

# New Data Sets From Rand

## NEW RESEARCH DATA FILES AND DOCUMENTATION

The **RAND** Corporation is releasing files of research data gathered in the Health Insurance Experiment (HIE), a major project conducted by RAND from 1974 to 1982. The experiment was a large-scale, controlled trial in health care financing, funded by the U.S. Department of Health and Human Services. Its purpose was to assess the effects of different types of insurance on patient health and health care delivery, including both fee-for-service (FFS) and health maintenance organization (HMO) modes.

Over 8,200 persons were enrolled in the experiment in six sites: Dayton, Ohio; Seattle, Washington; Franklin County and Fitchburg, Massachusetts; and Georgetown County and Charleston, South Carolina. Each family was assigned to one of a number of insurance plans that varied by coinsurance rate, delivery mode, and maximum out-of-pocket expenditure. Data were collected on enrollees' use of health care services and state of health throughout their term of enrollment, three or five years. A large subset of the collected data is being released in 67 fully documented files.

Documentation for each file includes a data tape and hardcopy codebook published as a RAND Note.

- \* The machine-readable tape contains three datasets: data in Statistical Analysis System (SAS) format, data in character format, and a dictionary for the character-format file.
- \* The codebook provides an overview of the experiment, outlines the characteristics of the file and related files, and describes each variable. Variable descriptions include response codes and response frequencies or summary statistics.

Codebooks can be ordered from the RAND Publications Department; indicate the desired titles and N numbers. Data files can be ordered from the RAND Data Facility. For each file desired, indicate the file reference number.

### "MASTER SAMPLE SERIES"

**MS1: Eligibility-Family Changes File.** Data on insurance status, changes in eligibility status, and family relationships. No other file provides this information.

**Sample:** Insured enrollees (including members of HMO control group) and persons of interest, a subset of adjunct enrollees. Total of 9,142 records, one per person.

**Representative Variables** (of the total of 54): Insurance status, start/end dates of insurance coverage, reason for losing coverage, length of time insured, family identifier, individual's relation to family/household head.

**Codebook:** N-2264/1-HHS, Vol. 1: *Codebook for Eligibility-Family Changes File*, by S. M. Polich and C. d'Arc Taylor, May 1986.

**MS2: Full Sample Demographic File.** Baseline, enrollment, and demographic data. No other file provides this information.

**Sample:** All HIE participants—insured enrollees, adjunct enrollees, and baseline-only participants. Total of 26,148 records, one per person.

**Representative Variables** (total, 62): Date of baseline interview, sex, age, race, marital status, education, income, occupation, health insurance, welfare status, hospitalizations, family doctor, medical/dental visits and expenses, assigned experimental insurance plan.

**Codebook:** N-2264/2-HHS, Vol. 2: *Codebook for Full Sample Demographic File*, by S. M. Polich et al., May 1986.

**MS3: Supplemental Data File.** Sample data needed for specific analyses. No other file provides this information.

**Sample:** Overall sample is the same as that in the full sample demographic file; specific sample differs by variable. Total of 26,148 records, one per person.

**Representative Variables** (total, 19): Changes affecting HMO and FFS-HMO analyses, enrollment refusals, identifier for mothers of newborns, revised death date.

**Codebook:** N-2264/3-HHS, Vol. 3: *Codebook for Supplemental Data File*, by S. M. Polich and C. d'Arc Taylor, June 1987.

#### "CLAIMS LINE-ITEM SERIES"

**LI1-LI14: FFS Claims Line-Item Files** (14 files). Detailed FFS claims data, including data for HMO enrollees who used medical/dental services in FFS sector. Number of variables in each file shown parenthetically below.

- LI1. Hospital inpatient services (39)
- LI2. Inpatient physician procedures billed by institutions (36)
- LI3. Drugs prescribed by physicians (52)
- LI4. Supplies prescribed by physicians (44)
- LI5. Services rendered by physicians (53)
- LI6. Drugs sold by physicians (63)
- LI7. Supplies sold by physicians (53)
- LI8. Injections administered by physicians (69)
- LI9. Outpatient services billed by institutions (45)
- LI10. Services rendered by dentists (33)
- LI11. Drugs prescribed by dentists (24)
- LI12. Drugs purchased (37)
- LI13. Supplies purchased from pharmacies (17)
- LI14. Supplies purchased from nonpharmacy suppliers (19)

**Sample:** Insured enrollees who filed claims. Total of 603,998 records, one per line item.

**Representative Variables:** Diagnoses (multiple), provider ID, inpatient/outpatient procedures/services, drugs prescribed and sold, dosage instructions, symptoms, relation to employment/accident, treatment history, charges, supplies bought.

