

A RAND NOTE

THE EPISODES-OF-ILLNESS PROCESSING SYSTEM

Joan Keesey, Emmett Keeler, William Fowler

January 1985

N-1745-1-HHS

Prepared for

The U.S. Department of Health and Human Services

Rand

1700 MAIN STREET
P.O. BOX 2138
SANTA MONICA, CA 90406-2138

The research reported herein was performed pursuant to Grant No. 016B-80 from the U.S. Department of Health and Human Services, Washington, D.C.

The Rand Publications Series: The Report is the principal publication documenting and transmitting Rand's major research findings and final research results. The Rand Note reports other outputs of sponsored research for general distribution. Publications of The Rand Corporation do not necessarily reflect the opinions or policies of the sponsors of Rand research.

PREFACE

This work is part of the Rand Health Insurance Study (HIS), sponsored by the U.S. Department of Health and Human Services. It documents a system to process medical expenditure claims into episodes of illness.

The first version of the episode-of-illness programs was created by David Cates, Tom Davis, Andrea Henninger, Joan Keesey, Helene Mills, and Sherry Trees under the direction of William Fowler. The idea for grouping medical charges by episode and dating the episode to when the patient might know how much it would cost comes from R-1514-OEO/NC, *Deductibles and the Demand for Medical Services: The Theory of the Consumer Facing a Variable Price Schedule Under Uncertainty* by Emmett Keeler, Joseph Newhouse, and Charles Phelps. The specific rules for linking episodes were made up by Emmett Keeler, with much help from George Goldberg, M.D., who also created the procedures and lists used by Rand and Glen Slaughter and Associates, the HIS Administrative Service Subcontractor, to divide various types of services, drugs, and supplies into acute, chronic, or elective categories.

In 1979, the system was revised to fix some minor problems uncovered by empirical work by Naihua Duan and Emmett Keeler on data from the first version. The revision also permitted more flexible analysis of scope-of-coverage questions by specifically tagging certain types of expenditures whose coverage would be most controversial in a national plan. The rules were revised by Emmett Keeler with help from Joan Keesey and Bonnie Scott, M.D. Joan Keesey rewrote the programs to incorporate the revisions.

Explanation and documentation in the first version of this note refer to the revised version of February 1979. The 1979 version was used in R-2829-HHS, the April 1983 *JAMA* paper by Keeler and Rolph, and the 1984 paper by Leibowitz et al., in *Pediatrics*.

The programs were again revised in 1984 to improve linkup of drugs to episodes, and to give additional information on each episode. Joan Keesey rewrote the programs and documentation to reflect these changes. This version will be used to create the Episodes Public Use Files.

CONTENTS

PREFACE	iii
FIGURES	vii
TABLES	vii
Section	
I. INTRODUCTION	1
Potential Advantages of Episodic Analysis	1
How Economic Episodes are Defined	3
Other Objectives of Processing	4
II. Overview of the System	5
How Claims are Coded and Entered in the HIS Data Base ...	5
The Claims Inventory File (CIF)	12
The Episodes-of-Illness Programs	14
III. Description of Programs Used for the Episodes-of-Illness	
Analysis	22
Introduction	22
Program: REIMB.V120	25
Program: SPLIT.V130	27
Program: SPECM.V230	29
Program: HOSP.V310	40
Program: HOSP.V320	43
Program: UPDAT.V330	44
Program: MRGy.V410	44
Episodes-of-Illness: EPIS.V420	45
Program: REPRy.V440	58
Program: SAMPy.V450	60
Program: MTCHy.V450	60
Program: MTCHy.V460, MTCHy.V470	61
Program: CORRy.V480	61
Program: SAMPy.V490	62
Program: AGGRy.V520	62
Program: REPRy.V520	63
Program: AGGRy.V540	63
Program: REPRy.V550	64
Appendix	
A. REASON CODES AND EXPLANATIONS	65
B. LIST OF ANNIVERSARY DATES	67

FIGURES

2.1. Physician MER	6
2.2. MER data	8
2.3. Flow chart of episodes-of-illness processing	16

TABLES

3.1. Definition of Variables Created for Episode-of-Illness Algorithm and Stored in the Array "VARPAD(15)" of the Claim Inventory Records	23
3.2. Relationship Between Type of MER and the Variable, DOC_SEQ .	24
3.3. Drug Table	31
3.4. Supply Table	34
3.5. Potential CLM_TYPE/DSI Assignment	41
3.6. Episode-of-Illness Record Structure	59
A.1 MER Worksheet Coding Instructions	66
B.1 List of Anniversary Dates	69

I. INTRODUCTION

This Note is designed to document the methods and assumptions used to process the expenditure claims of families and organize them into episodes which reflect the economics of participants' medical decisions.

The rest of this section gives the potential advantages of shorter than annual analysis, and the rationale for economic episodes. The second section gives an overview of what happens to the claims from their arrival at Glen Slaughter and Associates (GSA)¹ to their emergence as a series of dated and classified episodes.

The third section gives a more detailed description of the assumptions and procedures.

POTENTIAL ADVANTAGES OF EPISODIC ANALYSIS

The Family Medical Protection Plans given to participants in the HIS have an annual accounting period, so it is natural to analyze expenditures to the participants on an annual basis. We will get unbiased and reasonably precise results about annual expenditures for each plan in the experiment. However, we collect information from the participants much more frequently. In particular, we have detailed reports from each expenditure claim they file. Can we use this information to improve our results?

There are many potential advantages of studying shorter than annual time periods. First, most medical services are purchased by sick people during an episode of illness. The episode is the natural sales unit for analyzing the effects of price, and annual analysis using average or marginal out-of-pocket price is biased if a deductible is present (Keeler et al., op. cit.). Second, with shorter time periods the sample size increases. For example, there are twelve times as many monthly observations as there are annual. If the periods are too short, almost no one will have any expenditures in any given period, but many

¹Glen Slaughter and Associates act as Rand's subcontractor in the administration of the insurance plans and in the collection of related data.

participants will have had expenditures in six or more different months. If we can analyze the shorter time periods, the increased sample size must improve precision.

Third, the situations arising through the year are natural experiments for different-sized deductibles. A deductible is the amount a policyholder must pay before insurance takes over; deductibles are of great interest to economists because they can preserve some consumer market incentives at a low cost in risk. To keep the number of different plans small, we limited the plan types to one individual deductible at \$150, and three family deductibles at 5, 10, and 15 percent of income. During the course of the year, individuals and families with varying expenses up to that point will find themselves with different remaining amounts of deductible. By analyzing their behavior at those points, we may be able to better extrapolate our results to insurance plans that were not in our design.

Since all the plans have a limit on out-of-pocket expenditures, any family that exceeds that amount has all medical services free for the rest of the year. It is interesting to see how much families try to take advantage of their "sale" condition. Understanding "sales" is especially important to analysis of scope-of-coverage questions, since many of the medical services whose coverage is most debatable--eyeglasses, hearing aids, psychiatric services, dentures, and elective surgery--should be strongly influenced by sales. Sales are important in another way to the experiment--those people with the best Family Health Protection Plans may feel that the experiment itself is a long sale. Shorter than annual analysis of the other plans can provide additional evidence bearing on catch-up demand at the beginning and end of the experiment.

Finally, the detailed study of shorter time periods can be used to group expenditures into chronic, acute, well, and elective categories. In theory, decisions about these categories will be made differently, and breaking them down for analysis may permit better modeling and more precise estimates of demand.

HOW ECONOMIC EPISODES ARE DEFINED

We generally assume in our analysis that participants act as if they understand their insurance policies. That is, we assume that they know their share of expenses as well as which medical services are covered. (This assumption is tested directly in the experiment.) Since each plan has a deductible (except full coverage), the price depends on the amount of medical services already purchased. This information is provided by GSA with each claim processed, and is thus available to participants even if they haven't kept a record themselves.

However, the rational economic person doesn't base his actions on the date he receives the bill, but on the date he knew the expense would be incurred. Suppose a woman learns she is pregnant and will deliver before the end of the year. The hospital bill alone at that time will exceed her individual deductible. She doesn't have to wait to get her new glasses, or to go more readily to the doctor for mild acute conditions. Since she knows the deductible eventually will be exceeded, she can act as if it already has been.

Thus, for purposes of economic analysis, we try to date each charge to the earliest time that the participant would have known he would spend it. Continuing expenditures for chronic conditions such as diabetes or hypertension are presumably anticipated from the start of the accounting year and are dated to the first office visit for that condition. Expenses carried over from an episode in a preceding year are dated to the first of the year. Dental work is dated to the first in a series of procedures or the preceding examination. We distinguish elective surgery and well-care from acute care by the criterion of deferrability, calling it acute if it loses most of its value if deferred.

We assume that the whole cost of the episode is known to the patient at the time of the first charge. This is not always true, but claims records do not reveal how much the doctor told the patient, or indeed, how much the doctor knew about the course of the episode at that time.

We can examine the sensitivity of this assumption by comparing the results with those under two extreme assumptions. At one extreme, the participants know at the beginning of each year whether or not they will exceed the deductible (i.e., we assume that all people who do exceed the deductible knew that they would from the beginning); at the other extreme, we could analyze charges by billing date assuming that patients have no foreknowledge whatsoever.

OTHER OBJECTIVES OF PROCESSING

The HIS is designed to estimate the demand for medical services under various national health insurances. This will probably be done by a simulation of family expenditures for a variety of families. Such a simulation will be easier if expenditure events are few, and independent of each other. By collecting expenditures in episodes, we reduce their number, and make them more likely to be independent. The reduction from expenditures to episodes may also save money in storage and analysis.

Participants are asked to submit claims for expenses such as orthodontia that are not covered by the Family Health Protection Plans. In processing the claims, we should ensure that these expenditures are separated out from the one we want to analyze. Some of the expenditures we cover, such as eyeglasses or outpatient psychiatric services (up to a limit), may or may not be included in National Health Insurance. This gives them a special interest, and by tagging them, we make their study easier.

The exact dating of episodes makes them useful in studies of short-run causes of illness, such as air pollution. We can use diagnosis to focus on episodes that are particularly likely to follow a hypothesized cause closely.

Finally, the claims have to be coordinated with a record of participants. Some people will not have filed any claims during the course of a year, and we will want to include them together with their zero episodes. Others will have been dropped from analysis or died, and it will be helpful to mark their episodes accordingly.

II. OVERVIEW OF THE SYSTEM

The goal of processing claims is the transformation of several different claim types (called MERS for "Medical Expense Reports") into and through the HIS data base ending up with analytically useful data files. The process consists of the entry of MER information onto data entry coding sheets and into machine readable records at GSA, the translation of the machine readable records into the HIS data base's generalized record structure (called RMIT for "Rand Master Input Tape"),¹ machine verification of transferred data and entry into the HIS data base, the processing of RMIT data into a user-oriented Claims Inventory File (CIF), and a series of programs converting the CIF into episodes of illness. We will discuss these steps in order.

HOW CLAIMS ARE CODED AND ENTERED IN THE HIS DATA BASE

There are four types of MERS reflecting the kinds of services for which subjects may make a claim: (1) hospital, (2) physician and supply, (3) pharmacy, and (4) dental. Each of these MER types collects some information found on all of the others (for example, name, date of service, provider number, etc.). Some of the MERS collect information common to some but not all other MERS (for example, diagnosis is collected on physician and hospital MERS but not pharmacy MERS). Finally, all MERS collect information unique to themselves. A physician MER is shown as Fig. 2.1. As can be seen, MERS are organized on the basis of line items, which are specific services for which charges are filed. In principle, a MER may contain any number of line charges. Most information on a MER refers to a specific line item. Note that providers are asked to itemize services, and for each service give the amount charged, the date of service, and the diagnosis(es) to which the service is related. For the related diagnoses, the provider gives the

¹For a comprehensive description of the RMIT format and other related subjects, see Rand publications *Rand Master Input Tape (RMIT): Description* (RCC-1550/11), and *Rand Master Input Tape Generator System (RMITGEN): System Description* (RCC-1550/12) by Helene Mills.

FAMILY HEALTH PROTECTION PLAN

PHYSICIANS, DOCTORS, SUPPLIERS AND OUTPATIENT MEDICAL EXPENSE REPORT

(Use this form for all outpatient charges: clinics, surgery, emergency, etc.)
MAIL TO: FAMILY HEALTH PROTECTION PLAN, P.O. BOX 2878, Oakland, CA, 94604

PART 1: PATIENT INFORMATION

1. Last Name of Patient		First	M.I.	2. Sex	3. Age	4. Patient's Family No.
5. Patient's Address		City, State, Zip Code			6. Patient's Individual No.	

7. What Was The Major Reason or Symptom For This Visit To The Doctor?

Was illness or injury
Employee related? YES ☐ NO ☐

Was illness or injury
Accident related? YES ☐ NO ☐

10. Date of Injury or Accident: / /

11. Describe how and where accident occurred:

7A. Date You First Noticed This Symptom: (For illness or Accident)

12. Name of Doctor, Supplier or Outpatient Facility

13. Has the Patient Ever Visited This Doctor, Supplier or Outpatient Facility Before?
YES ☐ NO ☐

14. I authorize any holder of medical or other information about the patient to release to the Family Health Protection Plan or its intermediaries any information needed for this or related medical reports. I permit a copy of this authorization to be used in place of the original. In conformance with the Family Health Protection Plan Enrollment Agreement, all health care benefits covering the Patient are hereby assigned to the Family Health Protection Plan.

Signature of Adult Participant or Signature of Minor Participant: _____ Print Adult's Name: _____ Date Signed: _____

SIGN HERE

PART 2: DOCTOR OR SUPPLIER TO FILL IN ITEMS 15 THROUGH 23 PLEASE PRINT OR TYPE

15. Full Name of Referring Doctor: IF NONE, WRITE NONE.

16. Full name(s) of Providers to Whom You Referred Patient for Consultation, Lab Tests, or Other Services: IF NONE, WRITE NONE.

17. Describe the Primary Problem or Diagnosis Which Brought the Patient To Your Office and Any Other Problem(s) for Which You Supplied Treatment
Please List Primary Problem or Diagnosis on Line A.

