How-To Guide: Sub-County Heart Attack (AMI) Hospitalizations

Provided by CDC's Environmental Public Health Tracking Program

August 2021

PURPOSE AND USE OF THIS DOCUMENT

This document describes the steps to prepare hospitalization data for acute myocardial infarction (AMI or Heart Attack) for dissemination by CDC'S Environmental Public Health Tracking Program.

HOW-TO GUIDE

	Description
Measures	Number of Hospitalizations of Heart Attack among Persons >= 35 Years of Age Crude Rate of Hospitalizations of Heart Attack among Persons >= 35 Years of Age Age-adjusted Rate of Hospitalizations of Heart Attack among Persons >= 35 Years of Age
Data Source(s)	Inpatient hospitalization admissions
NCDM Data Requirements	• Health outcome: Acute myocardial infarction (AMI) • Census tract/County/State of residence • Geocoding precision • Age group • Hospital admission year Exclude: • Transfers • Out-of-state residents • Admissions to federal facilities Optional: • Inclusion of admissions of residents to out-of-state hospitals
Definitions Relevant to Indictor	*Acute Myocardial Infarction (AMI):* Irreversible death of heart muscle because of prolonged loss of blood supply; ICD-9-CM: 410 and ICD-10-CM: I21-I22 as the primary diagnosis. *Admission Date*: The date of the hospital admission; month, day, and year. Month and year of admission are required in data submitted to CDC. *Discharge Date*: The date of discharge from hospital. *Duplicate Records*: More than one record for the same person with the same hospital admission data (e.g., where sex, date of birth, admission date, and zip code have exactly same information).

	Description
	Event/Event Year: A hospital admission for the health outcome of interest during specific calendar year. Event year is based only upon admission year, even when discharge year is different. *Hospital Transfers*: Generally, a patient discharged from one facility and readmitted to a second facility on the same day (within 24 hours). *Hospitalization/Hospital Admission*: Condition of being placed (admission) or treated as a patient in an acute care hospital for treatment as an inpatient. Treatment as an outpatient is not considered to be hospitalization. To be considered as inpatient hospitalization, a minimum stay is required (often over 23 hours). *Multiple Admissions*: Second or subsequent admission for the same person for the same primary diagnosis code but on a different date and related to a separate event within a given year. Multiple admissions are considered separate events (generally at least 48 hours apart). *Out-of-State Admissions*: When a resident of the grantee state is admitted to a hospital located in another state (usually an abutting state). *Primary Diagnosis Code:* The first discharge diagnosis field(s) of the coded clinical record (i.e., primary or principal diagnosis). *ICD-9-CM*: Prior to October 1, 2015, diagnosis codes are represented by ICD-9-CM codes (the International Classification of Diseases, 9th Revision, Clinical Modification). AMI is classified as a primary diagnosis code of 410. *ICD-10-CM*: As of October 1, 2015, diagnosis codes are represented by ICD-10-CM codes (the International Classification of Diseases, 10th Revision, Clinical Modification). AMI is classified as a primary diagnosis code of I21–I22. *Resident*: Any person with a residential address in the county/state of the grantee at the time of the hospital admission. *Sub-county*: Sub-county (geographic resolution below county) is how the Tracking Program describes data at a finer resolution than county data. This includes census tract data and data using the Tracking Program's set of minimum population geographies, created by aggregating census tracts.
HTG Requirements and Cautions	• This How-to Guide provides instructions for the development of the dataset for submission to CDC and for calculating the required and optional measures. The Sub-County Hospitalization Data Dictionary should be referred to for the standardized definitions and notations of the variables to be submitted to CDC. The data file should be converted to the .XML file format and the required header inserted into the XML file, according to the Schema found on SharePoint. • An additional How-to Guide (entitled "How-to Guide for Measure Calculation") is available that provides instructions for calculating sub-county measures. Recipients should use the additional How-to Guide to ensure calculation of measures is consistent with those used by CDC. • *Data Source:* The data source is an individual level state inpatient hospital admission data based on primary diagnosis at an acute care facility. Please consult your data steward and data managers to understand the variables and coding system, specifically for race and ethnicity variables.

