Enrollment

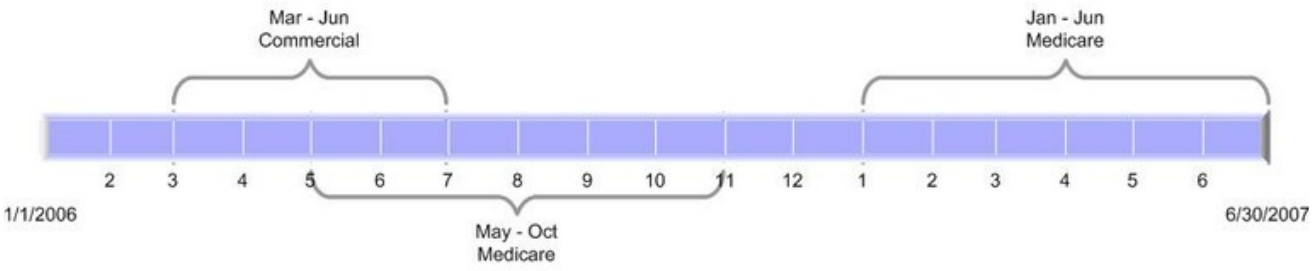
Records of periods during which we have information on the health care of the people involved.

1. Overview

The **prior** version of the enrollment data table contains a list of people and periods, during which there is a formalized relationship between the person and HMO whereby the HMO will either *provide* or *pay for* the bulk of the person's medical care. People who are only nonexclusive patients of the HMO (e.g., a person whose only contact w/the HMO is showing up at a vaccination clinic, where they pay out of pocket, or even one who receives substantial care paid for by a third party, but has no particular affinity for the HMO) should **not** have appeared in this file (though their utilization should appear elsewhere in the VDW).

In this version, we extend this definition to also embrace--at sites where appropriate--people/periods during which there is no formal relationship, but for whom the site can make a credible claim to have a reasonably complete record of the person's medical treatment. These people/periods will be discriminable from those with a formal relationship by the value found in the new enrollment_basis field described below.

Each record represents a period of time during which the information on the included variables was true. As many records as are necessary should be added to represent changes over time in, say, insurance type(s), or completeness of data capture. There may be many contiguous records for a single period of enrollment, in order to account for changes in the character of the relationship. However, periods **may not overlap** one another--there should be one and only one record covering any given day and MRN. See Figure 1 below for an illustration of the concept.



Should result in 4 records:

MRN	enr_start	enr_end	ins_commercial	ins_medicare
111	03/01/2006	04/28/2006	Y	
111	05/01/2006	06/30/2006	Y	Y
111	07/01/2006	10/31/2006		Y
111	01/01/2007	06/30/2007		Y

Figure 1: A person with the following 3 coverage periods should be represented as shown in the table at the bottom of the graphic.

Site Data Managers should refer to the discussion and code for achieving this on the CRN Portal .

No effort should be made to close gaps of any length in enrollment periods that appear in the indigenous data.

Note that some sites have added site-specific variables that are not part of the specifications. At these sites, end-users may well find multiple records for contiguous time periods, over which none of the official specification variables change. End-users will likely be interested in the %CollapsePeriods macro for reducing these sets to single record per period (and closing up enrollment gaps of a user-specified length).

2. Enrollment Table Data Dictionary

Variable name	Type (Length)	Preferred Format	Description	Valid Values	Comments
MRN		n/a			