18. Type of Problem (check one)

<input type="checkbox"/> Acute Well Care (or pregnancy)	<input type="checkbox"/> Flare-up of Chronic Chronic (not flare-up)	<input type="checkbox"/> Initial Visit for this episode Repeat Visit for this episode
<input type="checkbox"/> Acute Well Care (or pregnancy)	<input type="checkbox"/> Flare-up of Chronic Chronic (not flare-up)	<input type="checkbox"/> Initial Visit for this episode Repeat Visit for this episode
<input type="checkbox"/> Acute Well Care (or pregnancy)	<input type="checkbox"/> Flare-up of Chronic Chronic (not flare-up)	<input type="checkbox"/> Initial Visit for this episode Repeat Visit for this episode
<input type="checkbox"/> Acute Well Care (or pregnancy)	<input type="checkbox"/> Flare-up of Chronic Chronic (not flare-up)	<input type="checkbox"/> Initial Visit for this episode Repeat Visit for this episode

19. Treatment History (omit if well care or pregnancy)

KEY: Place of Service Codes: O = Doctor's Office; IL = Independent Laboratory; H = Patient's Home; IH = Inpatient Hospital; NH = Nursing Home or SNF; ER = Emergency Area; OH = Outpatient Hospital, including Hospital Clinic and Outpatient Surgery; SC = School Clinic; CC = Company Clinic; OL = Other Location, Including Other Non-Hospital Clinic.
Type of Visit Codes: 1 = Minimal Service; 2 = Brief Examination; 3 = Limited Examination; 4 = Intermediate Examination; 5 = Extended Examination; 6 = Comprehensive Examination; 7 = Unusually Complex Examination. SEE DETAILED INSTRUCTIONS ON REVERSE SIDE
For Inpatient Services, Omit 18, 19, and 21.

20	A. Date Of Service	B. Place of Service Use code above	C. Describe Each Medical or Surgical Procedure and Other Service or Supplies Furnished For Each Date, Including Specific Lab Tests and the Specific Name of Any Drug Injected.	D. Type of Office Visit Use code above	E. Relate Treatment to Problem by Ref. to 17 A, B, C or D above	F. Charge	21. Were Any Drugs Prescribed? Were any Supplies Prescribed or Suggested? <input type="checkbox"/> Yes <input type="checkbox"/> No
1							A. If yes, specify drug(s) and/or supply(ies): B. Relate to Problem by Reference to 17 A, B, C or D above
2							
3							
4							
5							

22. Name and Address of Doctor or Supplier

23. Social Security or Provider Tax ID Number

24. TOTAL CHARGE

25. AMOUNT PAID, IF ANY

26. BALANCE DUE

27. I hereby certify that the services and/or supplies listed above have been provided on the date(s) shown.

PROVIDER'S SIGNATURE

Date Signed: _____

28. I hereby authorize payment directly to the above-named provider of the benefits otherwise payable to me but not to exceed the charges shown. I understand that I am financially responsible for any charges not covered by the Family Health Protection Plan.

ADULT PARTICIPANT'S SIGNATURE

Date Signed: _____

HIEI #971 REV. 1-77

MAIL TO FHPP

127

Fig. 2.1 -- Physician MER

type of problem (acute, chronic, chronic flare-up, or well-care), and the treatment history (initial or repeat).

Information received on the MER is hand-copied onto data entry coding sheets. There are one or more types of coding sheets for each MER type. In general, different types of coding sheets associated with a single MER type are used to code different kinds of service. For example, physician MERs are coded onto a drug, injection, and supply sheet and an outpatient supplement sheet as well as a physician service sheet. This is done because different kinds of information are collected for each of those types of items (e.g., a National Drug Code (NDC) is collected for an injection line item but not, say, for a physician examination). At this point, related data that do not appear on the MER itself (such as plan type) are entered into the data entry sheets, as well as information derived from MER data (e.g., as shown in Fig. 2.2, HICDA codes for diagnoses, and California Relative Value Studies (CRVS) codes for services are looked up).

Information on the data entry coding sheet is coded into machine readable records and translated into the RMIT format. RMIT records are entered into the HIS data base after machine verification of specifications. The HIS data base is organized by Health Insurance Experimental Instruments (HIEI), which are sometimes called documents. HIEIs are the data collection forms given to subjects at different times throughout the experiment. Each individual data collection form type is assigned an identification number. The HIS data base is accessed by these numbers. In the case of MERs, however, HIEI numbers reference the data entry coding sheet rather than the MERs themselves, although HIEIs do reflect the four basic MER types (that is, coding sheets are used for one and only one MER type).

Each instance of a line charge is divided into a number of logical records types (the RMIT structure being the physician record type). These record types are numbered sequentially from one to a predefined number of each HIEI (e.g., the physician outpatient HIEI has four record types). The different record types contain different kinds of data. For example, record type one contains information common to all MERs (i.e., person identifiers, provider identifiers, date of service), while

FAMILY HEALTH PROTECTION PLAN **PHYSICIANS, DOCTORS, SUPPLIERS AND OUTPATIENT MEDICAL EXPENSE REPORT** (Use this form for all outpatient charges: clinics, surgery, emergency, etc.)

1. Last Name of Patient DOE		First John	M.I. 	2. Sex M	3. Age 29	4. Patient's Family No. FN000001
5. Patient's Address 100 Anystreet		City, State, Zip Code Anytown, State 00000	6. Patient's Individual No. PN000001			
7. What Was The Major Reason or Symptom For This Visit To The Doctor? COLD		8. Was Illness or Injury Accident-Related? <input checked="" type="checkbox"/> YES <input type="checkbox"/> NO	10. Date of Injury or Accident //		11. Describe how and where accident occurred: 	
7A. Date You First Noticed This Symptom: (For Illness or Accident) 1/4/80		9. Name of Doctor, Supplier, or Facility Smith, S.	12. Has the Patient Ever Visited This Doctor, Supplier or Outpatient Facility Before? YES <input type="checkbox"/> NO <input checked="" type="checkbox"/>			

PHYSICIAN DATA ENTRY															
LINE	PL	TYP	INDIVIDUAL NO.	IND	1ST NAME	PROVIDER NO.	PROV. NAME	PAY	SYMPTOM	SYMPTOM	SYMPTOM	EMP/ACC	DATE OF INJURY	PV	PAYMENT DATE
A	1	08	PN000001		JOHN	PV000001	SMITH, S.	13	12			NW			W020580

17. Describe the Primary Problem or Diagnosis Which Brought the Patient to Your Office and Any Other Problem(s) for Which You Supplied Treatment. Please List Primary Problem or Diagnosis on Line A.

A.	Nasopharyngitis	<input checked="" type="checkbox"/> Acute <input type="checkbox"/> Well Care (or pregnancy)	<input type="checkbox"/> Flare-up of Chronic <input type="checkbox"/> Chronic (not flare-up)	<input checked="" type="checkbox"/> Initial Visit for this episode <input type="checkbox"/> Repeat Visit for this episode
----	-----------------	--	---	--

Diagnosis Codes
Nasopharyngitis

460

LINE	DIAGNOSIS CODE R/O	DIAGNOSIS CODE	SYMP/PROB DATE	DATE 1ST CONSUL	THC
B 1	A	460	+	01048001	1

HIEI 111
RECORD TYPE 1

KEY
Place of Service Codes: O = Doctor's Office; IL = Independent Laboratory; H = Patient's Home; IH = Inpatient Hospital; NH = Nursing Home or SNF; ER = Emergency Area; OH = Outpatient Hospital, including Hospital Clinic and Outpatient Surgery; SC = School Clinic; CC = Company Clinic; OL = Other Location, including Other Non-Hospital Clinic.
Type of Visit Codes: 1 = Minimal Service; 2 = Brief Examination; 3 = Limited Examination; 4 = Intermediate Examination; 5 = Extended Examination; 6 = Comprehensive Examination; 7 = Unusually Complex Examination. **SEE DETAILED INSTRUCTIONS ON REVERSE SIDE**
 For Inpatient Services, Omit 18, 19, and 21.

20	A	Date Of Service	B	Place of Service Use code above	C	Describe Each Medical or Surgical Procedure and Other Service or Supplies Furnished For Each Date including Specific Lab Tests and the Specific Name of Any Drug Injected.	D	Type of Office Visit Use code above	E	Rate Treatment as Problem by Ref to 17 A, B, C or D above	F	Charge
1		1/05/80		0		Office Visit.		1	A			20.00

CRVS Codes
Office Visit

90000

LINE CHARGE	NON-COVERED	REAS.	REIMBURSEMENT
20.00	1.00	031	1.00

LINE	DATE OF SVC	PL	CRVS CODE	CRVS MODIFIERS	CRVS UNITS	CPT CODE	MOD	DIAG.	REL
01	010580	1	90000		1.00	0		1	A B C D Z

HIEI 111
RECORD TYPE 4

Figure 2.2 -- MER data

record type two through (N) typically contains line-item-based information. There is a record type one for each HIEI occurrence, and a record type two through (N) for each appropriate line item. All record types (and HIEI numbers) derived from one MER will contain common identification information (namely, claim number).

Figure 2.2 illustrates how MERs are broken down into data entry coding sheets (and HIEIs), and logical record types.

The coding decisions on three variables are crucial to later episodes of illness processing. These variables are the reasons for noncoverage of charges (when all or part of submitted claims are not covered), the codes allocating charges to diagnoses, and the codes for treatment history of services. Coding decisions are made by human clerks following a set of guidelines developed and continually updated by George Goldberg, M.D. Cases not yet covered by the guidelines are referred to Dr. Goldberg and the Rand Claim Policy Committee for decisions. Even with the guidelines, the clerks must use common sense and their experience. About 2 to 4 percent of the MERs are independently recorded as a validity check, and about once a month a sample coding is checked by Dr. Goldberg.

While it would be possible in theory to automate these coding decisions, the determination of whether a service for a particular condition (which may be described differently in two claims) is part of an old episode or the start of a new, or whether a certain condition is acute or chronic depends heavily on the context. Thus, this human help makes possible what even so is a very complicated machine system. We now discuss the details of coding each of the three variables.

The line charge represents the amount submitted to Family Health Protection Plan (FHPP) by the provider or participant for payment. This charge may not be fully reimbursed for two reasons that are important to distinguish. First, the claim may be an attempt to get payment for a service not deemed by the insurance plan to be medically necessary, e.g., certain hospital amenities, more than an annual dollar limit on eyeglasses, or orthodontia. For the purposes of analysis, such expenditures are not considered to be medical expenditures, and we would like to subtract them out. Second, the reimbursements may be reduced by

the coinsurance that the insured is obliged to pay. Here the line charge represents medical expenditures, and despite the fact that they are not fully paid for by insurance, we want to include them in the analysis. The means for making the distinction is the reason for the noncoverage code, which is provided whenever the line charge is reduced. A list of these reasons is provided in Appendix A. In Fig. 2.2, coinsurance (reason 31) cuts the bill in half.

The second crucial variable is the diagnosis for which the service(s) were rendered. On the physician MER there are four possible concurrent diagnoses, while on the inpatient MER there are seven. Each diagnosis is given a sequence identifier (from A to D for physician claims and from A to H for inpatient claims). This sequence number is used to link individual line charges and services performed to specific diagnoses. One diagnosis is assigned as primary by the physician, and this is assigned the sequence identifier "A".

There are three parts to a complete diagnosis. There is an actual diagnosis, a diagnosis qualifier (called the rule out), and a secondary diagnosis when required by the rule-out code.² For example, a diagnosis might be written as "influenza rule out pneumonia,"³ for which the codes for influenza, "rule out," and pneumonia are entered. HICDA codes⁴ are used for diagnoses.

Each line charge must be related to one or more of the diagnoses as shown in Fig. 2.2. The provider is asked to do this and ordinarily will do so. However, it is not always done and is occasionally done in an obviously incorrect way. In this case, the clerks allocate services to diagnoses following guidelines and decisions that have been developed over time. In addition, if the provider does not apportion charges to services, but bills as one lump, they divide the lump into charges following the CRVS values of the listed services.

²The five possible code values are (1) not a rule out, (2) rule out, (3) possible, (4) with, (5) not.

³This means the physician is considering the possibility that pneumonia may exist, but cannot yet conclude if it is "ruled in" or "ruled out."

⁴Commission on Professional Hospital Activities (1973), *Hospital Adaptation of the ICDA (International Classification of Diseases Adapted for Use in the United States)*, Ann Arbor, Mich.

The final set of important variables is the treatment history coding for the diagnoses found on the MER. There are 11 categories that specify the type of problem and whether or not a MER has been submitted before for the same episode of the particular problem. The categories are:

- initial acute
- repeat acute
- acute (not specified as to initial or repeat)
- initial chronic
- chronic flare-up (initial)
- chronic flare-up (repeat)
- chronic (not specified as to initial or repeat)
- well-care
- unobtainable
- not applicable

The treatment history should be filled in by the provider, but has proved to be incorrect or omitted, on occasion. For example, a 1976 study by GSA of a sample of 104 physician charges found 19 treatment history codes blank, and another 12 incorrect (GSA memo SR-465, June 17, 1976). Particularly if the patient switches providers, the earlier history of the condition may not be known to the provider. For these reasons, we decided to have clerks at GSA inspect all treatment history codes, and change them when necessary. The assignment is done by reference to an MER "history card" maintained by GSA for each subject in the study (to determine whether the problem is initial or repeat)⁵ and to lists of diagnoses. The diagnoses lists are categorized as chronic, well-care, and usually acute (outpatient). There are supplementary rules that are used to assign a treatment history in difficult cases. These lists and rules are in the GSA claims manual.

⁵The most difficult problems are for charges for acute conditions that occur after a medium interval. The minimum interval that results in a new episode being coded is about three weeks, but this is modified by the provider's code and the overall history. The clerks also notice and link probably related diagnoses, such as those in the upper respiratory cluster, when these occur close together.