	Description
	• *Complete Dataset Guidelines:* The Tracking Network's NCDM are based upon date of admission because of the goal of relating a hospitalization event with an environmental event. Most hospitalization data (inpatient and outpatient) are released in annual discharge-based datasets; sometimes quarterly files are also released. Because the NCDM is based on admission date, it is necessary to have the dataset of the year of interest as well as that for the subsequent year (or first quarter of the subsequent year) in order to capture admissions that were discharged in the subsequent year. For example, 2005 data should not be submitted prior to receipt of either the first quarter 2006 or annual 2006 discharge dataset from the data steward. Some discretion on this rule is allowed if a program can show that inclusion of the subsequent year's data does not impact the data for the year of interest to a degree that would require re-submission. Re-submission due to incomplete data should be avoided. • *Duplicate Records:* This How-to-Guide presumes that the user has removed duplicate records (see definitions for more information), while keeping multiple admissions. • *Out-of-State Admissions*: Admissions of residents to out-of-state hospitals should be included when available but are not required to be included. For states with significant out-of-state admissions, it is preferable to wait until the out-of-state data are available for inclusion so as to avoid the need for re-submission of more complete data in the future. However, some consideration of timeliness is also appropriate; if out-of-state data are overly delayed then submission without them is acceptable. It is noted that some states must include out-of-state admissions of its residents. Use the Metadata Creation Tool (MCT) to acknowledge the disposition of these admissions and provide any additional information about out-of-state data. • *Federal Facilities:* Admissions to federal facilities, such as Veteran's Hospitals, are not included. Be certain to inform CDC if your state requires that your dataset includes admissions to federal facilities so that the measures can be appropriately footnoted. *Transfers:* Hospitalizations due to transfers between acute care hospitals ***must be excluded*** from the data and counts/measures to be generated (AMI ONLY). An algorithm to exclude transfers has been developed and is included in this how-to guide as an appendix. Use the MCT to capture if and how transfers were excluded.
Step #1	From state inpatient hospital admission data with duplicate records already removed, select all hospital records that meet the following criteria: • **Admitted** during the year(s) of interest • State of residence is your state • Date of admission is not missing Note: Steps #2 – #7could be performed in a different order so long as you retain all necessary variables needed until step #8.
Step #2a (ICD-10-CM)	Flag all admissions where primary diagnosis code is "I21.* to I22.*" by creating a variable (*for example Ishospital*) that takes the value of 1 if admission is due to diagnosis codes targeted; else its value is 2.

	Description
	Select cases having any of the following ICD-10-CM codes as **a primary diagnosis**: **ICD-10-CM Code Description** I21.*: Acute myocardial infarction I21.0* : ST elevation myocardial infarction (STEMI) of anterior wall I21.01: STEMI involving left main coronary artery I21.02: STEMI involving left anterior descending coronary artery I21.09: STEMI involving other coronary artery of anterior wall I21.1: STEMI of inferior wall I21.11: STEMI involving right coronary artery I21.19: STEMI involving other coronary artery of inferior wall I21.2*: STEMI of other sites I21.21: STEMI involving left circumflex coronary artery I21.29: STEMI involving other sites I21.3*: STEMI of unspecified site I21.4*: Non-ST elevation myocardial infarction (NSTEMI) I21.9*: Acute myocardial infarction, unspecified I21.A*: Other type of myocardial infarction I21.A1: Myocardial infarction type 2 I21.A9: Other myocardial infarction type I22.*: Subsequent STEMI and non-STEMI I22.0: Subsequent STEMI of anterior wall I22.1: Subsequent STEMI of inferior wall I22.2: Subsequent NSTEMI I22.8: Subsequent STEMI of other sites I22.9: Subsequent STEMI of unspecified site (note: '*' includes all sub variation codes) Retain only admissions identified as targeted diagnosis codes.
Step #2b (ICD-9-CM)	Flag all admissions where primary diagnosis code is "410.*" by creating a variable (for example Ishospital) that takes the value of 1, if admission is due to diagnosis codes targeted; else its value is 2. Select cases having any of the following ICD-9-CM codes as a primary/principal diagnosis: Primary diagnosis ICD-9-CM code = 410.* 410.0* Acute myocardial infarction of anterolateral wall 410.1* Acute myocardial infarction of other anterior wall 410.2* Acute myocardial infarction of inferolateral wall 410.3* Acute myocardial infarction of inferoposterior wall 410.4* Acute myocardial infarction of other inferior wall 410.5* Acute myocardial infarction of other lateral wall 410.6* True posterior wall infarction 410.7* Subendocardial infarction 410.8* Acute myocardial infarction of other specified sites 410.9* Acute myocardial infarction of unspecified site Note: '*' includes all sub variation codes Retain only admissions identified as targeted diagnosis codes.

