Metadata Sample Content
Applies to Version 2.0.1 of the MCT

This document is a sample metadata record for use with data submission. To help with submission, required elements (indicated by *) have been completed with general information. You may change or add to these fields as necessary. Please review and complete all required fields.

IF YOU HAVE ANY QUESTIONS ABOUT THIS SAMPLE METADATA, CONTACT _____

*Means this element is required

I. IDENTIFICATION TAB

A. CITATION PAGE

MCT FIELD	Common Language Question Equivalent or Interpretive Statement	Sample Content
* PUBLICATION DATE	What is the official public date of release of these data? The 'date of release' refers to the date the data are sent to CDC. Since the MD record has to be created prior to sending data to CDC, please enter the date you anticipate uploading to their server.	
*TITLE	By what official name is this data set known when referenced by the data steward? For datasets that are not supplied to the CDC for the national EPHT portal, use the name provided by the data steward. For EPHT national datasets the title is standardized and provided by CDC. In the Distribution Section you can enter the data steward title in the RESOURCE DESCRIPTION field.	"State" Daily Emergency Department Visits YYYY-YYYY
URL	If these data, as described by this metadata document, are available online, what is the grantee web address (e.g., grantee portal, data steward site where the data is being held, etc.) that would take the requester directly to the data? This should be a URL that will take a user directly to the download or view location. Do not reference SharePoint – this site is not available to the public.	

| * NATIVE DATASET ENVIRONMENT | In what format(s) are the raw data stored and manipulated (e.g. MS Access, SAS, Excel)? Note: Most datasets are standardized to a CSV or other flat file but there may be exceptions. Include the number of records and file size (uncompressed). Information about the zipped file size goes into the distribution section (again in the RESOURCE DESCRIPTION field). | |

B. DESCRIPTION PAGE

MCT FIELD	Common Language Question Equivalent or Interpretive Statement	Sample Content
* ABSTRACT	(Generally the What, Where, and Who) Provide a brief summary of the content of the data found in this dataset. The summary should include a statement about the time frame covered by the data, the affected population covered by the data (e.g., for DW the data covers Public Water Supplies), geographic coverage (e.g. entire state by county). What does a record describe? To where are the data referring? About whom does the data refer? (Children with elevated lead?, asthmatics? Etc.) Please try to limit the abstract to 500 words.	This dataset contains daily, county level counts of respiratory-related emergency department (ED) visits for one year. An ED visit is counted in this dataset if a respiratory disease code (ICD-9 codes 460-519) is listed as the first-listed/primary/principal diagnosis. Respiratory outcomes include asthma, COPD, acute respiratory infections, pneumonia, and all other respiratory outcomes. Counts are provided by sex and by 3 age groups: 0 to <19, 19 to <65, and 65+. ED visits for non-residents of the reporting state are excluded.
* PURPOSE	(Generally the Why) Why was this dataset compiled and who is the target audience for these data? What are people supposed to learn from these data (e.g. data compiled to show the trend in incidence of carbon monoxide hospitalization throughout New Jersey between 2000 and 2005). Legislation that mandates the collection of these data can also be listed or	This data set will be used to conduct a county level time-series analysis to estimate the associations between air pollution and ED visits for respiratory outcomes.

	referenced here.	
SUPPLEMENTAL INFORMATION	(Caveat Section) May be a good place to state a data update policy (e.g. Since data are routinely updated and corrected throughout the year, it is recommended that data for all years be requested when requesting data and not just data for the most recent year.). In addition, may be a good place to denote how missing data are identified and how special situations were handled (e.g. how are non-detects identified or coded in the dataset). The date a dataset is considered "final" by the data steward should be included in this field. This is also the location to add additional information clarifying the date entered in the 'Currentness' field. There can be a statement here as to how to properly acknowledge the originator/source of the data. Note: This can be your state's or department's official indemnification statement on the use of data. The data usability information goes into **Liability** (what it can/cannot be used for)	
* PROGRESS	This is a pull down menu in the tool and contains the fields: Complete, Planned, In Work. By agreement, all datasets submitted to the Tracking Network must be marked "Complete".	Complete
* UPDATE FREQ.	How often are the publically available data updated on your grantee portal?	None planned

	Choices from this pull-down menu include: Continually, Daily, Weekly, Monthly, Annually, Unknown, As Needed, Irregular, None Planned.	

