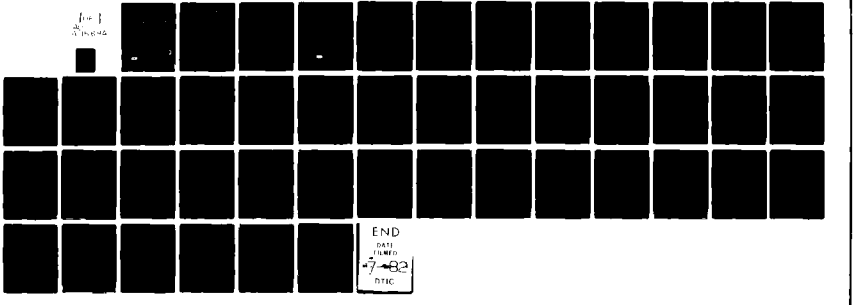
AD-A115 684 PUBLIC RESEARCH INST ALEXANDRIA VA F/G 6/5
AVAILABILITY OF MATCHABLE EMPLOYMENT AND HEALTH DATA,(U)
MAR 82 P FELDMAN J-9-M-1-0163
UNCLASSIFIED CRC-461 NL



CRC 461 / March 1982

12

AVAILABILITY OF MATCHABLE EMPLOYMENT AND HEALTH DATA

Paul Feldman

The Public Research Institute
A Division of the Center for Naval Analyses

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE	READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER CRC 461	2. GOVT ACCESSION NO. AD-A115 684 / 3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) Availability of Matchable Employment and Health Data	5. TYPE OF REPORT & PERIOD COVERED 6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) Paul Feldman	8. CONTRACT OR GRANT NUMBER(s) J-9-M-1-0163
9. PERFORMING ORGANIZATION NAME AND ADDRESS Center for Naval Analyses 2000 No. Beauregard Street Alexandria, Virginia 22311	10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS
11. CONTROLLING OFFICE NAME AND ADDRESS Asst. Secretary for Policy, Evaluation and Research U.S. Department of Labor Washington, D.C. 20210	12. REPORT DATE March 1982 13. NUMBER OF PAGES 38
14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office)	15. SECURITY CLASS. (of this report) Unclassified 15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

This research contribution does not necessarily represent the opinion of the Department of Labor.

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

data bases, employment, health, insurance, salaries, social security, state employment data unemployment, work

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

It is difficult to find reliable industrial and occupational histories of individuals who have been identified as suffering from particular diseases, such as cancer. State Employment Security Agencies (SESA) maintain records of employment of all workers covered by Unemployment Insurance.

To determine the availability of useful data maintained by these agencies, SESA in twelve states were visited and the characteristics and availability of archived data were verified. A telephone survey was also conducted to determine the characteristics and availability of death certificates and other health data in those states.

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE
S/N 0102-LF-014-6601

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

20. The period and form of storage of archives varied from state to state; Texas and New Mexico maintain data on microfilm going back more than forty years. Colorado has records on computer tape going back to 1968. Computer tape storage in other states covers periods ranging from two to ten years. Details about the characteristics and availability of matchable health and unemployment data from twelve states are presented in this report.

CRC 461 / March 1982

AVAILABILITY OF MATCHABLE EMPLOYMENT AND HEALTH DATA

Paul Feldman

Prepared under contract J-9-M-1-0163 for:

Asst. Secretary for Policy, Evaluation and Research
U.S. Department of Labor
Washington, D.C. 20210

The Public Research Institute

A Division of the Center for Naval Analyses

2000 North Beauregard Street, Alexandria, Virginia 22311

ABSTRACT

It is difficult to find reliable industrial and occupational histories of individuals who have been identified as suffering from particular diseases, such as cancer. State Employment Security Agencies (SESA) maintain records of employment of all workers covered by Unemployment Insurance.

To determine the availability of useful data maintained by these agencies, SESA in twelve states were visited and the characteristics and availability of archived data were verified. A telephone survey was also conducted to determine the characteristics and availability of death certificates and other health data in those states.

The period and form of storage of archives varied from state to state; Texas and New Mexico maintain data on microfilm going back more than forty years. Colorado has records on computer tape going back to 1968. Computer tape storage in other states covers periods ranging from two to ten years. Details about the characteristics and availability of matchable health and unemployment data from twelve states are presented in this report.

TABLE OF CONTENTS

LIST OF ILLUSTRATIONS

LIST OF TABLES

INTRODUCTION

One of the first steps in studying occupationally related disease is to find data associating health status and occupation. Typically, sets of data on the health of individuals and on the occupations of individuals can be found, but unless the separate sets contain a common identifier, their usefulness in such studies is very limited.

This report describes the characteristics and availability of certain data sets on employment and health that can be matched on the basis of social security numbers (SSN). The employment data are created and maintained by State Employment Security Agencies (SESA) pursuant to operation of Unemployment Insurance basic records on the employment of individuals use the Social Security Number (SSN) as an identifier.

Information about the characteristics and availability of employment data maintained in active or archived files of SESA was first gathered in telephone interviews and then confirmed in visits to the state offices. Information about health data has been gathered by telephone interviews but not checked in actual visits.

BACKGROUND: SESA RECORD-KEEPING REQUIREMENTS

Unemployment insurance is a federally coordinated system of independent SESAs. The state agencies collect payroll taxes from each employer to establish a fund specific to the firm, and they use the revenues to pay benefits to the firm's workers who are out of work. Most states base the tax rate for an individual firm on recent claims against its fund. This requires that they maintain records of receipts credited to the employer and benefits debited to him.

The benefit paid to an unemployed worker is based on his prior earnings, duration of employment, and prior claim experience. This establishes another record-keeping requirement. About one-third of the states, called wage request states, rely on employers to keep the records; benefit eligibility is determined, at the time of the claim, by asking employers for employment information about the claimant. The other two-thirds of the states, called wage reporting states, keep the records themselves; they require employers to report, quarterly, the employment and earnings of <u>all</u> covered workers, and they consult these reports to determine benefit eligibility when an individual presents a claim. Table 1 lists wage request and wage reporting states. As a source of useful data for occupational disease studies, the wage reporting states are most valuable.

Claims usually involve payments over many weeks; to remain eligible over time, claimants must show evidence of searching for work and willingness to accept a "suitable" job. Record requirements associated with

claim processing vary from state to state, but they may involve maintaining information on the claimant's actual occupation to provide a basis for judging the suitability of job offers.

TABLE 1

WAGE REQUEST STATES

(No Listing of Workers by SSN)

Hawaii
Massachusetts
Michigan
Minnesota
Nebraska
New Jersey
New York
Ohio
Rhode Island
Utah
Vermont
Wisconsin

WAGE REPORTING STATES

(Listing Workers by Employer by SSN)

Alabama
Alaska
Arizona
Arkansas
California
Colorado
Connecticut
Delaware
D.C.
Florida
Georgia
Idaho
Illinois
Indiana
Iowa
Kansas
Kentucky
Louisiana
Maine
Maryland
Mississippi
Missouri
Montana
Nevada
New Hampshire
New Mexico
North Carolina
North Dakota
Oklahoma
Oregon
Pennsylvania
Puerto Rico
South Carolina
South Dakota
Tennessee
Texas
Virginia
Washington
West Virginia
Wyoming

Appeals and/or audit procedures require states to archive data for up to 7 years. Archived data are usually microfilms of the original forms although, occasionally, the original forms or computer tapes are kept.

The basic forms used for filing quarterly reports vary slightly across states and over time, but they generally look like what's shown in figures 1 through 4, which are currently in use in South Carolina. On the form shown in figure 1, employers are asked to give the SSN, name, and wages paid to each worker employed during the quarter. Figures 2 through 4 provide information about the employer. Note that the employer's SIC is not recorded in these forms. SIC numbers are usually assigned by a division in the SESA set up for that purpose.

PROJECT CONCERNS

To determine the availability of these data, a thorough interrogation of selected state agencies was undertaken. The states were chosen by NCI on the basis that studies in those states were currently under way or were contemplated. The states were:

California
Colorado
Connecticut
Delaware
Georgia
Louisiana
New Mexico
North Carolina
Oklahoma
Pennsylvania
South Carolina
Texas
Washington.

All of the states listed above, except for Connecticut, were visited and the presence of archived data verified. Connecticut refused to authorize a visit.

FINDINGS: CHARACTERISTICS AND AVAILABILITY OF SESA DATA

Several of the states have held archived data longer than required by law. Two, Texas and New Mexico, have records going back more than 40 years. Others have records going back beyond the archive requirement but not quite so far. Some data, particularly for recent years, are in computer readable form. Generally, the older records are on microfilm. Table 2 summarizies our findings, and the next section explains the results in detail.