	Description
Step #3 AMI HOSP Only	Remove transfers between hospitals following the Transfers Exclusion Protocol (Appendix A): Note that exclusion of transfers is required for AMI hospitalizations. 1. Create a unique patient identifier using a set of variables such as, but not limited to, date of birth, sex, and zip code of residence 2. Identify those with multiple admissions using the following variables: a. Primary diagnosis code b. Unique patient identifier c. Admission date (admission hour and date, if available) d. Discharge date (discharge hour and date, if available) 3. Identify admissions within 24 hours of a previous discharge for AMI (same day or following day) 4. Exclude the transfer admission (the second admission) from the dataset used for data submission and presentation
Step #4	If data were previously geocoded, skip to step #5. Otherwise, geocode addresses for remaining records according to the Environmental Public Health Tracking Geocoding Standards document.
Step #5	If a geocoding precision variable is not already present, create a new variable called "GeoCodingPrecision." Assign values to GeoCodingPrecision according to the Environmental Public Health Tracking Geocoding Standards document. There are five possible values for GeoCodingPrecision: HPCT, LPCT, UPCT, C, or S (see Sub-County AMI Hospitalization Data Dictionary). GeoCodingPrecision must have one of these values. If you receive previously geocoded data and have not arranged for geocoding precision information ahead of time, your records will likely be set to "UPCT."
Step #6	Retain, at least, the following variables. Additional variables may be necessary depending on your state's data. The actual names of the variables may differ. Please consult your data steward and data managers to understand the variables and coding system. • State of residence • County of residence • Census tract of residence • Date of admission • Date of discharge • Date of birth or age at time of admission • Geocoding precision code
Step #7	Create "AgeGroup" variable. Create AgeGroup variable using either patient's date of birth and date of admission or age at time of admission. The base format for AgeGroup is by 5-year age groups beginning 0-4 and ending with 85+ resulting in 18 age groups plus one for unknown. Hospitalization counts must be submitted to CDC by these 5-year age groups coded from 1 to 19 (see Sub-county Asthma Hospitalization Data Dictionary).

	Description
Step #8	Create summary variable "HospEvents" and summarize data by the following variables coded according to the Sub-County AMI Hospitalization Data Dictionary: • AgeGroup • CensusTract (patient's census tract as 11-digit FIPS code showing state, county, and tract FIPS, e.g., ssccctttttt (06067001101)) • County (patient's county of residence as 5- digit FIPS code) • GeoCodingPrecision • YearAdmitted Do not expand dataset to include all combinations of these variables where HospEvents equals zero. CDC will expand data and fill in zeros after data are validated. If missing combinations of these variables should not be interpreted as zero (for example, county X didn't report data in year Y), then please include this information in your metadata.
Step #9	Create the following variables and code according to the Sub-County AMI Hospitalization Data Dictionary: HealthOutcomeID 2 = Acute Myocardial Infarction
Step #10	Create new variable called "RowIdentifier." RowIdentifier should be a sequence of numbers from 1 to the number of rows in your dataset. Note: this will be included in the .XML file, but it is not included in the Sub-County AMI Hospitalization Data Dictionary.
Step #11	Order the variables according to the schema: • RowIdentifier • AgeGroup • CensusTract • County • GeoCodingPrecision • HealthOutcomeID • HospEvents • YearAdmitted
Step #12	Note: CDC has created tools for XML conversion and header creation. These tools can be found at (*https://cdcpartners.sharepoint.com/sites/NCEH/EHHE/tracking/Resources/Pages/Metadata-and-Data-Submission.aspx*). Create separate data files for each year of data. Convert dataset to .XML file. The data file should be converted to the .XML file format and the required header inserted into the .XML file, according to the schema found on SharePoint. Insert your state FIPS code in the XML header and make sure the Metadata Control Number is in the .XML file. This completes the required steps for data submission.

