sulfate concentrations. Possible changes in meteorological conditions and mode of smelter operations were neglected. Acknowledgment is not given in the discussion and summary that the critical concentrations relating to health effects are nothing more than estimated concentrations.

It is questionable whether or not long-term exposures should have been attempted for Magna, based on only one year's record of observations that are abnormal because of the smelter strike. It would certainly have been appropriate to have mentioned that only estimated long-term data were available and indicated their degree of uncertainty in the discussion and summary.

Further, we find many errors on Page 2–37, Table 2.1.A.14. It seems that this table should have never been included in the report. Aside from the misuse of the diffusion model (discussed in Chapter IV) this table lists suspended sulfate values for Magna for the years 1940–1970, that are not the same as listed in Table 2.1.A.16, on page 2–39. The values are estimated by a simple ratio from the smelter emission rates, but this is not explained. On page 2–39 a regression equation is used for the same purpose. All of the sulfate concentrations under the heading CHESS are estimated observations except those for the year 1971. This has not been properly indicated, e.g. by the use of parentheses.

On pages 2–37 emission rates are not sulfur dioxide rates as indicated but emission rates in tons of sulfur per day. This means that the sulfur dioxide emissions were twice the values listed. It also means that the dispersion model estimates are incorrect. However, the listed estimated concentrations in Magna and Kearns, which are based on a simple ratio between observed concentrations in 1971 and some emission rate for 1971, whatever it might be, are not changed.

Note that the regression equation for suspended sulfates, Salt Lake City, (pages 2–39) which is:

$$SS=0.101(TSP)-3.65$$

is quite different than that which can be obtained from Table 2.1.4, i.e.:

$$SS=0.065(TSP)+1.93$$

SO_2 exposures were derived by multiplying the yearly smelter emission of SO_2 by the ratio of the 1971 measured annual average SO_2 concentration to the 1971 SO_2 emission rate (193 tons/day).

Estimates of suspended sulfates were derived from the estimates of SO_2, using the following regression equation for 1971:

$$SS=0.09(SO_2)+6.66$$

The annual TSP exposures were derived by multiplying the yearly smelter production of copper by the ratio of the 1971 measured annual arithmetic mean TSP concentration to the 1971 copper production rate (260,000 tons/year).

Smelter emissions of sulfur dioxide in the early 1940's were roughly three times greater than they were after 1956 although copper production has remained more or less constant. The method for estimating suspended sulfate, which is based on sulfur dioxide estimates leads to very high values in the 1940's whereas the total suspended particulate are estimated lower in 1940 than in 1971. The procedure used produced very high ratios between SS and TSP for the earlier years. For example, the 1940 ratio (34.6/63) is 0.55. This ratio is so large that it is obviously questionable.

The audacity of the estimates can be seen in Figure 2.1.17. The lowest value, which occured in 1971, is extrapolated all the way back to 1940, reaching unusually high annual average concentrations of more than one part per million. Considering the effects of wind direction, which would result in low concentrations much of the time because the smelter stack plume would not be blowing toward the town, such an annual average would result in short-period concentrations many times

the annual average. It is questionable that such high concentrations ever occurred. If they did, they would be well-remembered, and living conditions in Magna would be different than in 1971. Such unreasonably high estimates should have been further investigated before being presented.

The grossness of the estimates made overrides other shortcomings in this study pertaining to exposure that might be mentioned. However, more carefully made estimates would have required considerably more work, including obtaining meteorological records and details of smelter operations affecting plume behavior over the period of years studied. Such a large effort may not have been worthwhile considering the inexactness of some of the other aspects of the study. Nevertheless, a study of this nature seems to call for actual observations, more accurate estimates, or considerably less exactitude in its conclusions.

2. Frequency of Acute Lower Respiratory Disease in Children: Retrospective Survey of Salt Lake Basin Communities, 1967–1970

The same comments apply to this study as for the preceding study on the prevalence of disease symptoms in adults. Inadequate recognition is given to the fact that only estimated concentration data are being used in the discussion and summary.

3. Aggravation of Asthma by Air Pollutants: 1971 Salt Lake Basin Studies

In this study, daily entries in a dairy were used to determine weekly asthma attack rates. A statistical relationship was then determined between the attack rates (weekly) and observed air pollution concentrations (averaged weekly). Participants lived within a 2-mile radius of air monitoring stations.

Daily exposure of asthmatics in a community such as Magna, which is close to the smelter, are poorly characterized by a single monitoring station. On a given day, one side of the community could be much more affected by the smelter stack plume than the other, and high concentrations from looping or fumigation might affect one neighborhood but not others. The study inadequately assesses the effects of peak exposures and episodes.

This report does not make clear that the minimum temperatures used were from the Salt Lake City airport. The assumption seems to have been made that temperature was uniform over the entire study area. This is not true because of the differences in elevation and the effects of the mountains, and the lake. Perhaps the differences were not important, but they should have been considered. It is not clear why days were stratified by minimum rather than mean temperature.

Minimum temperatures occur during the early morning when people are generally indoors and perhaps in bed. When temperatures are low, windows are generally closed. Also, lower minimum temperatures are correlated with other meteorological phenomena that could also affect asthma attack rates, e.g., lower humidity and lower wind speed. Further there may be a correlation with wind direction. A lower than average minimum temperature probably is also associated with a strong temperature inversion which would be conducive to lofting the smelter stack plume. Because of the many questions raised, the findings pertaining to temperature merely suggest further study and have no general application.

Near the middle of the left hand column, page 2–89, the following sentence appears. "The shut-down of operations by the strike was accompanied by a pronounced improvement in air quality and a reduction in asthma attack rates that occurred sooner and were larger than seasonal reductions observed in the more distant study communities some 2 weeks later." Here there is a lack of appreciation of the natural climatic differences that exit in the Salt Lake Basin. Some effects of summer weather could easily be delayed two weeks before reaching Ogden. The average date for the last killing frost in Ogden is about May 6, whereas the average date of the last killing frost at Saltair (the climatic station nearest to Magna with a long record) is about April 12.

On Page 2–76 (near middle of page, right hand column) the smelter is not "5 miles north of Magna."

On page 2–81 the first graph in Figure 2.4.1 is incorrectly drafted. After the 17th week the broken line should be solid and the solid line broken. The temperature curve should appear as in the graph for the high exposure community.

Figure 2.4.2, page 2–81, shows a weakness in the argument that the sulfur dioxide concentrations are responsible for the asthma attack rate. In the High Exposure Community the attack rate starts up at the 18th week as the sulfur dioxide concentrations approach zero, or near zero, and remain very low for about six weeks. It is noted that this same graph shows the highest SO_2 peak occurring at the 9th week, which seems to begin about May 9. The graph on page 2–16 seems to show the peak in April.

In Figure 2.4.4, page 2–82, with respect to the High Exposure Community, it may be noted that the sulfate concentrations are not particularly well-correlated with the sulfur dioxide concentrations plotted in Figure 2.4.2, on the preceding page. The highest sulfate reading occurs in the 3rd week, whereas the sulfur dioxide levels build up to a peak in the 9th week.

On page 2–87, left hand column, it is stated that a threshold concentration of 1.4 $\mu g/m^3$ was calculated for suspended sulfates for the higher temperature range. In Figure 2.4.4 all of the plotted concentrations are greater than this value. Considering the background of suspended sulfates generally observed, this low threshold value seems to have no practical significance.

The third paragraph that appears in the right hand column, page 2–89, probably applies to Magna, however, this is not made clear. There is a possibility that the paragraph could be given broader interpretation than actually intended since the last three sentences seem to refer to conditions in urban areas generally. The paragraph probably should have been divided into two separate paragraphs. However, the main fault with the paragraph is that important conclusions are drawn that are not supported by information presented elsewhere in the report. It says "excess asthma attributable to sulfur dioxide might be expected 5 to 10 percent of summer days", "total suspended particulates could occur on up to 5 percent of summer days and 30 percent of fall and winter days", and "excesses due to suspended sulfates are likely to occur on 10 percent of fall and winter days and 90 percent of summer days." Assuming that the stated relationships

between concentrations and temperature are true, the report does not explain how the percentages of days were obtained. The study covered only 26 weeks, but these conclusions apply to an entire year. The percentages given seem to be rough estimates since they appear to be given only to the nearest 5 or 10 percent. The percentages might have been obtained from daily values for the minimum temperature, pollutant concentrations and asthma attack rate; but it is not clear how they were obtained.

