

Medicare Chronic Care Improvement Program

National sample file of Medicare beneficiaries

WARNING

All uses of data on this CD are covered by a Data Use Agreement with the Centers for Medicare & Medicaid Services. The Data Use Agreement specifies the purposes for which these data may be used, persons responsible for the data's confidentiality and security, and the period during which data may be retained. Violations of this Data Use Agreement may be punishable by criminal penalties, including fines or imprisonment.

Files on this CD

- CCIP National Sample Data Dictionary.xls, an Excel file that describes the contents of the claims and other files. The data dictionary applies to both disease cohorts (diabetes and/or CHF; COPD). Note: when printing this document, in the "Print" dialog box choose "Print what: entire workbook" in order to print all the tabs with correct pagination.
- Claims and eligibility/demographic records for a random national sample of 10,000 Medicare beneficiaries with diabetes and/or CHF. Selection criteria are further described below.
- Claims and eligibility/demographic records for a random national sample of 10,000 Medicare beneficiaries with COPD. Selection criteria are further described below.
- For each cohort, there are three types of files:
 - The .csv (comma delimited) files are the claims and eligibility records, labeled by cohort and type
 - The .rtf (rich text file) files give the file format and layout for each .csv file
 - The .sas (SAS) files provide SAS import statements for each .csv file

Diabetes and/or CHF cohort sample record counts

Data Type	Records
Other Part A procedure codes	13,319
Other Part A diagnosis codes	274,469
Part A	86,047
Part B	752,502
DME	48,131
Member Demographics	10,000

COPD cohort sample record counts

Data Type	Records
Other Part A procedure codes	13,097
Other Part A diagnosis codes	270,663
Part A	84,347
Part B	770,460
DME	157,992
Member Demographics	10,000

Sample file construction

Data in the files on this CD were extracted by the following process:

- Source files consisted of fully-adjudicated Medicare claims for dates of service in CY2002, plus associated eligibility information from the Medicare Beneficiary Database (MBD) and Eligibility Database (EDB) eligibility databases. These data are maintained in the CSC National Repository. Extracts used claims paid through June, 2003 (six months of run-out).
- Persons with selected chronic conditions were identified according to criteria given in the Medicare Chronic Care Improvement Program (CCIP) national sample data dictionary. Two cohorts were created: (1) all persons with either diabetes or congestive heart failure (CHF); (2) all persons with chronic obstructive pulmonary disease (COPD). Persons meeting the criteria may appear in both cohorts. 5,025,250 persons met the criteria for the diabetes/CHF cohort, while 1,738,691 persons met the criteria for the COPD cohort.
- For persons identified by the criteria, claims were extracted according to criteria given in the data dictionary.
- Using the CMS HCC risk adjustment grouper (available for download from the CMS website, www.cms.hhs.gov/healthplans/rates, download hccsoftware.zip), HCC community risk scores were assigned to each person.
- The subset of persons in the cohorts suitable for the CCIP was defined as follows:
 - COPD cohort: HCC score greater than or equal to 1.34 (n = 1,206,437)
 - Diabetes/CHF cohort: HCC score greater than or equal to 1.35 (n = 3,015,870) [for more information see www.cms.hhs.gov/medicarereform/ccip]
- From each cohort subset as defined in the above bullet, 10,000 persons were chosen at random for the sample file, and all of their claims for CY2002 (incurred) were extracted.
- Only selected fields from the claims are included in the sample file, as described in the data dictionary.
- Pursuant to privacy regulations, identifying information about beneficiaries has been altered or suppressed, as follows:

- The beneficiary ID number has been replaced by a substitute, called hic_key. The hic_key substitute ID number will link claims for the beneficiary, across Part A, Part B, DME, and member demographics files (linking of other procedure and other diagnosis files is described below).
- Dates of service have been suppressed; all dates of service are in CY2002.
- Geographic information has been suppressed.
- Provider ID information has been suppressed; provider type and provider specialty information has been retained.

File linkages

Part A data are contained in three files:

- The Part A claims file contains most of the Part A data, including the principal diagnosis and principal procedure codes
- The Part A other procedure codes file contains additional procedure codes. The linking field between the Part A claims file and the Part A other procedures codes file is ICN
- The Part A other diagnosis codes file contains additional diagnosis codes. The linking field between the Part A claims file and the Part A other diagnosis codes file is ICN.

Part A, Part B, DME, and member demographics file records link on the field hic_key.

Caveat

Claims data are necessarily limited in their usefulness in health services research and analysis. Many studies document the utility and limitations of claims data; examples are listed below. Further, it cannot be warranted that claims in the sample file represent 100% of all medical services for the beneficiary: for example, services may be provided but paid by the beneficiary or otherwise not billed to Medicare; claims could be missing because they were not fully adjudicated at the time of the extract; services could be provided but not covered by the Medicare benefit (notably, prescription drugs); etc.

Cooper GS; Yuan Z; Stange KC; Dennis LK; Amini SB; Rimm AA; The sensitivity of Medicare claims data for case ascertainment of six common cancers [see comments] Medical Care 37 5 1999 May 436 444

McClish DK; Penberthy L; Whittemore M; Newschaffer C; Woolard D; Desch CE; Retchin S; Ability of Medicare claims data and cancer registries to identify cancer cases and treatment American Journal of Epidemiology 1997 Feb 1;145 3 1997 Feb 1 227 233

Newschaffer CJ; Bush TL; Penberthy LT; Comorbidity measurement in elderly female breast cancer patients with administrative and medical records data *Journal of Clinical Epidemiology* 1997 Jun;50 6 1997 Jun 725 733

Steinwachs DM; Stuart ME; Scholle S; Starfield B; Fox MH; Weiner JP. A comparison of ambulatory Medicaid claims to medical records: a reliability assessment. *Am J Med Qual* 1998 Summer;13(2):63-9).