	Description

Step #13	Please check the following if you submit your .XML file to the gateway and it provides error messages (e.g., incorrect census tract IDs).
	• Census tracts should be geocoded to 2010 Census boundaries for 2019 and prior to ensure you are submitting valid 2010 census tracts.
	• Census tracts should be geocoded to 2010 Census boundaries for 2020 to 2029 to ensure you are submitting valid 2020 census tracts.
	• Census tracts should have the correct county ID for the county they are in (i.e., the 11-digit FIPS code should have the tract match to the given county and state).
	  ○ For example, 06067001101 shows the state (06 CA), the county (067 Sacramento County), and the tract 001101 (11.01 tract). If the county is incorrectly assigned, this will be rejected.
	• There should not be duplicates of a census tract for a given stratification (e.g., census tract 06067001101 should only have one record for a given year, month, age group, sex, and geocoding precision).
	  ○ For example, you could have the following, which would be valid:
	  ▪ 06067001101, M, high
	  ▪ 06067001101, M, low
	  ▪ 06067001101, F, high
	  ▪ 06067001101, F, low
	For data geocoded to the tract level, you should only have either a mix of high and low precision ***or*** unknown precision. This should apply across all stratifications in a given year but could change from year to year. You should not have a mix of high, low, and unknown precision for census tracts within a given year.

This is the end of the necessary steps for creating NCDMs to submit to CDC. Please see the How-to Guide for Measure Calculation for guidance on calculating measures for your state's portal.

APPENDIX A

Hospitalization Transfers Exclusion Protocol

Transfers Exclusion Sub-Team
Hospitalization Workgroup
Environmental Public Health Tracking

July 16, 2016

Sub-Team Members

Robert Knorr (MA)
John Braggio (MD)
Kirk Bol (CO)
Shuqin Xing (ME)
Tom Lambert (NH)

Table of Contents

Overview

Recommendations are presented for the identification and exclusion of inpatient hospitalizations for acute myocardial infarction (AMI or heart attack) only. A transfer is defined as admission to an acute care hospital within 24 hours following discharge from a different hospital. After the exclusion of transfers, pilot testing in five states found that crude hospitalization rates for AMI were overestimated from 2 to 13% statewide and up to 30% for some areas within states. Rates may be overestimated more in rural areas though some non-rural areas were also observed to have significant overestimation. A method for identifying and excluding transfers is provided.

Sincere thanks are extended to members of the sub-team, some of which are no longer working on EPHT and whose expertise will be missed.

Purpose

To exclude from the Environmental Public Health Tracking (EPHT) databases all inpatient admissions with a primary discharge diagnosis of acute myocardial infarction (AMI or heart attack - ICD-9-CM code 410.*, ICD-10-CM: I21.*-I22.*) who meet the definition of a transfer patient. This protocol addresses transfers of inpatients for AMI only. It does not address transfers for asthma or carbon monoxide admissions nor emergency department visits.

Operational Definition of a Transfer

A transfer is defined as a patient who is discharged from one acute care hospital as an inpatient and admitted to another acute care hospital as an inpatient within 24 hours of discharge (i.e., same date of discharge or next day), where both admissions have the same primary discharge diagnosis. Patients initially visiting an ED then admitted as an inpatient to the same hospital are not considered transfers under the EPHT definition.

Justification

Individual-level data was analyzed by five recipients (Colorado, Maine, Maryland, Massachusetts, and New Hampshire) to meet the following objectives:

1. To determine if the transfer of heart attack patients from one acute care facility to another adversely impacts statistics on the number of hospitalizations due to counting the same patient at least twice for the same incident/event;
2. To determine, if the impact appears meaningful, the feasibility of excluding transfers based upon the availability of necessary variables accessible by recipients from their data stewards; and
3. If feasible, to specify the steps for recipients to exclude transfers.