Presumably daily average concentration levels of specific pollutants were used in the construction of the "hockey stick" curves shown on pages 2–86 and 2–88. The discussion implies that "24-hour levels" were used, but the precise nature of the air quality data used in the threshold analyses is not made clear.

There could be various reasons not explored by the study why the thresholds for asthma attacks were lower on warmer days. One of these is that there may be more plume looping on warmer days. This might result in localized, short-period, high concentrations, but relatively low average concentrations.

The validity of scientific work can be tested by the repeatability of results. In this and the other CHESS studies there were factors affecting asthma attack rates that were not considered and whose effects are unknown. Such factors are: time spent outdoors, percentage of time windows are open, temperature change, relative humidity, etc. The incompleteness of the study and the lack of understanding of the causes of the asthma attacks suggest that it might be repeated with significantly different results.

Short-term exposures to concentrations much higher than average annual or weekly concentrations could have occurred in the communities studied that were near large sources of air pollution such as smelters. There exists the possibility that asthma attacks could be triggered by brief-duration high concentrations. Such exposures could have been determined only inadequately by the procedures used in the study. The report does not make clear why more attention was not devoted to peak concentrations.

4. Human Exposure to Air Pollutants in Five Rocky Mountain Communities, 1940–1970

On pages 3–7 through 3–12 beginning with the second column; paragraph near middle of page, which begins "By comparing . . .". There is not a simple relationship between average daily pollutant emissions and average annual pollutant concentrations because the receptor area is often now downwind. Also, some consideration should have been given to determining if the years for which data are available were representative meteorologically.

(Page 3–11) Second paragraph, left hand side of page. Information obtained during this investigation indicates that the ratio 1.63 ± 0.21 should be 1.42 ± 0.21. (The value 1.63 is the upper limit of this ratio.)

(Page 3–12) Emission ratios of particulate and sulfur dioxide for 1971 are omitted from this report. Therefore, it is not possible to verify the ratios given here.

(Page 3–12) According to information obtained during this investigation, the two values for TSP listed as 99.5 for the years 1971–70, should be 98.1 for both years.

5. Prevalence of Chronic Respiratory Disease Symptoms in Adults: 1970 Survey of Five Rocky Mountain Communities

Four communities were studied: Bozeman, Helena (Low), Helena (High), Anaconda, and Kellogg. These communities are all in mountainous areas and subject to terrain effects, some much more than others. Average pollution patterns are irregular so that exposures in a particular community will vary considerably. It is particularly important whether the community is somewhere downwind of a source, such as Magna, or on either side of it, as might be the case in Helena (High), or perhaps Kellogg. In these communities assumptions should not have been made about exposure without some effort of sampling and/or meteorological modeling to determine what the actual exposure might be.

In this study long-period concentrations of sulfur dioxide and suspended sulfates are estimated. Also, the most important finding, which was that excess bronchitis occurred with 2–3 year exposure to sulfur dioxide concentrations of 177 to 374 $\mu g/m^3$ and suspended sulfate concentrations of 7.2 to 19.9 $\mu g/m^3$ (in the presence of low total suspended particulates), covers such a wide range of concentrations it has hardly any practical significance.

On page 3–31, in the Summary is stated "Metallic sulfates may well have accounted for the findings of excess bronchitis." The nature of the sulfates is nowhere mentioned in the report. The smelter is apparently the source of only some of the sulfates, and the nature of none of the sulfates has been determined.

6. Frequency of Acute Lower Respiratory Disease in Children: Retrospective Survey of Five Rocky Mountain Communities, 1967–1970

Sulfur dioxide levels are crudely estimated from sulfation plate (or candle) data and smelter emission rates. Suspended sulfates are estimated from suspended particulate data using a ratio found in East Helena, Helena and Magna. It was assumed that the estimated values prevailed throughout the community. This is unlikely to have been true considering the topographic effects in the Rocky Mountain Communities. A major shortcoming in this paper is a failure to make clear the inaccuracies that might be associated with estimating the pollution levels.

There is an irregular distribution of annual average sulfur dioxide concentrations in the vicinity of a smelter. These differences affect concentrations in East Helena and can be confirmed by examination of a report on the Helena Valley Environmental Pollution Study.[1]

During 1965–66, 66 measurements of suspended sulfate indicated an average concentration of 7.9 $\mu g/m^3$. The observed total suspended particulate during this period of time was 96 $\mu g/m^3$. Hence the ratio of suspended sulfate observed to particulate during this period was .082. A series of 25 measurements made during the time period April to August 1968 indicated a suspended sulfate to total suspended particulate ratio of 0.057.

Then to determine a ratio range, measurements made during the period June through October at two stations of the Helena Valley study were used. The original information is as follows:

---

[1] Helena Valley, Montana, Area Environmental Pollution Study, Environmental Protection Agency Office of Air Programs, Research Triangle Park, North Carolina, January 1972.

Station	Location from stack		Suspended particulate ($\mu g/m^3$)	Suspended sulfate ($\mu g/m^3$)	Ratio
	Deg.	Miles			
1	34	0.8	108	3.5	0.032
3	112	.4	59	4.1	.069

The difference between these ratios is .037, as stated. Station One was the only station in the Study actually located in the town of East Helena.

These observations were made during the warmer months of the year when there would be generally be no frozen ground or snow cover possibly affecting the suspended particulate concentrations. It is not stated when the other observations used in the CHESS study were taken. However, from all of these data came the ratio of suspended sulfate to suspended particulate that was used for East Helena, which was 0.063 plus or minus 0.022.

The ratio of suspended sulfate to suspended particulate determined for Anaconda, which was 0.111 plus or minus 0.057, was determined by taking an average of the East Helena ratio (0.063 plus or minus 0.022) and the Magna ratio of 0.159 (Page 3–11). It is not made clear why this particular ratio was chosen.[2] From Table 2.1.2 a ratio of 0.178 might also have been determined, using a suspended sulfate concentration of 9.6 $\mu g/m^3$ and a suspended particulate concentration of 53.9 $\mu g/m^3$. It may also be noted that the ratio for Magna is much larger than for other locations. No explanation is given for including the Magna value in the average. From the applications of these ratios to estimates of suspended particulate concentrations came the conclusions pertaining to the health effects of sulfates. It is stated that "we conclude that excessive acute lower respiratory illnesses can be expected among asthmatic and nonasthmatic children who are exposed for longer than 2 years to elevated annual average sulfur dioxide levels (177 $\mu g/m^3$) accompanied by elevated annual average suspensed sulfate levels (7.2 $\mu g/m^3$) in the presence of low levels of suspended particulates (65 $\mu g/m^3$)." This statement refers to findings in Anaconda. No mention is made that concentrations of suspended sulfate might range from 3.51 $\mu g/m^3$ to 10.92 $\mu g/m^3$ because of the uncertain nature of the ratio alone.

The following statement from the Helena Valley Study (Page 45) is of interest. "These results (i.e., average concentrations ranging from 2.9 to 4.5 $\mu g/m^3$ for the period of the study) are considerably below the national average and serve to point out that there is no problem from sulfates. It should be noted that the sulfate content is evenly distributed among the five sampling stations, which indicates that there is no significant increase in concentration in the vicinity of the East Helena industry."

Suspended sulfate observations made by the Montana State Department of Health and Environmental Sciences suggest that average annual concentrations of 4 or 5 $\mu g/m^3$ occur in Anaconda [See discussion of Estimate Suspended Sulfate Concentration, 7.2 $\mu g/m^3$ (Anaconda), Chapter II]. Assuming that suspended sulfate concentrations of about 7 $\mu g/m^3$ did occur in the Rocky Mountain communi-

[2] Referenced publication contains no information pertaining to Magna.

ties where certain health effects were observed, the source of the suspended sulfate is inadequately determined. The study findings are much too incomplete to call for the stringent control of suspended sulfates as has been done on page 3–51.

