Article from

Predictive Analytics and Futurism

July 2016
Issue 13

Seasonality of Mortality

By Kyle Nobbe

Figure 1: Scatterplot of F&P Excess Deaths by Month

WARNING: Brace yourself for freshman year statistics. Modeling techniques such as neural networks, generalized linear models or random forests will not be found in this article.

Seasonality profoundly impacts mortality and that impact is widely felt throughout the life insurance industry. A variety of demographic, socio-economic and geographic factors influence the degree and direction of seasonal mortality. Tim Rozar explored these factors in great detail in 2012,[1] but the exceptional nature of excess mortality in the early part of 2015 begged an analysis to be conducted to further deepen our understanding.

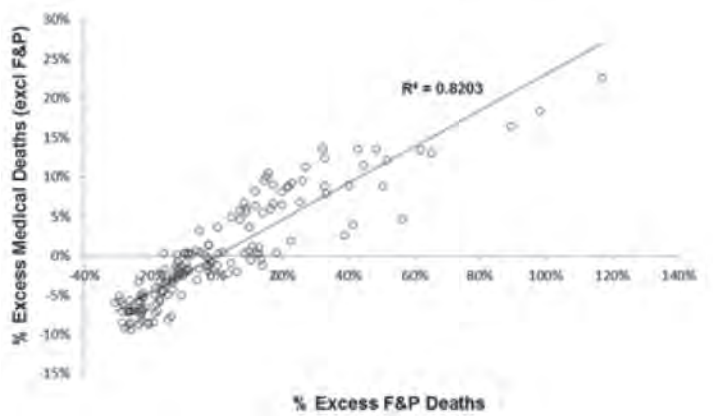Key outcomes from this research demonstrated:

1. the strong correlation between influenza and pneumonia (F&P) related deaths to other causes of death, and

2. how F&P deaths are strong predictors of total population mortality using a simple linear regression.

F&P CORRELATION WITH OTHER CAUSES

Medical research demonstrates how individuals with pre-existing conditions[2] (such as diabetes, heart disease, etc.) are more likely to have serious complications when diagnosed with the flu or pneumonia. Knowing this fact, we wanted to test how excess F&P deaths in a given period correlated with other causes of death. To do this, ICD-10 codes were grouped into specific causes of death (such as F&P, unnatural, cancer, etc.) using the Centers for Disease Control and Prevention's Multiple Cause of Death Data. Figure 1 demonstrates how excess F&P deaths in a month correspond to an excess of all other medical deaths.

Clearly, a strong connection is evident between months with elevated F&P mortality and elevated mortality of other medical causes. Note that unnatural causes of deaths were excluded from the chart above due to their reverse seasonal nature of mortality. For a more granular look at which causes of death correlate the highest with excess F&P mortality, see Table 1 (below).

Respiratory, cardiovascular, neurological, diabetes and other medical causes of death total about 70 percent of total U.S. population deaths, so a significant percentage of U.S. mortality is highly correlated to F&P mortality (which is about 2 to 3 percent of all U.S. deaths depending on the year).

PREDICTING TOTAL POPULATION MORTALITY

With the knowledge that F&P mortality correlates with the majority of other causes of deaths, we built a linear regression to predict total population mortality. We leveraged the CDC's National Center for Health Statistics (NCHS) data for this analysis

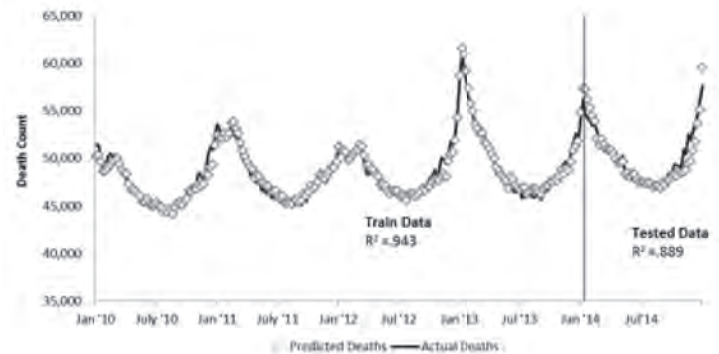**Figure 2: Prediction of Total Population Mortality**

Table 1

CORRELATION BETWEEN EXCESS F&P MORTALITY AND OTHER CAUSES OF DEATHS						
Respiratory (excl F&P)	Cardio	Other Medical	Neuro	Diabetes	Cancer	Unnatural
0.95	0.90	0.87	0.87	0.87	0.37	-0.34

to incorporate more recent years of data. We fit the regression to weekly death counts between October 2009 and December 2013, which produced the following regression formula:

*Predicted Deaths = 31,648.07 + 4.077 * (Weeks since 10/1/2009) + 14.558 * (F&P Deaths)*

We tested the regression line on all 2014 deaths. Although the goodness of fit declined on the test data, the prediction was still strong. Figure 2 demonstrates the predictive power of using F&P deaths to predict to population mortality. I suspect the predictive power would only increase if cancer and unnatural deaths were stripped out.

IMPLICATIONS AND CHALLENGES

Monitoring flu season is widespread across numerous industries, and often the data and results are publically available. This includes hospitalizations, mortality, social media activity and search engine analytics. Insurance companies have an opportunity to better understand their business and stay ahead of potential epidemics thanks to the robustness of this data.

On the downside, there is still a lot of work to do. For starters, bridging the gap between population mortality and insured mortality can present challenges such as reporting lag, age standardization and underwriting wear-off, just to name a few. Additionally, flu and pneumonia forecasting is still in its infancy. Consider Columbia University's Prediction of Infectious Diseases model, which won the CDC's Predict the Influenza Season Challenge. The lead Columbia researcher commented, "Much work remains to improve the science of flu forecasting."

The life insurance industry has barely scratched the surface of this topic. I have no doubt advancements will continue to be made and it is imperative the industry be at the forefront of this important topic.

DATA SOURCES

- Centers for Disease Control and Prevention (CDC). Multiple Cause of Death 1999-2014 on the CDC Wide-ranging On-Line Data for Epidemiologic Research (WONDER) Online Database, released 2015. Data are compiled from data provided by the 57 vital statistics jurisdictions through the Vital Statistics Cooperative Program, *http://wonder.cdc.gov/mcd.html.*

- National Center for Health Statistics (NCHS). Weekly Mortality Surveillance Data, *http://www.cdc.gov/flu/weekly/nchs.htm.*

Kyle Nobbe, FSA, MAAA, is assistant actuary, global research and development, at RGA Reinsurance Company in Chesterfield, Mo. He can be reached at k*nobbe@rgare.com*

REFERENCES

[1] Rozar, Timothy. RGA. [Online] May 2012. *https://www.rgare.com/knowledgecenter/Documents/AW_SOA2012.pdf.*

[2] CDC. Influenza - People at High Risk. [Online] *http://www.cdc.gov/flu/about/disease/high_risk.htm.*

[3] Columbia University. Columbia Prediction of Infectious Diseases. [Online] *http://cpid.iri.columbia.edu/.*

[4] CDC. Flu News & Spotlights. [Online] *http://www.cdc.gov/flu/news/predict-flu-challenge-winner.htm.*