C. TIME PERIOD INFO PAGE

MCT FIELD	Common Language Question Equivalent or Interpretive Statement	Sample Content
* DATE TYPE	Options from the pull-down menu: • Single date • Multiple dates within a year or series of non-consecutive years(generally does not apply to EPHT at this point in time) • Range of dates (i.e., data spanning a year or a multi-year period as is the case for almost all EPHT submissions) Most EPHT datasets are a range of dates unless there are gaps between years (e.g., 2000-2008)	Range of dates
SINGLE DATE	What is the one date covered by this entire dataset (e.g. November 1, 2008)	
MULTIPLE DATES		
Date 1	What is the first date covered by these data in this dataset (November 1, 2008)	
Date 2	November 15, 2008	
Date 3	November 30, 2008	
* RANGE OF DATES	This field is used to denote the ENTIRE date range that can be found in this dataset. So, if you are providing data	

	covering the years 2000 though 2008, the range would be indicated as: January 1, 2000 – December 31, 2008 even if data are only collected on a monthly or bi-annual basis.	
* CURRENTNESS	Since many datasets are updated with corrected information, provide the time frame for which the data are believed to be accurate (e.g. these data, spanning the years 2000-2007, are known to be complete and accurate as of November 1, 2009). The two choices from this pull-down menu are: **Time Period End Date** or **Publication Date**. Use **Time Period End Date** for original source data. Use **Publication Date** for data that is secondary or has undergone processing by the data steward.	Time Period End Date

D. GEOGRAPHIC AREA PAGE

MCT FIELD	Common Language Question Equivalent or Interpretive Statement	Sample Content
NOTE: Data are normally submitted for an entire state at this time, so these bounding coordinates capture the "four corners" of your state (ie. northernmost point, southernmost point, easternmost point, westernmost point). (see your GIS person for the values or accept the ESRI defaults used by the tool for each state). If this dataset only covers one county or other coverage area, please provide the "four corners" coordinates for that area. By default, when a user of the MCT logs into the system, coordinates for that state are automatically placed into the bounding coordinates boxes. These values need only be changed when bounding coordinates for a geographic area other than the entire state are desired. All values should be recorded in Decimal Degrees.		
* WEST COORDINATE	-95.774699	
* EAST COORDINATE	-89.098842	
* NORTH COORDINATE	40.6136399	
* SOUTH COORDINATE	35.9954790	

E. KEYWORDS PAGE

MCT FIELD	Common Language Question Equivalent Interpretive Statement	Sample Content
* THEME	The theme field denotes the overall topic of the data (e.g. health, environmental, ICD 9, SNOMED etc.) Note: CWG Teams are developing best themes for each content area – this will be incorporated into the tool as it becomes available. Guidance for selecting the theme will be provided by your EPHT representative	PH_DiseaseClassification_ICD-9CM

	If no specific theme is being referenced, select 'NONE' from the choices provided.	
* THEME KEYWORD	If you are using a specific theme, ascertain that the keyword chosen is a valid keyword for that theme via a keyword dictionary. If no specific theme is being used, determine what is/are the most logical word/words that one would use when searching for these data on your website or when searching the internet. *CWG Teams are developing best themes for each content area – this will be incorporated into the tool as it becomes available (via links).	Extrinsic asthma NOS; 49300
* THEME	The theme field denotes the overall topic of the data (e.g. health, environmental, ICD 9, SNOMED etc.) Note: CWG Teams are developing best themes for each content area – this will be incorporated into the tool as it becomes available. Guidance for selecting the theme will be provided by your EPHT representative If no specific theme is being referenced, select 'NONE' from the choices provided.	None
* THEME KEYWORD	If you are using a specific theme, ascertain that the keyword chosen is a valid keyword for that theme via a keyword dictionary. If no specific theme is being used, determine what is/are the most logical word/words that one would use when searching for these data on your website or when searching the internet.	Respiratory disease; disease of respiratory system; health effects; health outcome; asthma; respiratory disorders; asthma attack; obstructive lung disease; reactive airway disease; Asthma emergency department visits; Asthma emergency room visits