7. Prevalence of Chronic Respiratory Disease Symptoms in Military Recruits: Chicago Induction Center (Paragraph 4.2)

Exposure estimates in this study are extremely crude. In the summary the following statement is made, "Available evidence indicates that exposures lasting 12 years or more to ambient air pollution characterized by elevated annual average levels of sulfur dioxide (96 to 217 $\mu g/m^3$), suspended particulates (103 to 155 $\mu g/m^3$) and suspended sulfates (14 $\mu g/m^3$) were accompanied by significant increases in the frequency of chronic respiratory disease symptoms."

The 96 $\mu g/m^3$ value is the average urban core value for 1969–70, which ranges from 54 to 138 $\mu g/m^3$, whereas the 217 $\mu g/m^3$ is an average value for five suburban communities for the year 1969. Going back 12 years concentrations were much higher. During the period 1960 through 1965, the lowest value was 222, and there was a high of 344 in 1964. For the five suburban communities there was data only for one other year. It averaged 183 $\mu g/m^3$. The 14 $\mu g/m^3$ concentration for sulfates is for a period of 7 years, not 12 as stated. It basically represents data for the Chicago core area, with some scattered observations from East Chicago and Hammond, Ind. The average concentrations for the city should be somewhat less than in the core area. Use of the core area value would generally result in an overestimate.

It is difficult to characterize exposures lasting 12 years for the entire Chicago area. Either this should have been done in very general terms, nonquantitatively, or a greater effort should have been made to present more representative estimates.

The assumption is being made that sulfate observations made at a central urban location in Chicago, averaged with a few observations from East Chicago and Hammond, Ind. are generally representative of the entire Chicago area.

(Page 4–8) Referring to the Chicago area the following statement is made: "Each sampler location, identified by a station name in Figure 4.1.2, represents the central business-commercial district of that particular area." This statement is not true. Practically all, if not all the samplers are located on the roofs of school buildings in an effort to obtain representative community values. They were not located deliberately in business commercial districts and do not slightly overestimate area-wide concentrations as suggested.

(Page 4–23) In reading this paper about the prevalence of chronic respiratory disease symptoms in military recruits, questions arise about the actual locations from which the men came and the local pollution levels to which they might have been exposed. Some rural occupations result in high exposures to dusts, plant allergens, etc.

(Page 4–35) (Summary) The 12-year value for suspended sulfates should be 16 micrograms per cubic meter, not 14, as stated. Also, it appears that the concentrations of sulfur dioxide and suspended particulate are for only the period 1969–1970 and not for 12 years as is stated. (See Table 4.1.A.6)

8. Prospective Surveys of Acute Respiratory Disease in Volunteer Families: Chicago Nursery School Study, 1969–1970

On page 4–41, in Table 4.3.1, it is not clear where the sulfur dioxide data for the years 1959–63 come from. The Chicago network, which would have provided community data, was not operating effectively until 1964.

On the same page the suspended sulfate data are probably representative for the core area but are high to be used as an average for the city as a whole.

A serious weakness in this study is that the communities are ranked Intermediate, High, and Highest according to a ranking that was determined by suspended particulate values, whereas the most important finding pertains to sulfur dioxide. Referring back to Table 4.1.A.1, it can be seen that a considerably different ranking would have resulted if the communities had been ranked according to sulfur dioxide concentrations. In Table 4.3.1, it may be noted that during the study that the "High" community had the lowest concentration of sulfur dioxide.

Also note in Table 4.1.A.1 that the Highest communities include GSA, which happens to be on the south edge of the Chicago Loop area. This station probably contributed considerably to the high concentration of sulfur dioxide attributed to the Highest community during 1969–1970, yet it is very nonrepresentative of a nursery school. Also, note that the Highest stations include Carver, which for some reason ranks highest because of suspended particulate concentrations whereas the sulfur dioxide concentrations are relatively low.

Sulfates are not considered in the summary of this study, which seems to focus on sulfur dioxide without quantitative considerations of suspended sulfate levels.

(Page 4–54) In the first paragraph of the Summary, the following statement appears: "It is also possible that more recent lower air pollution levels contributed to increased rcspiratory illness." On page 4–51 the following statement is found. "Acute respiratory morbidity was significantly lower among families living in neighborhoods where sulfur dioxide levels had been substantially decreased." These two statements are contradictory and require clarification. The first statement is remarkable. It can be interpreted to mean that some air pollution is good for you. Did the authors intend to say this? Such an important finding is inadequately supported by the contents of the report.

9. Human Exposure to Air Pollution in Selected New York Metropolitan Communities, 1944–1971

An overusage of estimated data can be found on page 5–19. The following two statements appear: (Left hand column, middle paragraph) "Measured values for suspended sulfates for 1956–1970 were available from the Manhattan 121st Street Station, and these values were used for citywide values." (Last paragraph on page) "The observed annual ratios of suspended sulfate to dustfall for New York City were used to estimate the suspended sulfate levels in Queens and Bronx."

10. Prevalence of Chronic Respiratory Disease Symptoms in Adults: 1970 Survey of New York Communities

Three communities were compared: Riverhead, Long Island, a low exposure community, Queens, an intermediate exposure community, and the Bronx, a high exposure community. Parents of all children attending certain elementary schools located within 1.5 miles of an air monitoring station in each community were asked to participate in the study. Each child was given a questionnaire to be filled out by his parents and returned.

Regarding exposure, were the concentrations measured at the monitoring stations generally representative? Assuming a person remains reasonably near the station, in this case within 1½ miles, and breathes the outside air, the station measurements would be generally representative for long-term average exposure. Maps of annual concentrations which are for sulfur dioxide and suspended particulate matter, show reasonably uniform concentrations across the study areas. However, as has been mentioned in the report (5.1) Human Exposure to Air Pollution Selected New York Metropolitan Communities, 1944–1971, by Thomas D. English, et al., the Queens Community lies about 1 mile west of the John F. Kennedy International Airport. The effect of this airport and the various other possible sources of air pollution that could have affected particular local areas were not determined.

The fact that the CHESS monitoring sites were the same as used in the city air pollution control programs suggests that the sites were picked and are being used because they seem to be generally representative.

More important than the representativeness of the monitoring site locations in this study is the proper interpretation of the effects of the greatly reduced pollution levels during the period 1969–1971. It is not meaningful to draw conclusions from sulfur dioxide exposures ranging from 144 to 404 $\mu g/m^3$ and sulfate exposures ranging from 9–24 $\mu g/m^3$, as was done in this study. The implication is that health effects can be caused by the lowest concentrations mentioned, and this is not shown in the study. Also, it is stated that annual sulfur dioxide levels of 50 to 60 $\mu g/m^3$ (accompanied by annual average suspended sulfate levels of about 14 $\mu g/m^3$ and annual arithmetic mean total suspended particulate levels of about 60 to 105 $\mu g/m^3$) could be associated with such effects. These are levels that were measured in 1971, whereas in the study there seems to have been no way to have differentiated between the effects of pollution in 1971, or that might have occurred during some earlier time. It is not reasonable to infer that lower pollution levels are responsible for the observed health effects.

11. Prospective Surveys of Acute Respiratory Disease in Volunteer Families: 1970–1971 New York Studies

In this study families were telephoned once every two weeks and questioned about possible health effects. The families resided within 1 to 1.5 miles of the air monitoring stations.

In the discussion it is stated that acute lower respiratory disease morbidity can be attributed to exposures to 2 to 3 years involving annual average sulfur dioxide levels of 256 to 321 $\mu g/m^3$ (accompanied by elevated annual average levels of total suspended particulate of

97 to 123 $\mu g/m^3$ and annual average suspended sulfate levels of 10 to 15 $\mu g/m^3$). These values are average values for the period 1966–1970, a five year period, and not period of 2 to 3 years as indicated. Also, they are the averages for the Bronx and Queens, respectively, and therefore do not represent a range of concentrations that would have occurred in any particular community, as implied. For example, the sulfur dioxide concentrations in the Bronx ranged from 184 to 472 $\mu g/m^3$ and in the Queens from 131 to 420 $\mu g/m^3$, during the five year period. Three year averages are 174 to 247 $\mu g/m^3$, and two year averages, lower still.