Targeting of Acute Myocardial Infarction

Studies have shown that air contaminants, such as fine particulates and ozone, can increase the risk of heart attack and respiratory disease, as reflected in hospitalization and mortality rates. Valid conclusions regarding possible causal relationships between outcome hospitalization rates with air pollution data in specific geographic areas rely on outcome data that represent individual outcome events. Individuals admitted as an emergency or inpatient at an acute care hospital for a heart attack often require specialized medical care, such as angioplasty. When warranted, patients are often transferred to another facility within hours or days of an event depending upon the level of personal risk, in order to obtain the necessary specialized care. Consequently, hospitalization rates that include transfers of heart attack patients could be overestimated because of the bias introduced from counting the same individual/heart attack event twice. In contrast, asthma hospitalizations rarely require specialized medical care that requires transfer to another facility. As examples, in Maryland, about 3% of hospitalizations for asthma lead to a transfer; in Massachusetts, less than 2% of asthma hospitalizations were transferred; and in Colorado, less than 1% were transferred.

The percent of patients that are transferred for heart attack or asthma varies by state. However, the number of transfers among asthma admissions is consistently small. By contrast, the number of transfers among heart attack admissions is consistently at least 3-5 times greater than those of asthma admissions in the same state.

Transfers from an emergency department are not addressed in this protocol because it was determined through the Nationally Consistent Data Measures (NCDMs) selection process that separate measures for ED visits due to heart attack will not be informative to public health practice. This is because most all ED visits for heart attack are subsequently admitted as inpatients and captured there as inpatient hospitalizations. Therefore, the goal of this protocol is to focus on excluding only transfers among heart attack patients admitted as inpatients and to not exclude transfers of asthma patients.

Other outcomes were not evaluated by the sub-team, but no recommendation is made to assess these. It is believed that among current NCDMs, AMI hospitalizations seem to the outcome most potentially impacted by the inclusion of transfers because of the clinical justification for transferring an AMI patient.

Impact of Transfers on AMI Hospitalization Rates

Each participating recipient determined the number of transfers by first establishing a unique patient identifier for all admissions and then utilizing available variables within the data steward's database to extract admissions that met the above definition of a transfer.

The table below presents the results of the sub-team analyses showing the numbers of transfers identified and the statewide hospitalization rates before and after the exclusion of transfers.

Table: Number of transfers and statewide hospitalization rates before and after their exclusion.*

Recipient (year of hospital data)	2010 Population (≥ age 35)	Number of AMI Admissions	Number of Transfers Identified**	Percent of Admissions Identified as Transfers	Crude Hospitalization Rate Including Transfers (per 10,000)	Crude Hospitalization Rate Excluding Transfers (per 10,000)	Percent Change
Colorado (2010)	2,589,611	5,418	122	2.2	20.9	20.4	2.4%
Maine (2008)	793,402	4,457	398	8.9	56.2	51.2	8.9%
Maryland (2009)	3,101,186	9,309	945	10.2	30.0	27.0	10.0%
Massachusetts (2008)	3,605,677	14,936	1,580	10.6	41.4	37.0	10.6%
New Hampshire (2004)	761,650	2,841	377	13.1	37.3	32.4	13.1%

*NOTE: The numbers and rates of AMI hospitalizations shown above were generated for evaluation purposes only using special selection criteria and should not be used to represent the true number or rate of hospitalizations.
**Using the recommended algorithm.

As shown in the table, percent transfers varied from 2.2 to 13.1% by state. However, there were variations within each state that are not shown, including by year and smaller geographical areas.

Some states observed significant variation in the number of admissions and the number of transfers across years. This variation might be due to changes in health care utilization patterns (as observed in Massachusetts due to new health care laws) or administrative changes in coding variables used in identifying transfers by reporting hospitals. The trend over time was usually toward fewer admissions and a smaller percentage of transfers. For example, in 2000, Massachusetts identified 18% of admissions as transfers but the proportion was only 11% in 2008. For Maine, the proportion was 12% in 2001 but 9% in 2008.

Importantly, the proportion of admissions that were the result of transfers was not consistent across regions/counties of residence within a state. The percent of hospitalizations identified as transfers across Maryland regions (four regions) ranged from 10 to 17 percent. Across Massachusetts counties (14 counties), the percent of transfers ranged from 1 to 30 percent. In New Hampshire counties (10 counties), the range was from 10 to 23 percent.