THE CLAIMS INVENTORY FILE (CIF)

Content and Structure

Because claims data are stored with great efficiency in the HIS data base, considerable preprocessing of claims is necessary before the episode of illness work file can be developed. The primary problem is that data from various documents appear in divided, logically incomplete records. Although there are links between these records, their logical structure is not useful for either analytical efficiency or for understanding the structure of claims data.

The establishment of a standard logical view was the most important aspect of making the claims data usable. We tried to reconstruct the logical view of the claims as they were filled out by the provider. The smallest useful unit of observation on the claim is the line item, and this unit was chosen to be reflected in the CIF record.⁶ The CIF record provides a logically complete and normalized view of claims. There is one record per line charge. The records themselves contain data that are common to all line charge occurrences within a claim (that is, identification information, dates of service, plan type, etc.). The CIF records, therefore, are easily aggregated into claims, or to any other unit that is required by the analysis.

Time Frame

Claim data from GSA are transferred to Rand on a continuous basis. The claims are entered into the HIS data base without regard to either the date of service or date of payment. Claims are entered on the basis of a release number that reflects the data entry process rather than the submittal or payment process. This is a problem for users of claims data because the time period in which the claim is submitted is vital to the analysis. This is especially true in the episodes-of-illness

⁶Even though each line-item-based variable (such as diagnosis and treatment history code) is connected with individual line items, all other (nonrelated) entries are placed in each of the CIF records for a particular MER as well. The related entries are connected with the line charge by a set of relationship codes (1 = related, 2 = not related, -1 unknown or missing). This allows access to other diagnoses within the same claim, even those that were not related to the specific (or perhaps any) line charge.

processing. The primary temporal unit for economic analysis of claims data has been the accounting year.

The accounting year is a twelve-month period during which medical expenditures covered by the Family Health Protection Plans are counted toward deductibles and coinsurance rates. The accounting year begins on an anniversary date. The date that subjects are enrolled in the study acts as the anniversary. There are several possible anniversary dates for each site, but each member of a family has the same anniversary date. The possible anniversary dates for each site are listed in Appendix B.

Claims inventory files are organized by site and accounting year. That is, there will be one CIF for each site year that contains all claims submitted during that accounting year in that site. Because of delays in the claims submittal and payment processes, the first set of episode files did not include all claims that were eventually processed. There are slightly higher expenditures in the final episodes files than were reported in R-2829-HHS.

Families and the Eligible Sample

Defining family constructs within a large sample of families is difficult because most definitional schemes will produce many ambiguities and exceptions. To deal with this problem, the HIS has a number of ways of grouping individuals into families, each designed to meet a particular analytical or accounting need. For the episode-of-illness processing, and most of the other economic analyses of claims, the relevant family unit is called the insurance family.

The insurance family includes those individuals that are covered under a single policy and spend toward a single maximum dollar expenditure (MDE).⁷ The insurance family unit reflects family membership as it is generally understood, with one major exception. The exception is that a single insurance family relationship is maintained until the beginning of the next accounting year, even when the family splits into more than one unit or otherwise changes composition.⁸ Thus,

⁷For a definition of MDE, see Rand publication *Rules of Operation for the Rand Health Insurance Study*, R-1602-HEW, M. E. Brown, L. Clasquin, May 1977.

while a divorce might mean that a single family becomes two distinct entities living at separate addresses, this change will not be reflected in the policy under which they make claims or in the MDE they spend toward until the next anniversary.

Discrepancies caused by the use of the insurance family as the unit of analysis have generally been resolved by deleting split families. Where analysis has been carried out on the split family sample, cases are either resolved by hand on a case-by-case basis or by defining a number of special rules.

THE EPISODES-OF-ILLNESS PROGRAMS

Basic Concepts

The goal of generating episodes of illness is to distinguish among chronic, elective, and acute episodes and then link up different claims that pertain to the same illness. The basic process defined in the original design documents (SM-2526, SM-3005) lists the following steps:

1. List all claims (for an individual) by date.
2. Classify hospital procedures into acute, elective, or pregnancy.
3. Classify all office visits into chronic, well-care, acute, and chronic flare-up.
4. Relate physician visits to hospital claims by date of service.
5. Define episodes by starting with the latest claim and linking it to previous claims by diagnosis and treatment history.
6. Classify dental claims on the basis of a procedure list, and link to previous dental claims by procedure and treatment history.
7. Classify pharmacy claims by type of drug and number and spacing of claims. Link drugs and supplies to episodes.
8. Classify pregnancy and early pediatric services as a separate episode.

⁸There were a few exceptions to this rule early in the experiment, but these exceptions were dropped by spring of 1975.

9. Derive individual episode summary data and add individuals with no episodes.

Final specifications, of course, require much more detail than this, plus many hours of defining and checking specific program processes by hand. (Revisions and their explanation are in SM-3629, SM-4108, SM-5008, SM-5034, SM-4966, SM-5475, SM-5887, SM-6126, SM-7483, SM-10,650, SM-11,650.) With some exceptions, this document will cover detail to the system flowchart level, identifying the functions and processes of each program or segments of the system. The exceptions will be for those processes that created the episode records themselves, as well as those closely related functions and processes.

The system level flowchart appears in Fig. 2.3. Before describing the system on a process by process level, we would like to make some general comments. Developing the episodes-of-illness processing system was not easy. Seven individual programmers contributed to the final system, and many man-months went into its completion. The complexity of the process and the number of programmers working on the project dictated a structured, highly modularized approach to design and implementation. A high-level design was produced and each programmer was responsible for specific segments. Each of these segments represented a separate process, which resulted in the development of one or more programs. The program that actually created the episode-of-illness records was considerably bigger and more complex than the others and required three programmers, each of whom was responsible for one or more program module. This method saved time, but required careful scheduling, adequate documentation, and effective program and personnel interface. The modular approach allowed these problems to be solved with reasonable success.

The claims inventory file, while considerably simplifying claim processing, is a very complex file with eighteen different record types (each of which matches a HIEI-RMIT record type combination). Each of these record types was given a sequence number (called the document sequent number) so that sorting can be done on different types of services. The many record types made program interacting difficult,

particularly when design and implementation were attempted in parallel. The problem was mitigated by specifying that the CIF records would not be altered throughout processing until the final episode creation program, except in a fifteen-variable vector defined within the record to store transformed, calculated, and relocated variables used in the process. This sometimes required additional program development (as shown below), but proved to be highly useful for documentation, debugging, and simplifying program interfaces.