	char (varies)		An arbitrary identifier unique to an individual within a site.	Any, so long as they uniquely identify individual people.	Used to link people across files within a site. May or may not contain the official local indigenous person identifier (e.g., "medical record number"). **Regardless of whether it contains the official local identifier, this variable should never leave the site.** Projects needing to move individual-level data should create a study-specific person identifier and substitute it for MRN on any data that is to move (See, e.g., the %DeIDDset() standard macro.)
enr_start	numeric (4)	mmddyy10	Beginning date of the enrollment period.	SAS Date	Dates setting out the limits of the period described by this record. Sites with exact date information on starts/stops should use them, those that don't should use the first and last of the month.
enr_end	numeric (4)	mmddyy10	End of the enrollment period.	SAS Date	Periods that are ongoing at the time the file is created should bear the last day of ~~the month prior to the update~~ known, actual enrollment at the time of the update. This will vary across sites depending on the latency of their source data. For example, if the source data feed is realtime, enr_end should be set to &sysdate; if source data is updated weekly, enr_end should be the last day of the week prior to the vdw update, and so on.
ins_medicaid	char(1)	n/a	Flag indicating whether the person had any Medicaid coverage during the period.	Y Yes N No U Unknown	See appendix B for a detailed description of Medicaid coverage.
ins_commercial	char(1)		Flag indicating whether the person had any Commercial coverage during the period.		
ins_privatepay	char(1)		Flag indicating whether the person had Private Pay coverage during the period.		This is defined broadly as the covered person bearing all (or most) of the cost of the premium.
ins_statesubsidized	char(1)		Flag indicating whether the person had any State Subsidized coverage during the period.		Individuals who are eligible for a State-subsidized health insurance plan and have elected to join the participating HMO under a risk contract with the State health insurance program. These programs are typically designed for persons who are too wealthy to qualify for Medicaid coverage but who cannot afford to purchase health insurance on their own, usually because they do not

			have access to group health insurance.
ins_selffunded	char(1)	Flag indicating whether the person had any coverage through an Employer group that insures itself.	Larger companies may seek to save costs by directly bearing the risk of their employees' medical expenses, rather than transferring it to an insurer/HMO. In these cases, the HMO generally retains the role of *provider* of care, and takes on the duties of a claims administrator.See also Appendix B
ins_highdeductible	char(1)	Flag indicating whether the person had any coverage in a high-deductible plan.	A form of catastrophic health insurance defined in US Internal Revenue Service Publication 969 . This type of coverage is required to take advantage of Health Savings Accounts. See Note 1 , and the detailed description in Appendix B .
ins_medicare	char(1)	Flag indicating whether the person had any Medicare coverage, including Medicare Working Aged	Included for cases where the particular parts of Medicare are not known, but it is known that *some* Medicare coverage is in place. If any of the specific Medicare-part vars are set to 'Y', this should also = 'Y'.See also, the detailed description in Appendix B .
ins_medicare_a	char(1)	Flag indicating the person had Medicare Part A coverage during the period.	Medicare Hospital insurance. See the detailed description in Appendix B .
ins_medicare_b	char(1)	Flag indicating the person had Medicare Part B coverage during the period.	Medicare outpatient insurance.See the detailed description in Appendix B .
ins_medicare_c	char(1)	Flag indicating the person had Medicare Part C coverage during the period.	"Medicare Advantage" (nee Medicare+Choice) plans. Captitated benefits covering in- and outpatient care.See the detailed description in Appendix B .
ins_medicare_d	char(1)	Flag indicating the person had Medicare Part D coverage during the period.	Medicare prescription drug coverage.See the detailed description in Appendix B .
ins_other	char(1)	Flag indicating whether the person had an Other type of coverage during the period.	Intended to cover any type of insurance not adequately categorized by one of the other Ins_* variables.
plan_hmo	char(1)	Had at least some coverage under an HMO plan	HEDIS defines HMO as: An organized health care system

					that is accountable for both the financing and delivery of a broad range of comprehensive health services to an enrolled population. An HMO is accountable for assessing access and ensuring quality and appropriate care. Practitioners affiliated with the health care system render health care services. In this type of organization, members must obtain all services from affiliated practitioners and must usually comply with a predefined authorization system to receive reimbursement. See also the detailed description in Appendix A .
plan_pos	char(1)		Had at least some coverage under a point-of-service plan		HEDIS defines POS as: An HMO with an opt-out option. In this type of organization, members may choose to receive services either within the organization's health care system (e.g., an in-network practitioner) or outside the organization's health care delivery system (e.g., an out-of-network practitioner). The level of benefits or reimbursement is generally determined by whether the member uses in-network or out-of-