The question whether exclusion of transfers affected the interpretation of statistically significant differences between county and state rates was evaluated by Massachusetts and New Hampshire. The interpretation changed for one Massachusetts county and two New Hampshire counties, where previously these counties had statistically significantly higher rates than the state. Because heart attack patients are usually transferred in order to receive specialized cardiac procedures, the geographic areas most impacted are usually those where local hospitals are not able to provide these services. Often these are more rural areas. However, the predominantly rural areas were not the only areas that had the higher transfer rates. In Massachusetts, for example, some suburban Boston counties and some rural counties had the highest percentages of transfers.

Feasibility of Excluding Transfers

In order for a transfer to be identified, first either a single variable must be present in the hospitalization dataset that represents a unique patient identifier or a set of variables must exist, which together construct a proxy unique identifier. This process likely already has been established by all recipients in order to prepare datasets for data submissions to CDC and for state portals that exclude duplicate admissions.

Single variables that may be available include medical record number and encrypted social security number. However, medical record numbers are often hospital-specific and, therefore, only unique for a patient while they are patients of that hospital. Consequently, use of only medical record number is likely to under-identify unique patients who are admitted to more than one specific hospital. Encrypted social security number may not be available to all recipients. When available, it often is found to be missing for a large number of patients.

Recipients usually have available date of birth, sex, and zip code of residence that can be used as a set to identify a unique patient. Substitute variables, such as age for date of birth, may be used. However, the sub-team did not evaluate these, and the accuracy of uniquely identifying patients will likely be reduced using substitute variables.

Once unique patients are identified, the next step is to identify whether a transfer had occurred. The sub-team evaluated several variables that could be used to identify transfers. Initially it was believed that transfers could be identified through application of the "admission source" and/or "discharge type" variable, which are most always collected by data stewards, although the variable names may differ. The purpose of these variables is to identify where an admission came from (e.g., home or another hospital) or where the patient was discharged to (e.g., home or another hospital). It was determined that the number of transfers identified using only these codes usually would be much higher than identified when accounting for unique patient identifier and dates of admission/discharge. The reasons are unclear but often, when reviewing discharge records of patients with these codes, an actual transfer to or from another hospital was not confirmed. In some cases, a transfer was confirmed but had occurred a number of days before the second admission. In other cases, reliance on these codes could result in missing transfers because not all likely transfers were identified as a transfer by these codes. Therefore, both the sub-team and the National Association of Health Data Organizations (NAHDO) considers these codes to be unreliable for the purpose of identifying patients that meet the EPHT definition of a transfer.

In the absence of a single, reliable variable that can be used to identify transfers, a combination of variables were evaluated as a proxy indicator of a transfer. The variables focused on were the unique identifier, admission date (M/D/Yr), discharge date (M/D/Yr), and primary diagnosis code. These variables are available to a majority of recipients.

Some additional variables were found to be helpful in identifying a transfer. These were admission and discharge hour. However, while data stewards usually collect these variables, they may not be available to recipients. When available, their primary value is in determining if a second admission occurred within 24 hours of the last discharge. This becomes more important in specific circumstances, for example, if a discharge occurs early in the day prior to a second admission the following day. In this case, the longer the period following the first discharge, the greater the likelihood that the second admission is not a true transfer but a second event.

The transfer exclusion variables set (unique patient identifier, admission date, discharge date, and primary diagnosis) was evaluated compared with the use of other individual variables, such as admission source, discharge type, and medical record number. All comparisons yielded substantially greater numbers of possible transfers from the transfers exclusion variables set than from the other variables. Further, the transfer exclusion variable set is believed to more accurately identify true transfers and not miss transfers than other methods. This conclusion could only be deduced from the data available since no medical records at hospitals were

reviewed to obtain conclusive evidence of which admissions were the result of transfers. The underlying assumption of the conclusion is that patients with same or next day second admissions at different facilities for the same primary diagnosis would be the result of same event.

The underlying assumption may be incorrect if some admissions flagged as transfers are not actual transfers but independent admissions. The resultant number of heart attacks would then be underestimated. However, it seems unlikely that these multiple admissions would be the result of separate events. About 20% of patients with a heart attack are observed to have a recurrent attack soon afterwards. Most are over a period of 3-10 days after the first attack, and these appear usually related to the underlying pathology and not a separate environmental event.