Form UCE 120
Revised 10-1-72

SOUTH CAROLINA EMPLOYMENT SECURITY COMMISSION, COLUMBIA, S. C.

Page No. 1

EMPLOYER QUARTERLY REPORT OF EMPLOYEE WAGES

Enter in space below Employer's Name, Address and South Carolina Account Number

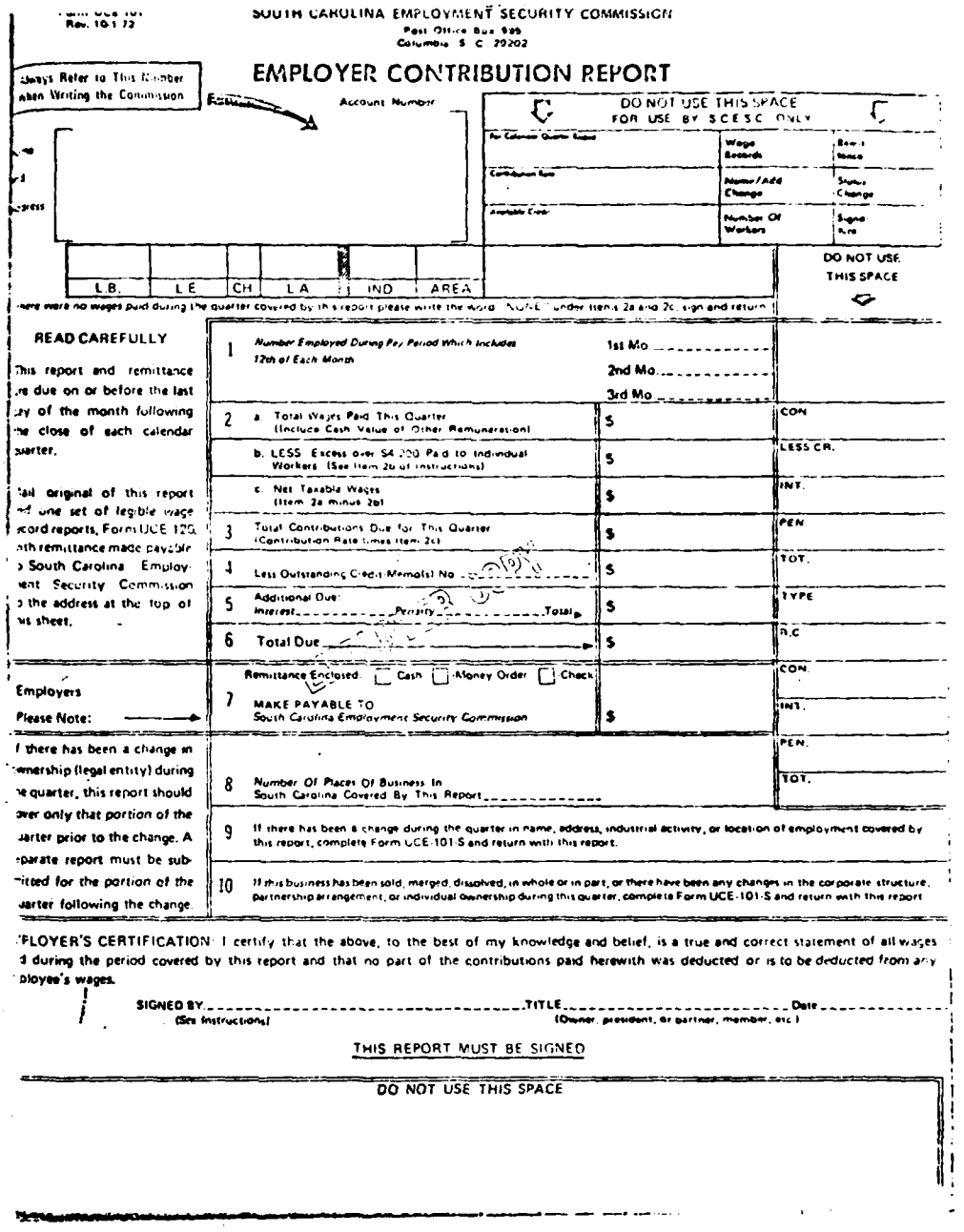
2. DATE QUARTER ENDED	4. TOTAL PAGES OF THIS RETURN INCLUDING CONTINUATION SHEETS ATTACHED
3. CONTRIBUTION RATE	5. TOTAL NUMBER OF EMPLOYEES LISTED ON RETURN

If there is not enough space to list all employees below, use Continuation Sheet (Form UCE-120 A). Each Item Continuation Sheet

EMPLOYEE'S SOCIAL SECURITY ACCOUNT NUMBER (6)			NAME OF EMPLOYEE (Please type or print) (7)	TOTAL WAGES PAID To Employee During Quarter (Before deductions) (8)		Non-Taxable Wages Paid To Employee During Quarter (in Excess of $4,200) (9)	
000	00	0000		Dollars	Cents	Dollars	Cents
10. Totals for this page							
11. TOTALS FOR THIS RETURN – Total Wages Paid for Quarter							
12. TAXABLE WAGES PAID THIS QUARTER (Column 8 Minus Column 9) $ _____ —▶ Item 2c. UCE 101				TOTAL WAGES Item 2a. UCE-101		NON TAXABLE WAGES Item 2b. UCE 101	

EMPLOYER COPY

FIG. 1: EMPLOYER QUARTERLY REPORT OF EMPLOYEE WAGES



SOUTH CAROLINA EMPLOYMENT SECURITY COMMISSION
Post Office Box 995
Columbia S C 29202

Rev. 10-1-72

EMPLOYER CONTRIBUTION REPORT

Always Refer to This Number When Writing the Commission

Account Number

DO NOT USE THIS SPACE
FOR USE BY SCESC ONLY

For Calendar Quarter Ended		Wage Records	Rate/Issue
Contribution Rate		Name/Add Change	Status Change
Available Data		Number Of Workers	Signature

LB.	LE	CH	LA	IND	AREA

DO NOT USE THIS SPACE

If there were no wages paid during the quarter covered by this report please write the word "NONE" under items 2a and 2c, sign and return.

READ CAREFULLY

This report and remittance are due on or before the last day of the month following the close of each calendar quarter.

Mail original of this report and one set of legible wage record reports, Form UCE 120, with remittance made payable to South Carolina Employment Security Commission to the address at the top of this sheet.

1	*Number Employed During Pay Period Which Includes 12th of Each Month*	1st Mo. / 2nd Mo. / 3rd Mo.	
2	a. Total Wages Paid This Quarter (Include Cash Value of Other Remuneration)	$	CON
	b. LESS: Excess over $4,200 Paid to Individual Workers (See Item 2b of Instructions)	$	LESS CR.
	c. Net Taxable Wages (Item 2a minus 2b)	$	INT.
3	Total Contributions Due for This Quarter (Contribution Rate times Item 2c)	$	PEN.
4	Less Outstanding Credit Memo(s) No.	$	TOT.
5	Additional Due: Interest / Penalty / Total	$	TYPE
6	Total Due	$	R.C
	Remittance Enclosed: ☐ Cash ☐ Money Order ☐ Check		CON.
7	MAKE PAYABLE TO South Carolina Employment Security Commission	$	INT.
			PEN.
8	Number Of Places Of Business In South Carolina Covered By This Report		TOT.
9	If there has been a change during the quarter in name, address, industrial activity, or location of employment covered by this report, complete Form UCE-101-S and return with this report.		
10	If this business has been sold, merged, dissolved, in whole or in part, or there have been any changes in the corporate structure, partnership arrangement, or individual ownership during this quarter, complete Form UCE-101-S and return with this report.		

Employers Please Note: If there has been a change in ownership (legal entity) during the quarter, this report should cover only that *portion* of the quarter prior to the change. A separate report must be submitted for the *portion* of the quarter following the change.

EMPLOYER'S CERTIFICATION: I certify that the above, to the best of my knowledge and belief, is a true and correct statement of all wages paid during the period covered by this report and that no part of the contributions paid herewith was deducted or is to be deducted from any employee's wages.

SIGNED BY _____ TITLE _____ Date _____
(See Instructions) (Owner, president, or partner, member, etc.)

THIS REPORT MUST BE SIGNED

DO NOT USE THIS SPACE

FIG. 2: EMPLOYER CONTRIBUTION REPORT

Form UCE 101 S
Rev. 10-1-72

SOUTH CAROLINA EMPLOYMENT SECURITY COMMISSION
Post Office Box 995
Columbia, S. C. 29202

EMPLOYER'S REPORT OF CHANGE

Account Number | For Calendar Quarter Ended

Name
and
Address

I – CHANGE IN NAME, ADDRESS OR BUSINESS LOCATION

If there has been a change in trade name, corporate name, address or business location, enter the information required below.