On page 5–16 the dustfall concentrations shown in Figure 5.1.21 seem to be greater than would be obtained from the data presented in Figure 5.1.16.

On page 5–36 (Table 5.2.1) the values in this table seem to come from Table 5.1.A.8. The values in the column headed 1949–58 are, except for dustfall, for shorter time periods. For example, the values for Queens come from data for the years 1956–58.

On page 5–45 (Summary) we find that since the concentration data base comes from Table 5.1.A.8, the long term exposure values represent a period of less than 20 years.

Further, it is stated that there is a distinct possibility that increased susceptibility to acute lower respiratory illness is maintained or induced by exposures involving annual average sulfur dioxide levels of 51 to 63 $\mu g/m^3$ (accompanied by annual average total suspended particulate levels of 63 to 104 $\mu g/m^3$ and annual average suspended sulfate levels of 13 to 14 $\mu g/m^3$). The 51 to 63 $\mu g/m^3$, is a range resulting from two different analyses of samples (see page 5–53). It represents uncertainty in measurement techniques rather than a range of exposure as would be interpreted. These concentrations and the suspended sulfate concentrations of 13 to 14 $\mu g/m^3$ happen to have occurred in the Intermediate I and the Intermediate II communities during 1971. This particular study as conducted could not have differentiated between the effects of these levels of pollution and the effects of higher levels that occurred earlier.

Only average annual concentrations were considered and not peak or episode concentrations.

12. Aggravation of Asthma by Air Pollutants: 1970–1971 New York Studies

Panelists who lived within a 1.5 mile radius of three monitoring stations in communities identified as Low, Intermediate I, and Intermediate II, because of their average air pollution concentrations, recorded asthma attacks each day in a diary for a period lasting 32 weeks, October 1970–May 1971. From a statistical association between asthma attack rates, 24-hour average concentrations from the monitoring stations, and daily minimum temperatures from airports near the study communities, it was concluded that 24-hour suspended sulfate levels of 12 $\mu g/m^3$ on cooler days (T_{min} equal to 30 to 50°) and 7.3 $\mu g/m^3$ on warmer days (T_{min} greater than 50°F) were thresholds for the induction of excessive asthma attacks. No firm evidence could be found to associate elevations in sulfur dioxide (100 to 180 $\mu g/m^3$ on 10 percent of days) with excessive asthma attack rates on either cold or warmer days.

Regarding exposure levels, there is much less assurance that daily average levels throughout a community would be more or less uniform than would be the case with annual average levels. More monitoring stations might have been operated, or mobile stations used, to determine how pollution exposure varied from location to location. The determination of such differences in air pollution concentrations might have been important, but probably more important is that the other factors (in addition to the observed air pollutants) that could have caused or contributed to the asthma attacks were not examined. It would not be worthwhile to refine the information on the distribution of the air pollutants studied, unless a greater effort were made to study all of the various possible causes of the asthma attacks more thoroughly.

The study focused on the effects of minimum temperature. The possible effects of other meteorological variables could also have been explored. Of particular interest would be the effects of sudden, large temperature changes.

It is not made clear why minimum instead of average, or even maximum, temperatures were picked for correlation. Generally there would be less actual exposure to minimum temperature, which usually occurs about sunrise, than to warmer temperatures. Asthmatics would generally be expected to protect themselves from colder temperatures, staying indoors and keeping windows closed, whereas on warmer days they might be more subject to exposure to outdoor air with its assortment of possible allergens. There are diverse reasons why temperature might be an important factor determining asthma attack rates. No attempt was made in the study to provide an explanation.

It is expected that there would be noticeable temperature differences between Riverhead (the Low community) and Queens (the Intermediate I, community). Although it is stated that the temperatures come from nearby airports, the temperature curves plotted in Figure 5.4.1 seem to be identical for both communities. It may be noted that a different curve is plotted for the low community Figure 5.5.2.

(Figure 5.4.4) Although at a glance it appears that for the Intermediate community that the "Attack Rate" and the "Suspended Sulfate" curves are similar, close inspection shows that more often than not, they are out of phase. Between the 2nd and 3rd week the attack rate (AR) curve continues down as the suspended sulfate (SS) curve starts up, between the 10th and 11th week the AR-curve continues down after the SS-curve starts up, between the 14th and 16th week the AR-curve goes up while the SS-curve continues down, between the 19th and 20th week the AR-curve starts up while the SS-curve continues down, and again on the 27th week the AR-curve rises a week before an increase in the suspended sulfate concentrations. In all, three of the five increases in attack rate precede, rather than follow, increases in suspended sulfate concentrations.

13. Frequency and Severity of Cardiopulmonary Symptoms in Adult Panels. 1970–1971 New York Studies (Paragraph 5.5).

Symptom diaries were maintained daily for the 32-week period October 8, 1970 through May 22, 1971, by four panels, depending on state of health. The panelists were distributed in three communities

and lived within 1.5 miles of air pollution monitoring stations. It was concluded that elderly panelists in the low exposure community reported higher symptom rates on days when sulfate levels exceeded 10 $\mu g/m^3$. There seemed to be good evidence of a threshold effect between 6 and 10 $\mu g/m^3$, with a greater morbidity excess on warmer days.

Since suspended sulfates seem to be more uniformly distributed than a pollutant such as sulfur dioxide, the concentrations determined by monitoring should be generally representative of outdoor exposure ard in most cases indoor and outdoor average exposures would be expected to be similar. The question not answered by this study is whether or not the panelists are also being exposed to some other causative agent, or stress factor, that might happen to correlate with the sulfate concentrations. It, and not the suspended sulfate concentrations, might be the cause of the observed health effects.

(Page 5–91) (Figure 5.5.3) The low value of sulfur dioxide that began at the 19th week and continued until the 24th week are suspected of not being true values. Near the end of the last paragraph on the preceding page it is suggested that meteorological conditions may have been responsible. A careful study of the meteorological conditions and fuel usage would be necessary to determine if these might have caused the persistent low concentrations. However, a scanning of the daily local climatological data shows no obvious reason for the reported low values.

Furthermore, the minimum temperature curve for the Low community in Figure 5.5.2 is not the same as given in Figure 5.4.1.

The New York Department of Air Resources also reported a large drop in concentrations following the mid-winter peak at the Queens (Intermediate I) monitoring station, but reported values were never as low, and a period of low values was not followed by a rise as shown in the Figure. Further, the low values shown, which are about 25 $\mu g/m^3$, or .01 ppm or less, are quite low for the New York metropolitan area. Average weekly low values two or three times this value would generally be expected for a comparable period.

14. Ventilatory Function in School Children: 1970–1971 New York Studies (Paragraph 5.6).

Pulmonary tests were made in three elementary schools in communities with different air pollution levels, and there were four rounds of testing, November-December 1970, January 1971, February-March 1971, and April 1971. The children lived within 1.5 miles of a particular air monitoring station. The Queens monitoring station is on top of a school where the testing was done. However, the Bronx station is on top of a "court house in the center of a busy commercial area" (page 5–6) and may not be close to the school. For the Riverhead community it is not made clear whether or not the school and the monitoring station are at the same location or near each other. It is assumed that the schools in Riverhead and the Bronx were within 1½ miles of the monitoring stations, but this is not actually stated.

It was concluded that 9 or more years exposure to annual sulfur dioxide levels of an estimated concentration of 131 to 435 $\mu g/m^3$ (accompanied by suspended particulate levels of about 75 to 200 $\mu g/m^3$) and suspended sulfate levels of about 5 to 25 $\mu g/m^3$ can be

associated with a small but significant impairment in ventilatory function. These values are from Table 5.6.2, and are the extreme high and low values listed. There is an implication here that the low concentrations, 131 $\mu g/m^3$ for sulfur dioxide and 5 $\mu g/m^3$ for suspended sulfates represent threshold values. Actually they are only annual average concentrations for the years 1969 and 1970. The observed health effects may have been the result of exposure to much higher concentrations in other years, or to some other cause.