			network services. Common uses of 'POS' include references to products that enroll each member in both an HMO (or HMO-like) system and in an indemnity product. A POS product is also referred to as an 'HMO swing-out organization', an 'out-of-organization benefits rider to an HMO' or an 'open-ended HMO'. See also the detailed description in Appendix A .
plan_ppo	char(1)	Had at least some coverage under a preferred provider organization plan	HEDIS Defines PPO as: PPOs take responsibility for providing health benefits-related services to covered individuals and for managing a practitioner network. They may administer health benefits programs for employers, either by assuming insurance risk or by providing only administrative services. See also the detailed description in Appendix A .
plan_indemnity	char(1)	Had at least some coverage under a traditional indemnity plan.	Indemnity insurance gives an insured the right to reimbursement of covered healthcare costs from the insurer. There is typically no pretense that the care involved is 'managed'. See also the detailed description in Appendix A .
DRUGCOV	char(1)	Flag indicating whether the insurance coverage included at least some payment/coverage for prescription drugs.	Defined as *any* drug coverage during the enrollment period. See Note 2 .
outside_utilization	char(1)		Y

			Is this person/period likely to have substantial health care utilization outside the HMO system, under circumstances that would not result in a claim being submitted to the HMO?	N U	Yes, we expect the utilization capture for this person/period is incomplete. No, there is no reason to suspect we have incomplete capture of utilization. completeness of capture is unknown.	See Note 3 . **The capture intended here does not include lab results or vital signs data.** Because those data areas are not feedable from claims data, and because the number of studies that rely on those data areas are relatively few, they are defined outside the scope of this variable. Note that having data solely from claims is not a bar to using an 'N' value here, so long as the claims feed the appropriate VDW files.
enrollment_basis	char(1)		The basis for the claim (if we are making one--see outside_utilization above) that we have something like complete capture of the medical record in the VDW.	G I B P	Geographic Basis Insurance Basis Both Insurance and Geographic bases Non-enrollee Patient	See Note 4. The only sites expected to have any records with enrollment_basis values of 'G' or 'B' are Marshfield and Geisinger. The non-enrollee patient value 'P' is new as of March 2012, and ▪ **does *not* make the same claim of reasonably-complete-data-capture that the other values are intended to convey** (subject to modification by the outside_utilization variable). ▪ Thus these people/periods **are not usable for applications that require a defined population.** ▪ Adding records for non-enrollee-patients is *optional*--some sites will add these records to the enrollments table, and others will not. ▪ Users should check the individual site implementation pages (reachable by clicking the cells in the Enrollments row of the implementation

					overview page) to see which of their sites have added these records. ▪ In particular, *do not assume that just because you don't find these records in your home Enrollments file that other sites where your program will be run have not added them.* Please see Note 4abelow for further comments on the new non-enrollee-patient records.
PCC	char (varies)		The clinic to which the patient is paneled in administrative record.	Same as utilization.facility_code	See Note 5 .
PCP	char (varies)		The clinician to which the patient is paneled in administrative data.	Same as utilization.provider.	See Note 5 .

3. Notes:

Note 1: High Deductible Health Plans

These plans came into vogue around 2006, and were aimed at stemming premium cost increases. They are characterized by deductibles far in excess of those typically offered in traditional managed care plans--to the point where they resemble catastrophic coverage plans with a thin veneer of preventive/routine care coverage. The rules for when deductibles apply and how they are calculated within/between family members can be quite complicated. Please document any variations on individual/family deductible structure.

Note 2: DrugCov

If the member has **any** drug coverage, meaning that the insurer is at risk for any portion of the cost associated with drug dispensings, this should be coded as 'Y'. It should of course follow from this that coverage for any particular drug is not guaranteed.

Note 3: outside_utilization

One of the primary advantages that HMORN member orgs enjoy in doing research is our ability to enumerate populations who are *at risk* for various things--suffering illness, utilizing healthcare services, etc. This is what allows us to e.g., compute rates of illness, and to identify sets of "unexposed" people for use as control subjects in studies.