**Codebook:** N-2347/1-HHS, Vol. 1: *Codebooks for Fee-for-Service Claims*, by C. E. Peterson et al., June 1986.

**LI15-LI25: HMO Claims Line-Item Files** (11 files). Detailed services provided or reimbursed by the HMO. Number of variables in each file shown parenthetically below.

- LI15. Hospital inpatient services (34)
- LI16. Inpatient physician services (41)
- LI17. Drugs prescribed by physicians (52)
- LI18. Supplies prescribed by physicians (45)
- LI19. Services rendered by physicians (51)
- LI20. Drugs dispensed by physicians (57)
- LI21. Supplies dispensed by physicians (44)
- LI22. Injections administered by physicians (66)
- LI23. Outpatient services provided by institutions (42)
- LI24. Drugs dispensed (32)
- LI25. Supplies dispensed (13)

**Sample:** HMO participants, a subset of insured Seattle enrollees. Total of 177,566 records, one per line item.

**Representative Variables:** Diagnoses, provider ID, procedures performed, drugs/supplies, imputed charges.

**Codebook:** N-2347/2-HHS, Vol. 2: *Codebooks for Health Maintenance Organization Claims*, by C. E. Peterson et al., August 1986.

**LI26-LI29: FFS Claims for FFS-HMO Comparison** (4 files). Detailed Seattle FFS claims data with imputed charges for physician services to enable dollar comparisons with HMO data. Number of variables in each file shown parenthetically below.

- LI26. Hospital inpatient services (35)
- LI27. Inpatient physician procedures billed by institutions (32)
- LI28. Outpatient services rendered by physicians (50)
- LI29. Injections administered by physicians (64)

**Sample:** Insured Seattle FFS enrollees who filed claims. Total of 70,991 records, one per line item.

**Representative Variables:** Diagnoses, provider ID, imputed charges, procedures performed.

**Codebook:** N-2347/3-HHS, Vol. 3: *Codebooks for Seattle Fee-for-Service Claims for Comparison with Health Maintenance Claims*, by C. E. Peterson, M. Nelsen, and D. L. Wesley, October 1986.

**"AGGREGATED CLAIMS SERIES"**

**AC1: FFS Annual Expenditure File.** Claims data aggregated by year for insured FFS enrollees. Also covers FFS dental usage by HMO enrollees.

**Sample:** Insured enrollees. Total of 25,740 records, one per enrollee per year.

**Representative Variables** (total, 23): Number per year of the following: hospitalizations, physician and nonphysician visits, mental health visits, and dental visits; annual expenditures for inpatient, outpatient, mental health, and dental services.

**Codebook:** N-2360/1-HHS, Vol. 1: *Codebook for Fee-for-Service Annual Expenditures and Visit Counts*, by C. E. Peterson, M. Nelsen, and E. S. Bloomfield, May 1986.

**AC2-AC4: FFS Visit Files** (3 files). Claims data aggregated by outpatient, inpatient, and dental visit for insured FFS enrollees. Covers FFS dental visits by HMO enrollees. Number of variables in each file shown parenthetically below.

- AC2. FFS outpatient visits (46)
- AC3. FFS inpatient visits (55)
- AC4. FFS dental visits (16)

**Sample:** Insured enrollees. Total of 148,123 records, one per enrollee-provider-date of service.

**Representative Variables:** Type of visit, providers, visit dates, procedures, diagnoses, charges.

**Codebook:** N-2360/2-HHS, Vol. 2: *Codebooks for Fee-for-Service Visits—Outpatient, Inpatient, and Dental*, by C. E. Peterson et al., June 1986.

