Death

Information on the fact and causes of death.

Contents

Overview

The Death table contains one record per person in the VDW Demographics table for whom you have some belief that the person may be dead. The unique primary key of the table is MRN. The MRN, Confidence, and Sources variables are the only variables required to be non-null for inclusion in the table. Although records with unknown death dates (DeathDt) have fewer uses than those with known dates, they should still be included in the file.

The Cause of Death table contains one record per person per known cause of death. The unique primary key is the combination of MRN, COD, Dx_CodeType, and CauseType. All variables on the table are required to be non-null. All MRN in the Cause of Death table should have a corresponding record in the Death table, but not all MRN in the Death table will necessarily have a record in Cause of Death.

When death sources provide conflicting information, site data managers should make local determinations as to which source to use, collating information when possible, and reflecting their confidence in the observation using the Confidence variable.

Mortality Specifications V3.1

Death Table

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.)
deathdt	num(4)	mmddyy10	Date of death	Any SAS date	
dtimpute	char(1)	n/a	How DeathDt was imputed	M Month of date imputed D Day of date imputed B Both month and day imputed N Not imputed	
source	char(1)	n/a	Source of the information on this record	S State Death files N National Death Index T Tumor data B	Please see the site/dataset implementation pages for details on any site-specific codes.

Variable name	Type (Length)	Preferred Format	Description	Valid Values		Comments
					Social Security Admin	
					Others as locally defined	
confidence	char(1)	n/a	How you rate the accuracy of the observation based on source, match, # of reporting sources, discrepancies, etc.	E F P	Excellent Fair Poor	

Cause of Death Table

Variable name	Type (Length)	Preferred Format	Description	Valid Values		Comments
MRN	Same as the MRN variable above.					
COD	char(6)	N/A	Cause of death	A valid ICD diagnosis code		Please include decimal points in your ICD codes
Dx_CodeType	Refer to the Dx_Codetype variable on the Diagnosis file for type, length, and value set					
causetype	char(1)	n/a	The type of cause of death.	I U C O	Immediate/Primary Underlying Contributory Other	
source	Same as the source variable above.					
confidence	Same as the confidence variable above.					

Notes

Some sources, particularly State Death files, may have additional approvals and restrictions on their use.

V3.1 implementation verification program

The program to used to verify V3.1 status is here.

V3.2 Change to replace SOURCE with SOURCE_LIST and add new values

Summary of the change

The following specification was accepted for implementation at the VIG 2013 Mid-Year meeting in Seattle. The change includes:

1. replace SOURCE variable with SOURCE_LIST variable
2. add allowed values E, P, M, O to SOURCE_LIST
3. disallow local values

DEATH V3.2

Name	Type (length)	Description	Valid Values	Comments
SOURCE_LIST	char(12)	A list of all sources of data that report this death regardless of whether there are discrepancies in other variables (e.g. death date)	A concatenated string of letters indicating all sources of death / cause of death for this record using the following letter abbreviations: B=Social Security Admin N=National Death Index S=State Death files T=Tumor data **E=Encounter data** **P=Patient data** **M=Membership data** **O=Other** Sources should be listed in descending order of reliability of the source.	All values should be UPCASE. The reliability of a source in indicating death status may be different from its reliability to provide other data, e.g., death date. Example: State data, Tumor data, and a local source all indicate the person's death. State data is considered most reliable, followed by Tumor, and the local source being considered least reliable. SOURCE_LIST = "STO"

CAUSE OF DEATH V3.2

Variable name	Type (Length)	Description	Valid Values	Comments

| Source_List | Same as the Source_List variable on the Death table. | The *value* of Source_List on a cause of death may be different than the value on the Death table. For example, if a person's death was reported from both State and Tumor but a cause of death of suicide was reported only by the State source, the Source_list variable on Death would take a value of "ST" where as the suicide cause of death record would take a value of "S". |

Implementation guidelines for new source values

Value	Description	Definition/implementation guideline
E	Encounter data	All encounter-based death data, typically from "expired" discharge disposition. The encounter can be any type of encounter that documents discharge disposition, for example, acute inpatient hospitalization or non-acute institutional stays, especially hospice. The data source can be encounter data from owned facilities, from allied outside facilities, or claims/referral data.
P	Patient data	Any patient status data that is not tied to a specific encounter, including EMR patient information tables, death notes and other provider notes, and manually reviewed patient status. May or may not include date of death information.
M	Membership data	Membership/enrollment data that includes mortality status, e.g., termination of enrollment due to death. Would typically come from an organization's health plan/insurance function. May or may not include date of death information.
O	Other local sources	All local values will be grouped/replaced by 'O'. Other local values are no longer allowed.

Note: Sites with locally-defined source values at a greater level of detail than the above can choose to retain the greater level of detail in locally-defined variables.

Implementation Date

The target implementation date is May 31, 2014.

Milestone QA program to check implementation

To be created.

Implementation Status by Site

To be created.

Original proposal
The original Proposal and Request for Comment page is here (*NOTE: the linked proposal is different from the final specification accepted for implementation*).

(DRAFT) Implementation Guidelines

A draft version of implementation guidelines are presented here. Please review and comment or provide feedback to the Mortality ad hoc workgroup (contacts: Tyler Ross and Daniel Ng).

Comments (14)

Butani, Amy Dec 22, 2009 02:54 PM
If a specific COD is listed as both CAUSETYPEs I and C... should both be kept?
>Ross, Tyler Feb 28, 2013 05:41 PM
>Good question that is not clear -- in our proposed overview, this is still a question (i.e. the comment asking whether CauseType should be part of the unique key)

parsons, william Jun 10, 2010 10:59 AM
I am unsure why the data are divided into two tables when 1 will suffice. All the information in the death table can be stored in the cause of death table - there will be as many rows per person as causes of death; only one row per person will have value 'I' in field 'Cause,' and fields DeathDt and DTImpute can be added.