New Trade Name_____

New Corporate Name_____ Changed by charter Amendment ☐ Yes ☐ No

New Mailing Address_____
Street City State Zip Code

New Business Location_____
Street City State Zip Code

By_____ Title_____ Date_____

II – CHANGE IN OWNERSHIP OR DISCONTINUANCE OF BUSINESS

If your business was discontinued, or if a change in ownership occurred during the period covered by this Contribution Report enter the information required below. SEPARATE REPORTS MUST BE FILED BY DIFFERENT OWNERSHIPS. (For each ownership such separate report should cover only that part of the quarter for which the particular ownership operated.)

BUSINESS DISCONTINUED WITHOUT SUCCESSOR ON_____
Date

CHANGE OF OWNERSHIP: *(Enter date and type of change)*

Exact Date of Change_____

☐ Entire Business Sold
☐ Partial Sale only, not Out-of-Business
☐ Partner Added
☐ Partner Withdrew
☐ Partnership Dissolved
☐ Corporation Formed
☐ Corporation Dissolved
☐ Merger
☐ Other Change, explain_____

Explain any change in nature of business activity_____

New Owner's Name_____

New Business Name_____

New Owner's Mailing Address_____
Street City State

I certify that the above information is true and correct to the best of my knowledge and belief.

By_____ Title_____ Date_____

My future mailing address will be_____
Street City State Zip Code

FIG. 3: EMPLOYER'S REPORT OF CHANGE

Form UCE - 101a
Revised 10/1/72

SOUTH CAROLINA EMPLOYMENT SECURITY COMMISSION
P. O. BOX 995, COLUMBIA, S. C. 29202
MULTI-UNIT AND MULTI-AREA BREAKDOWN REPORT

Account Number

THIS COPY FOR EMPLOYER FILES

For Calendar Quarter Ended

Contribution Rate

L.B.	L.F.	CH.	L.A.	IND.	AREA

Individual Employing Units / County or City	Do Not Use		Number Employed During Pay Period Which Includes 12th of Each Month			Wages Paid in Quarter		Contributions Due for This Quarter
	Industry	Area	1st Mo.	2nd Mo.	3rd Mo.	Total Wages Paid This Quarter	Net Taxable Wages	
I	II	III	IV	V	VI	VII	VIII	IX
TOTALS (As shown on Employer's Contribution Report, Form UCE 101)			Item 1, UCE-101			Item 2a, UCE-101	Item 2c, UCE-101	Item 3, UCE-101

FIG. 4: MULTI-UNIT AND MULTI-AREA BREAKDOWN REPORT

TABLE 2

ARCHIVED EMPLOYMENT DATA IN SELECTED STATES

State / Records per quarter	Quarterly wage reports beginning (year)			Employer file beginning (year)			Claims file beginning (year)		
	Tape	Film	Paper	Tape	Film	Paper	Tape	Film	Paper
California 13 million	1975	1979	1977	1976	--	--	1975	1977	--
Colorado 1.5 million	1967	1975	N/A	1967	1967	N/A	1971[a]	None	1979
Delaware 250,000	1976	None	1976	1974	None	1967	1976	None	1977
Georgia 3 million	1976	None	1976	1974	1974	1974	1971	None	1980
Louisiana 4 million	1980	1976	1976	1980	1969	1976	1976	1971	1980[a]
New Mexico 350,000	1976	1942	N/A	1967	1967	1942	1978	1976	1976
North Carolina 2.9 million	1980	None	1977	1976	1972	1976	1979	1971	1979
Oklahoma 1.4 million	1979	1975	1975	1971	1971	1977	1975	1975	1980

[a]Approximately

TABLE 2 (Cont'd)

ARCHIVED EMPLOYMENT DATA IN SELECTED STATES

State	Quarterly wage reports beginning (year)			Employer file beginning (year)			Claims file beginning (year)		
Records per quarter	Tape	Film	Paper	Tape	Film	Paper	Tape	Film	Paper
Pennsylvania 5 million	1972	1974	1981	1972	N/A	N/A	1968	None	1980
South Carolina 1.4 million	1980	None	1977	1971	1973	1975	1978	1976	1981
Texas 5.6 million	1976	1938	1980	1976	1978	--	1977	--	1976
Washington 2 million	1980	1980	1973	1978	1974	1974	1980	1970	1974

All states that were visited agreed to provide data from the Quarterly Wage Reports, the Employer file, and the claim file to the Labor Department in Washington, D.C. if requested to do so by the Office of Reseach in the Unemployment Insurance System. Subsequent use of the data would be the responsibility of the Labor Department.

Data Characteristics

Several pieces of information are necessary for constructing an occupational history of an individual worker. Most of these can be found in two master files kept by every SESA we visited.

The quarterly wage reporting system is standard across wage-reporting states. Employers in these states are required by law to file a status report with the SESA. This status report is used to determine which employers are required to provide their employees with unemployment insurance. While requirements differ by state, a very high percentage of employers must provide coverage.

Each qualifying employer is assigned a state identification number--usually just a sequential number assigned when the status report is processed. This number is not the same as the employer's Social Security identification number. That is important because, while it is possible to form matching files that contain important variables, the data cannot be matched directly with information about the firms that may be available from the Social Security Administration.

The Employer's Quarterly Wage Reports contain five important pieces of information: the employer's name and state identification number and each employee's name, SSN, and quarterly earnings. The employer's address is usually also on the form. This address may be the firm's local address, the address of its headquarters, or simply the address of an accountant who handles the employer's correspondence.

Some of the states microfilm these forms as they are submitted to the SESA. Some keep the original forms for a number of years. One or the other or both is always done for reports covering the time period specified by the state's Statute of Limitations; this is required by law. In addition, all states now keep the most recent 5 or 6 quarters of data on easily accessible disk or magnetic tape files. As a rule, 5 quarters is all the time needed to establish a claim and therefore generally all that is ever reviewed by the SESA.

As shown in table 2, some states do have more data in computer-readable form. The computerized files do not usually include the alphabetic information on the quarterly wage report, but they all carry every variable of interest to us: employer ID number, employee SSN, and employee's total quarterly wages. Details about individual states are presented in the appendix.

Each of the states has a second file that is constructed from the information in the employer's original status report. This file is called either the Employer Master File or the Employer Address File. In each state, it can be referenced by employer ID number; in some states, it is also sequenced alphabetically. The information in this file also includes one or more of the addresses--local, headquarters, or accountant--discussed previously. In some cases, the SESAs do not know which type of address is on the file; they may vary by employer. Other states require one or more of the addresses and can identify them.

The employer file contains the employer's SIC code. Some states (particularly for earlier years) use only 3 digits, others use 5. The majority of the states currently classify employers by a 4-digit SIC code.

There is a third file in some states, commonly known as the Predecessor-Successor File, which allows a firm to be traced through time despite ownership, name, or structural changes. The states that do not keep this file usually code the information within the Employer Address File.

The last file of interest in each state is the claim file. This file usually contains demographic information about a claimant: race, sex, age, etc. More importantly, it generally includes information about an employee's occupation. While this specific information is recorded for only those workers that file a claim, claimants are a high proportion of the total working population.

The claim file can also be valuable in constructing the employee's continuous work history, which is important if we should want to construct a longitudinal file that tracks workers over time.

While these data are a 100 percent sample of workers, there is always a problem of unusable records. Errors in SSN reporting, intentional or unintentional, often occur on either typed or handwritten quarterly wage reports. Sometimes SSNs are not reported at all. Many times, these errors can be corrected by matching the employee name. But these, too, often will not match due to marriage or other name changes. And, of course, there will be errors in other data items. From our interviews with various SESA employees, we estimate that between 3 and 5 percent of all records should be expected to be erroneous.

Data Availability

Every SESA we visited is willing to release its data. None, however, was willing to discuss the time or money needed. All requests for data must come directly from the Department of Labor, and all negotiations must be handled by them. In the course of this project, we discovered that Texas was planning to destroy some of its microfilmed

records. At the request of the Unemployment Insurance Service (UIS) in Washington, records for 1938 to 1959 were turned over to us at no cost. Those records are currently in storage at CNA in Alexandria, Virginia.

Appendix A provides a detailed description of data available by state, including the timespan covered by the states' records and the media on which the records are stored. The most advanced storage form (magnetic tape or disk) will be the only one discussed unless additional (i.e., older) data are available on a less advanced storage medium (microfilm or paper).