**AC5-AC6: FFS Episode Files** (2 files). Claims data aggregated by treatment episode for FFS enrollees. Description and expenses for each episode (individual file); episode counts and expenses per year (annual file). Number of variables in each file shown parenthetically below.

- AC5. FFS individual episodes (19)
- AC6. FFS annual episodes (42)

**Sample:** Insured FFS enrollees. Total number of records: 99,001 in individual file, one per episode; 21,094 in annual file, one per enrollee per year.

**Representative Variables:** Episode description, start/end dates, diagnosis, expense limit at beginning of year, remaining expense limit at start/end of episode, number of episodes per year by type, expenses per episode type per year.

**Codebook:** N-2360/3-HHS, Vol. 3: *Codebooks for Fee-for-Service Treatment Episodes and Annual Episode Counts*, by C. E. Peterson, C. d'Arc Taylor, and E. S. Bloomfield, June 1986.

**AC7: HMO and Seattle FFS Annual Expenditure File.** Claims data aggregated by year for insured HMO and Seattle FFS enrollees.

**Sample:** Insured Seattle enrollees. Total of 11,221 records, one per enrollee per year.

**Representative Variables** (total, 33): Number per year of the following: hospitalizations, physician and nonphysician visits, mental health visits, imputed expenditures.

**Codebook:** N-2360/5-HHS, Vol. 5: *Codebook for Health Maintenance Organization and Seattle Fee-for-Service Annual Expenditures and Visit Counts*, by C. E. Peterson et al., December 1986.

**AC8-AC9: HMO and Seattle FFS Visit Files** (2 files). Claims data aggregated by health care visit for insured HMO and Seattle FFS enrollees. Number of variables in each file shown parenthetically below.

AC8. HMO and Seattle FFS outpatient visits (45)

AC9. HMO and Seattle FFS inpatient visits (53)

**Sample:** Insured Seattle enrollees. Total of 61,597 records, one per enrollee-provider-date of service.

**Representative Variables:** Type of visit, providers, visit dates, procedures, diagnoses, imputed charges.

**Codebook:** N-2360/4-HHS, Vol. 4: *Codebooks for Health Maintenance Organization and Seattle Fee-for-Service Visits—Outpatient and Inpatient*, by C. E. Peterson, M. Nelsen, and D. L. Wesley, December 1986.

#### **"HIE REFERENCE SERIES"**

**RF1: Codes** (no tape file). N-2349/1-HHS, Vol. 1: *Codes Used in HIE Claims—Diagnoses, Symptoms, Procedures, Drugs, and Supplies*, by M. Nelsen and C. A. Edwards, May 1986. Defines all codes used in HIE claims data, both line-item and aggregated. Includes standard and HIE-created codes: diagnosis (H-ICDA-2), CRVS, supply, reason for visit/symptom, NDC, generic drug, drug therapeutic, American Dental Association procedure. \*\*

\*\*-ICDA-2 refers to the second version of the hospital adaptation of International Classification of Diseases Adapted for Use in the United States; CRVS refers to codes for medical and surgical procedures taken from California Relative Value Studies; NDC refers to National Drug Code.

**RF2: HIE Provider File:** Information about the physicians, hospitals, dentists, and other providers of services to HIE enrollees.

**Sample:** All providers cited in HIE data. Total of 22,658 records, one per provider identifier.

**Representative Variables** (total, 26): Provider type, provider specialty, linking identifier.

**Codebook:** N-2349/2-HHS, Vol. 2: *Providers Cited in HIE Data*, by S. M. Polich, M. Nelsen, and D. L. Wesley, June 1987.

**RF3: User's Guide** (no tape file). N-2349/3-HHS, Vol. 3: *User's Guide to HIE Data*, by C. d'Arc Taylor, S. M. Polich, C. E. Peterson, and E. M. Sloss, August 1987. Analytic possibilities and

limitations of HIE data; suggestions for choosing analytic subsamples and linking data across series, files, years, sample groups, and sites for particular analytic purposes.

**"MEDICAL HISTORY QUESTIONNAIRE SERIES"**

**MH1A-MH3A: Adult Form A** (3 files). Data from self-administered questionnaire on health status, attitudes, and habits. Number of variables in each file shown parenthetically below.