FLOW CHART OF EPISODES OF ILLNESS PROCESSING

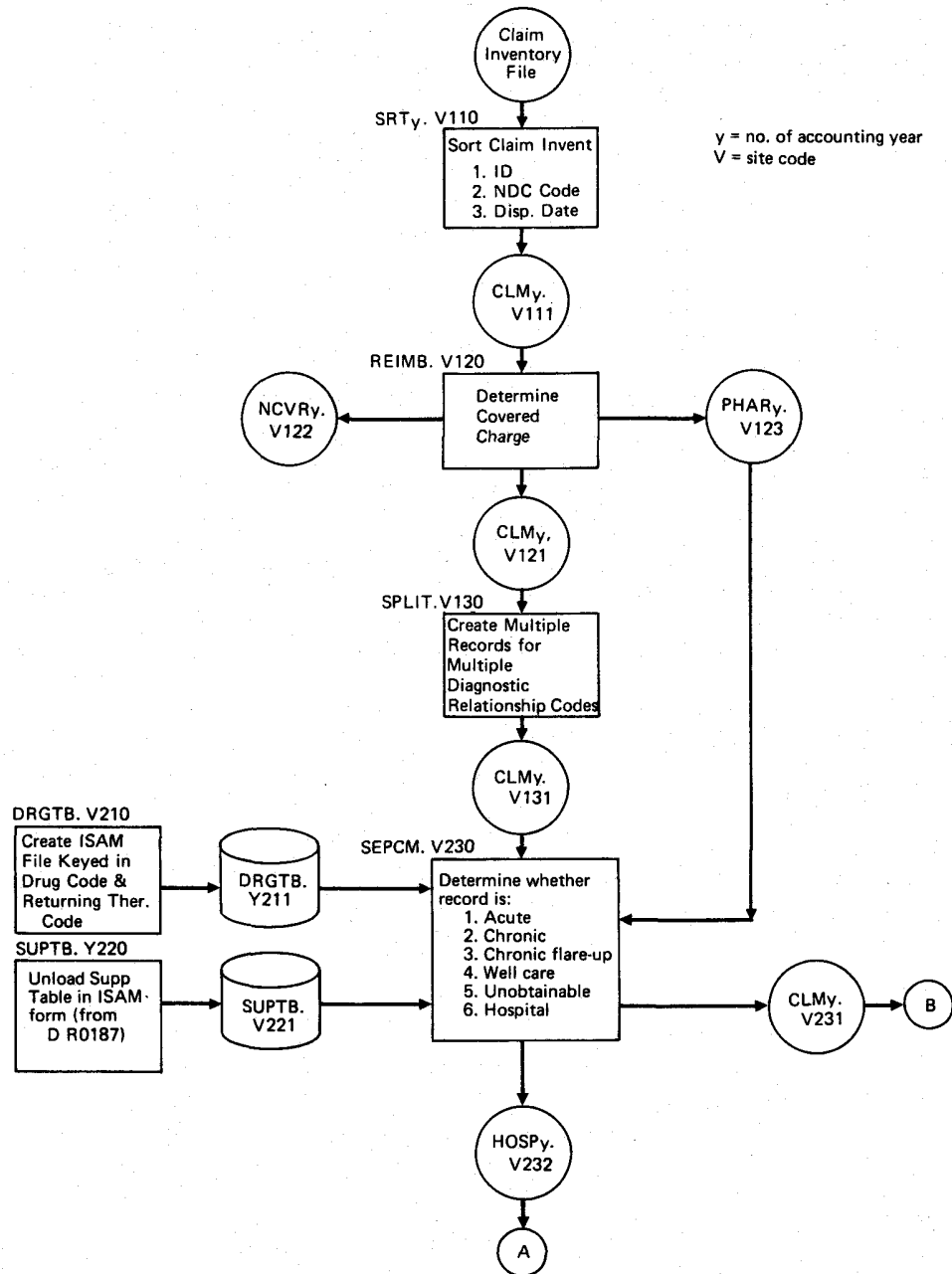


Fig. 2.3

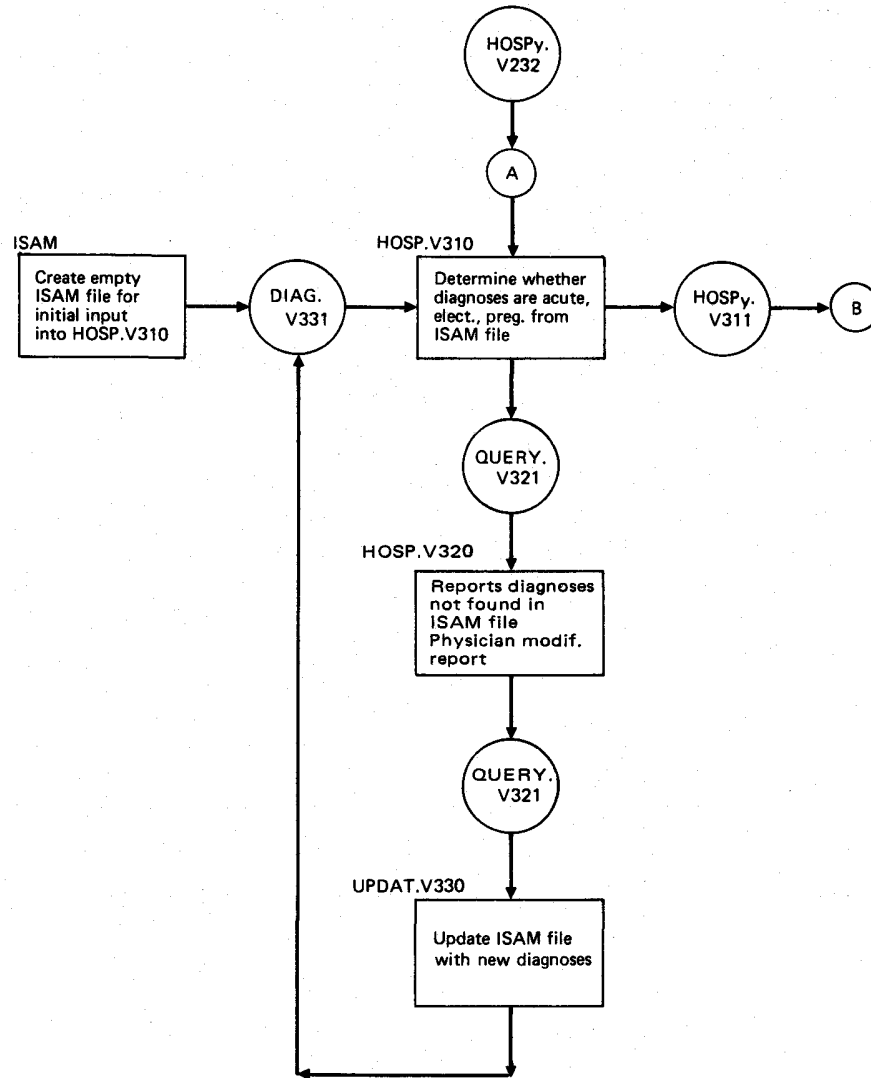


Fig. 2.3 -- continued

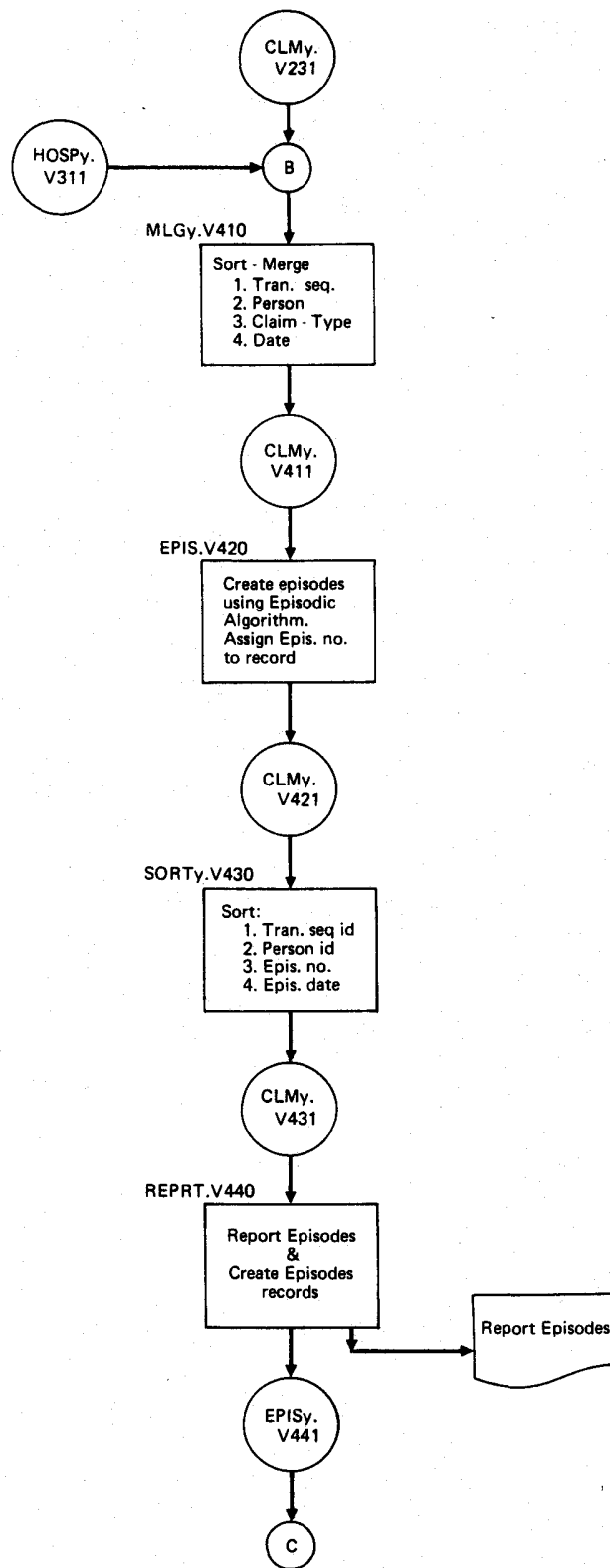


Fig. 2.3--continued

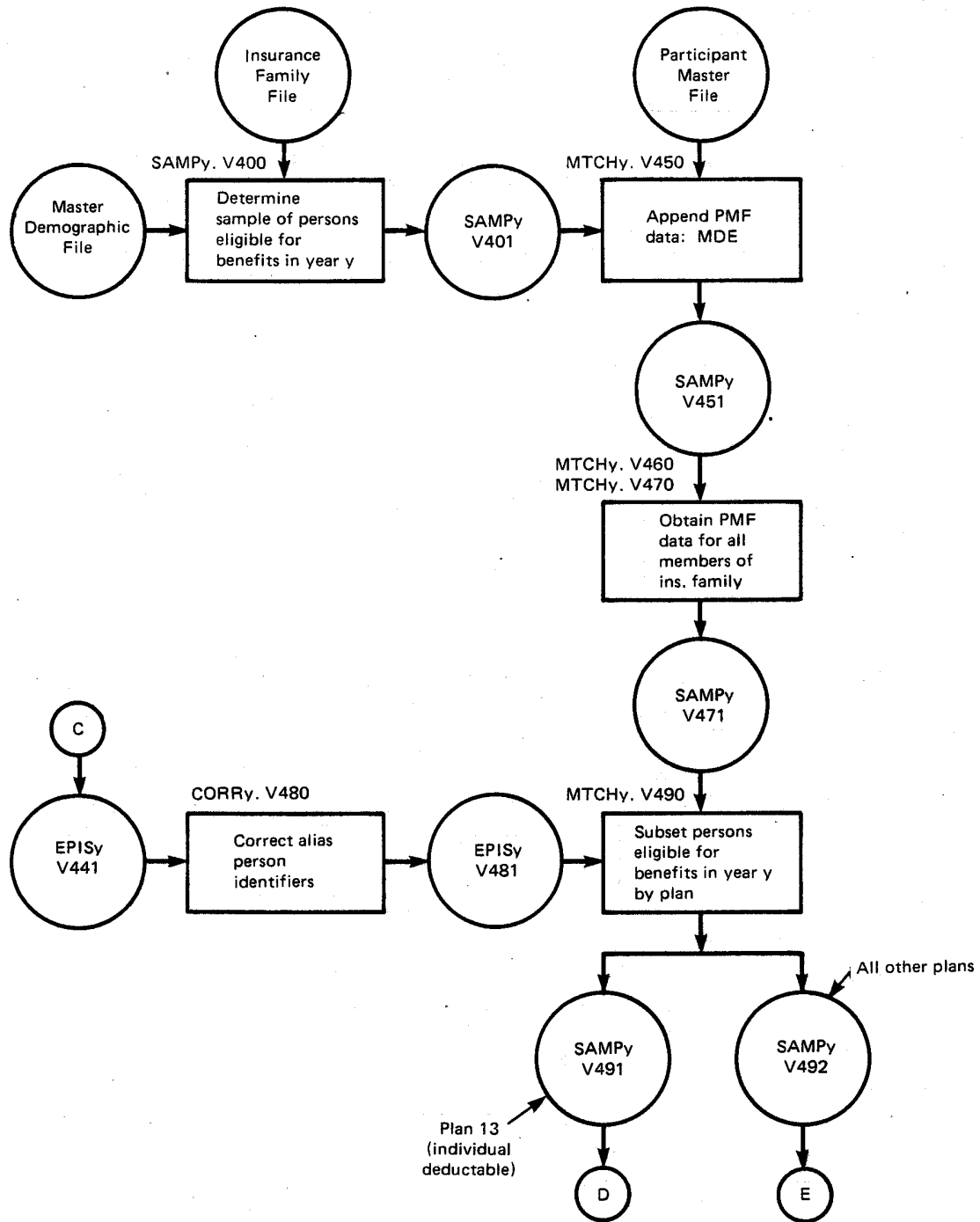


Fig. 2.3--continued

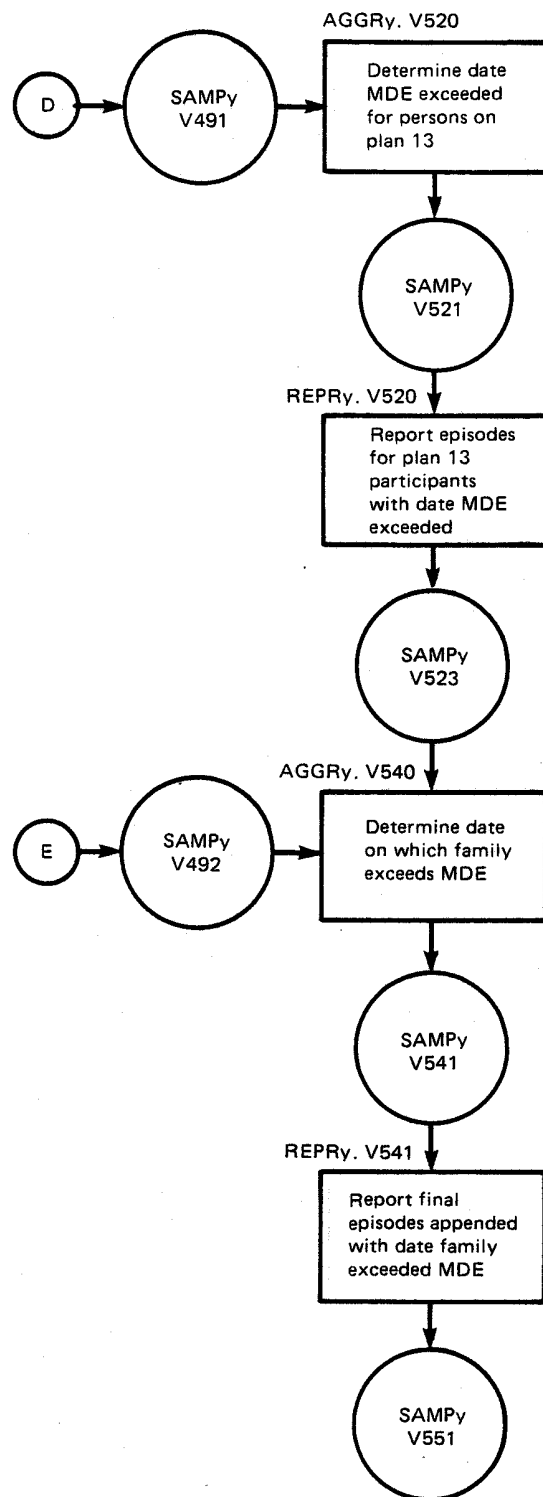


Fig. 2.3--continued

III. DESCRIPTION OF PROGRAMS USED FOR THE EPISODES- OF-ILLNESS ANALYSIS

INTRODUCTION

This chapter describes the programs that are executed in order to process the claim inventory records into episodes. The steps required to accomplish this linkage are briefly summarized below:

1. Noncovered charges are subtracted from the line charge as the analysis focuses on covered charges only.
2. A given line charge may be divided into multiple line charges such that each new line charge is related to a single diagnosis.
3. The claim type (acute, chronic, chronic flare-up, etc.) is determined for each line charge.
4. The line charges are linked together and assigned an episode number; the rules used for linking are described in detail in the description of the program, EPIS.V420.
5. The line charges are aggregated into episodes and a file is generated containing one record per episode per person.
6. A sample of all persons eligible for benefits during any part or all of a specified accounting year is generated and appended with the HIS variables:
 - a. Insurance Family Identifier--identifies participants sharing a common Maximum Dollar Expenditure.
 - b. HIS Insurance Plan.
 - c. Maximum Dollar Expenditure (MDE).
7. Episodes are subset for eligible participants and the date on which a family exceeded its MDE is determined as a function of plan, insurance family identifier, and MDE.

Each program assigns a value to at least one variable in the VARPAD array of each claim inventory record. Table 3.1 describes this array of variables.

Table 3.1

DEFINITION OF VARIABLES CREATED FOR EPISODE-OF-ILLNESS
ALGORITHM AND STORED IN THE ARRAY "VARPAD(15)"
OF THE CLAIM INVENTORY RECORDS

VARPAD (K) Name	Definition
1 CLMS_SEQ	Insurance family ID--allows aggregation of persons within an MDE-related family
2 NEW_REAS	Reason for new ID (tracking file)
3 OLD_REAS	Reason for old ID (tracking file)
4 ANN_DATE	Beginning date (YYYYMMDD) of accounting year of claim
5 ACC_YEAR	The number of the year (1-5) of the person's participation in the experiment when the claim was made
6 LINE_CHRG	Covered charges before copayment is subtracted, V5558-V5559. (Assigned in program, REIMB.V120)
7 MDE_FLG	= 1, if reason noncovered (V5560 = 44 or 45); = 0, otherwise. (Assigned in program, REIMB.V120)
8 SPLIT_CHRG	The amount of the covered charge apportioned to the specified diagnosis. (Assigned in program, SPLIT.V130.)
9 CLM_TYPE	= ab, where a = 1, hospital b = 0, dental exam = 2, physician = 1, acute = 3, dental = 3, chronic flare-up = 4, pharmacy = 4, well-care/elective hospital = 5, supply = 5, unobtainable = 6, pregnancy = 7, abortion (Assigned in programs, SEPCM.V230 and HOSP.V310.)
10 THC	Treatment history code. This variable classified the charge as initial or repeat and as acute, chronic, chronic flare-up or elective. (Assigned in programs, SEPCM.V230 and HOSP.V310.)
11 DIAG	Diagnosis corresponding to SPLIT_CHRG. (Assigned in programs, SEPCM.V230 and HOSP.V310.)
12 EPIS_DATE	Date of episode. (Assigned in program, EPIS.V420.)
13 EPIS_NO	Episode number. (Assigned in program, EPIS.V420.)
14 CATEG	= 1, 2, 3, 4 for mental health, hearing, vision or pregnancy related services. (Assigned in program, EPIS.V420.)

Table 3.2

RELATIONSHIP BETWEEN TYPE OF MER AND THE VARIABLE, DOC_SEQ

DOC_SEQ	Document Number and Record Type	Structure Name
1	Document 109 Record Type 2 Hospital Inpatient Charge	INPAT2
2	Document 109 Record Type 5 Hospital Adjustment Record	INPAT5
3	Document 106 Record Type 1 Hospital Inpatient Service	INMED1
4	Document 111 Record Type 2 Drugs Prescribed by Doctor (no line charge)	PHY1
5	Document 111 Record Type 3 Supplies Prescribed by Doctor (no line charge)	PHY2
6	Document 111 Record Type 4 Service Given by Doctor	PHY3
7	Document 111 Record Type 5 Physician Adjustment Record	PHYAD
8	Document 105 Record Type 1 Drugs Given by Doctor	PHYDIS1
9	Document 105 Record Type 2 Supplies Given by Doctor	PHYDIS2
10	Document 105 Record Type 3 Injection Given by Doctor	PHYDIS3
11	Document 113 Record Type 1 Outpatient Service Charge	PHYO
12	Document 107 Record Type 2 Dental Services Given	DENTAL2
13	Document 107 Record Type 3 Dental Drugs Prescribed (no line charge)	DENTAL3
14	Document 107 Record Type 5 Dental Adjustment Record	DENTAL5
15	Document 108 Record Type 2 Drugs Purchased	PHARM1
16	Document 108 Record Type 3 Supplies Purchased at Pharmacy	PHARM2
17	Document 108 Record Type 5 Pharmacy Adjustment Record	PHARMAD
18	Document 112 Record Type 2 Supplies Purchased	SUPPL2
19	Document 112 Record Type 5 Supply Adjustment Record	SUPPL5

The rules often refer to the document sequence number (variable, DOC_SEQ) of a record. This variable identifies the charge and claim document (MER) that the record comes from, as shown in Table 3.2.

Data elements are referred to with the notation Vnnnn, where nnnn is a unique numerical identifier for a data item in a claim inventory record. These numbers are taken from the HIS data base.