HEALTH DATA

For health data to be used in association with employment data, the key factor is that they contain individual social security numbers and that they be population based, preferably nationally but possibly at the state level. We shall discuss the availability of four types of data that contain SSNs and would report long latency illnesses. These are death certificate data maintained by the states, the National Death Index, Social Security Disability Insurance data, and Medicare records.

Death Certificates

Death certificates are official state documents, filed and maintained by state agencies. Although most of the states have requested that the SSN be recorded on the certificates for at least the last 30 years, some SSNs are missing. Most states have also requested that primary and secondary causes of death be recorded for at least a few decades.

Actual certificates have been archived since early in the 20th century. More recent certificates are often copied on microfilm, and for recent years, selected items from the certificates are available on computer tape. Table 3 shows what death certificate data have been archived in the states we checked for employment data. Appendix B contains additional details about death certificate data and points of contact in each of the 12 states. Except for California, death certificates are not public record, but they can be obtained with proper authorization from the states.

National Death Index

The National Center for Health Statistics (NCHS) maintains computer tapes of death record information submitted by the states. However, only since the beginning of 1979 have SSNs been part of the Index. The Index picks up 14 variables off a standard death certificate, but does not code cause of death. In order to use the Index, a user submits a request for a search of the Index; the search seeks to match a given user request record with a Death Index record. Both records must

TABLE 3

ARCHIVED DEATH CERTIFICATE DATA IN SELECTED STATES

State	Date Computerized (for 100% sample)	SSN added	Primary/ Secondary Cause of Death	Data Furthest Back	SSN added	Primary/ Secondary Cause of Death
California	1960	1960	No, just underlying	1905 - certificates and microfilm	1940s	Yes
Colorado	1959	1975	1959-79 no, just underlying: 1980 yes (they hope to go back and code multiple causes for 1975-79, too)	1907 - certificates and microfilm (but not very complete records until 1920)	1940	Yes
Delaware	1980	1980	No, just immediate	1913 - certificates and microfilm	1949 (but in 50s many may be missing)	Yes
Georgia	1961	1979-80	Yes since 1979 (before 1979 just underlying)	1919 - certificates and microfilm	1950s (perhaps 1940s too)	yes

TABLE 3 (Cont'd)

ARCHIVED DEATH CERTIFICATE DATA IN SELECTED STATES

State	Date Computerized (for 100% sample)	SSN added	Primary/ Secondary Cause of Death	Data Furthest Back	SSN added	Primary/ Secondary Cause of Death
Louisiana	*	*	*	1918	*	*
New Mexico	1964-65	1979	No, just underlying	1920s - certificates and microfilm	1936-37	Yes
North Carolina	1968	1968	1968-75 just underlying: 1975 forward code all causes	1913	1930s	Old records, just single cause--newer records (post WW II) all causes
Oklahoma	1975	1979	No, just underlying	1908 - certificates (but not complete records until 1950) 1970 - microfilm (they hope to go back further)	1968	Yes

* Refused to give information by telephone.

TABLE 3 (Cont'd)

ARCHIVED DEATH CERTIFICATE DATA IN SELECTED STATES

State	Date Computerized (for 100% sample)	SSN added	Primary/ Secondary Cause of Death	Data Furthest Back	SSN added	Primary/ Secondary Cause of Death
Pennsylvania	1979	1979	Cause of death not coded - only yes or no (i.e., yes means cause(s) is listed on certificate)	1906 - certificates 1978 - microfilm	1941	Yes
South Carolina	1969	1969	1969-79 - just underlying 1980 - multiple	1915 - certificates and microfilm	1938	Yes, since 1930
Texas	1964	1964	Yes	1903 - certificates and microfilm	1940s	Yes
Washington	1967	1979	No, just underlying	1907 - certificates 1960s (perhaps further back) - microfilm	1950s (perhaps 1940s)	Yes, from 1960s on (before 1960s only single cause listed)

satisfy at least one of several conditions. Among these is a match on both records of SSN and last name, both of which are obtainable from SESA records. Once matched, the user can obtain the name of the state where death occurred, when death occurred, and the state death certificate number. This is the only information NCHS can provide in accordance with its contract with the states. A user can then contact the appropriate state and request information concerning causes of death and any other death certificate information desired. For further information, see User's Manual - The National Death Index, U.S. Department of Health and Human Services, Hyattsville, Maryland, August 1981.

Social Security Disability Insurance (SSDI) and Medicare Records

SSDI data are derived from SS-831 forms--disability determination forms. They are completed by hand in the various state agencies and sent to Baltimore. Although some of the forms get lost, close to 100 percent of them do arrive in Baltimore. The forms contain a written description for each SSDI applicant of the nature of the disability and occupation, as well as SSN. The forms themselves currently go back to 1978; periodically, the oldest forms are destroyed.

From this "100 percent sample," 20 percent are selected (through a stratified random sampling procedure) for electronic data processing (primarily for statistical research purposes). The handwritten diagnosis (primary and secondary) of the nature of the disability is given a 4-digit code according to a standard medical classification scheme. Occupation, as recorded on the form, is coded according to a standard occupation classification scheme. These data, along with SSN and other information for the 20 percent sample, go back to the late 1960s. (For a complete description of the file, see the publication "Continuous Disability History Sample Restricted Use Data File: Description and Documentation," SSA, Office of Research & Statistics, ORS Publication No. 024 (1-78). A description of the data fields is attached as exhibit 1.)

According to SSA procedures, SSNs are scrambled for use outside SSA, which would prevent matching with employment date. It may be possible, however, to provide SSA with certain SSNs for matching employment data with SSDI information from the 20 percent sample or possibly the 100 percent sample of SS-831 forms.

It is also possible that special arrangements could be made to match medicare records with employment data. Medicare records contain SSNs and cover more than 95 percent of the over-65 population.

FEASIBILITY OF USING UI DATA FOR EPIDEMIOLOGIC STUDIES

The UI Data described above can be used in population based epidemiologic studies concerned with occupational exposure to carcinogens. The general procedure will be to match individual health records

EXHIBIT 1

"CDHS Restricted Use Data File Format

Field Number	Data Field Description	Field Length	Position
1	Social Security Number	9	1-9
2	Data of Application	4	10-13
3	Date of Birth	4	14-17
4	Multiaction Code	1	18
5	Sex	1	19
6	Alleged Onset Date	4	20-23
7	Category	1	24
8	Race	1	25
9	Marital Status	1	26
10	Number of Children	2	27-28
11	State Agency Basis Code	4	29-32
12	Data Disability Period Began	4	33-36
13	Medical Re-Exam Date	4	37-40
14	Date of Release from E and A	4	41-44
15	Type of Action	1	45
16	Class of Adjudicative Action	2	46-47
17	Statutory Blind (established)	1	48
18	Short Term Occupation Code	1	49
19	Mobility	1	50
20	Occupation	6	51-56
21	Education	2	57-58
22	Previous Action	1	59
23	State	2	60-61
24	County	3	62-64
25	Primary Diagnosis	4	65-68
26	Secondary Diagnosis	4	69-72
27	Weight Factor	3	73-75
28	Payment Center Office Code	1	76
29	Current Primary Insurance Amount	4	77-80
30	Computation and Insured Status Code	1	81
31	Military Service Code	2	82-83
32	Railroad Code	1	84
33	Date of Death of Primary Beneficiary	4	85-88
34	Monthly Payment Amount	5	89-93
25	ZIP Code	5	94-98
36	Beneficiary Indentification Code	2	99-100
37	Ledger Account file Code	2	101-102
38	Initial Date of Entitlement	4	103-106
39	Current Date of Entitlement	4	107-110
40	Date of Suspension of Termination	4	111-114
41	Special Action Code	1	115
42	Monthly Benefit Payable	4	116-119
43	Date of Entitlement to HIB	4	120-123
44	HIB Entitlement Code	2	124-125

vi"

of cancer patients--death certificates or other sources of medical information--with longitudinal employment records, using the SSN as the basis for the match.

The most important attribute of these data is that they provide an inexpensive way of selecting cohorts of workers by SSN by employer and SIC. One of the drawbacks of the data is that demographic detail in the file will cover only a (substantial) subset of workers: those who have claimed benefits at some time. For those who have never claimed, the file can be augmented with age, race, sex, and fact of death by submitting SSNs to the Social Security Administration. Details of such an augmentation would have to be worked out with SSA but their response to a preliminary inquiry was that such augmentation would be possible under the Privacy Act.

Cohort Selection

If a particular employer within a state is the basis for study, a 100 percent sample of workers employed by that employer would be drawn based on the employer's ID number appearing in the worker's record. The ID is assigned by the UI system and is unique. To find the ID for a known firm, at a given time, we would look up the firm name in the Employer Address File. Of course, the firm might have changed names before or after that date, but there exists, in most states, a "Predecessor-Successor File" so that firms can be tracked through name changes and mergers.