Implicit in this ability is the notion that we have a generally complete record of a patient's health care in electronic data, enabling us to make inferences from an absence of data in a given person's record. For example, if a given person has no record of a colonoscopy during a period of enrollment, we would like to be able to conclude that they *did not in fact have a colonoscopy* . We would not like to have to worry--as many other healthcare organizations do--that the person has had a colonoscopy elsewhere, under circumstances that would not result in a claim or other notification to the HMORN member organization for inclusion in administrative data.

That said, at nearly every HMORN site there are enrollees whose relationship to the organization is known to be of a type that does not give much assurance that the electronic data are complete. The forms these relationships take are many and varied, and yet most every site has had to come to terms with those whose enrollment should not be taken to assure completeness of the medical record. Rather than try and take account of the various circumstances and insurance products at the various sites, and harmonize them to the point where we might define variables from which completeness can be derived, **we have chosen to put the logic for the derivation in the hands of the Site Data Managers** . They are closest to their sites' data, are familiar with local custom, and their good judgment is the basis for many other important decisions. In general, the question SDMs should ask themselves in populating this variable is "would I exclude this person/period from the denominator if I were calculating an incidence rate of a disease, procedure or pharmacy fill?"

Please note the negative nature of the logic here--a record with outside_utilization = 'N' is **not** a guarantee that data capture is in fact complete. It is merely a statement that the local SDM does not know the capture to be incomplete--a much weaker statement. This is due to the difficulty entailed in determining (for every person, over every period of enrollment) the degree of data capture. We do not expect SDMs to derive this variable from what is found empirically in utilization data--this is only for people/periods for whom we can tell from indigenous enrollment data that their capture is incomplete. SDMs at those rare sites that do not have any of these known-partial-capture enrollees, they should feel free to assign an 'N' to every record.

This variable is meant to serve an important need to research projects, but projects that are very sensitive to the need for complete data capture should consult with their sites about the degree to which this variable is suitable, or whether study-specific measures need to be taken.

A Special Note about Lab Result and Vital Signs Data Capture

In general, people/periods for whom the only source of data is claims should have outside_utilization = 'N'. That is, good capture of claims data in utilization and pharmacy is enough to have an 'N' value. However, at most sites large proportions of this group will also have no records in the vital signs or lab results files (and at Group Health at least, the Tumor file as well). Thus, projects requiring complete data capture in these files will have to make special arrangements with the sites to develop such samples.

Note 4: enrollment_basis

Across HMORN sites, there are two different bases for the claim that we have a complete medical record for someone in administrative data. The traditional basis is the insurance function served by the HMORN member organization. For cases where the member organization is not itself *providing* the care, the fact that it is responsible for *paying for* the care gives the actual providers a strong incentive to inform the organization about the particulars of the care so they can be paid for it.

In (at this writing at least) two cases, however, there is another possible basis for the proposition that an organization has complete data for a person. These organizations--Marshfield Clinic and Geisinger--are so dominant in the geography where they operate that they have a near monopoly on most medical services. At those sites, the mere fact that a person lives in a particular location gives excellent assurance that the vast bulk of their care will be captured in HMORN member data systems.

This variable exists so that these sites can include their geography-based samples in studies using VDW data, and so individual studies that need to can discriminate the two types of relationship to the HMORN member organization.

Note 4a: Non-Enrollee-Patients

Users should program defensively, assume that one or more of their sites will have these records, and decide to in- or ex-clude them as appropriate for their project.

Deciding on enr_start and enr_stop date values for the non-insurance bases is difficult. At this writing we do not have code, guidelines or recommendations for e.g., imputing periods on the basis of specific contacts w/the site's facilities. We do aspire to developing empirically-derived probable periods-of-engagement, likely building on the work Craig Wood presented at the 2010 annual meeting.

While we encourage implementing sites to include all available information regarding e.g., the insurance and plan types of these non-enrollee patients, users should expect a higher rate of 'unknown' values in the descriptive variables on these records.

Note 5: Primary Care Clinic/Physician

These data items are intended to come from administrative data (e.g., a membership and billing system) and **do not purport to describe the location or provider of primary care actually received** . That is, there is no effort to empirically assign PCP/PCC values from observed utilization data. Providers and facilities listed in these variables *may* exclusively provide primary care or may be a part of a facility that provides an array of care.