15. Ventilatory Function in School Children: 1967–68 Testing in Cincinnati Neighborhoods (*Paragraph 6.1*).

This study included a pair of public elementary schools in each of six neighborhoods differing in socioeconomic level, race, or pollution exposure. All children in one or two classrooms of the second grade of the elementary schools were asked to participate in the study to achieve sample sizes of 60 to 75 children in each of the six study sectors. Ventilatory performance as measured by a spirometer was obtained 12 times from each child: once weekly in the months of November 1967 and February and May 1968. The tests were administered on Tuesday and Wednesday mornings.

Air monitoring stations were placed in locations within three blocks of each school to provide samples representative of the air quality in the neighborhood served by the school. No information is reported on the distances of the homes of the children from the school. Apparently it was assumed that the home environment and the school environment were the same. Indoor soiling index and sulfur dioxide observations were taken in the schools, but results are not reported. It is reported that it was determined that indoor and outdoor sulfur dioxide, soiling index, and suspended particulate levels measured over the 24-hour or 4-hour period directly preceding pulmonary function tests did not consistently correlate with the test values.

Details of this lack of correlation are not given, but it was concluded that "ventilatory performance of children thus did not appear to be acutely affected by variations in pollutant levels on the day of the test." Possible exposures over intermediate periods, say three days or one week, prior to testing were not considered. Conclusions seem to be based on possible long-period exposures, probably over a lifetime.

Concentrations of sulfur dioxide were low (less than 52 $\mu g/m^3$) in all areas, so health effects were attributed to particulate pollutants independent of atmospheric levels of gaseous sulfur dioxide.

Average sulfate levels during the period of the study were observed to be between 8.9 and 10.1 $\mu g/m^3$, in the polluted lower middle white community, but previous average exposure was estimated to be 10.7 to 12.1 $\mu g/m^3$, based on the National Air Surveillance Network station. The average suspended sulfate level in the clean white sectors was 8.3 $\mu g/m^3$, a relative difference of 13 percent. (The largest differences in area exposure were in the concentrations of suspended particulates. Levels of total suspended particulates were 131 $\mu g/m^3$ in polluted sectors and 61 to 92 $\mu g/m^3$ in clean sectors.

In reading this paper one wonders about the psychological interaction between the children and the team members administering the tests, who could anticipate the outcome of the experiment. The curves for the black children in Figure 6.1.3, are particularly interesting.

There was only 0.1 $\mu g/m^3$ difference between the average sulfate levels between the Clean Lower Middle Black and the Polluted Lower Middle Black, yet there was a large difference in ventilatory function that disappeared as the study progressed. The discussion acknowledges a culturally determined response of these children to the white interviewers. The effect of the interviewers on all of the children may have been greater than was realized.

It was concluded that the observed area effect on ventilatory performance was probably attributable mostly to suspended sulfates. There is no particular reason to challenge the estimated average exposure of the children to suspended sulfate concentrations. However, it should be noted that the sulfate level in the "clean" school #4, was 9.1 $\mu g/m^3$, more than the 8.9 $\mu g/m^3$ concentration in the polluted schools #8, #11, and #12. The sulfate concentration differences between the schools is acually very slight. It seems far-fetched that the observed average differences in ventilatory performances were caused by a difference in suspended sulfate concentrations of only 1.2 $\mu g/m^3$.

C. SUPERFICIAL AND PERFUNCTORY TREATMENT OF METEOROLOGICAL INFORMATION

A discussion of meteorology is included in the CHESS report on the studies in the Salt Lake Basin, Chicago, and New York. The treatment is so superficial and perfunctory that there is hardly any usefulness. Little attempt has been made to tie the meteorological information with information subsequently presented, or to use it to make clearer the possible reasons for the study findings.

The air flow over the Salt Lake Basin is much more complicated than implied by the wind information included on pages 2–4 and 2–5. Winds in the vicinity of the Oquirrh Mountains, where the smelter is located are significantly different from those at the Salt Lake City airport. Also, the wind rose for Hill Air Force Base shows the effect of a nearby canyon and is not indicative of air trajectories affecting the Ogden area generally. The meteorological discussion on these pages could have made clearer the fact that Ogden is rarely, if ever, affected directly by the smelter plume, and that Salt Lake City is infrequently affected.

Similarly, Chicago winds are inadequately explained by the Midway and O'Hare Airport wind rose; and the reader remains unsure of the wind conditions at Riverhead even though John F. Kennedy and LaGuardia wind roses are given.

Even so, the wind rose data, no matter how carefully given, is of practically no value for interpreting the CHESS studies because source-receptor orientations are generally unknown and were not a factor considered in developing the CHESS conclusions.

D. INSUFFICIENT EXPLORATION OF POSSIBLE RELATIONSHIPS BETWEEN METEOROLOGICAL CONDITIONS AND ASTHMA ATTACK RATES

With regard to meteorological correlations, the CHESS studies for the most part only explored a possible relationship between asthma attack rates and minimum temperature. In designing studies of asthma attack rates it would seem that steps would have been taken

to investigate other meteorological factors also. Of particular interest would be the effect of large and sudden temperature changes. Some other factors that could affect asthma attack rates are relative humidity, frontal passages with their change of air mass, wind direction and speed, precipitation, type and amount, frozen or wet ground, and possibly atmospheric pressure.

A community study of asthma attack rates would also be expected to take note of any air pollution episodes resulting from atmospheric stagnation that occurred during the course of the study. If there were none, notice would be taken of this also.

E. FAILURE TO CONSIDER PEAK AND EPISODE CONCENTRATIONS

The CHESS studies do not adequately consider the short-period exposures to high concentrations that would occur in the smelter communities. It is conceivable that two different communities might have similar annual average concentrations of sulfur dioxide whereas one of them might be subjected much more to short-period high level concentrations that could trigger health effects, asthma attacks in particular. Three causes of much higher than average concentrations are plume looping, plume fumigation associated with the dissipation of a surface-based temperature inversion, and episode conditions caused by a shallow mixing layer and light winds, perhaps with fog.

In order for the CHESS studies to be complete much more needs to be known about the short-period exposures to high air pollution concentrations in the smelter communities.

F. USE OF A SINGLE MONITORING STATION TO DETERMINE THE EXPOSURE OF A COMMUNITY

Although it may be reasonable to assume an average concentration based on a single monitoring station in a large metropolitan area without any well-defined sources of air pollution, such an assumption may not be reasonable in a mountain community with a single large source such as a smelter and possible topographic effects. For example, air pollution concentrations in East Helena, Anaconda, and Kellogg are likely to be rather non-uniform. There is no assurance that the air quality values being used by CHESS are average, or that some families were not subjected to much higher or lower levels of pollution.

One would expect in community studies such as conducted under CHESS that more effort would be made to determine if a single monitoring station is sufficiently representative, particularly in areas where there is but one major source of air pollution. More use could have been made of meteorological techniques in the CHESS studies to estimate exposures and the representativeness of monitoring stations, including the wider application of simple dispersion models. Where there were particular sources, wind direction frequency information should have been examined. Generally, a visit by a meteorologist to a monitoring station location would have been helpful in determining its suitability. Temporary or mobile monitoring stations might have been used to determine if there were significant concentration differences from one location to another within a community.

G. FAILURE TO ESTABLISH SIMILARITY OF EXPOSURE AND STRESS FACTORS BETWEEN COMMUNITIES IN THE SAME STUDY, EXCLUDING THE EXPOSURE TO SPECIFIC POLLUTANTS

No account seems to have been taken in the CHESS studies of actual exposure inside of buildings with various kinds of ventilating or air conditioning systems, or from air at various elevations above ground. Neither was there any effort to determine if there were any significant differences in vegetation. Any of these things might have caused the actual exposure to air pollutants and allergens to be quite different from that estimated on the basis of usual ambient air monitoring. Further, the general stress factors affecting health are no doubt different in the Bronx than they are in Riverhead, or perhaps between Ogden and Magna, where, for example, homelife might have been upset by the smelter strike.