The resulting sample would include all workers who had ever worked for that employer from the earliest date covered by the data to the present. If a whole industry is the focus, rather than a single employer, the sample could be drawn on the basis of the SIC; or all employers in a given SIC could be listed and a subset selected for study based on geography or some other known characteristics of the firms.

Matching

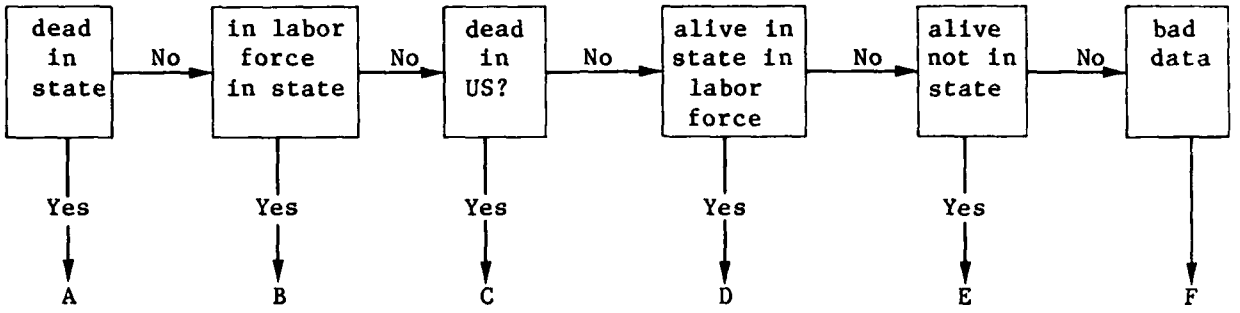
After the cohort has been selected, it can be matched with death or health records that also contain SSNs.* There are many logical approaches to this task; one approach is shown below.

* An alternative would be to match all SSNs in the entire population to a comprehensive set of death records, e.g., a state file of death certificates.

1. Check on state wage history file for employees of selected plant. Of these people, there are five possible outcomes:

   (1) Worker is dead in state
   (2) Worker is still in the labor force in state
   (3) Worker is dead anywhere outside of state
   (4) Worker is alive but not in the labor force in state
   (5) Worker is alive not in state

   A plan of attack is as follows:



A. Checking each individual against state death certificate information will produce SSNs of those who died in state. Then check:

   - age, occupation
   - cause of death
   - length of employment in plant.

B. If not dead in state, individual may still be in labor force, either on the wage file or unemployed in state. Check for a more recent worker file. If found, note length of employment at plant, incidence of illness, and more recent employment.

C. If not dead or employed in state, individual could be dead elsewhere in US:

   - Use Social Security death-searching data, matched against SSNs not already matched in A or B.
   - If any found, check age at death from SS data.

- From Claim files, check occupation, length of employment at plant.

- If possible, note state address (even a zipcode) at time of death from SS data (if provided).

At this point, it would be useful to know cause of death. It might be possible to obtain death certificate data from the necessary states, if it exists.

D & E. If individual has not turned up in above checks, he is alive, but not employed in state.

- Length of employment in plant and wage could be noted.

- Without Social Security data, it would be difficult to track the whereabouts of someone alive and not employed in state. It may be possible to get a current status from the Social Security Administration.

- If health data had been obtained (see B), a check there would produce those not employed who are or had been ill. Chances are that those who are alive, either in state or elsewhere, but who are not working are receiving some sort of benefits. Possible matches of worker data with disability files, medicare data, and Workmen's Compensation data should be explored.

F. If any individuals fail all the above tests, then the data were incorrect or incomplete somewhere along the way.

Cost

Once the basic records have been collected and the data set structured, the major costs of using it will be for

- Securing matchable health data

- Programming and processing the matches

- Checking unaccounted for leakage from the sample

- Conducting whatever analysis is to be done.

Until some experimental work along that line has been carried out, it will be difficult to estimate costs, although it is easy to predict that the costs will be small relative to other methods of finding and following cohorts.

APPENDIX A

ARCHIVED UI DATA

APPENDIX A

ARCHIVED UI DATA

CALIFORNIA

The wage file, a 100 percent sample of the 13 million (average) employees in the state is maintained on tape to the second quarter of 1975. The wage file will go to a 5-year purge cycle at the end of 1982.

The Employer Address File contains about 960,000 employers per quarter and is stored to 5 years on tape. Inactive employers are kept for an additional 3 years. The file contains a 4-digit SIC (3 plus ownership code), in addition to the location address, home address, and bookkeeping address.

Claims payment information is available to the second quarter of 1975. This will soon go on a 5-year purge cycle.

A thorough data set is maintained on a 10-year cycle for 10 percent of the population. This Historical Archive Data File tracks all employees with a terminal digit 5 SSN. Of particular importance is that the file contains wages, SSN, 4-digit SIC of the main employers for a given year, claims information (age, race, sex, ethnic background included), and ESARS information. This file has already been used for TRA studies.

COLORADO

Colorado's machine readable data go back further than those in any other state. Computerized Wage Report Files date back to 1967, which is when Colorado became a wage-reporting state. Currently, there are approximately 1.5 million workers per quarter. Microfilming of these reports was not begun until 1976.

The Employer Address File contains about 148,000 employers of which 76,000 are currently active. This file was created in 1966. 18,000 inactive employers were purged to another tape and archived in 1978. In 1979, another 8,000 were purged, but this tape has been destroyed. It's possible, however, that this file could be recreated using the microfilmed Employer Address File that also dates back to 1966. The file carries both a local and headquarter's address for each firm and a 4-digit SIC code.

The Colorado UI Claims File is also on tape. The 70-reel file contains all claims data back to 1970. It's possible that the data actually cover claims back through 1967, giving us completeness in all files; however, a search for a claim that old has never been required. Approximately 8 percent of the working population files a claim each year, giving Colorado additional demographic information and a Dictionary of Occupational Titles (Department of Labor) code.

DELAWARE

Delaware's current wage file, which can be accessed by employer or employee, tracks back 6 quarters on disk and tape. A 3-digit SIC code (merged from the Employer Address File) already appears on the records. Almost all of the 250,000 workers in the state can be traced to the first quarter of 1976 on wage record files that exist for each quarter. Workers who didn't originally balance (sum of wages in firm didn't equal total reported) do not appear on these tapes and must be located on hardcopy reports. No microfilming is done.

The Employer Address File contains 30,000 employers, 12,000 of which are active. Inactive files are not purged, and the file is complete since 1974. Folders exist for all employers back to 1968.

The Master Claim File contains earnings information and is on tape since 1976. Additional descriptive data of the 53,000 yearly claimants are available for 18 months on tape and 4 years in hardcopy form.

GEORGIA

Georgia's Wage Report File goes back to the first quarter of 1977. No microfilming is done; the actual reports are archived, but cover the same time period. Georgia receives reports on approximately 3 million workers per quarter. Georgia officials believe that their records are cleaner than most. Wage reports are balanced against tax reports; therefore, errors in reported wages (to within $100) and employer ID number are cross-checked. Errors in SSN are not detectable, but are thought to be less than 1 percent.

The Employer Address File contains all currently active employers--slightly over 100,000--plus all employers who became inactive within the past 7 years. This file is updated constantly and is also believed to be relatively error-free. Each employer records contains a 4-digit SIC code and a mailing address that could be the local headquarters or accountant's address at the option of the reporting employer.

The Claims File includes all claims processed from 1972 to the present on magnetic tape. In each quarter, old claims (1 year since benefit year) are purged to a separate tape, but these tapes are still available. This file contains a 9-digit DOT code.

LOUISIANA

Louisiana's data files are extremely well organized. Their active Quarterly Wage File is kept on disk for easy access. The file is sorted by employee SSN. Within each employee record is the employer name, ID number, mailing address, and total wages for each employer in the most current 5-quarter period. At the start of a new quarter, the oldest quarter's data are put onto microfilm. Here they are sorted by the last

4 digits of the employee's SSN and each record contains only employee SSN, employer ID, year and quarter, and total wages. This microfilming of purged records began in the fourth quarter of 1976.

There are approximately 73,000 employers currently active in Louisiana. The employer file is also on disk and includes all employers who have been active within the past 3 years. Sorted by employer ID number, the file includes the employer name, mailing address, federal account number, predecessor-successor information, inactive date where applicable, total employees and wages for the most current 5 quarters, and a 6-digit SIC code. Employers who have become inactive since 1968 have been purged to a history tape that is still accessible.