	*CWG Teams are developing best themes for each content area – this will be incorporated into the tool as it becomes available (via links).	
* THEME	The theme field denotes the overall topic of the data (e.g. health, environmental, ICD 9, SNOMED etc.) Note: CWG Teams are developing best themes for each content area – this will be incorporated into the tool as it becomes available. Guidance for selecting the theme will be provided by your EPHT representative If no specific theme is being referenced, select 'NONE' from the choices provided.	Topic Category
* THEME KEYWORD	If you are using a specific theme, ascertain that the keyword chosen is a valid keyword for that theme via a keyword dictionary. If no specific theme is being used, determine what is/are the most logical word/words that one would use when searching for these data on your website or when searching the internet. *CWG Teams are developing best themes for each content area – this will be incorporated into the tool as it becomes available (via links).	Health
* CATEGORY	Category has been included as a theme to describe broadly the category of your data, and you need to include category as one of your themes. There are currently three broad categories which describe EPHT data: Environmental Hazard, Environmental Exposure, or Health Effect. There are, however, more than three	Health Effects

	categories listed on the menu of the MCT. At the present time, creators of metadata should choose one of the following category identifiers to describe their dataset: Environment, Environmental Hazard, Exposure, Health, Health Effects. For datasets that could fall into multiple categories the question should be posed to the CWG Team responsible for data.	
* PLACES	By what geographic coding system are these data denoted (e.g. FIPS, GNIS, other)? The choices from the pull-down menu include: FIPS 5-2 (state), FIPS 6-4 (county), ISO 3166-1 (country), ISO 3166-2 (country subdivision). Current NCDMs use FIPS state or FIPS county.	FIPS 6-4 (county)
* PLACES KEYWORD	Note your state name, abbreviation, and FIPS code (e.g., Washington, WA, 53).	

F. SECURITY PAGE

MCT FIELD	Common Language Question Equivalent or Interpretive Statement	Sample Content
* SECURITY CLASSIFICATION SYSTEM	If your state or agency has a security classification system, list the formal name of this system here (this is a free text field). For publicly available data on the National portal	

	or a state portal, enter "None".	
* CLASSIFICATION	Are these data available to the public or are they restricted to those individuals who meet a certain "handling" criteria (e.g. choices from the pull-down menu are "Unclassified", "Restricted", "Top Secret", "Secret", "Confidential", "Sensitive" or "None"). "Unclassified" is the handling restriction for data available on the public portal.	Restricted
* SECURITY HANDLING DESCRIPTION	Describe, if applicable, the manner in which these data are to be stored once received by the requester (e.g. store dataset on secure server). Default entry into this free text field should be "None".	These data should be securely stored with restricted access.
* ACCESS CONSTRAINTS	If these are publicly available data, write 'NONE' in this space. If a certain criteria must be met before obtaining these data, describe the criteria in detail here (e.g. only researchers who have an approved protocol from an IRB may access these data). This field should include both legal (liability) information and non-legal access constraints. Information that pertains to dissemination of the data should be included in the Distribution tab. For all datasets available on the Public Portal this field should be "None".	These data should be accessible only to approved CDC staff and their contractors (contractors require access for purposes related to data management only).
* USE CONSTRAINTS	There should be a statement here stating that the	These data are for use in the analysis of daily,

	user must understand the metadata content before attempting to understand, interpret and use the data on the portal. Additional questions to be addressed in this field include: How should this dataset be used? And not used? Can it be linked to other datasets? Can these data be used for commercial purposes? Can these data be used to form a basis for additional health studies or some remediation actions? What are the constraints for data interpretation? This is also the field where messaging information should be included.	respiratory ED visits and air pollution only.

