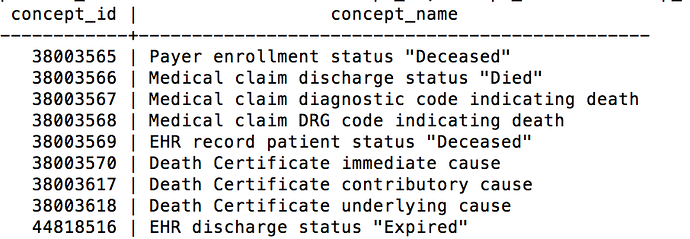
Contents

[Death type concept ids hierarchy - E. Burrows 1](#_Toc444771516)

[Death Table -One Record Per Person Constraint – E. Burrows 2](#_Toc444771517)

# Death type concept ids hierarchy - E. Burrows

http://forums.ohdsi.org/t/death-table-addition-of-death-type-concepts/508

[[](http://forums.ohdsi.org/uploads/default/109/cacc0444c21e9efb.png)](http://forums.ohdsi.org/uploads/default/109/cacc0444c21e9efb.png" \o "Screen Shot 2015-04-23 at 2.09.00 PM.png)

[Screen Shot 2015-04-23 at 2.09.00 PM.png906x314 68.5 KB](http://forums.ohdsi.org/uploads/default/109/cacc0444c21e9efb.png" \o "Screen Shot 2015-04-23 at 2.09.00 PM.png)

We are finding that we need the hierarchy represented for Death Certificate also present for the EHR record type to describe the cause of death for records.

Example:

EHR Record immediate cause  
EHR Record contributory cause  
EHR Record underlying cause

Would it be possible to add these as valid concepts for the Death Type Concept id?

# Death Table -One Record Per Person Constraint – E. Burrows

**burrowseApr '15**

The OMOP v5 documentation for the Death table indicates that a person can only have up to one record.

In our network, we are finding that this conflicts for two reasons:

In our source data data a patient can have **more than one cause of death recorded**. This means for cause of death, several ICD9 codes may be associated with one patient which is indicative of more than one row being represented in the death table.

We expect external source data (insurance vendors) to assert cause of death, which would mean this could also be numerous rows based on the amount of ICD9 codes. While the ICD9 codes could be the same we would still see more than one row because we would **utilize a different Death Type concept id to indicate that the originating source of the data is different from the ICD9 codes coming from the EHR**.

Are other networks experiencing this issue?

**jenniferduryeaJen DuryeaApr '15**

Thanks for bringing this up @burrowse! We have the same issue with analyzing SEER Medicare data. Patients could have 3 death records per patient - 1) the death date recorded by Medicare Enrollment; 2) the death date recorded by the SEER registry; and 3) ICD9 codes within the patient's claims. For this purpose, we would love to have multiple death records per patient.

We have created a lite ETL specification for SEER Medicare data and have decided to use a hierarchy to pick one death record for each patient. This is not ideal. We would love to report all sources of death dates and give the analyst the freedom to choose which source to use for analysis.

In the next couple of weeks, I believe we are going to start to create a proper ETL spec for SEER Medicare in the OHDSI community and I know this issue is going to come up again.

**Christian\_ReichApr '15**

Friends:

As usual, I will take the perspective of the analyst. The analyst doesn't want to deal with several deaths or figure out which one is the real one. You need to decide during the ETL when the patient most likely died.

As far as the cause of death goes: This is a condition. So, the question is, should we record those conditions at the same day as the death? Or is there a use case where we explicitly need the cause?

**Mark\_DaneseApr '15**

The most obvious use case for cause of death is for cardiovascular death. I just completed an analysis of CPRD data measuring time to non-fatal cardiovascular event, cardiovascular death, or non-cardiovascular death. Also, this is very common in oncology -- cancer death vs. all-cause death s used in relative survival. So, we definitely need to have cause of death in the CDM. Cause of death is an ICD9 or ICD10 code in the data sources with which I am familiar, so it might be possible to link death to a condition. Maybe just need a way to indicate condition type (cause of death) for the condition occurrence table, and put it in there.

As for multiple sources of data, I agree that it can be handled as part of the ETL. We typically pick the earliest of the available dates. The challenging in putting multiple death dates in the CDM is that we then have to identify the source of the death date so the investigator can make a decision. That seems like it is not worthwhile. If there are substantial differences among sources, it sounds like a data quality issue, and not a CDM storage issue.