For the most current 5 years, up to two claims per person are kept on an active disk file, accessed by claimant SSN. In addition to a benefit history, this file also contains the standard BLS demographic information, including a 9-digit DOT code. The 3-digit SIC code of the claimant's last employer is included. Most claim files back to 1 November 1970 also appear on microfilm.

The machine readable data only cover the past 5 years because the current data processing system was installed in 1976. As new data are put on line, the older data will not be destroyed. The information will be kept for at least 7 years and perhaps longer.

NEW MEXICO

The Wage Report File on tape goes back to 1976; microfilmed quarterly wage reports go back to 1940. Presently, there are about 350,000 employees covered by the file.

The Employer Master File on tape covers the 28,000 employers currently active, plus all employers who have become inactive since 1968, when all files were computerized. The Employer Master File, which includes all employers active since 1940, is on index cards, sorted by the employer's state identification number. Microfilming of this file is believed to have coincided with computerization. The file includes both a local street address and a headquarters address. Since about 1970, a 5-digit SIC code has been assigned; prior to that, the states were only required to carry a 3-digit code.

There are 3 years of claim data on the computerized claims file; 5 years of claims are on microfilm.

NORTH CAROLINA

North Carolina's quarterly wage file is stored on tape for only 6 quarters after which it's purged. There are 2.9 million wage records on the file, none of which are microfilmed. The wage reports are kept in-house for 6 quarters and then sent to a warehouse where they are held

for an additional 3 years. The reports are balanced against total wages, and the error rate is extremely low.

The Employer Address File contains a 4-digit SIC code in addition to the old 1967 code. Home office address, ownership codes, and total taxable wages are also included. The 110,000 employers are covered for 6 quarters, with inactive files remaining on an additional 4 years. Microfilming of the employer's reports goes back monthly to 1973.

Approximately 10 percent of the workers in North Carolina appear on the claimant file. This includes such descriptive characteristics as race, sex, and year of birth. The claim file is on tape for 2 years, but goes back to 1971 on microfilm.

The middle two numbers of the Employer ID number (which appears on the wage file) represents the county of employment. This may be of particular interest in future analyses.

OKLAHOMA

Quarterly wage reports on tape go back to the first quarter of 1979. There are approximately 1.4 million records per quarter. In addition to wages, the file contains the 2-digit major SIC of the worker's last employer. Microfilmed wage records go back to the third quarter of 1975. State law requires some form of record to be kept a minimum of 6 years. Most microfilm is therefore on a 6-year purge cycle.

Employer information is kept on an address file, which contains 63,000 current employers, plus inactive files for a period of 10 years. The file, which is updated every year, can be accessed on microfilm since 1975.

Claims data is stored on tape from the first quarter of 1976 through the current quarter. There are approximately 200,000 workers who can have their descriptive characteristics and claims records traced on this file. Microfilm provides an additional year of data.

PENNSYLVANIA

Quarterly Wage Files are kept on tape for 7 years. At the time of our visit (November 1981), the file included all wage reports from 1973 to the present time. This file, which covers about 5 million workers per quarter, is set up in much the same way as Louisiana's Wage File. The active file contains one record per worker (sequenced by SSN) that includes data on every one of his employers over the most current 5 quarters. At the beginning of each quarter, the oldest quarter is purged to a history tape. This data is kept in another format on microfilm. The original documents are filmed, the films are proofread, then the originals are destroyed. The films, too, are kept for 7 years.

The Employer Address File includes all employers active during the 7-year period. This file includes a street (local) address, a home office address, and a 4-digit SIC code. Pennsylvania keeps track of proprietorship or name changes in third file, the Predecessor-Successor, Cross-Reference file.

The Claim File is a computerized record of all claims from August 1968 to the present. It includes the major industry SIC code (2-digit) of the employer for the major portion of the claimant's relevant work history, as well as the employer ID number.

SOUTH CAROLINA

South Carolina is much the same as North Carolina and Georgia. Computer-readable information in South Carolina is limited. The 1.4 million wage items are kept to the second quarter of 1980. The system is on a 6-quarter purge cycle. No microfilming of any employees wages is done at the South Carolina SESA. Quarterly wage reports are kept 2 years in-house, after which they are sent to a warehouse for an additional 2 years.

The Employer Address File (53,000) is extremely complete, containing the employer's ID, address, and 4-digit SIC code back to 1972. This includes inactive employers. There are no plans to purge the file. The same information is kept on microfilm back to 1974.

Claims information is maintained on two separate files, both of which will soon cover 5 years. Combined, the files include SSN, FIPS code, education, address, and other descriptive data. Claims records only began being entered into the system in 1977.

TEXAS

Texas' quarterly wage reports are kept on tape back to the first quarter of 1977. Currently, there are approximately 5.6 million wage records quarterly. All wages were microfilmed from 1939 through 1980, much of which we have taken custody of.

The state's Employer Master File presently lists 272,000 employers. This file contains a 4-digit industry code, employer's address, tax information, and predecessor-successor data. The file is on tape back to 1977 and includes inactive employers for a period of 3 years. Last year, the state began microfilming all employer records as they were processed.

Claims data contain such descriptive data as sex, birth date, race, phone numbers, and home address. Sixteen quarters are maintained on tape, with an additional year being trackable on hard copy. No microfilm of benefits is available.

WASHINGTON

Seven quarters of wage reports are kept on an active magnetic tape file. Each quarter, the tape is rewritten, adding the most recent quarter and deleting the oldest. The original tapes have been kept since the third quarter of 1975, giving us machine-readable data back through the second quarter of 1974. These history tapes are seldom, if ever, mounted to read; therefore, the condition of the tapes is questionable. While Washington now uses packed 6250 BPI tapes, the history tapes are on tapes of earlier vintage and inferior quality, compounding the readability problem. Wage reports have only been microfilmed since the second or third quarter of 1980. The original reports, however, are archived in a warehouse for 7 years. All records from 1974 to the present are currently intact. Washington now covers approximately 2 million workers per quarter.

The current computerized employer address file covers only those employers active in the past 3 years. This file, including 97,000 currently active employers, keeps both a local and a headquarter's address for each employers. The SIC code is four digits. A file covering employers active within the past 7 years is available on microfilm. This file is in the process of being put into an automated system on disk packs; it is due to be completed in March 1982. This system will allow random access to 7 years of employer address data, accessible by name or by employer ID number.

The machine-readable Claim file only covers the third quarter of 1980 to the present, but the file has been microfilmed since 1970. The file has always kept a DOT code for each claimant, but it was not the full 9-digit code until 1978.

APPENDIX B

HEALTH DATA

APPENDIX B

HEALTH DATA

CALIFORNIA

Paper Certificates and Microfilm

Filed death certificates have been maintained by the state since 1905. Certificates are filed by year, by month, and by county. SSNs have been recorded since the 1940s. Multiple causes of death have been listed for decades and are currently recorded as immediate, underlying a), b), c), and contributory. A space for occupation has also been available for a long time. All certificates filed since 1905 have been filmed.

Computerized Data

Since 1960, all certificates have been put on tape with SSNs. A national coding system is used to code underlying cause of death only. The records are organized like the paper certificates and merged onto data tapes with a common data format. The cost of obtaining computerized death certificates is $25/thousand records for the first hundred thousand and $10/thousand records for each hundred thousand thereafter. There are approximately 3 million certificates on tape.

Accessibility of Data

Death certificates are public records in California. Thus, no special authorization is needed to obtain records. For further information, contact:

Bureau of Vital Statistics
410 N Street
Sacramento, California 95814
(916) 445-2684

Summary

Data furthest back - 1905 (paper and film)
SSN added - 1940s
Multiple causes of death - yes

Date computerized - 1960
SSN computerized - 1960
Multiple causes of death coded - no, just underlying.

COLORADO

Paper Certificates and Microfilm

Filed death certificates have been maintained by the state since 1907. Certificates are filed by year, by month, and by county. SSNs have been recorded since 1940. Multiple causes of death have been recorded for a long time and occupation since 1968. All certificates filed since 1900 and up to 1979 have been filmed; Vital Records is about to film 1980.

Computerized Data

Since 1959, all death certificates have been put on tape. However, SSNs were not computerized until 1975. Between 1959 and 1979, only underlying cause of death was coded; since 1980, multiple causes have been coded. Vital Records is hoping to go back and code multiple causes for the 1975-79 period. The records are organized sequentially within years. From 1975 forward the data tapes are very clean.

Accessibility of Data

Although not public record, death certificates can be obtained with proper authorization and guarantees of confidentiality. Contact:

Vital Records
4210 East 11th Avenue
Denver, Colorado 80220
(303) 320-8474

Summary

Data furthest back - 1907 (paper and film)
SSN added - 1940
Multiple causes of death - yes

Date computerized - 1959
SSN computerized - 1975
Multiple causes of death coded:
--1959-79 - no, underlying only
--1980 - yes, multiple coded
--1975-79 - expect to go back and code multiple causes of death.