MH1A. Dayton adults at enrollment, form A (364)

MH2A. NonDayton adults at enrollment, form A (373)

MH3A. Adults at exit, form A (408)

**Sample:** Adults (14 and older) when enrolling and when completing assigned term three or five years later. Includes insured enrollees, Dayton control group, and PEG-period-only participants. Total of 9,058 records, one per completed questionnaire.

**Topics:** height and weight, general health, eating habits, sleep and exercise, seat belt use, smoking and drinking, general well being, social activities, life events, symptoms, health perceptions, medical opinions, medical and dental care, effectiveness of health care.

**Codebook:** N-2485/1-HHS, Vol. 1: *Codebooks for Adults at Enrollment and Exit, Form A*, by C. A. Edwards et al., August 1986.

**MH1B-MH3B: Adult Form B** (3 files). Data from self-administered questionnaire on verifiable physical limitations and specific medical disorders. Number of variables in each file shown parenthetically below.

MH1B. Dayton adults at enrollment, form B (282)

MH2B. NonDayton adults at enrollment, form B (480)

MH3B. Adults at exit, form B (490)

**Sample:** Adults (14 and older) when enrolling and when completing assigned term three or five years later. Includes insured enrollees, Dayton control group, and PEG-period-only participants. Total of 9,914 records, one per completed questionnaire.

**Topics:** Vision, hearing, hay fever, teeth and gums, fluoride treatment, acne, thyroid, joints, heart/lung ailments, hypertension, stroke, stomach, kidney/bladder, cholesterol, anemia, diabetes, cancer, surgical conditions (hemorrhoids, hernia, varicose veins), physical/activity limitations, sleeping pill use, missing limbs, antibiotic allergy, effectiveness of health care, immunization, gall bladder/tonsil surgery, female organs, medical care, medical appliances, future health expenses, transportation for health care.

**Codebook:** N-2485/2-HHS, Vol. 2: *Codebooks for Adults at Enrollment and Exit, Form B*, by C. A. Edwards et al., October 1986.

**MH4A-MH6B: Child Forms A and B** (6 files). Data from two parent-completed questionnaires.

Form A pertained to health status, attitudes, and habits. Form B pertained to verifiable physical limitations and specific medical disorders. Number of variables in each file shown parenthetically below.

- MH4A. Dayton children at enrollment, form A (85)
- MH4B. Dayton children at enrollment, form B (63)
- MH5A. NonDayton children at enrollment, form A (151)
- MH5B. NonDayton children at enrollment, form B (224)
- MH6A. Children at exit, form A (147)
- MH6B. Children at exit, form B (235)

**Sample:** Children 5-13 years old, at family's enrollment and exit three or five years later. Includes insured enrollees, Dayton control group, and PEG-period-only participants. Total of 6,958 records, one per completed questionnaire.

**Topics:** Form A: height, weight, general health, fluorides, diet, immunizations, safety practices, learning, getting along, general well-being, symptoms, behavior problems. Form B: teeth, fluoride treatment, eyesight, hearing, ear infections, asthma, hay fever, eczema, anemia, lead poisoning, kidney/bladder infection, bedwetting, cancer, convulsions, tonsils, antibiotic allergy, missing limbs, medical appliances, future health expenses.

**Codebook:** N-2485/3-HHS, Vol. 3: *Codebooks for Children at Enrollment and Exit*, by C. A. Edwards et al., November 1986.

**MH7A-MH9B: Infant Forms A and B** (6 files). Data from two parent-completed questionnaires. Form A pertained to health status and development; form B pertained to verifiable physical limitations and specific medical disorders. Number of variables in each file shown parenthetically below.

- MH7A. Dayton infants at enrollment, form A (76)
- MH7B. Dayton infants at enrollment, form B (28)
- MH8A. NonDayton infants at enrollment, form A (98)
- MH8B. NonDayton infants at enrollment, form B (122)
- MH9A. Infants at exit, form A (94)
- MH9B. Infants at exit, form B (134)

**Sample:** Infants (0-4 years old) at family's enrollment and exit three or five years later. Includes insured enrollees, Dayton control group, and PEG-period-only participants. Total of 3,334 records, one per completed questionnaire.