Also, what is the actual name of the 'Cause of Death' table, i.e., without spaces?
>Ross, Tyler Feb 28, 2013 05:43 PM
>A person is only allowed one record in the death file but may have numerous causes of death. By normalizing the structure in to two tables, we make extraction easier and save on space.
>
>The actual table name is left to the discretion of the site. Please see the StdVars file for the agreed upon standard variable names however.

Hart, Gene Feb 15, 2012 05:46 PM
On a VOC call on Feb 14, 2012 there seemed to be consensus on using only the confidence values E and F on a multisite program trying to find deaths. Some sites have many poor match records and others have very few.
>Ross, Tyler Feb 28, 2013 05:45 PM
>Happy Valentine's Day!
>
>The Workgroup realizes that there is a lot of inconsistency in how sites are coding the confidence variable. We are working to provide some more guidance on this topic, but it's quite challenging to make rules that accommodate all scenarios since there are so many different circumstances. Hopefully we will have better guidance on this in the near future.

Ng, Daniel Mar 02, 2012 11:06 PM
Our death data goes way back and we have ICD8 and older. Since the spec only mentions ICD10 and ICD9, we are going to code ICD8 = C, ICD7 = D, etc.

This is outside spec, but IMO it is better to do this than have missing values.

Liu, Julie Feb 28, 2013 06:29 PM
Death table
* source variable
  1. do you have a code for other sources that are not defined in the existing code values? Or should we leave it blank for other sources?
  2. We have situation where one MRN could be identified in different source files (e.g., SSA and state death files), how should we populate the source variable?
>Ross, Tyler Feb 28, 2013 07:50 PM

Thanks Julie for the questions. The SOURCE variable as is currently defined allows sites to add values for their sources beyond Tumor, NDI, and State. The proposal to add "B" for Social Security does not change this. So for example, if you use your EMR as a source, you can choose any value (other than T, N, S, and the proposed B) to indicate that at your site. The variable is still expected to be populated however.

Regarding your second question, the specifications do not say what to do in this circumstance and we know that people are handling this differently. We did talk about this issue in a recent workgroup meeting, but have not come up with a solution or proposal at this time. Right now, you're forced to choose one value. Our proposals neither cause this problem nor fix the problem.

Schmidt, Mark Feb 28, 2013 06:36 PM

For death "source", the proposed mod still forces a single choice, without clear guidance on hierarchy, eg does NDI trump state?

An alternative would be to have separate flags for NDI, State, SSA, Tumor, Local/Other...

For cause of death, similar issues. Do we pick cause data from the "best" source? What if Tumor has a different code than appears on NDI? May just need additional guidance in the dictionary on these questions.

Otherwise agree with the changes.

Ross, Tyler Feb 28, 2013 08:02 PM

Thanks Mark -- this comment is similar to Julie's question #2 above. We recognize this as a problem and discussed it on the Workgroup. A series of flags is an intriguing solution that we will talk about and could help us in our quest to help figure out CONFIDENCE as well. As currently proposed, we certainly don't fix this problem, but I don't think we're making it worse. Most of the WG time has been spent these months talking about Confidence and Source, but we still have a lot to figure out to fix them. We wanted to move forward if possible with these proposed changes while continuing to work on confidence and source.

Krajenta, Rick Mar 01, 2013 07:43 AM

I support A and B as is and think they will solve some issues with our data use restrictions. I think option C should actually be expanded. Possibly survey sites death sources to group and standardize the ones that cross many sites. Many good sources of death would include discharge status for patients that expire in the hospital. Some studies require death certificte review and those findings could be fed back to the SDM to incorporate into the files.

Ross, Tyler Mar 04, 2013 01:51 PM

The workgroup talked at length about source. We did a metadata survey to find out what sources people had. Summarized results are here:
https://appliedresearch.can[...]aSurvey_2012Dec_Summary.xls

In our conversations that ensued, we struggled to come up with additional standard values since there's a lot of gray between what is a claim, a discharge summary, an EMR source etc. We then questioned the value of the source variable in total and in particular around these more gray areas and particularly since the spec allows only one source to be named. Plucking out social security admin as source seemed straightforward, but reconciling other sources didn't seem worth the squeeze at this point. We recognize adding SS Admin doesn't resolve the underlying issues around source, but still think it's worth adding it as an incremental improvement.

Saylor, Gwyn Mar 01, 2013 05:59 PM

I think that the unique primary key for the COD table is MRN, COD, DX_Codetype. (The same DX_Codetype value shouldn't appear for an individual more than once.) So, I wouldn't include CauseType. Perhaps the workgroup could establish a hierarchy of CauseTypes that win out if a DX_Codetype is duplicated. (rock, paper, sizzors?) With this thought in mind we need to note that there should be no more than one underlying cause of death (Causetype=U) for any individual in the table. I think listing some basic rules out would be very helpful to sites.

How are the records that have no date of death being used? I can see that this information is useful for not including them on a mailing list. What else? Just curious.

Change B is fine. Some additional notes explaining what one can expect to see from "official sources" like State and NDI would be nice. ICD-9 cause of death codes are used for 1998 and earlier deaths. (Except those really early deaths that have ICD-7 and ICD-8!) ICD-10 is used starting in 1999. While the ICD-9-CM codes used for clinical purposes have 5 digits (not counting the .) the ICD-9 cause of death codes have 4 digits maximum (as with ICD-10). ICD-9 cause of death match up with ICD-9-CM codes through the 4th digit. ICD-9 cause of death codes all start with a digit.