DELAWARE

Paper Certificates and Microfilm

Filed death certificates have been maintained by law since 1913. In 1949, SSNs were added. Multiple causes of death have for a long time

been coded as immediate, seconday, tertiary, contributory, and a space for occupation provided. Certificates are filed sequentially. Microfilm exists back to 1913.

Computerized Data

Data tapes were not created until 1980. They contain SSNs and code only the immediate cause of death (underlying is not coded).

Accessibility of Data

Although not public record, death certificates are obtainable with proper authorization. Contact:

Mr. Michael R. Richards
Assistant Registrar
Director of Vital Statistics
Bureau of Vital Statistics
Jesse S. Cooper Building
Dover, Delaware 19901
(302) 736-4721

Summary

Data furthest back - 1913 (paper and film)
SSN added - 1949
Multiple causes of death - yes

Date computerized - 1980
SSN computerized - 1980
Multiple causes of death coded - no, just immediate.

GEORGIA

Paper Certificates and Microfilm

Filed death certificates have been maintained by the state since 1919. Certificates are filed by year and separated on a weekly basis by county. SSNs have been recorded at least as far back as the 1950s. Multiple causes of death are recorded, as well as usual occupation. Microfilm also goes back to 1919.

Computerized Data

All certificates have been put on tape since 1961. However, SSNs were only added in 1979. Prior to 1979, only underlying cause of death was coded; since then multiple causes have been coded according to a national coding scheme. Records are stored sequentially by year.

Accessibility of Data

Death certificates are not public record, but are accessible with authorization. Contact:

Mr. M. Lavoie, Director
Vital Records
Room 2174
47 Trinity Avenue, S.W.
Atlanta, Georgia 30334 (-1202)
(404) 656-6696

Summary

Data furthest back - 1919 (paper and film)
SSN added - 1950s
Multiple causes of death - yes

Date computerized - 1961
SSN computerized - 1979
Multiple causes of death coded:
--1961-79 - only underlying
--1979-present - multiple.

LOUISIANA

The Registrar's office was unwilling to release information to us by telephone other than that Louisiana became a death certificate registration state in 1918. The office is willing to respond to written requests for information from federal agencies. Address requests to:

Mr. Stanley Brown, Registrar
P. O. Box 60630
New Orleans, Louisiana 70160

NEW MEXICO

Paper Certificates and Microfilm

Filed death certificates have been maintained by the state since about 1920. Paper documents are filed by year and by county. SSNs have been recorded since 1936-37. Multiple causes of death have been recorded since the 1920s with a space provided for occupation. Death certificates filed between 1927 and 1980 have been microfilmed. The Vital Statistics Bureau is currently working on filming those filed between 1920 and 1927.

Computerized Data

Since 1964-65, certificates have been put on tape. SSNs were computerized in 1979. Only the underlying cause of death is coded. The records are sequential.

Accessibility of Data

Although death certificates are not public record, they are accessible with authorization. Contact:

Mr. Michael Ammann
State Registrar
Vital Statistics Bureau
P. O. Box 968
Sante Fe, New Mexico 87504 (-0968)
(505) 827-2587

Summary

Data furthest back - 1920s (paper and film)
SSN added - 1936-37
Multiple causes of death - yes

Date computerized - 1964-65
SSN computerized - 1979
Multiple causes of death coded - no, just underlying.

NORTH CAROLINA

Paper Certificates and Microfilm

Filed death certificates have beem maintained by law since 1913. Certificates are filed by year, by county, and by month. SSNs have been recorded since the 1930s. Old death records simply state cause of death. Since the 1940s, multiple causes (immediate, underlying, and contributory) have been recorded. Occupation is also recorded. Microfilm back to 1913 exists in archives.

Computerized Data

All certificates, along with SSNs, have been put on tape since 1968. From 1968-75, only the underlying cause of death was coded. Since 1975, multiple causes have been coded. The records are stored in the same order as the paper certificates.

Accessibility of Data

Data certificates are not public record, but are accessible with authorization. Requests should include type of data required, purpose

for requesting, study methodology, and assurances of confidentiality. Contact:

Edward R. Warren
State Registrar
Vital Records Branch
P. O. Box 2091
Raleigh, North Carolina 27602
(919) 733-3000

Summary

Data furthest back - 1913 (paper and film)
SSN added - 1930s
Multiple causes of death - yes, since 1940s

Date computerized - 1968
SSN computerized - 1968
Multiple causes of death coded:
--1968-75 - just underlying
--1975-present - multiple.

OKLAHOMA

Paper Certificates and Microfilm

Filed death certificates have been maintained by the state since 1908; however, it was not until the 1950s that the state required registration of deaths. Certificates are filed by year, by certificate number (i.e., sequentially). SSNs were added in 1968. Multiple causes have been recorded for a long time, as well as occupation. Currently, only certificates back to 1970 have been microfilmed, but the state is continuing to microfilm further back.

Computerized Data

Good data tapes exist only as far back as 1975. SSNs were not coded however until 1979. The tapes code only the underlying cause of death.

Accessibility of Data

Death certificates are not public record, but are accessible with proper authorization. Contact:

Roger Pirrong
State Registrar
1000 N.E. 10th Street
Oklahoma City, Oklahoma 73105
(405) 271-4542

Summary

Data furthest back - 1908 (paper), 1970 (film)
SSN added - 1968
Multiple causes of death - yes

Date computerized - 1975
SSN computerized - 1979
Multiple causes of death coded - no, just underlying.

PENNSYLVANIA

Paper Certificates and Microfilm

Filed death certificates have been maintained by the state since 1906. Certificates are filed by year, by county, and then alphabetically. SSNs have been recorded since 1941. Cause of death is recorded as immediate cause a), b), c), and a space is provided for usual occupation. Death certificates filed since 1978 have been filmed.

Computerized Data

Since 1979 death certificates have been put on tape with SSNs. Cause(s) of death is not coded on the tapes; only a yes or no is provided. Yes means cause(s) of death is listed on the certificate, no means it is not. The records are organized by year.

Accessibility of Data

Death certificates are not public record, but records are available with authorization. Contact:

Mr. Charles Hardester
Division of Vital Records
P. O. Box 1528
New Castle, Pennsylvania 16103
(412) 656-3138

Summary

Data furthest back - 1906 (paper), 1978 (film)
SSN added - 1941
Multiple causes of death - yes

Date computerized - 1979
SSN computerized - 1979
Multiple causes of death coded - no, cause(s) of death is not coded; only yes or no (i.e., yes means cause(s) is listed on paper certificate.

SOUTH CAROLINA

Paper Certificates and Microfilm

Filed death certificates have been maintained by the state since 1915. Certificates are organized by year, by month, and by county. Since 1938, SSNs have been recorded. Prior to 1930, only cause of death was recorded. In the 1930s, principal, related, and contributory causes were added. Since the 1960s (possibly before), causes of death have been recorded as immediate, due to, as a consequence of, etc. A space for occupation has always been provided. All certificates filed since 1915 have been filmed.

Computerized Data

Since 1969, all certificates have been put on tape with SSNs. From 1969-1979, only the underlying cause of death was coded. Since 1980, the multiple causes have been coded, according to a standard national coding system. The records can be ordered as requested.

Accessibility of Data

Death certificates are not public record. However, the records are accessible with authorization. Contact:

Mr. Murray Hudson, Director
Office of Vital Records and Public Health Statistics
D.H.E.C.
2600 Bull Street
Columbia, South Carolina 29201
(803) 758-5511

Summary

Data furthest back - 1915 (paper and film)
SSN added - 1938
Multiple causes of death - before 1930 - only underlying, since 1930 - multiple

Date computerized - 1969
SSN computerized - 1969
Multiple causes of death coded - 1969-79 - only underlying, 1980 - multiple.

TEXAS

Paper Certificates and Microfilm

Filed death certificates have been maintained by the state since 1903. Certificates are filed by year, by month, and in alphabetical county order. SSNs have been recorded since the 1940s. Multiple causes of death appear to have always been recorded, as well as occupation. All certificates filed since 1903 have been filmed.

Computerized Data

Since 1964, all certificates have been put on tape with SSNs. A national coding system is used to code multiple causes of death, and a standard code is used for occupation. The records are organized by year, by month, and by county.