The programs are described in their order of execution. For those who like them, we have included a flow chart of the programs (Fig. 2.3).

PROGRAM: REIMB.V120

I. Introduction

People may file claims for quasi-medical services, such as orthodontia, that are not covered by our policies. In addition, some medical services, such as hospital amenities or eyeglasses, are covered only up to certain limits. A list of the reasons for noncoverage is given in Appendix A. Our analysis focuses on covered charges, so the noncovered charges must be subtracted out.

II. Purpose

This program calculates the covered charge--the charge (V5558) minus the noncovered portion of the charge (V5559)--for claim inventory records with DOC_SEQ = 1, 3, 6, 8-12, 15, 16, 18. Certain records which are intended to show corrections to earlier payments (DOC_SEQ = 2, 7, 14, 17, 19) are dropped from the file at this stage since the corrections have already been made to the line charge by GSA. Records having no line charge (DOC_SEQ = 4, 5, 13) are mainly records of prescriptions and are written unchanged on the output file.

III. Input

FILI--the claim inventory file sorted in the job step, SRTy.V110, by:

1. ID
2. NDC Code (V5589)
3. Dispense Date (V5603).

IV. Output

FILI--all claim inventory records except adjustment records or those records where the entire line charge was not covered. The records are updated with the covered remainder, LINE_CHRG.

FILB--claim inventory records in which the line charge was not covered: V5559 = V5558.

FDRUG--a file containing a record for each pharmacy record (DOC_SEQ = 15) sorted by the following fields:

1. ID
2. NDC Code (V5589)
3. Dispense Date (V5603)

V. Method

1. If the reason code (V5560) is one of the following, the covered charge, LINE_CHRG, is set equal to the total charge (V5558). The following codes refer to coinsurance deductions, which are *not* subtracted from the charge for analysis.

V5560 = -1, Missing
= 30, Deductible not met
= 31, Participants' coinsurance portion
= 34, Prepayment made
= 37, Discount
= 38, Not covered prepayment and deductible
= 39, Not covered prepayment and coinsurance
= 40, Discount and deductible not met
= 41, Discount and coinsurance
= 44, Services obtained outside of Grp. Hlth. Coop.
= 45, Plan benefit is 5 percent of covered charge.

2. If the reason code (V5560) is equal to 44 or 45, the variable, MDE_FLG, is set equal to 1.

3. For all other reason codes, the charge is reduced by the noncovered portion (V5559) and the result is saved in the VARPAD array as the variable, LINE_CHRG, unless the noncovered portion (V5559) is equal to the charge (V5558), in which case the record is written on FILB and omitted from FILO.

4. There is a second noncovered amount for the following claim inventory records:

DOC_SEQ	Description
1	Hospital Inpatient
9	Supplies Given by Doctor
18	Supplies Purchased

The above process is repeated for these claims inventory records using the following variables:

Second noncovered amount	DEI5562
Second reason code	DEI5563

PROGRAM: SPLIT.V130

I. Purpose

A single charge for an office visit may include services for old and new or chronic and acute conditions. This program divides these charges, so that each charge is connected with a single diagnosis. Physician claims (DOC_SEQ = 4, 5, 6, 8, 9, 10, 11) are examined for multiple diagnoses related to the charge, LINE_CHRG. Where multiple diagnoses exist, multiple records are created--one per diagnosis. The charge is apportioned between the records, the first or primary diagnosis receiving twice the value allocated to the remaining diagnoses. If there is only one diagnosis, SPLIT_CHRG is set equal to LINE_CHRG. Two variables in the VARPAD array are assigned values:

DIAG--diagnosis corresponding to apportioned charge,
SPLIT_CHRG--charge apportioned to diagnosis, DIAG.

II. Input

CLAIMS--claim inventory file--output from REIMB.V120.

III. Output

SPLTCLM--claim inventory file with one record (line charge) per relevant diagnosis for all records with DOC_SEQ = 4, 5, 6, 8, 9, 10, 11. Other records are copied to the output file with the variable, SPLIT_CHRG, set equal to the reimbursement, LINE_CHRG.

IV. Method

1. The VARPAD variable, SPLIT_CHRG, is set equal to the variable, LINE_CHRG, for all claim inventory records.
2. Physician records (DOC_SEQ = 4, 5, 6, 8, 9, 10, 11) are examined to determine the number of relevant diagnoses--diagnoses related to the line charge.
 - a. A diagnostic relationship code (V5596, V5597, V5598, V5599) equal to "1" indicates that the corresponding diagnosis (V5522, V5525, V5528, V5531) is relevant. The number of pertinent diagnoses is determined from the number of diagnostic relationship codes equal to 1, or if the diagnostic relationship codes are missing, from the number of nonmissing diagnoses.
3. One record is created for each relevant diagnosis; the charge, LINE_CHRG, is apportioned between the records; the first or primary diagnosis receives twice the amount allocated to remaining diagnoses; the apportioned reimbursement is saved as the VARPAD variable, SPLIT_CHRG. The corresponding diagnosis is saved in this VARPAD array as DIAG.
4. The diagnostic relationship codes in the new record are altered to point to one diagnosis only--that one saved as the VARPAD variable, DIAG.

PROGRAM: SEPCM.V230

I. Purpose

SEPCM separates claim line charges and assigns a CLM TYPE on the basis of this categorization:

Category	Claim Type
Dental Exams + cleaning	0
Acute	1
Chronic	2
Chronic Flare-up	3
Well-care	4
Unobtainable	5
Adjustments ¹	6
Hospitalizations	7

After assigning the claim type, SEPCM writes the line charge to the hospital file, if a hospital record, or to the general claim file, if other than hospital. The claim file is split between hospital and nonhospital to allow for further categorization of hospital records in programs following SEPCM. After that refinement of hospital CLM_TYPE, the two files are merged for episode number generation.

SEPCM also assigns a general document type indicator:

1. Hospital
2. Physician
3. Dental
4. Pharmacy
5. Supply

This field is used to sort the file correctly for episode number generation and is encoded in the CLM_TYPE variable. For example, an acute physician record will have a CLM_TYPE of "____21," the "2" indicating physician, the "1" indicating acute.

In addition to this categorization of each line charge, SEPCM moves the treatment history code (THC), if any, from the data portion of the record to the header, making it more readily accessible. It also drops

¹Currently, adjustments are dropped in program REIMB.V120 above.

any physician, pharmacy, and supply records having drug or supply codes indicated as inapplicable.

II. Input

CLAIMS--includes individual line charges extracted from hospital, physician, dental, pharmacy, and supply claims that have been submitted by participants in the study. The file is sorted by:

1. Household
2. Family
3. Person
4. NDC (drug) code
5. Dispense date

PHARM--includes a copy of every pharmacy record in the CLAIMS file, also sorted by:

1. Household
2. Family
3. Person
4. NDC (drug) code
5. Dispense date

PHARM is used in conjunction with the CLAIMS file to determine the refill pattern for a particular drug by a participant.

DRUGTBL--(index-sequential random access organization) includes a record for each drug therapeutic code that may appear in the claims data, identified by its THC and a text description. The access key is therapeutic code. See Table 3.3.

SUPPTBL--(index-sequential random access organization) includes a record for each supply code that may appear in the claims data, identified by its THC and a text description. The access key is supply code. See Table 3.4.

Table 3.3

DRUG TABLE

Thera- peutic Code	Treatment History Code[a]	Description
1	2	Digitalis glycosides
2	2	Antiarrhythmic agents
3	2	Antianginal agents
4	2	Peripheral vasodilators
5	1	Agents used in hypotension and shock (with outpatient)
6	2	Antihypertensive agents
7	2	Diuretics
8	3	Agents used to treat deficiency anemias[a]
9	3	Anticoagulants[a]
10	2	Blood, blood components, and blood substitutes (when outpatient)
12	2	Agents used to treat hyperglycemia
13	2	Agents used to treat hyperlipidemia
14	2	Vitamins and sources of vitamins
15	2	Miscellaneous nutritional agents
16	2	Replenishers and regulators of water, electrolytes, nutrients
17	2	Blood calcium regulators (when outpatient)
18	1	Local anesthetics (when outpatient)
19	1	General anesthetics (when outpatient)
20	1	Adjuncts to anesthesia (when outpatient)
21	3	Strong analgesics[a]
22	3	Mild analgesics
23	1	Narcotic antagonists (with outpatient)
24	3	Agents used to treat migraine
25	2	Agents used in gout
26	2	Antirheumatic agents (chronic as a group)
27	3	Sedatives and hypnotics
28	3	Antianxiety agents
29	2	Antipsychotic agents
30	3	Centrally acting skeletal muscle relaxants
31	2	Anticonvulsants (with outpatient)
32	3	Antidepressants
33	2	Anorexiant
34	2	Analeptics (with outpatient)
35	3	Adrenal corticosteroids
36	2	Androgens and anabolic steroids
37	2	Estrogens, progestagens, oral contra., and ovulatory agents

Table 3.3
DRUG TABLE (cont.)

Thera- peutic Code	Treatment History Code[a]	Description
38	2	Anterior pituitary and hypothalamic hormones
39	2	Thyroid hormones and antithyroid agents
40	2	Antidiuretics (with outpatient)
41	2	Bronchodilators
42	3	Nasal decongestants
43	3	Expectorants and inhalants
44	3	Antitussive agents
45	3	Antihistamines
46	3	Cold remedies
47	3	Therapeutic gases
48	1	Penicillins
49	1	Cephalosporins
50	1	Erythromycin and derivatives
51	1	Lincomycin and clindamycin
52	1	Polymycins
53	3	Tetracyclines
54	1	Chloramphenicol and derivatives
55	1	Sulfonamides
56	3	Nitrofurans
57	1	Aminoglycoside antibact agents and miscellaneous antimicrobial drugs
58	2	Antituberculous agents
59	2	Antileprosy agents
60	3	Antifungal agents
61	1	Antimalarial agents
62	1	Amebicides
63	2	Antitrypanosomal agents
64	1	Antitrichomonal agents
65	1	Anthelmintics
66	1	Scabicides and pediculicides
67	3	Antiseptics and disinfectants
68	3	Dermatologic agents
69	2	Agents used to treat glaucoma
70	2	Mydriatics and cycloplegics
71	1	Anti-infective and anti-inflam agents used in ophthalmology
72	3	Miscellaneous ophthalmic preparations
73	1	Topical otic preparations
74	2	Antiparkinsonism agents

Table 3.3

DRUG TABLE (cont.)

Thera- peutic Code	Treatment History Code[a]	Description
75	2	Agents used in myasthenia gravis (with outpatient)
76	1	Radiopaque media (when outpatient)
77	1	Miscellaneous diagnostic aids (when outpatient)
78	3	Antispasmodics
79	3	Antacids
80	1	Antidiarrheals
81	2	Laxatives and agents affecting fecal consistency
82	3	Anorectal preparations
83	2	Miscellaneous gastrointestinal agents
84	1	Emetics
85	1	Antiemetics
86	2	Stimulants of gastrointestinal and urinary tracts (when outpatient)
87	1	Oxythotics (when outpatient)
88	2	Antineoplastic agents
89	1	Vaccines, toxoids, and serologic agents
90	3	Chelating agents
91	1	Enzymes used as drugs
92	3	Miscellaneous drugs

[a]The values assumed by the Treatment History Codes are defined as follows:

THC	Description
1	Acute
2	Chronic
3	Acute or chronic
4	Chronic or elective

Table 3.4
SUPPLY TABLE

Supply Code	Treatment History Code[a]	Description
1101	3	Artificial foot
1102	3	Artificial leg
1103	3	Artificial arm
1104	3	Artificial hand
1105	3	Artificial eye
1201	4	Bra for prosthesis, reg.
1202	4	Bra for swimming
1203	4	Bra for sleeping
1204	3	Breast prosthesis, reg.
1205	3	Breast prosthesis, light
2101	3	Hospital bed, manual
2102	3	Hospital bed, electric
2103	1	Waterbeds (for burn patients)
2104	3	Trapeze
2105	3	Table, hospital
2106	4	Exercycle
2107	3	Floor stand for traction
2201	3	Bedpans
2202	3	Bedside commode chair
2301	3	Mattress
2401	3	Cart, nec
2402	3	Wrench
3101	3	Wheelchair, manual
3102	3	Wheelchair, electric
3103	3	Automatic drive for multiple sclerosis
3104	3	Wheelchair, manual with elevated leg
3201	3	Hearing aid
3202	2	Batteries
3301	3	Pacemakers
3401	3	Bath chair lift
4101	4	Monofocal lens
4102	4	Tinted monofocal lens
4103	4	Mono w/prisms
4104	4	Mono w/prisms tinted
4111	4	Bifocal lens
4112	4	Tinted biofocal lens
4113	4	Bifocal lens w/prisms
4114	4	Bifocal lens w/prisms tinted
4121	4	Tri-focal lens
4122	4	Tinted tri-focal lens
4131	4	Contact lens, soft

Table 3.4
SUPPLY TABLE (cont.)

Supply Code	Treatment History Code[a]	Description
4132	4	Contact lens, hard
4141	4	Lens, nec
4142	4	Sunglass lens
4151	4	Glasses, nec
4201	4	Frames, metal
4202	4	Frames, plastic
4203	4	Frames, nec
4204	4	Temple for glasses
4301	4	Glasses case
4302	1	Certificate for repair of glasses
5101	1	Crutches
5102	3	Walker
5103	3	Cane
5104	3	Helping hand
5111	3	Corset
5112	1	Rib belt
5113	1	Metal arm splint
5114	1	Wrist splint
5115	1	Finger splint
5121	4	Shoe
5122	3	Arch support
5123	3	Dennis Brown splint
5124	3	Inlay in shoe (to redistribute weight)
5131	3	Traction equipment
5133	3	Water bag, sand bag for traction
5134	3	Ankle weight
5201	3	Braces, back
5202	3	Braces, leg--short and long
5203	1	Neck supports
5204	1	Clavicle strap
5205	3	Sacral support belt
5206	3	Knee brace
5211	3	Pad for brace
5212	3	Strap for brace
5301	1	Body cast
5401	1	Adjustments--braces and splints
5901	1	Crutch--rented
6101	3	Diaphragm
6102	3	IUD

Table 3.4
SUPPLY TABLE (cont.)