Still another shortcoming was that no information was collected on possible exposures of individuals when they went outside of the study community.

The CHESS reports do not carefully distinguish between outdoor concentrations such as might be measured at a monitoring station and true exposure, which might be significantly different. Throughout the report the concept of dosage is never used. Whole communities were assumed to have the same exposure and presumably were considered to have received the same dosage. A more careful study might have considered the actual community dosages of air pollutants, which might be significantly different, even though the potential exposures are the same.

H. IMPRECISENESS OF MONITORING STATION LOCATIONS

Monitoring stations used in the CHESS studies are not precisely located within the community with which they are identified. In some instances it may not be necessary to know the exact geographic location, but it is not true for a community such as Anaconda. The Rocky Mountain Studies report leaves completely indefinite where the Anaconda station was located with respect to the smelter, or the community itself. It may even be possible that more than one station is involved in the determination of the exposure estimates. Similarly the reader of this particular report has no idea where the monitoring stations in the two Helena areas were located. Perhaps the general nature of the epidemiological studies is an excuse for omitting details, but a follow-up effort to verify some of the CHESS results would require the precise location of certain monitoring stations from which exposures were determined.

I. INEXACT LOCATIONS OF RESIDENCES OF INDIVIDUALS STUDIED

Assuming that an attempt was made to apply meteorological modeling techniques to better determine exposure in the CHESS studies, there would be a severe limitation on what could be accomplished because the locations of residences of participating individuals within the study area are never given except to say that they are within a certain radius of a monitoring station.

There are two problems. First, one cannot be sure that the population was all within the stated radius as claimed. Inquiries made in New York and elsewhere indicate that some of the children studied rode school buses for greater distances and that for various reasons some other participants in studies lived outside of what might be considered the effective radius of an air monitoring station. Second, and probably much more important, are the uncertainties with respect to populations in the vicinity of single large sources such as the smelters in the Rocky Mountain communities. Insufficient information has been given about the distance and direction that participants lived from the smelter stack, and perhaps where they live with respect to significant topographic features. When determining possible exposure, it is particularly important to know if participants lived in different directions from the stack because of the effects of wind direction.

In general, the conclusions from the CHESS studies are weakened because there is a lack of assurance that the populations were actually located as claimed, and in the smelter communities that the populations were so located with respect to a stack, or stacks, that it is feasible to represent exposure with a single monitoring station per community. The presentation of maps that showed, at least in a general way, the location of the residences of participating individuals would have helped to engender more confidence in the estimated exposures.

These comments, many of which are minor and have various orders of significance in total implication, are an indication of the many observations which cumulatively support the need for careful peer review of each research project. In the limited time available for analysis of historical research of this nature, it is difficult to determine the net impact of these errors on the final conclusions in the study but such errors certainly tend to reduce the credibility of the research results as published.

APPENDIX B

Legislative History of the CHESS Program

The programs now referred to as CHESS studies evolved during a period of increasing concern about air pollution. This same period produced significant changes in government organization to cope with new air quality legislation and substantial increases occurred in the funding levels in support of this legislation. In order to place this evolution in some legislative perspective, the following brief history has been prepared. The information is summarized and citations are primarily (except as noted) from House Appropriations Hearings held during the period 1964–1975.

By 1964, well before the 1974 publication of the 1970–71 CHESS studies, air pollution health effects had been under investigation within the Department of Health, Education and Welfare's (DHEW) Bureau of Disease Prevention and Control, National Center for Air Pollution Control. The Center was conducting clinical and laboratory studies in humans and experimental animals. The plans of the Center included a comprehensive program of research on the health effects of exposure to oxides of sulfur alone and in combination with particulates. In addition to clinical and laboratory studies, epidemiological studies were planned for the St. Louis-Philadelphia areas as well as a continuation of a statewide study in Alabama to assess the relationships of air pollution to the prevalence of pulmonary emphysema. The budget proposals in 1964 were the first to be specifically related to the Clean Air Act which had been adopted in 1963. At that time also, there was a concern within the scientific community about the conversion of SO_2 to sulfate and the potential health effects of such sulfates. (Estimated funding levels for 1965 were $2.1 million for the medical and biological studies and $1.5 million for the epidemiological studies).

In 1965, a spokesman for the Air Pollution Control Center indicated during hearings that the 1963 Clean Air Act had opened "a new era of opportunity to cope with the growing national problem of air pollution. . . . There is overwhelming evidence that air pollution is a serious threat to public health and welfare as well as an economic burden costing the Nation several billion dollars annually." In the request for appropriations at that time, the Center indicated an intent to continue the same programs supported in the previous year. Again, studies of the acute and chronic effects of inhalation of oxides of sulfur were specifically cited as receiving support. Human experiments to study the effects of these pollutants under controlled laboratory conditions which had been accelerated in 1965 were designated for expansion in 1966. Additional surveys to identify field sites with varying levels of sulfur oxides and to clarify the effects on human health of single pollutants or single-source pollutants

were identified as priority fields for investigation. A small increase to support these additional studies was requested.

The field studies identified for continuation in the 1967 program included the Alabama air pollution-emphysema studies and the detection of asthma epidemics; the epidemiology of air pollution of asthma in New Orleans, and the relationships between air pollution and absenteeism. It was also indicated that in 1967 there would be a retrospective survey of the effect of air pollution on human health through the use of Social Security disability records, Veterans' Administration records, and insurance records. A study of nursing home residents as a high-risk group was to be initiated and a study to survey air pollution in San Antonio was planned to place special emphasis on the synergistic effect of man-made air pollution and aeroallergens. A mortality study in large cities to determine if excess mortality could be correlated with air pollution was considered for support.

Studies scheduled for continuation in 1968 included the Alabama studies mentioned previously; the New Orleans studies; the continuation of the excess mortality studies; and the designation of Chicago as one area for studying the correlation of disease with air pollution.

Plans for 1969 again emphasized the need for continuation of studies on several air pollutants including sulfur oxides. By this time, the testimony being provided to the Appropriations Committees was citing as justification for some of the work the new requirements specified in the 1967 Air Quality Act. The need for data to develop air quality criteria was being emphasized as a prime goal for the continuation of many of the epidemiological studies. The criteria for sulfur oxides had been published. By July 1, 1968, the reorganization plans of the Administration had resulted in the establishment of the Consumer Protection and Environmental Health Service, and the responsibility for air pollution programs was assigned to the National Air Pollution Control Administration (NAPCA). Objectives of this reorganization included increased autonomy and a higher visibility within DHEW for the air pollution control programs. Progress cited during the 1969 hearings included the publication in the Federal Register of February 11, 1969 of the air quality criteria and related information on control technologies for sulfur oxides and particulate matter. The discussions also directed attention to the increased level of activity required to generate data on health and economic effects of pollutants. Plans for leasing facilities at Research Triangle Park (RTP), North Carolina, for NAPCA were under way and an increase of almost $11 million for this transfer was requested. Among the many justifications offered in support for the concentration of air pollution research at RTP and the move to that location of personnel from DHEW Washington and Cincinnati facilities was the planned construction of the new National Institute of Environmental Health Sciences at RTP. It was believed that a consolidation of NIEHS and NAPCA at RTP would facilitate general planning and centralization of related research programs. A continued expansion of the general air pollution monitoring network was supported.

Hearings in 1970 indicated that further reorganization of the Environmental Health Service had occurred with the splitting out of the Food and Drug Administration. Fiscal Year 1971 was designated

as a key year for developing Federal-State machinery for effective national control of air pollution. A majority of NAPCA personnel were located at RTP and close liaison with NIEHS was cited as real and viable. By this time, NAPCA was referring to the epidemiological studies which were being correlated with air monitoring collection as a health surveillance network (HESN). The effort was described as a "medical research effort destined to assess changes in the health of the population as a result of air pollution control program activities. It is also specifically directed toward conducting studies in unique air pollution types so we can have the kind of data we require for the proper development of air quality criteria." A Chattanooga, Tennessee study of nitrogen dioxide emissions was described as a specific example of the type of study intended to provide data for further air quality criteria development. At this time the HESN program was being highlighted among the several types of epidemiological studies and funding levels for FY 1970 were estimated for this program at about $400,000 with an expected growth in 1976 to about $2.3 million. Other epidemiological studies had been initiated or were planned including studies of air pollution in New York and Seattle, and in several university grants. In general, the HESN program involved cities purportedly selected for high to low exposure gradients to specific pollutants to correlate with appropriate health indicators, as a means to measure health effects. The HESN program had as its objective the construction of a framework of information in which selected health parameters such as pulmonary functions, acute respiratory disease and chronic respiratory disease could be evaluated at one or more points in time. The network was to be initiated on a nine-city basis in fiscal year 1970 and was to be expanded to a 12-city basis in fiscal 1971.