Accessibility of Data

Death certificates are not public record, but records are available with authorization. Contact:

Mr. W. B. Carroll
State Registrar of Texas
Bureau of Vital Statistics
Texas Department of Health
1100 W. 49th
Austin, Texas 78756
(512) 458-7451

Requests for data should include guarantees of confidentiality, how the data will be maintained, and the reason why death certificates are being requested.

Summary

Data furthest back - 1903 (paper and microfilm)
SSN added - 1940s
Multiple causes of death - yes

Date computerized - 1964
SSN computerized - 1964
Multiple causes of death coded - yes.

WASHINGTON

Paper Cerificates and Microfilm

Filed death certificates have been maintained by the state since 1907. Certificates are filed by year and by county, chronologically.

At least since the 1950s, SSNs have been recorded. Before the 1960s, only cause of death was recorded; since then, multiple causes have been recorded as immediate cause a), b), c), contributory conditions, etc. Occupation has been listed since the 1960s. All certificates filed since at least the early 1960s (and perhaps the 1950s) have been filmed.

Computerized Data

Since 1967, Vital Records began putting deaths on computer tape. However, it was not until 1979 that SSNs got coded. Only the underlying cause of death is coded on tape, and an occupation code is provided. Tapes are in chronological order by year.

Accessibility of Data

Although not public record, death certificates are accessible with proper authorization. Contact:

Mr. Tom Steinburn, Registrar
Vital Records
P. O. Box 9709
Olympia, Washington 98504
(206) 753-5944

Summary

Data furthest back - 1907 (paper); 1960s (film)

SSN added - 1950s, maybe 1940s

Multiple causes of death coded - only since the 1960s

Date computerized - 1967

SSN computerized - 1979

Multiple causes of death coded - no, only underlying.

INDEX TO PRI PUBLICATIONS

73-1	*The Retail Price of Heroin: Estimation and Applications*, George Brown and Lester Silverman, May 1973.
73-2	*The Use of Longitudinal Data to Assess the Impact of Manpower Training on Earnings*, Louis Jacobson, July 1973.
74-1	*The Effect of Unemployment Insurance and Eligibility Enforcement on Unemployment*, Arlene Holen and Stanley Horowitz, April 1974.
197-75	*Earnings Losses of Workers Displaced from the Steel Industry by Imports of Steel*, Louis Jacobson, August 1975.
260-76	*The Effects of Effluent Discharge Limitations on Foreign Trade in Selected Industries*, James Jondrow, David Chase, Christopher Gamble and Nancy Spruill, February 1976.
264-76	*The Labor Market Effects of Unemployment Insurance; Summary of Findings*, Christopher Jehn, March 1976.
312-76	*Voucher Funding of Training: A Study of the G.I. Bill*, David O'Neill and Sue Goetz Ross, October 1976.
CRC 308	*An Evaluation of the GNP Deflator as a Basis for Adjusting the Allowable Price of Crude Oil*, James M. Jondrow and David E. Chase, February 1977.
CRC 313	*Losses to Workers Displaced by Plant Closure or Layoff: A Survey of the Literature*, Arlene Holen, November 1976.
CRC 339	*The Economic Effects of Environmental Expenditures on the Constructions Industry*, James Jondrow, David Chase, Christopher Gamble, Louis Jacobson, Robert Levy, Bruce Vavrichek, September 1979.
CRC 349	*The Unemployment Insurance Tax and Labor Turnover: An Empirical Analysis*, Frank Brechling and Christopher Jehn, April 1978.
CRC 353	*The Tax Base of the U.S. Unemployment Insurance Tax: An Empirical Analysis*, Frank Brechling, April 1978.
CRC 367	*The Quit Rate as a Measure of Job and Pay Comparability*, Frank Brechling and Louis Jacobson, August 1979.
CRC 385	*Earnings Loss Due to Displacement*, Janet Thomason and Louis Jacobson, August 1979.
CRC 386	*Do Finances Influence Airline Safety, Maintenance, and Service?* David R. Graham and Marianne Bowes, April 1979.
CRC 388	*The Economics of Research and Development*, Lawrence Goldberg, October 1979.
CRC 414	*Taxes on Factors of Production: Their Effects on Factor Proportions and Inflation*, Marianne Bowes, Frank Brechling, Kathleen Classen Utgoff, and Bruce Vavrichek, December 1979.
CRC 419	*Labor Adjustment to Imports Under Rational Expectations*, Robert A. Levy and James M. Jondrow, September 1980.
CRC 423	*Earnings Losses of Workers Displaced by Plant Closings*, Arlene Holen, Christopher Jehn and Robert Trost, December 1981
CRC 431	*Evaluating Tax Systems for Financing the Unemployment Insurance Program*, Marianne Bowes, Frank P.R. Brechling, and Kathleen P. Utgoff, June 1980.
CRC 440	*Using Quit Rates to Set Compensation Levels in the Public Sector*, Kathleen C. Utgoff, January 1981.
CRC 441	*The Value of Stable Employment as Inferred From Market Wages*, Robert P. Trost, February 1980.
CRC 442	*Cost Differences in Public and Private Shipyards: A Case Study*, Marianne Bowes, October 1981.
CRC 451	*The Effect of UI Administrative Screening on Job Search*, Louis Jacobson and Ann Schwarz-Miller, June 1982.
CRC 452	*The Availability of Administrative Data to Analyze Trade Adjustment Assistance and Displacement*, Janet Thomason and Louis Jacobson, September 1981.
CRC 461	*Availability of Matchable Employment and Health Data*, Paul Feldman, March 1982.
PP 165	*Effects of Trade Restrictions on Imports of Steel*, James M. Jondrow, November 1976.
PP 166	*Why It's Difficult to Change Regulation*, Paul Feldman, October 1976.
PP 169	*Earnings Losses of Workers Displaced from Manufacturing Industries*, Louis S. Jacobson, November 1976.
PP 170	*A Time Series Analysis of Labor Turnover*, Frank P. Brechling, December 1976.
PP 175	*Public Drug Treatment and Addict Crime*, D.B. Levine, N. Spruill, and P.H. Stoloff, March 1977.
PP 192	*Effects of Unemployment Insurance Entitlement on Duration and Job Search Outcome*, Arlene Holen, August 1977.
PP 193	*A Model of Unemployment Insurance and the Work Test*, Stanley A. Horowitz, August 1977.
PP 194	*The Effect of Unemployment Insurance on the Duration of Unemployment and Subsequent Earnings*, Kathleen P. Classen, August 1977.
PP 195	*Unemployment Insurance Taxes and Labor Turnover: Summary of Theoretical Findings*, Frank Brechling, August 1977.
PP 198	*The Distributional Effects of Unemployment Insurance*, Kathleen P. Classen, September 1977.
PP 202	*Why Regulation Doesn't Work*, Paul Feldman, September 1977.
PP 203	*Efficiency, Distribution and the Role of Government in a Market Economy*, Paul Feldman, September 1977.
PP 232	*Can Policy Changes Be Made Acceptable to Labor?* Louis S. Jacobson, August 1978.
PP 233	*An Alternative Explanation of the Cyclical Pattern of Quits*, Louis S. Jacobson, October 1978.
PP 234 Revised	*Does Federal Expenditure Displace State and Local Expenditure: The Case of Construction Grants*, James Jondrow and Robert A Levy, October 1979.
PP 238	*Unemployment Insurance and The Unemployment Rate*, Kathleen Classen Utgoff, October 1978.
PP 246	*Layoffs and Unemployment Insurance*, Frank Brechling, February 1979.
PP 266	*Taxes and Inflation*, Frank Brechling and Kathleen Classen Utgoff, November 1979.
PP 267	*The Response of State Government Receipts to Economic Fluctuations and the Allocation of Counter-Cyclical Revenue Sharing Grants*, Robert G. Vogel and Robert P. Trost, December 1979.
PP 282	*Labor Adjustment Under Rational Expectations*, James M. Jondrow and Robert A. Levy, December 1980.
PP 299	*Wage Leadership in Construction*, James M. Jondrow and Robert A. Levy, January 1981.
PP 300	*On the Estimation of Technical Inefficiency in the Stochastic Frontier Production Function Model*, James Jondrow and Peter Schmidt, January 1981.
PP 301	*Technical Change and Employment in Steel, Autos, Aluminum, and Iron Ore*, James M. Jondrow, Robert A. Levy and Claire Hughes, March 1981.
PP 302	*The Effect of Imports on Employment Under Rational Expectations*, Robert A. Levy and James M. Jondrow, April 1981.
PP 311	*An Evaluation of UI Funds*, Marianne Bowes, Frank P.R. Brechling, and Kathleen P. Classen Utgoff, May 1981.
PP 312	*The Optimum Speed Limit*, James Jondrow, Marianne Bowes, and Robert Levy, May 1981.


ILMED
8