Supply Code	Treatment History Code[a]	Description
6103	2	Condoms
6201	2	Colostomy bags
6211	2	Urinary bags
6212	2	Catheters
6399	2	Dialysis supplies, nec
6401	2	Feeding/Nasal gastric tubes
6501	3	Sphygmomanometer (blood pressure)
7101	3	Respirator (IPPB)
7102	3	Respirator, other or nec
7201	1	Vaporizer
7202	1	Humidifier
7203	1	Mist tent
7204	3	Compressor for aerosols
7205	3	Regulator
7206	3	Adaptor
7301	2	Oxygen tanks, portable
7302	2	Oxygen tanks, not portable or nec
7303	2	Oxygen
8101	2	Tablets for urine test (clinitest)
8102	2	Keto stix
8103	2	Keto-diastix
8104	2	Testape & diastix
9101	1	Ace bandage
9102	4	Jobst stockings, Ted stockings, etc.
9103	2	Stump socks
9104	2	Truss
9105	2	Sanipants
9106	2	Sanipants liner
9107	1	Arm sling
9108	1	Basket weave dressing
9110	2	Rubber finger sock
9111	4	Hot water bottle
9112	1	Ice pack
9113	4	Heating pad
9121	2	Surgical gloves
9122	3	Ear plugs
9123	3	Lumbo-sacral seat
9124	3	Foam pad
9125	3	Donut pillow
9131	1	Adhesive tape, band-aids, etc.
9132	1	Gauze
9133	2	Corn/callus padding

Table 3.4
SUPPLY TABLE (cont.)

Supply Code	Treatment History Code[a]	Description
9134	1	Adaptic-sterile, nonadhesive dressing
9135	1	Telfa pads
9136	1	Compresses
9301	2	Syringes, nec
9302	2	Syringes, insulin
9303	2	Syringes w/needles
9311	2	Needles, disposable, insulin or nec
9312	2	Needles, nondisposable
9501	1	Alcohol swabs, disposable
9502	1	Wash 'N' Dri, etc.
9503	1	Swabs, medicated
9511	1	Tucks
9521	2	Toothpaste
9601	4	Dropper
9701	3	Walker tote bag

[a]The values assumed by the Treatment History Codes are defined as follows:

THC	Description
1	Acute
2	Chronic
3	Chronic flare-up and one-time purchases for chronic problems
4	Elective (well-care).

III. Output

OCLAIMS--includes all claim line charges read from the CLAIMS file except for hospital records, and any records bypassed due to critical data being invalid or inapplicable. The line charges are identified by CLM_TYPE and document type.

HOSP--includes all hospital claims line charges. The CLM_TYPE/document type for all records is "71."

IV. Processing

Records are processed on the basis of their DOC_SEQ code. For a list of these codes and structure names associated with them, see Table

3.2. Five record types are processed:

1. Hospital
2. Physician
3. Dental
4. Drug
5. Supply

The rules of categorization for these record types are as follows:

Hospital--(DOC_SEQ 1-3). All records are written to the HOSP file with a CLM_TYPE of "71." The DSI is always "0."

Physician--(DOC_SEQ 4-11). Physician records contain two four-element arrays critical to SEPCM processing: the diagnosis relationship code (DRC) array and the THC array. The DRC array is searched and when a value of 1 is found, this indicates that the THC in the same relative position in the THC array contains the relevant THC. CLM_TYPE is assigned on the basis of the relevant THC.

THC	CLM_TYPE	Description
1, 3, 8	1	Acute
2, 4, 9	2	Chronic
5, 7	3	Chronic flare-up
6	4	Well-care
-1	5	Unobtainable

Dental--(DOC_SEQ 12). CLM_TYPE is assigned on the basis of the first position of the ADA procedure code:

ADA	CLM_TYPE	Description
blank or 1	0	Exam (well-care)
other than blank or 1	1	Acute

Dental Drug Prescriptions--(DOC_SEQ 13). CLM_TYPE is assigned exactly as physician record types.

Pharmacy--(DOC_SEQ 15). CLM_TYPE is assigned on the basis of the five therapeutic codes in the record. Each code is looked up in the DRUGTBL file (Table 3.3). If the assigned treatment history codes do not all indicate chronic or all indicate acute, a secondary search of the PHARM file is made to determine the refill pattern for the particular drug and person. If the drug was filled more than three times during a period of over 90 days, the record is considered chronic, otherwise it is considered acute.

The possible THC's returned from a search of the DRUGTBL are:

THC	Description
1	Acute
2	Chronic
3	Acute or Chronic
4	Chronic or Elective

Pharmacy Supplies--(DOC_SEQ 16). CLM_TYPE is assigned on the basis of the supply code which is looked up in the SUPPTBL file (Table 3.4). The THC's assigned to the supply codes are directly translatable to CLM_TYPE:

THC	CLM_TYPE	Description
1	1	Acute
2	2	Chronic
3	3	Chronic flare-up
4	4	Elective (well-care)

Supply--(DOC_SEQ 18). CLM_TYPE is assigned exactly as it is for Pharmacy DOC_SEQ 16.

V. DEIS Referenced

DOC_TYPE	DOC_SEQ	DEI	Description
Hospital	1-3	none	
Physician	4-11	V5574	THC (four-element array)
		V5596	Charge related to THC 1
		V5597	Charge related to THC 2
		V5598	Charge related to THC 3

		V5599	Charge related to THC 4
		V5600	Diagnosis relationship code not obtainable
		V5601	Supply code
		V5589	NDC (drug) code
Pharmacy	15	V5603	Date dispensed
		V5589	NDC code
		V5595	Therapeutic code 1
		V5658	Therapeutic code 2
		V5659	Therapeutic code 3
		V5660	Therapeutic code 4
		V5661	Therapeutic code 5
	16	V5601	Supply code
Dental	12	V5630	THC
		V5625	ADA procedure code
	13	V5630	THC
Supply	18	V5601	Supply code

VI. DSI Assignment

If data errors are found, making CLM_TYPE assignment impossible, this is reflected in the DSI associated with CLM_TYPE. The potential CLM_TYPES and DSIs that may be assigned to each record type are listed in Table 3.5.

PROGRAM: HOSP.V310

I. Introduction

The next three programs classify hospital episodes as either acute, elective, pregnancy, abortion, or mental health. Initially, George Goldberg, M.D. was given the three main discharge diagnoses for each episode and asked to classify the episode. The programs subsequently only ask him to classify those episodes with different diagnoses. Goldberg's SM-3555, October 6, 1976, includes early lists of decisions and discussion.

Table 3.5

POTENTIAL CLM_TYPE/DSI ASSIGNMENT

Structure Name	DOC SEQ	DOC Type	CLM Type	DSI	Reason or Comments
INPAT1	1	1	7	0	All hospital records written to hospital file for further processing
INPAT5	2	1			Not currently processed
INMED1	3	1	7	0	Same as INPAT1
PHY1	4	2	1-4	0	Assigned on basis of THC pointed to by Diag Rela Code (DRC)
			5	0	Diag Rel Unobt Code other than -1
			5	1	No DRC present (none set to 1)
			5	2	THC missing (-1)
			5	3	THC invalid (other than 1-7)
PHY2	5	2			Same as PHY1
PHY3	5	2			Same as PHY1
PHYAD	7	2			Not processed
PHYDIS1	8	2			Same as PHY1
PHYDIS2	9	2			Same as PHY1
PHYDIS3	10	2			Same as PHY1
PHY0	11	2			Same as PHY1
DENTAL2	12	3	0,1	0	Assigned on basis of ADA procedure code
			1	4	ADA procedure code missing (-1)
DENTAL3	13	3	1-4	0	Assigned on basis of THC
			5	2	THC missing (-1)
			5	3	THC invalid (other than 1-7)
DENTAL5	14	3			Not processed
PHARM1	15	4	1,2	0	Assigned on basis of match of therapeutic code to DRUGTBL file and if needed a match to PHARM file
			5	6	Therapeutic codes are missing (-1) or no matching codes were found in table
			5	7	Dispense date, used in match to PHARM file, is missing (-1)
PHARM2	16	4	1,4	0	Assigned on basis of matching supply code to supply table
			5	5	Supply code is missing (-1) or has no match in table
PHARMAD	17	4			Not processed
SUPPLY2	18	5	1-4	0	Assigned on basis of matching supply code to supply table
			5	5	Supply code is missing (-1) or has no match in table
SUPPLY5	19	5			Not processed

II. Purpose

This program assigns a value to the VARPAD variable, CLM_TYPE, for all hospital records, DOC_SEQ = 1, 2, 3. The claim type, CLM_TYPE, is a function of three diagnoses--(V5522, V5525, V5528) and may assume one of the following values:

```
CLM_TYPE = 11, acute-hospital
          = 14, elective-hospital
          = 16, pregnancy-hospital
          = 17, abortion-hospital
```

Also, a value is assigned to the VARPAD variable, CATEG, which describes the type of service:

```
CATEG = 1, mental health
       = 4, pregnancy related
```

III. Input

MASTER--all hospital claim inventory records (DOC_SEQ = 1, 2, 3); this file is output from the program, SEPCM.V230.

ISFILE--an ISAM file keyed on a 24 character string of three diagnoses--V5522, V5525, V5528. The two variables, CLM_TYPE and CATEG, are associated with each key. These variables have been assigned a value by a physician.

IV. Output

OUT1--claim inventory records updated with the variables, CLM TYPE and CATEG.

OUTKEY--file of keys (V5522, V5525, V5528) not found in the file, ISFILE.

OUTSORT--sorted version of the file, OUTKEY.

V. Procedure

1. The three diagnoses--V5522, V5525, V5528--are concatenated into a 24-character string that is used as the key in reading an ISAM file, ISFILE.
2. If the key is valid, the VARPAD variables, CLM_TYPE and CATEG, are set equal to the corresponding variables on the file, ISFILE.
3. If no record exists for the key, the key is written on the output file, OUTKEY.
4. All hospital records are written on the output file, OUT1.
5. Those keys written on the file, OUTKEY, will subsequently be examined by a physician, and the associated variables, CLM_TYPE and CATEG, will be evaluated. The keys and the corresponding variables will be added to the file, ISFILE.
6. This program will be reexecuted until all of the keys can be accessed in the ISAM file, ISFILE.

PROGRAM: HOSP.V320

I. Purpose

This program is a SAS program. Those diagnosis keys that have not previously been classified and so could not be accessed in the program, HOSP.V310, are reported, translated into English so that a physician can determine the claim type (CLM_TYPE) and type of service (CATEG) associated with the key--the concatenation of the three diagnoses: V5522, V5525, V5528.

The output file from HOSP.V310 (QUERY.V313) is updated using WYLBUR. The variables, CLM_TYPE and CATEG, are modified according to the physician's specification. This file is then ready to be used as input to UPDAT.V330.

II. Input

OUTSORT--sorted keys from HOSP.V310.

III. Output

A file containing the keys from the file, OUTSORT, with the associated variables, CLM_TYPE and CATEG.

PROGRAM: UPDAT.V330

I. Purpose

This program updates the ISAM file containing the variables, CLM_TYPE and CATEG, as a function of the hospital diagnoses, a 24 character string concatenated from the diagnoses: V5522, V5525, V5528. New keys and the associated newly assigned variables are added to the file, ISFILE.

II. Input

MASTER--new keys to be added to the ISAM file. ISFILE--ISAM file to be updated.

III. Output

ISFILE--updated ISAM file. OUT1--dump of ISAM records in the case that the ISAM file overflows.

PROGRAM: MRGy.V410

I. Purpose

This step merges the hospital claim inventory records (output from HOSP.V310) with all of the other claim inventory records (output from SEPCM.V230) and sorts the records by:

1. Insurance ID
2. Person ID
3. Claim type
4. Date of service

EPISODES-OF-ILLNESS: EPIS.V420

I. Purpose

This program links the claim inventory records into actual episodes. Episodes numbered from 1 to n and prefixed with three digits defining the kind and year of the episode will be defined for each person in the claim inventory file. The rules set forth here describe how each line charge is either linked to a previously examined line charge and assigned the associated episode number or assigned a new episode number. After the line charge has been linked to an episode number (old or new), it is written on the output tape; a subsequent program (REPRy.V440) aggregates the claim inventory file into one record per episode per person.

In order that the claims can be linked together in an episode hierarchy, where hospital outranks physician, which outranks other episodic types, each person's claims are sorted by claim type, which reflects that hierarchy. The variable, CLM_TYPE, assigned a value in the programs, SPECM.V230 and HOSP.V310, is defined:

CLM_TYPE = "ab" where

a = 1, hospital	b = 0, dental exam
= 2, physician	= 1, acute
= 3, dental	= 2, chronic
= 4, pharmacy	= 3, chronic flare-up
= 5, supply	= 4, well-care/elective hospital
	= 5, unobtainable
	= 6, pregnancy
	= 7, abortion.

Three variables in the VARPAD array (see Table 3.1) of each claim inventory record will be initialized in this program:

EPIS_DATE--date of the line charge--date of service unless modified by the algorithm.

EPIS_NO --episode number in the form, yabnnn, where

y = number of accounting year, 1-5.

a = 1, acute
= 2, chronic

= 3, chronic flare-up
= 4, well-care (including dental exam and
 elective hospital)
= 5, unobtainable
= 6, pregnancy.

b = 1, hospital, inpatient
 = 2, physician
 = 3, dental
 = 4, pharmacy
 = 5, supply
 = 6, hospital, outpatient

nnn = 001, 002, 003, ..., number of episode
 for given person

CATEG = 1, mental health service
 = 2, hearing service
 = 3, vision service, except eye injuries
 or cataract-related services
 = 4, pregnancy related

The program is modular. Depending on the claim type and claim document type (DOC_SEQ) specific procedures are accessed to carry out different claim-linking algorithmic rules.

II. Input

CLAIMS--the claim inventory file sorted by:

1. Insurance ID
2. Person ID
3. Claim type
4. Date of service

III. Output

OCLAIMS--the claim inventory file updated with the variables described above.