By appropriations hearing time in April 1971, a major Executive Reorganization had occurred (Plan No. 3 of 1970—December 2, 1970) and the air pollution program had been transferred from DHEW NAPCA to the newly created Environmental Protection Agency. At this time, Mr. Ruckelshaus, the first Administrator of EPA to appear before an appropriations committee, provided some summary data about the health effects surveillance studies which had now been given the new acronym CHESS (Community Health Effects Surveillance Studies). Despite this change in title, it was obvious that the CHESS studies were simply a further expansion and formalization of the HESN programs which in turn had evolved from earlier effort in epidemiology. Several of the comments offered by Mr. Ruckelshaus are of interest as they relate to the priorities associated with CHESS at that time:

> CHESS studies have already associated environmental pollution with a number of health problems. Air pollution was associated with significant increases in acute respiratory disesases among schoolchildren and their parents who also reported chronic respiratory disease symptoms more frequently. Both parents and children residing in polluted areas exhibited significant decreases in lung function not accounted for by other factors. Asthmatics were also adversely affected, reporting more attacks on days with increased air pollution.
>
> The CHESS concept is anchored to extensive environmental and health monitoring in sets of communities demonstrating an exposure gradient for specified pollutants. Three years of intense effort have been devoted to the design and field testing of the health impact indicators now a part of CHESS. The air quality aspects of CHESS will not be fully operational until 1973 and will evaluate the health effects of specific pollutant sets. In other words, the air quality aspects of

CHESS are designed to obtain the maximum amount of health effects information about the most important air pollutants. These effects will be continually monitored as pollution is controlled and can thus document the health benefits of abatement. Extending intensive CHESS air monitoring with limited health monitoring to a larger number of cities would probably be counterproductive in terms of health information.

In May, 1972, Dr. Stanley M. Greenfield, Assistant Administrator for Research and Monitoring, EPA, listed as the objectives for CHESS:

First, to evaluate existing environmental pollutant standards as these relate to health. Second, to measure pollutant burdens in exposed populations. And third, to determine the health benefits of environmental pollution control; that is decreased adverse health effects as a result of improved environmental quality.

Presently, major CHESS studies are being conducted in the following areas: New York City; Chattanooga, Tenn.; Birmingham, Ala.; Salt Lake City, Utah; and Los Angeles, Calif.

CHESS operations essentially involve four basic integrated functions; data collection; bioenvironmental measurement; information synthesis; and research and development.

Testimony at the time indicated that about $6.7 million was programed for the CHESS program for fiscal year 1973 with operations anticipated in 27–30 communities in five geographical areas. To add additional areas was estimated to require about $1 million and 15 staff persons per area. CHAMP (Continuous Health Air Monitoring Program) was described for the first time during these appropriation hearings as an integral portion of the CHESS program which would eventually be the total air measurement part of CHESS. As noted by Dr. Greenfield:

Comprehensive monitoring is required to relate changes in sensitive health indicators to existing environmental pollutant levels. From these data, pollutant effects that occur from very short exposure, that is, from 10-minute peak to 24-hour average values, can be documented and available for ambient air quality standard setting. To date, stations [CHAMP] in the Los Angeles Basin have been selected, partially equipped, and data is being transmitted to North Carolina [RTP]. The cost for initiating the more comprehensive CHAMP program, to have 27–30 completely equipped and installed stations, is $2,400,000. The in-house costs are $750,000 and the contract costs are $1,650,000. The contract includes $465,000 for operating the stations in Los Angeles for this year. The yearly operational costs for the 27 to 30 stations will be $1,240,000; $465,000 is in contract and $775,000 of in-house funds.

During the 1973 hearings, the question was raised as to whether the health effects being measured in the CHESS studies were actually based on medical diagnosis and not on individual reports of health on a questionnaire which "then some research clerk codes as a specific disease." Mr. Ruckelshaus replied:

To the fullest extent possible, the medical response to EPA health questionnaires are validated by appropriate techniques. It would not be possible to have a physician visit the home of thousands of respondents to ascertain if their responses to medical questions were correct. However, we can assure the committee that the questionnaires we employ are a representative subset of the responses received and have been appropriately validated against physician records. Further, the questionnaires used in the CHESS program were previously well validated with clinical examinations by competent investigators.

In more detailed hearings held in September 1973 (Committee on Interstate and Foreign Commerce) it was already being reported that the CHESS program had provided information that adverse health effects might be more closely associated with suspended sulfates than with SO_2 or total suspended particulates. The time period estimated

at that time to permit expansion and modification of health effects research on sulfuric acid and suspended sulfates in the ambient atmosphere was 5–10 years. A very extensive expansion of the study of oxides of sulfur with emphasis on sulphates, including sulfuric acid, was being recommended. The FY 74 program was estimated to be about $1.2 million with about $780,000 devoted specifically to the CHESS studies, the remainder directed toward supportive biomedical research, methods development, atmospheric chemistry and monitoring. Reference was made to the fact that the data from CHESS suggesting significant health effects from sulfates had been presented in a scientific conference as early as December 1972. The total air-health research program was summarized as having expanded from about $4 million in FY 69 to about $11 million in FY 74. During this hearing, EPA personnel pointed out that the National Academy of Sciences was reviewing a number of data bases on air pollution health research and would be providing reports.

At this time, the program at RTP was being affected by management adjustments involving the transfer of temporary personnel to permanent positions with some total reductions in personnel as a result of insufficient permanent positions to accommodate all temporary personnel. Thus, although the air-health effects program at RTP showed some growth in general trends, there actually appeared to be some slowing of growth and some adjustments in programs because of the impact of inflation and fewer permanent positions than had been anticipated. With regard to the National Academy of Sciences' review of the EPA research program on sulfates, it is of interest to note that shortly after these congressional hearings, the implications of the potential dangers of sulfates were discussed and acknowledged by several investigators at an Academy conference. At this same conference, however, Higgins and Ferris noted with regard to the EPA CHESS studies:

> This is not the place to criticize these studies in detail. A few general points only will be made. The need for a great deal of information in the shortest possible time has meant that the E.P.A. has been forced to attempt too much too superficially. A more deliberate approach with outside consultation might have led to a more solid body of knowledge. As it is, the studies have a number of deficiencies which make evaluation difficult. The samples studied, the response rates in certain categories, the methods and procedures which have been used and the analysis of the results can all be criticized. It is particularly disquieting, in view of these deficiencies, that there has been a tendency to select findings which point to an effect of pollution on health and ignore those which do not. Perhaps the wisest conclusion which can be drawn from these studies in relation to current standards is that taken in the aggregate they do not appear to justify any lowering of the current standards for SO_2 and particulates. But on the other hand neither do they suggest that any relaxation of these standards is justified.[1]

It is of interest for purposes of this report that the Academy in the final report to the Senate on the results of a special review of the health effects of air pollutants noted also with regard to the CHESS report:

The 1970–1971 CHESS studies have received a good deal of criticism, although it must be noted that many of the potential problems are discussed in detail in the CHESS document (EPA 1974). Specific problems with the two asthma studies

---

[1] Higgins, Ian T. T. and Benjamin G. Ferris. Epidemiology of Sulphur Oxides and Particles. In: Proceeding of the Conference On Health Effects of Air Pollutants. U.S. Senate. Committee on Public Works. Proceedings of the Conference on Health Effects Of Air Pollutants. 93d Congress, 1st session. November 1973. Serial No. 93-15. Washington, D.C., Govt. Print. Off. p. 247.

