Activity Takeaways:

1. The results of the activity are most likely very different.
   1. If only 5/10 are identified among the 4 individuals in the group, then we only agree on the cause 50% of the time.

The CDC enters these in by hand. They use the ICD-10 manual (international classification of diseases). This requires at least 9 individuals (think about it) to enter in information from various entities to get results. This is a minimum estimate. The true number of individuals entering data into these spreadsheets is likely much higher. Each step of the process introduces new room for error.

ICD-10 codes offer a solution to codifying causes of death. They can be grouped into larger categories (as shown in the ICD-10 Chapters in dashboard). However, they are vague.

* What does ‘diseases of the circulatory system’ mean to PH professionals?
* How would PH professionals use this information to design interventions?

To get specifics, you need a breakdown of what goes into the category. Where does this come from? The data cannot be both classified into ICD-10 (by someone else’s hands) and broken down by the specific disease of interest to us. But first, let’s think.

* Who has the final say?
* Who decides what the ‘correct’ classification is?

This depends. The correct classification is technically ICD-10 according to international standards. However, it is not precise enough to give PH professionals the granularity (or quality) of information they need to mobilize direct action against a disease. This is especially true at the local level where getting specifics requires collaboration with state and federal entities just to gather the information.

Alternatively, the people doing the classification by hand have the most say in whether it is correctly classified or not. They enter it in directly. The system (ICD-10) can exist with rigid standards but are individuals who are classifying these causes going to do so the same way for every possible death in Rock County (and in the U.S?)?

Consider the following:

We identify individuals that died of cancer and cancer related diseases. This answers the question, how many individuals died of cancer or cancer related diseases? Now, we have an insight into the number of individuals dying from cancer and cancer related causes. This is great! The next question you may ask then is, well, what kind of cancer did the individuals die from?

The answers to these two questions require a different kind of approach to solve. The first one can be solved using the ICD-10 method. In this approach, our data already contains a category of “*Neoplasms”* which is a larger group of cancers that includes any abnormal growth of tissue in or on the body. We can then find totals of this category for any location in the dataset, including at the county level. Although, this approach has its flaws.

Shortened TLDR Ex: individual has a carcinoma on the skin. It spreads to the rest of the body, specifically their brain and blood. Which cancer caused their death?

For example, an individual has a carcinoma on the skin. This individual does not seek treatment and it spreads to the rest their body. In this case, the malignant growth in epithelial tissue of the skin spreads. Before they know it, the cancer reaches the individual’s bloodstream at which point they are diagnosed with leukemia. With the cancer in the blood, it is not long before it spreads further to the individual’s brain. Shortly thereafter, the individual dies.

* Which cancer caused their death?

A good question is, how often does cancer type [x] spread to other parts of the body? And, what percentage of those whose cancer spreads, die from [x]? This might require data from hospitals and clinics in quantities large enough to be representative, truthful, and build confidence in the results. By the end of the analysis of these questions, we would have actionable insights that could led to prevention measures that are also applicable to a wider audience.

Based on this, what is the best path forward?

The best first step is to acknowledge what we do not know. Then, we can take steps to learn. In this mortality data example, we need clarity of the problem before we can solve it. To help with this, we should consider what questions we are trying to answer and what you might do with the results. In the case of death certificate review or developing a mortality surveillance system, we have a couple options depending on the questions we seek to answer.

Holistically – if we want to compare to other counties (in WI and beyond) we need ICD-10 which we can get from the CDC. We should acknowledge the limitations of the data and adjust our questions accordingly. An important disclaimer with the results of the question is that this dataset may not be actionable for public health professionals, but it can be informative.

Rock Only – if we want to know exactly how many people died of [x] in rock then we would need to use our own data sets and identify individuals’ causes of death to answer that question. The disclaimer to using this kind of method is that it would not be comparable across counties in WI.

So, what are your questions?