The underlying assumption may also be incorrect if some transfers are delayed beyond 1-day post discharge because of scheduling difficulties or other administrative reasons. The resultant number of heart attacks could then be overestimated since the delayed transfers would be considered separate heart attack events according to the operational definition of a transfer. However, most heart attack patients requiring specialty procedures would either remain hospitalized to stabilize their condition until transfer can occur or their admission to the second facility is for treatment for a different primary diagnosis.

Recommended Algorithm to Exclude Transfers

1. Create a unique patient identifier

A de-duplication process to identify unique patients is already implemented by all recipients in preparation for annual data submissions to CDC and for state portals. A standardized approach for identifying duplicate records and re-admissions has not been established by EPHT. Currently, recipients utilize a variety of methods to identify unique patient records. Usually, no single unique patient identifier reliably exists. Medical record numbers are usually hospital-specific and, therefore, only unique for a patient at a specific hospital. Therefore, this variable is not recommended for use in identifying unique patients. Another single variable, encrypted social security number, is supposedly a unique patient identifying number across all hospitals and may be available to some recipients. However, missing values are common and may not always be reliably generated to serve the intended EPHT purpose. Therefore, recipients should evaluate completeness of the variable and also ensure that the number is created by hospitals in such a manner that it can be considered a reliable unique identifier across all hospitals. Most often, sets of variables are used by recipients to identify all records belonging to a unique patient. These variables are usually, but not limited to, date of birth, sex, and zip code of residence. If these variables are missing or have missing values, other variables may be used as substitutes. For example, age may be substituted for date of birth and census tract/city/town or county may be substituted for zip code. However, the accuracy of identifying unique patients may be reduced using substitute variables.

The description of the methods used to identify unique patients and any inherent limitations should always be included in metadata and possibly described in portal footnotes. This is important because no method is expected to identify 100 percent of unique patients and because the methods used will vary amongst recipients. It is essential that those comparing rates between recipients be aware of any differences in methods used to generate the numbers the rates are based upon.

2. Identify and Flag Admissions Meeting the Definition of a Transfer

A straightforward approach to identify admissions meeting the definition of a transfer is recommended using the following variables:

a. Primary diagnosis code ICD-9-CM: 410.* or ICD-10-CM: I21.*-I22.*
   ('*' includes all sub variation codes)
b. Unique patient identifier (as discussed above)
c. Admission date (admission hour and date, if available)
d. Discharge date (discharge hour and date, if available)

Recipients should use the four recommended variables to (a) identify unique patients with a primary diagnosis of ICD-9-CM: 410.* or ICD-10-CM: I21.*-I22.*, (b) identify those with multiple admissions, (c) identify those admissions within 24 hours of a previous admission for ICD-9-CM: 410.* or ICD-10-CM: I21.*-I22.*(same day or following day), (d) exclude the transfer admission (i.e., the second admission) from data submissions and data presentations, and (e) note the methods used and any relevant limitations in the metadata and footnotes. No SAS code or specific algorithm is recommended because recipients will need flexibility based on differences in available variables.

Other variables may be available to recipients to help meet the transfer definition. These include admission and discharge hour, although not all recipients have access to these variables. Their use is in helping to consistently conform to the 24-hour rule. Most transfers occur the same day of the first discharge or the day following, so recipients with only day of admission/discharge are expected to include only a slightly larger number of transfers than those using admission/discharge hour.

Other variables which are commonly available in the data steward's data file specifically should **not** be used alone to identify transfers, even though that is, in fact, their intended purpose. These variables may go by different names but can be called discharge type or admission source. Their codes are intended to identify where a patient was discharged to (e.g., to home, to another hospital, to another type of facility, or if the patient had died) or where the patient was admitted from (e.g., from home, from a physician referral, or from another hospital). The codes are not considered reliable. Recipients that must consider using them should first establish their reliability and then, if possible, only use them in conjunction with the four recommended variables.

Recipients who cannot exclude transfers should submit hospitalizations to CDC with transfers included and follow guidance provided by the data dictionary/schema to note that transfers are not excluded. The inclusion of transfers must also be noted in footnotes and metadata.