II.   DATA QUALITY TAB

MCT FIELD	Common Language Question Equivalent or Interpretive Statement	Sample Content

This section describes how the data were manipulated from their raw format to the current state in which they have been made available. Each step in the manipulation process should have its own processing date and description. Consider steps such as conversion from one program to another, data cleaning, data aggregating, geocoding, deduplication, addition of data fields, additional computations, etc.). You can list multiple process steps. Examples of process steps shown below. These fields will vary by grantee. These fields describe how and when you manipulated the data prior to uploading it to the Portal. Do not reference EPHT SharePoint site.		
* PROCESS DATE	Date that the processing described below was completed – NOT when the data were sent to the portal. (e.g., November 12, 2009)	
* PROCESS DESCRIPTION	Process that occurred. (e.g., Data downloaded from server and record deduplication was performed using XXXX criteria.)	
* LOGISTICAL CONSISTENCY REPORT	This field deals solely with GIS data (e.g., are polygons closed and neat line-simplified, are there any missing polygons, percentage of missing features, etc.). For most users, this section is left as "None".	None
* COMPLETENESS REPORT	This field can contain information as to what happened during processing to create the dataset into its available form. Address the following points: How many records lost to deduplication, geocoding errors, incomplete record information, geographic boundary changes, etc? Identify any data that has been omitted from the dataset that you might logically expect to see and the reason for exclusion. This field can also contain information as to what	These data include emergency department visits of state residents discharged from all acute care hospitals in each county reported. Visits to the Veterans Administration Hospital and other military facilities are not included.

	percentage of data are missing, the accuracy of the data, the version of a specific coding system that was used (e.g. ICD-9 vs. ICD-10 or data based on 2000 census population vs. intercensal values). You can list the test(s) used to check for data inconsistencies here. If this metadata record is being created for data available on a GRANTEE PORTAL, add information regarding suppression and aggregation	

III. ENTITY AND ATTRIBUTES TAB

MCT FIELD	Common Language Question Equivalent or Interpretive Statement	Sample Content
* OVERVIEW	This field can be thought of as a brief overview of the primary attribute fields included in the dataset. This field should contain two or three sentences that describe what one record represents.	This dataset contains the following fields: State= Patient state of residence; AgeGroup= age group of individuals; County= Patient county of residence (FIPS 6-4 code); EDVisitDay= Day of ED visit; EDVisitMonth= Month of ED visit; EDVisitYear= Year of ED visit; Ethnicity= Ethnicity of individuals (Optionally reported); ExclusionMenthod= Variable used to identify and exclude transfers from dataset; HealthOutcomeID= Health outcome (1=Asthma, 5=COPD, 6=Acute respiratory infections, 7=Pneumonia, 8=All other respiratory outcomes); Count= Number of events; OutOfStateExclusion= Exclusion of residents admitted to hospitals located in another state; Race= Race of individuals (Optionally reported); RaceEthnicityReported= Race or ethnicity fields reported in source data; Sex=Sex of patient admitted; TransferExclusion= Exclusion of transfers from dataset
* DETAILED CITATION	This field can be thought of as a user guide and data dictionary to each data element. It may contain a complete data dictionary listing, a data key or a link to a data dictionary and supporting documents. Do not link to EPHT SharePoint in this field.	This dataset contains the following fields: State= Patient state of residence; AgeGroup= age group of individuals; County= Patient county of residence (FIPS 6-4 code); EDVisitDay= Day of ED visit; EDVisitMonth= Month of ED visit; EDVisitYear= Year of ED visit; Ethnicity= Ethnicity of individuals (Optionally reported); ExclusionMenthod= Variable used to identify and exclude transfers from dataset; HealthOutcomeID= Health outcome (1=Asthma, 5=COPD, 6=Acute respiratory infections, 7=Pneumonia, 8=All other

MCT FIELD	Common Language Question Equivalent or Interpretive Statement	Sample Content
	respiratory outcomes); Count= Number of events; OutOfStateExclusion= Exclusion of residents admitted to hospitals located in another state; Race= Race of individuals (Optionally reported); RaceEthnicityReported= Race or ethnicity fields reported in source data; Sex=Sex of patient admitted; TransferExclusion= Exclusion of transfers from dataset	

IV.   DISTRIBUTION TAB

MCT FIELD	Common Language Question Equivalent or Interpretive Statement	Sample Content
RESOURCE DESCRIPTION	What is the full and complete name of the dataset as it is stored at your location, its file size (in megabytes) and fomat (e.g. XML). Information as to the file's compressed size can be entered here. This field explains the format of the file that is available to end users.