**Topics:** Form A: height, weight, development, general health, fluorides, diet, immunizations, safety practices, symptoms. Form B: colds, ear infections, eczema, anemia, lead poisoning, cancer, convulsions, tonsils, antibiotic allergy, missing limbs, medical appliances, future health expenses, fluoride treatment.

**Codebook:** N-2485/4-HHS, Vol. 4: *Codebooks for Infants at Enrollment and Exit*, by C. A. Edwards et al., December 1986.

**“HEALTH STATUS AND ATTITUDE SERIES”**

**HS1-HS2: Adult and Child (2 files).** Data derived from medical history questionnaire on enrollees' state of health and attitudes toward health care at enrollment and exit. Number of variables in each file shown parenthetically below.

- HS1. Adults at enrollment and exit (136)
- HS2. Children at enrollment and exit (28)

**Sample:** Insured enrollees: adults (14 and older) and children (aged 0-13) at family's enrollment and exit three or five years later. Total of 5,871 records in adult file, 2,840 records in child file—one per person per file.

**Representative Variables:** Scales of physical health, mental health, social health, perceptions of general health; satisfaction with medical and dental care, cigarette smoking, alcohol consumption, weight, and exercise (adults only).

**Codebook:** N-2447/1-HHS, Vol. 1: *Codebooks for Adults and Children at Enrollment and Exit*, by E. M. Sloss et al., November 1986.

**“MEDICAL DISORDERS SERIES”**

**MD1: Adults.** Data derived from medical history questionnaire and medical screening examination on 17 disorders: acne, anemia, angina pectoris, chronic obstructive airway disease, congestive heart failure, diabetes mellitus, hay fever, hearing loss, hypercholesterolemia, hypertension, joint disorders, kidney disease and urinary tract infection, peptic ulcer disease, sleeping pill and tranquilizer use, surgical conditions, thyroid disease, and vision disorders.

**Sample:** Insured enrollees: adults (14 and older) at enrollment and exit three or five years later. Total of 5,871 records, one per person.

**Topics** (total number of variables, 286): Status and severity of disorder, impact of disorder, results of medical tests.

**Codebook:** N-2446/1-HHS, Vol. 1: *Codebook for Adults at Enrollment and Exit*, by B. H. Operskalski et al., February 1987.

**MD2: Children.** Data derived from medical history questionnaire and medical screening examination on four disorders: allergic conditions, anemia, middle ear disease and hearing impairment, and vision impairment.

**Sample:** Insured enrollees: children (aged 0-13) at family's enrollment and exit three or five years later. Total of 2,840 records, one per person.

**Topics** (total number of variables, 73): Status and severity of disorder, impact of disorder, results of medical tests.

**Codebook:** N-2446/2-HHS, Vol. 2: *Codebook for Children at Enrollment and Exit*, by E. M. Sloss et al., March 1987.



**“DENTAL EXAMINATIONS FILE”**

**DE1: Dental Examinations File.** Data from a dental screening examination on tooth decay and its consequences, and periodontal disease and its severity.

**Sample:** Insured enrollees: persons aged three and older in randomly selected subsample (50-75 percent) of families at enrollment; all persons aged three and older at exit three or five years later. Total of 7,317 records, one per person.

**Representative Variables** (total, 50): Number of decayed primary and permanent teeth, number of missing or extracted primary and permanent teeth, number of filled primary and permanent teeth, oral hygiene index score, and (for those 12 and older) periodontal disease index score.

**Codebook:** N-2506-HHS, Dental Examinations: *Codebook for Adults and Children at Enrollment and Exit*, by E. S. Bloomfield, L. Y. Weissler, and A. M. Bell, February 1987.

**“INSURANCE PREFERENCE FILES”**

**IP1-IP2: Insurance Preference Files** (2 files). Data from self-administered questionnaire on willingness to pay a higher health insurance premium in return for a lower annual out-of-pocket expense limit—three hypothetical premium-expense limit combinations. Number of variables in each file shown parenthetically below.

- IP1. Maximum-dollar-expenditure plans (21)
- IP2. Fixed-dollar-limit plan (25)

**Sample:** Heads of insured enrollee families (except HMO enrollees and enrollees assigned to receive free care) when completing assigned enrollment term. Total of 2,020 records, one per questionnaire recipient.