IV. Method

The procedures for each claim type will be described one by one in the order they are performed.

Procedure: ALL_HOSP

CLM_TYPE = 11, acute
 = 14, elective
 = 16, pregnancy
 = 17, abortion
DOC_SEQ = 1, 2, 3

1. If CLM_TYPE = 11 (acute) and there are previously established acute hospital episodes, link to episode if either:
 - a. There is a match on claim number, or
 - b. The date of service falls within the admission and discharge dates of the episode.
2. If CLM_TYPE = 14, 16, 17 and there are previously established elective hospital episodes, then link to episode if either of the above conditions is true.
3. If the line charge cannot be linked to an established episode, create a new episode using the procedure, BUILD_HOSP.

Procedure: BUILD_HOSP

The following structure is initialized for each new hospital episode:

1 HOSP

2 EPIS_NO	Episode number
2 CLM_NO	Claim number
2 ADM_DATE	Admission date, V5513
2 DIS_DATE	Discharge date, V5514

1. The VARPAD variable, EPIS_DATE (normally the date of service) will be modified under the following conditions:
 - a. If CLM_TYPE = 17 (abortion), the line charge will be dated 60 days prior to the date of service.
 - b. If CLM_TYPE = 16 (pregnancy), the line charge will be dated 200 days prior to date of service.

- c. However, if the new date precedes the anniversary date (ANN_DATE), the anniversary date will be used as the EPIS_DATE.
2. If the CLM_TYPE = 16 or 17, the episode will be identified as a pregnancy episode; HOSP_PREG_FLG will be set to 1, and the pointer, PREG_PTR, will point to the specified episode.
3. Cataract procedures will be flagged (CATARACT = 1) if either:
 - a. CRVS code (V5606) equals 66800-66999, or
 - b. Any of the three diagnoses--V5522, V5525, V5528--is equal to "y55.4" or "374.n."

Procedure: CHECK_HOSP

This procedure links physician records to hospital records using the following rules:

1. If the date of service is within seven days (14 days for elective hospital episodes) before the admission date or ten days after the discharge date, the claim will be linked to the hospital episode. It will be assigned the same episode number if the date of service falls exactly within the admission and discharge dates. If it falls outside the actual hospital stay, 5,000 will be added to the episode number. Therefore, to separate inpatient and outpatient charges:

yalnnn - hospital inpatient episode number
ya6nnn - hospital outpatient episode number.

2. If the physician record was identified as being a postoperative procedure (POST_FLG = 1), it will be linked to the first hospital episode.

Procedure: PHY_ACUTE FLARE

CLM_TYPE = 21, acute physician
 23, chronic flare-up physician
DOC_SEQ = 4, 5, 6, 8, 9, 10, 11

This procedure links acute/chronic flare-up physician records to hospital episodes (acute or elective) or to acute/chronic flare-up physician episodes by the following rules:

1. Check hospital episodes, using the rules described in the procedure, CHECK_HOSP.
2. Link to acute/chronic flare-up physician episode if the claim number matches a claim number in an existing physician episode.
3. Link to acute/chronic flare-up physician episode if the date of service is within 16/28 days of a previously processed line charge in an acute/chronic flare-up episode and either:
 - a. The treatment history code (THC) indicates that it is a repeat visit, i.e.,
Acute THC = 3, 8
Chronic flare-up THC = 7.
 - b. The diagnoses (or the first three characters of diagnoses in the formnnn.n) match.
4. If no match, create a new acute/chronic flare-up physician episode.
5. Using the procedure, BUILD_PHYS, initialize a structure for the record.

Procedure: BUILD_PHYS

DOC_SEQ = 4, 5, 6, 8, 9, 10, 11
CLM_TYPE = All

The following structure is initialized for each physician line charge:

1 PHYS

2 POINT	Points to preceding line charge
2 EPIS_NO	Episode number
2 CLM_NO	Claim number
2 DOC_SEQ	Type of document
2 THC	Treatment history code (V5589)
2 DIAG	Diagnosis (V5595)
2 DIAG_3	First three characters of diagnosis
2 DRUG	NDC code (V5589)
2 THER_CODE	Therapeutic code (V5595)
2 SUPPLY	Supply code (V5601)
2 SERV_DATE	Date of service
2 PRESC_SUP_DATE	Prescription/Supply Date (V5555)
2 PROV	Provider (V5502)
2 REF_PHYS	Referring physician (V5568)

Depending upon CLM_TYPE, the following pointers and counters are updated:

CLM_TYPE = 21 (Acute)

ACUTE_PHYS_CUR, points to new structure for acute line charge.

ACUTE_PHYS_CNT, number of acute physician line charges.

CLM_TYPE = 22 (Chronic)

CHRON_PHYS_CUR and CHRON_PHYS_CNT

CLM_TYPE = 23 (Chronic flare-up)

FLARE_PHYS_CUR and FLARE_PHYS_CNT

CLM_TYPE = 24 (Well-care)

WEL_PHYS_CUR and WEL_PHYS_CNT.

Cataract procedures will be flagged (CATARACT = 1) if either:

1. CRVS code (V5606) = 92312, 92316, 92352, 92353, 92358, 92371, 92395, or

2. SUPPLY code (V5601) = 4125 or 4127.

The record will be tagged as a vision service (CATEG = 3) if there is no evidence of cataract procedures (CATARACT = 0) and either:

1. CRVS code (V5606) = 92000-92499 or

2. Supply code (V5601) = 4101-4205.

Hearing services will be flagged (CATEG = 2) if either:

1. CRVS code (V5606) = 92551-92589, or

2. Supply code (V5601) = 3201 or 3202.

Mental Health services are flagged (CATEG = 1) if:

CRVS (DEI5606) = 908xx.

Procedure: PHYS_CHRONIC

CLM_TYPE = 22, Chronic physician

DOC_SEQ = 4, 5, 6, 8, 9, 10, 11.

This procedure attempts to link chronic records to chronic episodes, using the following rules:

1. Link to chronic episode if the diagnosis (or the first three digits of a diagnosis in the form nnn.n) matches the diagnosis of a previous line charge in an existing chronic episode.
2. Link to chronic episode if there is a match on claim number.
3. If no match occurs, create a new chronic episode and change the treatment history code (THC) to 2 (initial).
4. Initialize a structure for the chronic line charge, using the procedure BUILD_PHYS.
5. If the record was tagged as a vision or hearing service (CATEG = 2, 3), then build an additional structure in the "well-care" category for subsequent matching of "well-care" line charges.

Procedure: PHYS_WELL

CLM_TYPE = 24, Well-care physician
DOC_SEQ - 4, 5, 6, 8, 9, 10, 11.

This procedure processes well-care physician records.

1. Identify record as pregnancy related (PREG_RCD = 1) if the first three characters of the diagnoses are equal to "y06" or "y07" and the diagnosis is not equal to "y06.9" (pregnancy--but not primary reason for visit).
2. Identify record as postoperative (POST_FLG = 1) if first three characters of diagnoses are equal to "y12."
3. If a pregnancy record is identified (PREG_RCD = 1), the variable, CATEG, is set equal to 4 and the following tests are performed:
 - a. Check to see if a hospital pregnancy episode was established (HOSP_PREG_FLG = 1). If so, link record to hospital episode.
 - b. Check for previously established physician pregnancy episode (PREG_FLG = 1). If one exists, link to physician episode.
 - c. If no match, create new physician pregnancy episode and set PREG_FLG = 1.

4. If record is not a pregnancy related charge, link to hospital episode using the rules set forth in CHECK_HOSP.

5. Link to previous well-care episode if either there is a match on claim number or the date is within fourteen days of a previous well-care physician line charge.

6. If the record was identified as postoperative (POST_FLG = 1), link to acute physician episode if either there is a match on claim number or the date of service is within 28 days of the acute line charge.

7. If there was no link to an existing episode, create a new well-care episode.

8. Initialize a structure for the well-care line charge, using the procedure, BUILD_PHYS.

Procedure: PHYS_UNOBT

CLM_TYPE = 25, unobtainable physician
DOC_SEQ = 4, 5, 6, 8, 9, 10, 11.

This procedure processes line charges for which the claim type is unobtainable. It has been assumed that these charges are related to laboratory tests. SM-6126 reports a study of sampled "unobtainable" line charges, which explains the assumptions of this step. This procedure compares the unobtainable line charges with all line charges in existing episodes. The record is linked to that episode which gives "the best possible match," the match with the highest level number. The match conditions and their corresponding level number are:

Level No.	Match Condition
1	Closest date forward within fourteen days
2	Closest date backward within fourteen days
3	Same date
4	Same claim number or hospital match

If no link was made to an existing episode, a new acute physician episode is created. The claim type is changed accordingly. A structure is initialized for the record, using the procedure BUILD_PHYS.

Procedure: BUILD_DENT_EXAM

CLM_TYPE = 30
DOC_SEQ = 12

1. If the claim number matches the claim number of a previous dental exam, link to previous episode.
2. If the date of service is within fourteen days of previous examination, link to previous episode.
3. If no link was obtained, create a new episode.
4. Initialize the following structure for each line charge.

1 DENT_EXAM

2 POINT	Pointer to previous line charge
2 EPIS_NO	Episode number
2 CLAIM_NO	Claim number
2 CLM_TYPE	Claim type
2 SERV_DATE	Date of service
2 CAL_SDATE	Calendar date of service

Procedure: BUILD_DENT_ACUTE

CLM_TYPE = 31
DOC_SEQ = 12.

1. If the claim number matches the claim number of previous acute or pharmacy line charge, link to previous episode.
- 2-3. If the claim number matches the claim number of previous exam line charge, or if the date of service is within 31 days of the exam, then link to the exam episode; assign the exam date to the acute line charge, but change the prefix to indicate the acute portion of the episode. Therefore:

43nnn = Exam portion of the episode
13nnn = Acute portion of the episode

4. If the acute line charge is within thirty days of a previous acute charge and the first digits of the ADA procedure codes (V5625) are equal, then link to acute episode.

5. If the acute charge is within 120 days of a previous acute charge, with the first digits of the ADA procedure codes (V5625) equal and the tooth number (V5620) equal, then link to acute episode.

6. If no link, create new episode.

7. Build structure for all acute line charges:

1 DENT_ACUTE

2 POINT	Pointer to previous line charge
2 EPIS_NO	Episode number
2 CLAIM_NO	Claim number
2 CLM_TYPE	Claim type
2 THC	Treatment history code
2 SERV_DATE	Date of service
2 ADA_PROC	ADA procedure code (V5625)
2 TOOTH	Tooth number (V5620)

Procedure: BUILD_DENT_PHARM

DOC_SEQ = 13

1. If the claim number matches the claim number of any dental line charge (exam, acute, pharmacy) link to episode.

2. If no match, create new episode.

3. Initialize the following structure for each charge:

1 DENT_PHARM

2 POINT	Pointer to previous line charge
2 EPIS_NO	Episode number
2 CLAIM_NO	Claim number
2 CLM_TYPE	Claim type
2 THC	Treatment history code
2 SERV_DATE	Date of service
2 DRUG	NDC code (V5589)
2 THER_CODE	Therapeutic code (V5595)

Procedure: ALL_PHARM

CLM_TYPE = 41, 42, 44, 44, 45
DOC_SEQ = 15.

Pharmacy records are processed, using the following rules:

1. If the prescription date (V5650) is missing, proceed to rule 5 because rules 2, 3, 4 require this date.
2. If the prescription date precedes the beginning of the accounting year (ANN_DATE), use the beginning of the accounting year as the date.
3. Search all physician, dental, and pharmacy line charges for a match of prescription date with date of service of the previously processed line charge. If such a match is found, and either the NDC code (V5589) or the therapeutic code (V5595) also matches, then link to existing episode.
4. Search hospital, physician, dental, and pharmacy records according to CLM_TYPE for a match of prescription date with date of service. Flag (EPIS_FLG = 2) to indicate "date only" match. The acceptable categories for matching, depending upon CLM_TYPE, are:

CLM_TYPE of Pharmacy Record	Acceptable Matching Categories
41, 45	Hospital (acute and elective) Acute physician Chronic flare-up physician
42, 43, 44	Hospital (elective only) Chronic physician Chronic flare-up physician Well-care physician

5. If the prescription date is missing, or if no match has been obtained, use the dispense date (V5603) for the following rules.

Rules 6 and 7 differ from those reported in *The Episodes-of-Illness Processing System*, N-1745-HHS. They were changed in order to improve the linking of drugs to episodes. The changes were based on other HIS work which developed extensive rules for linking drugs to visits. See SM-10,600.

6. This step is performed for all unmatched chronic (2), chronic flare-up (3) or unobtainable (5) pharmacy records. Link to an existing episode if:

- a. Match on either NDC code (V5589), or therapeutic code (V5595) of physician, or dental record; match to closest record on time within 365 days preceding and 3 days after.
- b. Search all pharmacy records for match on therapeutic code or NDC code within preceding 120 days of dispense date (V5603),
- c. Search physician and pharmacy records for match 7 days before or 3 days after dispense date.

If the therapeutic and NDC codes are missing, or if no match is obtained and the CLM_TYPE is not unobtainable, create a new chronic pharmacy episode.

7. This step is executed for unmatched acute, well-care, and unobtainable pharmacy records.

- a. Search all physician, dental, and pharmacy records for a match on either therapeutic code or NDC code within 30 days before or 7 days after.
- b. Search all physician records for date match of 7 days before or 2 days after dispense date.
- c. Search for nonchronic pharmacy episode within preceding 7 days of dispense date or 3 days after.

8. The claim type of all "unobtainable" pharmacy records will be changed to that of the matching episode. If a new episode is created from an unobtainable pharmacy record, the claim type will be changed to 43, chronic flare-up.

9. Build structure for all pharmacy line charges:

1 PHARM_ALL

2 Point	Points to previous line charge
2 EPIS_NO	Episode number
2 DRUG	NDC code
2 THER_CODE	Therapeutic code
2 SERV_DATE	Date of service
2 CLM_TYPE	Claim type.