Note: There may appear to be some crossover between file information listed here and the information listed in the 'Native Dataset Environment' field (refers to the original dataset). However, this field contains information about the processed file that is being uploaded or being made available to an end-user. | |
| LIABILITY | This should contain a statement of liability against improper usage and endorsements of any kind. | |

	Generally speaking, each grantee should have a liability statement that indicates a department or data steward is not responsible if an individual misinterprets the presented information, links data in a manner other than that mentioned in the 'Use Constraints' field, somehow repackages and redistributes the available data, typographical errors are found, if data errors are found, etc.	
CUSTOM ORDER PROCESS	Describe the terms and conditions required to obtain these data. This is usually reserved for higher geographic resolution data that have access restrictions. If these data are fully available on the public or secure portal, a statement to this effect should be provided (e.g., The aggregated data are available on the CDC National EPHT Public Portal, for unsuppressed data please see data access policy document). Note: you can also link directly to appropriate documents or refer to the contact field, but do not put an email address in this field.	

V.   METADATA TAB

MCT FIELD	Common Language Question Equivalent or Interpretive Statement	Sample Content
* DATE CREATED	On what date was this metadata document finalized?	
* STANDARD NAME	EPHTN Tracking Network Profile Version 1.2 or FGDC Content Standard for Geospatial Metadata can be chosen from the pull-down menu.	EPHTN Tracking Network Profile Version 1.2

* ACCESS CONSTRAINTS	Are there any constraints to accessing or reviewing this metadata file? Note: Metadata provided to the Tracking Network is required to be open access. The default for this field is "None".	None
* USE CONSTRAINTS	Are there any use limitations for this metadata file? Note: Metadata provided to the Tracking Network is required to be open access. The default for this field is "None".	None

VI.   CONTACTS TAB

A.   MATRIX PAGE (Note: only MCT users see the Matrix Page. When this form is sent to data stewards you can highlight the fields they need to complete.

MCT FIELD	Common Language Question Equivalent or Interpretive Statement	Sample Content
* CONTACT 1 NAME	Corporate Contact Title of Data Steward	
* CONTACT 1 TYPE	Ex: NJDHSS Family Health Services	
CONTACT 2 NAME	Corporate Contact Title of Metadata Creator	
CONTACT 2 TYPE	Ex: NJ EPHT Metadata Coordinator	
CONTACT 3 NAME	Corporate Contact Title of Data Distributor	
CONTACT 3 TYPE	Distributor of the data	
CONTACT 4 NAME		
CONTACT 4 TYPE		

B.   Contact Fields (for originators, creator, distributor) – give us the contact info for:
   i. Metadata contact
      1. To whom the MCN/approval email goes to;
      2. Usually is a grantee contact

ii. Originator/Creator
  1. The contact to send questions to about the data;
  2. The contact that created the MD record;

Distributor contact
  3. The contact to obtain the data (e.g., distributor).

MCT FIELD	Common Language Question Equivalent or Interpretive Statement	Sample Content
* PERSON	Corporate Contact Title of Group responsible for providing dataset to EPHT program	
* ORGANIZATION		
* TITLE		
USERID		
HOURS	Hours during which they can be contacted	
INSTRUCTIONS	Special instructions (e.g. leave a voicemail and your call will be returned within X days)	
* PHONE NO. 1		
PHONE NO. 2		
* FAX		
* E-MAIL	Note: Corporate Contact E-mail Address ONLY. No Personal E-mail addresses.	
TDD/TTY		
* STREET ADDRESS		
* CITY		
STATE		
COUNTRY	U.S.A.	
* ZIP		