**Topics:** Family's expense limit during its last year in HIE; premium-expense limit combination for each of three hypothetical offers; degree of willingness to accept each offer.

**Codebook:** N-2508-HHS, *Codebooks for Insurance Preference Files: Relation between Expense Limit and Premium*, by E. S. Bloomfield, L. Y. Weissler, and A. B. Holland, October 1986.

**Order form**

Place a check by the HIE data tapes or publications you wish to order. If you want more than one copy, indicate the number desired. Send the completed form, along with your address, to CIS Business Office, The RAND Corporation, P.O. Box 2138, Santa Monica, CA 90406-2138.

Data Tapes*	Publications
<b>Master Sample Series</b>	
<input type="checkbox"/> MS1                   \$ 50.00	<input type="checkbox"/> N-2264/1-HHS           \$ 7.50
<input type="checkbox"/> MS2                   50.00	<input type="checkbox"/> N-2264/2-HHS           15.00
<input type="checkbox"/> MS3                   50.00	<input type="checkbox"/> N-2264/3-HHS           7.50
<b>Claims Line-Item Series</b>	
<input type="checkbox"/> LI1-LI14               150.00	<input type="checkbox"/> N-2347/1-HHS           25.00
<input type="checkbox"/> LI1-LI14 (SAS version only)   100.00	
<input type="checkbox"/> LI15-LI25             100.00	<input type="checkbox"/> N-2347/2-HHS           25.00
<input type="checkbox"/> LI26-29               50.00	<input type="checkbox"/> N-2347/3-HHS           15.00
<b>Aggregated Claims Series</b>	
<input type="checkbox"/> AC1                    50.00	<input type="checkbox"/> N-2360/1-HHS           7.50
<input type="checkbox"/> AC2-AC4              100.00	<input type="checkbox"/> N-2360/2-HHS           15.00
<input type="checkbox"/> AC5-AC6              100.00	<input type="checkbox"/> N-2360/3-HHS           10.00
<input type="checkbox"/> AC7                   50.00	<input type="checkbox"/> N-2360/5-HHS           10.00
<input type="checkbox"/> AC8-AC9              100.00	<input type="checkbox"/> N-2360/4-HHS           15.00
<b>HIE Reference Series</b>	
<input type="checkbox"/> RF1 (no tape exists)	<input type="checkbox"/> N-2349/1-HHS           20.00
<input type="checkbox"/> RF2                   50.00	<input type="checkbox"/> N-2349/2-HHS           7.50
<input type="checkbox"/> RF3 (no tape exists)	<input type="checkbox"/> N-2349/3-HHS           10.00
<b>Medical History Questionnaire Series</b>	
<input type="checkbox"/> MH1A-MH3A            50.00	<input type="checkbox"/> N-2485/1-HHS           25.00
<input type="checkbox"/> MH1B-MH3B            50.00	<input type="checkbox"/> N-2485/2-HHS           25.00
<input type="checkbox"/> MH4A-MH6B            50.00	<input type="checkbox"/> N-2485/3-HHS           25.00
<input type="checkbox"/> MH7A-MH9B            50.00	<input type="checkbox"/> N-2485/4-HHS           20.00
<b>Health Status and Attitude Series</b>	
<input type="checkbox"/> HS1-HS2               50.00	<input type="checkbox"/> N-2447/1-HHS           25.00
<b>Medical Disorders Series</b>	
<input type="checkbox"/> MD1                   50.00	<input type="checkbox"/> N-2446/1-HHS           25.00
<input type="checkbox"/> MD2                   50.00	<input type="checkbox"/> N-2446/2-HHS           15.00
<b>Other Files</b>	
<input type="checkbox"/> IP1-IP250.00 <input type="checkbox"/> N-2508-HHS	7.50
<input type="checkbox"/> DE1                   50.00	<input type="checkbox"/> N-2506-HHS             7.50
<input type="checkbox"/> All 67 files	1000.00

\*Except where noted, each tape contains three datasets: (1) data in Statistical Analysis System (SAS) format, (2) data in character format, and (3) dictionary for character-format data.