described above include the relatively poor response rate and high turnover, which may be inherent in diary studies of patients with this disorder; the presence of anomalous data, such as a decrease in the temperature-specific risk of asthma attacks associated with increasing sulfate concentrations on days when the minimum temperature was 30–50° F in the Salt Lake area; the failure to consider temperature change, rather than absolute temperatures; the lack of information on atmospheric allergens, which conceivably vary with pollution concentrations; and the tendency toward overanalysis and overinterpretation of the available data. Reasonable preliminary conclusions from these studies are that they do provide support for the association of air pollution with asthma attacks and that they further implicate suspended sulfates as an important deleterious component of polluted air. It does not appear to be necessary or reasonable at this time to draw firm conclusions concerning the 1970–1971 CHESS asthma studies, inasmuch as data have been collected and analyzed for later years and the reports will be forthcoming soon. If the more recent studies are able to replicate the 1970–1971 findings, this will provide much firmer support for the interpretations advanced by the CHESS investigators.[2]

In April, 1974 Mr. Train, now the Administrator of EPA, pointed out that he considered 1975 to be a critical year for the EPA air programs. He emphasized the need for considerable more research in:

Better defining the health effects of air pollutants, determining the causes and effects of atmospheric concentrations of sulfates and developing improved instrumentation and analytical methods for monitoring air pollutants.

Our proposed 1975 program shows a substantial increase of approximately $9 million over the funding levels of this area during 1975.

The overwhelming bulk of this increase, some $7 million, will be used to considerably accelerate our programs dealing with the health effects of air pollutants.

During detailed hearings in July 1975 held by the Subcommittee on the Environment and the Atmosphere, Committee on Science and Technology, Dr. Greenfield, as a former administrator of EPA research which included the CHESS programs, provided comment on this program from his new perspective as a private consultant:

If you take the report from CHESS for 1974, the thing that is forgotten is that this relates to the data that was collected in 1970–71. Where are the data for 1971–72 and 1972–73. These have never been released. Yet, if you go back and look at these data there are interesting ambiguities relative to the correlations between sulfates and these health effects and whether or not they show up again the following year. There are interesting ambiguities, for example, in the number of chronic bronchitics shown in the 1970–71 period that actually dropped out in the 1971–72, as if you have suddenly removed chronic bronchitis, and they don't remove from the population in that way, which raises serious questions as to whether or not the sampling techniques were really correct.

Finally, we come to the question of health effects, and potentially dangerous pollutants. I do not want to spend a great deal of time describing the EPA Community Health Environmental Surveillance System, the CHESS program. I'm sure that's been described here previously. But these studies have given rise to the most recent sulfur oxides health effects data which have implicated sulfates. Suffice it to say that this was a massive epidemiological program, probably the largest ever attempted, but suffers from many of the deficiencies inherent in all such studies. These deficiencies identified by the CHESS investigators, themselves, have included an [in] ability to adequately quantify the exposure of individuals to various pollutants simultaneously present. These deficiencies preclude the ability to unilaterally ascribe the observed adverse health effects to any single pollutant. It should be noted that an even more fundamental problem pervades those data that purport to demonstrate adverse health effects due to sulfates. The sulfates measured were total water soluble sulfates. No attempt was made to separate the various sulfate, and sulfite, compounds so as to determine which forms may be innocuous and which forms may be deleterious. Neutral sulfate compounds such as sodium and magnesium sulfate occur naturally.

---

[2] U.S. Senate. Committee on Public Works. Air Quality and Stationary Source Emission Control. A Report By the Commission on Natural Resources, National Academy of Sciences, National Academy of Engineering, National Research Council. March 1975. Serial No. 94-4. 94th Congress, 1st session. Washington, D.C., Govt. Print. Off. p. 81.

Are these compounds harmful? If not, what fraction of the measured sulfates represent a potentially harmful concentration? The problem is further compounded by the fact that nitrates may correlate better with the selected health indicators than sulfates, and this, incidentally, is shown more clearly in the 1971–72 data than in the 1970–71 data. That was what I mentioned earlier. In fact, in one analysis made of the data, the most consistent correlation occurred with suspended particulates including irritant respirable particulates.

This does not necessarily mean that some oxides of sulfur are not deleterious to health. Rather, it implies that the health effects observed are probably due to a complex multipollutant mix of irritant respirable particulates consisting of sulfates, nitrates, organic and inorganic acids, et cetera, the sources of which are as complex as the mixture itself. There is no question that we just do not understand the nature of this complex mixture pollutant well enough today to either assign a safe threshold level or ascribe a cause/effect relationship to any or all components of the mixture. In the absence of such ability it is almost impossible to stipulate an effective control strategy for these pollutants.

The March 1976 House Appropriation hearings duplicated much of the information which had been provided the previous year. The 1976 planned program was modified slightly to indicate that the objectives on air health research were intended to focus studies toward selected population subgroups to describe health effects of exposure to sulfur oxides, respirable particulates and other pollutants as well as refining the data on acid sulfate aerosols. The general impression is one of reduction and refinement rather than expansion of surveillance programs. This impression is strengthened by examination of the proposed 1977 program which does include a number of objectives related to clarification of sulfate air chemistry and health effects. Again, no information is provided with regard to anticipated completion of final analysis of the remaining three years of data from the 1971–74 CHESS data.

To briefly recapitulate this legislative history, the CHESS programs evolved from smaller sets of epidemiological studies initiated well before the establishment of EPA. The concept of longitudinal health studies in selected cities to try to correlate health effects with air pollutants was a well established requirement prior to 1970. The phasing together of health indicators with real time measurements of air pollutants received more attention as the air monitoring capability became more available. The history does show that the execution of such epidemiological studies requires considerable preplanning and coordination; some of the studies have persisted for considerably longer periods of time than had been estimated for the data gathering, analysis, publication and utilization in policy planning. Associated with the evolution of the surveillance type health studies was a sequence of reorganizations and establishment of new organizations which undoubtedly produced significant management problems in maintaining continuity for projects requiring long term commitment of manpower and management attention. The CHESS program thus was essentially a continuous program, including many specific studies not originally contemplated as a part of a total system of programs. Reorganizations within EPA, such as the consolidation of laboratories at RTP under a different management system and significant restructuring of data processing systems are identified as other temporary obstacles to effective progress. It now appears that the epidemiological studies of the health effects of air pollutants have been somewhat restructured and reduced in scope to focus on specific subsets of populations for more specific objectives.

It is evident that the February 1976 Los Angeles Times series of articles on the 1974 CHESS report were preceded by a number of public discussions of the total problem of the health effects of sulfur oxides and including criticisms of the significance of the analyses in these reports. The 1974 CHESS monograph is actually a simultaneous publication of a number of studies of health effects of sulfur oxides, including some not actually a part of the health surveillance system known as CHESS. Critical evaluation of these analyses has been difficult as compared with the usual type of scientific peer review. Access to the data had been somewhat limited, and there was little time provided for review outside the writing team. Unlike most reports in scientific literature, the CHESS papers were not referreed in the usual manner nor are the epidemiological experiments readily repeated by individual independent researchers as is the case with more controlled experiments. Rather significant resources in air monitoring and data processing as well as in statistical analysis and collection of health effects information are required and for this reason, the studies represent a unique contribution to the literature. Nevertheless, experts in the field had drawn attention to several of their own misgivings about the conclusions of the CHESS reports. As discussed herein in more detail, the conclusions from epidemiological studies frequently are difficult to support in isolation but must be examined in the context of many other variables.

The task of evaluating the CHESS program then must include an understanding of the historical issues which identified the oxides of sulfur as a potential hazard to health, the problems of conducting epidemiological studies, and then an examination of the various scientific processes and management structures which were brought to bear upon the problem during the past 6–8 years. As noted in the Introduction to the Monograph, the purpose of the program was "to evaluate existing environmental standards, obtain health intelligence for new standards, and document the health benefits of air pollution control." It is this last task of defining benefits that requires a high degree of precision and for which maximum assurance must be present that the data being considered are indeed valid and not overinterpreted.

○


A11600 768996