Limitations of the Recommended Transfers Exclusion Algorithm

Falsely classified multiple admissions or duplicate admissions are possible if different individuals have the same date of birth, sex, date of admission, and primary diagnosis.

Recipients may not all have the same data elements available; therefore, consistent application of the recommended algorithm by all recipients for either duplicating records or flagging transfers may not be possible. For example, some data stewards cannot share admission/discharge hour or admission/discharge day, therefore, making it more difficult to restrict the identification of transfers to those occurring within 24 hours of a previous discharge.

Some admissions may be misclassified as transfers or missed as transfers as a result of using only proxy indicators of a transfer and deterministic linkage. Further, it is possible that independent heart attack events within the 24 hour period between discharge and separate new admission can result leading to undercounting heart attack admissions.

Some transfers may be missed if the transfer occurred beyond 24 hours after discharge.

The exclusion of transfers algorithm does not address the issue of readmissions to the same facility or recurrent heart attacks.

Conclusion

It is acknowledged that the inclusion of hospital transfers can be important in assessing public health burden, especially in terms of health care costs, because transfers are usually more severe cases requiring more intensive and expensive care. However, it has also been recognized that the inclusion of transfers may potentially introduce significant bias in estimating disease rates used for public health surveillance and research including evaluating the relationships between rates and environmental factors. Consequently, the inclusion of transfers might be appropriate in answering a question such as whether air pollution quality or regulations affect health care costs (through more frequent hospital admissions by a single or multiple individuals) but might not be helpful in answering a question such as whether air pollution quality or regulations affect hospitalization rates (through more frequent outcome events as measured by hospital admissions by unique individuals).

The Transfers Exclusion Sub-Team sought to assess the feasibility of identifying transfers and to determine the potential impact on hospitalizations for acute myocardial infarction in five states (Colorado, Maine, Maryland, Massachusetts, and New Hampshire). One of three possible outcomes was the goal: (1) consider the development of a weighting factor to adjust rates that all recipients would use; (2) not require the exclusion of transfers but require messaging on the possible bias that including transfers might have; or (3) present a methodology for the required exclusion of transfers by all recipients.

Outcome #1 was not fully evaluated because the geographic variability of transfers across and within states appeared to be too great and the development of a weighting factor was beyond the scope of the sub-team, if feasible at all.

Outcome #2 and #3 were considered by evaluating, in the five pilot states, which data elements were necessary and available for identifying and excluding transfers, the level of work required to exclude transfers, and the potential impact on rates observed.

A small set of variables (date of birth, sex, zip code of residence, admission date, and discharge date) was identified that appeared to be readily available to most recipients and, together, could be used to identify unique patients and to flag likely transfers. In rare occurrences, substitute variables may be necessary (e.g., census tract/county of residence substituted for zip code). The work required to apply these variables was not judged to be significant because these variables are mostly used already by recipients to remove duplicates and multiple admissions, as required in the How-to Guide. The recommended transfers exclusion algorithm has limitations that include inconsistencies in the available data elements and in the reporting of data elements across states, the general lack of a single unique patient identifier, and the lack of a reliable and dedicated transfer variable in data steward datasets. However, it is well understood that there is imprecision in the accuracy of hospitalization rates, as well as in their application for disease surveillance. The exclusion of transfers has been recommended because it presents an opportunity to substantially minimize some of that imprecision.

It is acknowledged that it will not be possible to identify all transfers and that some admissions may be incorrectly flagged as transfers. Most importantly, the bias introduced by the inclusion of transfers is present within and across states, as well as across time. The required application of an algorithm to exclude transfers should result in improved estimates of hospitalization for heart attack that, consequently, will be more suitable than existing rates for comparisons and linkages with environmental data between counties and other geographic areas and between years. Recipients have the option to additionally display data on their portals that include transfers if such data are deemed useful for addressing other issues, such as public health burden.

CDC'S ENVIRONMENTAL PUBLIC HEALTH TRACKING PROGRAM

Contact us: trackingsupport@cdc.gov

Visit the Tracking Network today: www.cdc.gov/ephtracking

Follow us on social media:

- Twitter (@CDC_EPHTracking)
- Facebook (facebook.com/CDCEPHTracking)