Primary Care Physicians will often be MDs, but may include other types of providers such as a physician assistants or nurse practitioners if the medical plan panels patients to clinicians with these disciplines. Who the enrollee actually receives primary care from may be different than the PCP. The PCP value should link to the PROVIDER lookup table associated with the Utilization data structure.

4. New V3.1 enr_end definition for ongoing enrollment periods

The proposal to change enr_end values for ongoing enrollment periods was approved at the VIG mid-year meeting in November 2012. Refer to the change proposal document for the complete specification.

Change Enrollment.enr_end Values For Ongoing Enrollment Periods

The target implementation date for the change is found at the VIG Calendar page.

V3.1 Change to enr_end Definition Implementation Status by Site

Site Data Managers: please update with the following statuses.

- Not started
- Started
- Verification program results submitted
- Addressing issues (uncovered by program)
- Completed

Feel free to add additional comments and clarifications.

Site	Status	Date of last status update
EIRH	Complete	3/26/2013
Fallon	This will not result in a change in our enrollment file.	3/25/13
Group Health	This will not result in a change in our enrollment file.	2/26/2013
Geisinger	Completed	2/26/2013
Henry Ford	Implemented Not QA yet	2/26/2013
HealthPartners	Completed	4/09/2013
Harvard Pilgrim	Not started. Will update during regular yearly VDW update	4/09/2013
KPCO	Completed	4/8/13
KPGA	This will not result in a change in our enrollment file.	5/15/2013

KPH	This will not result in a change in our enrollment file.	3/1/2013
KPMA	Completed	6/15/2013
KPNC	This will not result in a change in our enrollment file.	2/25/13
KPNW	This will not result in a change in our enrollment file.	2/25/13
KPSC	Not started, but plan to do it in the near future.	4/11/13
LCFResearch	This will not result in a change in our enrollment file.	3/6/13
Marshfield	This will not result in a change in our enrollment file.	4/11/13
SWH		
PAMFRI	This will not result in a change in our enrollment file.	2/26/13

5. Future Directions:

The current definition of the outside_utilization variable is a retreat from a much more ambitious proposal--to have a suite of flags--one for each substantive dataset in the VDW, which would express an opinion on the completeness of the capture in each. That is generally considered a worthy goal, but far too difficult to take on speculatively. Our hope is that as projects arise who need more particular information (and can fund the effort required), general methods can be established at each site to do just that.

6. Contact information:

Please contact the VDW Enrollment and Demographics working group for questions and suggestions.

Related content

- v3 Enrollment: Appendix A: Plan Type Descriptions
- v3 Enrollment: Appendix B: Insurance Type Descriptions
- Enrollment/Demographics QA Workplan
- Change Proposal: Change Enrollment.enr_end Values For Ongoing Enrollment Periods
- Change Proposal: Add variable to document the expected end of periods that are ongoing at the time of file creation

Comments (1)

Pardee, Roy May 19, 2010 04:37 PM
On 19-may-2010 Gwyn Saylor of KP Colorado asked:
"I am assuming that the new variable State_Subsidized should not include Medicaid. Am I right?"

To which Paul Hitz replied:

"This is correct. At the Austin meeting we specifically stated that Medicaid would not be included/considered when setting this variable."

And Mark Hornbrook:

"CORRECT!! Medicaid is a joint FEDERAL/STATE program. The State-subsidized health plan option covers state initiatives to provide health insurance to the uninsured that are not simply Medicaid-expansions. These can include

pure State-subsidized individual health insurance (that is not open to Medicare, Medicaid, and commercially insured persons), as well as State high-risk groups for medically uninsurable persons (which will be unnecessary under national health reform) and for self-employed and small businesses. Washington State has the Basic Health Plan administered by the Washington State Health Care Authority, which runs the State employees health benefits program, that is open to anyone who can prove they do not have access to group health insurance. Higher income persons receive no subsidies for their premiums and pay the full actuarial premium for their age group."