Procedure: ALL_SUPP

CLM_TYPE = 51, 52, 53, 54, 55

DOC_SEQ = 16, 18.

1. Supplies used for cataract procedures will be flagged (CATARACT = 1) if the supply code (V5601) equals 4125 or 4127.

2. The record will be tagged a vision service (CATEG = 3) if there has been no evidence of cataract procedures (CATARACT = 0) and the supply code (V5601) is in the range, 4101-4205.

3. The record will be tagged as a hearing service (CATEG = 2) if the supply code (V5601) equals 3201 or 3202.

4. Link to existing physician episode if a match on referring physician (V5604) and supply code (V5601) occurs within 21 days before or two days after the supply line charges.

5. Search all hospital and physician records for a match on date only. Match according to claim type; see Procedure: ALL_PHARM, rule 4 for match categories.

6. Link to physician line charge within 21 days before or two days after the supply line charge.

7. Link to supply line charge within seven days before the supply record.

8. If no match was made, change claim type to chronic and create new supply episode.

9. Build structure for supply records:

1 SUPPLY_ALL

2 Point	Points to previous line charge
2 EPIS_NO	Episode number
2 SERV_DATE	Date of service.

PROGRAM: REPRy.V440

I. Purpose

This MARK IV program aggregates claim inventory records which have been assigned an episode number and date by the program EPIS.V420 into episode records--one record per episode per person. These records are subfiled and a report is generated.

II. Input

M40LD--claim inventory, output from EPIS.V420.

III. Output

M4SUBF1--episodes file--one record per episode per person. The structure for this new record type is shown in Table 3.6.

IV. Procedure

1. The claim inventory file is sorted in the job set SRTy.V420 by:
 - a. Insurance If
 - b. Person ID
 - c. Episode No. (last three digits)
 - d. Episode No. (ab--two-digit prefix)
 - e. Episode date.
2. Each new episode number or person initializes a new episode record. The line charges--VARPAD variable, SPLIT_CHRG--are summed over all records with the same person ID and episode number and saved as the variable, LINE_CHR.
3. The episode is dated (EPIS_DATE) as of the first record in the episode with the exception of outpatient hospital episodes (not pregnancy), which are dated as of the hospital admission date.
4. All other variables are initialized from the corresponding VARPAD variables in the first claim inventory record for the episodes number except the variables CATEG--category of service. If variable CATEG in any claim inventory record in an episode has a value other than -1, the value of CATEG in the episode record will be defined accordingly with one exception: an acute episode (episode prefix = 12)

Table 3.6

EPISODE-OF-ILLNESS RECORD STRUCTURE

DCL 1	EPISODE,		
2	PERSON	CHAR(8)	/* Person ID
2	CLM_SEQ	CHAR(8)	/* Claim SEQ.ID
2	PLAN	CHAR(11),	/* Plan
2	MDE	CHAR(11),	/* Maximum Dollar Expenditure
2	ACCT_YR	CHAR(11),	/* Begin.date acct.year
2	DIAG	CHAR(11),	/* Episode diagnosis or NDC code if drug episode
2	EPIS_NO	CHAR(11),	/* Episode number
2	EPIS_DATE	CHAR(11),	/* Begin.date of episode
2	FINAL_DAT	CHAR(8),	/* Final episode date
2	NO_ACT_YR	CHAR(3),	/* No. of account.year
2	CATEG	CHAR(11),	/* Mental Health,hear.vis.
2	DOC_SEQ	CHAR(11),	/* DOC.SEQ.initial claim
2	LINE_CHR	CHAR(11),	/* Total charges
2	MDE_STRT	CHAR(11),	/* MDE available at the beginning of the day of this episode
2	MDE_STRT	CHAR(11),	/* MDE available at the end of the day of this episode
2	DATE_EXCEED	CHAR(11),	/* Date individual or /* family exceeded MDE
2	DATE_FAM_EXC	CHAR(11),	/* Date family exceeded MDE for families on Plan 1 or 13

NOTE: For eligible persons having no episodes and therefore a dummy record, all fields will be set to ' -1999' with the exception of: PERSON, CLM_SEQ, ACCT_YR. All 11-character fields consist of two parts: 1. VAL CHAR (8), the value of the variable; 2. DSI CHAR(3), the data status indicator-- either ' 0' or '999'.

tagged as a vision or hearing service (CATEG = 2, 3) will be changed so that CATEG = -1. This reflects the assumption that injuries, as opposed to standard vision services, will be covered in any National Health Insurance Plan. Similarly, following the precedent of Medicare, we have considered vision services related to cataracts as likely to be covered, and so have not included them in the optional vision services category. (See SM-5887 for definitions.)

5. Chronic episodes (episode prefix = 22 or 25 (chronic supplies)) in which the variable CATEG = 3 (vision) indicates a vision service will be changed to well-care episodes. The episode number prefix will be changed to "42" or "45."

6. The first nonmissing diagnosis saved in the VARPAD array in the program, SPLIT.V130, is used as the episode diagnosis (DIAG). If the episode is a drug episode and there is no associated diagnosis, the NDC code (V5589) is saved as the variable, DIAG. If the episode is a hospital episode and no diagnosis has been determined, the first of the three hospital diagnoses (V5522) is saved as the episode diagnosis.

PROGRAM: SAMPy.V450

I. Purpose

This SAS program generates the sample file of all persons eligible for benefits during any part or all of year y appended with the Insurance Family Identifier, an identifier defining the participants sharing a Maximum Dollar Expenditure (MDE), and the insurance plan.

II. Input

HIS Sample File: K.K1375.AR062.DR0407.EGFAM.VF12.
Insurance Family File: F.F0825.AR376.DR0376.INSUR.V111.

III. Output

Episodes Sample File: A file containing one record per person eligible for benefits during year y appended with Insurance Family ID and plan.

PROGRAM: MTCHy.V450

I. Purpose

This MARKIV program appends the sample file with the Maximum Dollar Expenditure (MDE) from the Participant Master File (PMF) corresponding to the beginning of year y.

II. Input

PMF: DR0072.RAWm.V210 where $m = y - 1$.
Sample File: SAMPy.V401.

III. Output

Sample File appended with MDE.

PROGRAM: MTCHy.V460, MTCHy.V470

I. Purpose

Two MARKIV programs assign the MDE appended to the sample file in MTCHy.V450 to all members of the Insurance Family. In some cases newborns or new persons joining the family in the middle of the year are not on the PMF.

II. Input

Sample File appended with MDE--SAMPy.V451.

III. Output

Sample File appended with MDE including persons joining the family in the middle of the year.

PROGRAM: CORRy.V480

I. Purpose

This MARKIV program corrects alias person identifiers by recoding any occurrences of the alias or second identifier back to the person's initial identifier.

II. Input

Multiple ID File: File of person identifiers and associated alias identifiers.
Episodes File: EPISy.V441.

III. Output

Episodes File: One record per person per episode with corrected person identifiers.

PROGRAM: SAMPy.V490

I. Purpose

This MARKVI program matches the episodes file to the sample file and subsets all episodes for all persons eligible for benefits in year y. Two files are generated:

1. Episodes for participants on plan 13, the individual deductible.
2. Episodes for participants on all other plans.

II. Input

Sample File--SAMPy.V471.
Episodes File--EPISy.V481.

III. Output

SAMPy.V481--Episodes for persons eligible on plan 13.
SAMPy.V492--Episodes for persons eligible on all other plans.

PROGRAM: AGGRy.V520

I. Purpose

This program determines the date on which an individual participating on Plan 13 (Individual Deductible) exceeds his or her MDE (Maximum Dollar Expenditure).

The episodes file is sorted by person identifier and episode date. The costs to the individual of each episode are subtracted successively from the MDE until the MDE changes from positive to negative. When this change occurs, the date of the episode causing the change is saved as the date on which the MDE was exceeded.

The coinsurance rate for participants on Plan 13 is 0.95 and all inpatient services are fully covered; therefore, the cost of an episode to the individual is equal to 0.95 of the covered charges excluding inpatient charges. The MDE for each individual is \$150. The \$450 per family limit on Plan 13 is treated later.

II. Input

Episodes file for persons on plan 13--EPISy.V491.

III. Output

Episodes file appended under date on which individual exceeded MDE.

PROGRAM: REPRy.V520

I. Purpose

The episodes file for participants on plan 13 is sorted by Insurance Family ID, person, date, and episode numbers. A MARKIV program appends all of a participant's episodes with the date on which his or her MDE was satisfied and reports this information.

II. Input

Episodes file for plan 13 participants--SAMPy.V521.

III. Output

Episodes file appended with date MDE exceeded.

PROGRAM: AGGRy.V540

I. Purpose

This PLI program determines the date on which a family (identified by the Insurance Family ID) exceeds its Maximum Dollar Expenditure (MDE).

The episodes file is sorted by Insurance Family ID and episode date. The costs to the family of each episode are subtracted successively from the MDE until the MDE changes from positive to negative. When this occurs, the date of the episode causing the change is saved as the date in which the MDE was satisfied.

The actual cost of an episode to a family is a function of the family's plan and its associated coinsurance rate. There follows a list of HIS plans and their respective coinsurance rates:

Plan	Coinsurance Rate(s)
1-4	1.00
5-7	0.25
8-10	0.50
11 (Free Plan)	1.00
13 (Individual deductible)	0.95, 1.00 for inpatient
14-16	0.95
17-19	0.25, 0.50 for dental or mental health

Therefore, the actual cost of an episode to a family is equal to the sum of the covered charges times the appropriate coinsurance rate. After a family has spent an amount equal to its MDE, there is no charge to the family for covered medical expenditures. Exceptions to this are plans 11 (the free plan) and 13 (the individual deductible); for plan 11, the free plan, all covered charges are paid for by the plan and the date on which the MDE is satisfied is set to the first day of the family's accounting year. The MDE for plan 13, the individual deductible, is satisfied when the family has exceeded an MDE of \$450.

II. Input

Episodes for persons on plan 13--SAMPy.V523.
Episodes for persons on other plans--SAMPy.V492.

III. Output

Episodes appended with date Insurance Family satisfied MDE.

PROGRAM: REPRy.V550

I. Purpose

This MARKIV program reports the final episodes appended with the date the family satisfied its MDE.

II. Input

Episodes File--SAMPy.V541.

III. Output

Final Episodes File--SAMPy.V551.

Appendix A
REASON CODES AND EXPLANATIONS

Reason Code	Reason
01.	Inpatient hospital accommodations in a private room
02.	Inpatient hospital comfort items
03.	Inpatient hospital custodial care
04.	Cosmetic surgery not resulting from an accidental injury
05.	Psychiatric outpatient services in excess of fifty-two consultations per year
06.	Outpatient psychiatric services
07.	Outpatient personal care services
08.	Orthodontia not resulting from accidental injury
09.	Christian Science practitioner or Sanatorium not listed in the Christian Science Journal.
10.	Nonemergency transportation
11.	More than one eye or hearing examination during the accounting year
12.	More than one pair of eyeglass frames every two accounting years
13.	More than one set of eyeglass lenses during the accounting year
14.	More than one hearing aid during the accounting year
15.	Exceeds limit on eyeglass frames or hearing aids
16.	Repairs to eyeglass frames and hearing aids
17.	Diagnostic, screening, preventive, or rehabilitative services not otherwise specified in the Scope of Coverage
18.	More than one piece of medical equipment, appliance, or supply
19.	Prescription or supplies costing more than \$25.00
20.	Not medically necessary
21.	Duplicate line item
22.	Amount paid on another EOB
23.	Service prior to enrollment
24.	Procedure done twice
25.	Certificate of Benefits stipulations on service not met
26.	Prior authorization not approved
27.	Participant not eligible for dental care
28.	Blood credit
29.	Over-the-counter drugs
30.	Deductible not met
31.	Participant's coinsurance portion
32.	Service covered by Workmen's Compensation or Employer's Liability Laws

REASON CODES AND EXPLANATIONS (cont.)

Reason Code	Reason
<hr/>	
33.	Pass through
34.	Prepayment made
35.	Services covered by accident insurance policies
36.	Medicare paid
37.	Discount
38.	Not covered prepayment and deductible
39.	Not covered prepayment and coinsurance
40.	Discount and deductible not met
41.	Discount and coinsurance
42.	Paid by other insurance carrier
43.	Paid by agency other than insurance company
44.	Services obtained outside of Group Health Cooperative
45.	Plan benefit is 5 percent of covered charges
46.	Services obtained at GHC
47.	Allowance on over-the-counter drugs per illness per accounting year is met
48.	Services paid for by GHC
54.	Charge information unavailable--charge coded as one cent
55.	Discount plus plan benefit is 5 percent
56.	Medicaid paid
57.	Company physical provided as fringe benefit--charge coded as one cent but true charge unknown
58.	Workmen's Compensation--charge coded as one cent, but true charge unknown
59.	After termination date
60.	Claim in duplicate
61.	Not eligible participant
62.	Suspended
63.	No service
64.	Before enrollment date
65.	Claim filed after time limit
66.	No charge
67.	Underpayment
68.	Overpayment, deducted on another claim
69.	Overpayment, returned
70.	Overpayment, deducted on this claim, overpaid on another claim
71.	Billed in error--patient not seen
73.	Duplicate payment, recovered
74.	Duplicate payment, not recovered
80.	Prepayment no MDE involved
81.	Prepayment--part applied to the MDE

Appendix B

LIST OF ANNIVERSARY DATES
(All Sites)

Date		
<hr/>		
Dayton	11/1/74	
	12/1/74	
	1/1/75	
	2/1/75	
Seattle ^a	1/1/76	
	2/1/76	
	3/1/76	
	4/1/76	
	5/1/76	
	6/1/76	
	7/1/76	
	8/1/76	
	9/1/76	
Massachusetts	7/1/76	
	8/1/76	
	9/1/76	
	10/1/76	
South Carolina		
	5 Year Group	3 Year Group
	11/1/76	11/1/78
	12/1/76	12/1/78
	12/31/76	1/1/79
	1/31/77	2/1/79

^aDoes not include Group Health Cooperative of Puget Sound enrollees.