How Did COVID-19 Change the Dependence Structure Across Major Causes of Death?

NOVEMBER| 2023

LIVING TO 100
SOA INTERNATIONAL SYMPOSIUM

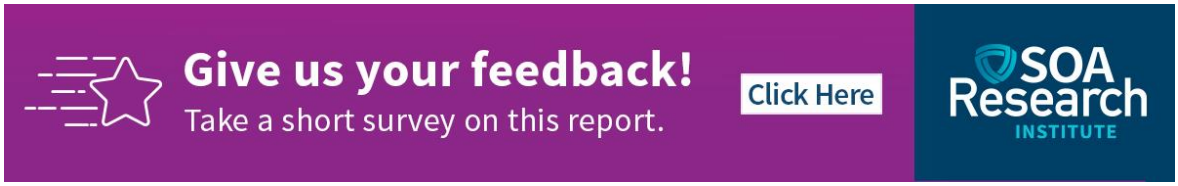
SOA Research INSTITUTE

How Did COVID-19 Change the Dependence Structure Across Major Causes of Death?

AUTHOR Han Li, Ph.D
Associate Professor, Centre for Actuarial Studies
Faculty of Business and Economics
The University of Melbourne, Victoria, Australia

SPONSORS Mortality and Longevity Strategic Research Program Steering Committee

Aging and Retirement Strategic Research Program Steering Committee



Caveat and Disclaimer
The opinions expressed and conclusions reached by the authors are their own and do not represent any official position or opinion of the Society of Actuaries Research Institute, Society of Actuaries, or its members. The Society of Actuaries Research Institute makes no representation or warranty to the accuracy of the information.

Copyright © 2023 by the Society of Actuaries Research Institute. All rights reserved.

How did COVID-19 change the dependence structure across major causes of death?

Abstract

Two years into the pandemic, COVID-19 continues to be one of the biggest killers with a current global toll over 6 million, exceeding that of cancer or malaria. At the same time, deaths from certain causes (such as dementia and diabetes) have increased while deaths from others (such as respiratory diseases) have decreased. To anticipate the impact of COVID-19 on future life expectancy, we need to carefully analyze the dependence structure across major causes of death, both pre and post pandemic. This paper quantifies and visualizes the joint dynamics across six major causes of death including *circulatory*, *dementia*, *diabetes*, *influenza*, *malignant*, and (chronic) *respiratory*, via the cutting-edge vine copula-based modeling. The empirical results are based on US weekly mortality data from January 2015 to March 2022. We find that the dependence structure across major causes of death does not change before and after the pandemic. On the other hand, we observe an increase in the overall dependence, which could help explain the extreme mortality levels experienced during the pandemic.

Keywords: COVID-19 pandemic; Dependence modeling; Cause-of-death mortality; Vine copulas.

1 Introduction

Cause-specific mortality modeling can provide valuable insights on how future mortality is likely to evolve (Alai *et al.*, 2018; Li *et al.*, 2019; Diaconu *et al.*, 2020). Under the current pandemic situation, without knowing how different causes interact with each other, it is challenging to understand the overall impact of COVID-19 on mortality, and to conduct scenario-based prediction on life expectancy.

It has been widely acknowledged that causes of death are not independent of each other, nor should they be considered in isolation. As a result of the pandemic, we have observed an increase in death counts from causes such as diabetes and circulatory diseases potentially attributed to common risk factors. On the other hand, lockdown restrictions and border closures reduced the spread of respiratory diseases and thus have led to lower number of deaths from influenza and other respiratory-related conditions.

In the paper, we quantify and visualize the dependence structure across 5 major causes of deaths, namely *circulatory*, *dementia*, *diabetes*, *influenza*, *malignant*, and *(chronic) respiratory*. Given the fact that these causes reacted to the COVID-19 pandemic differently and the magnitude, direction, and tail behaviour of dependence across the causes are also different, a flexible modeling framework is needed. We propose a vine copula-based approach to model the dependence across different causes of death both pre and post pandemic, while controlling the impact of COVID-19. We find that COVID-19 death count is a significant predictor for deaths from *circulatory*, *dementia*, *diabetes* and *influenza*. The estimated correlation is positive for *circulatory*, *dementia*, and *diabetes*, and negative for *influenza*, which is consistent with our observations.

US weekly death count data published by the Centers for Disease Control and Prevention (CDC) via its "COVID-19 Death Data and Resources" is considered for the empirical studies. We collect cause-specific death data for the period Jan 2015 - Mar 2022 including those from COVID-19. As a first step we remove seasonality and series correlations in the death counts data using time series models. The vine copula model is then fitted to the pre-pandemic data (2015–2019) as well as the post-pandemic data (2020–2022), and we compare the differences between the two selected vine structures.

We find no change in the overall dependence structure before and after the pandemic: *circulatory* remains a center-piece and a key "hub" connecting other causes of death. In the pre-pandemic era, it exhibits relatively strong positive association with all other causes expect for *influenza*. However, the copula model based on the pandemic experience shows a noticeable increase in the dependence level between *circulatory* and *dementia*, *diabetes*, and *influenza*. On the other hand, there has been a decrease in the pairwise dependence between *circulatory* and *malignant*. Finally, the dependence between *circulatory* and *respiratory* remains at a similar level before and after the pandemic.

The rest of the paper is organized as follows. In Section 2, we describe and visualize the US cause-specific death data during Jan 2015 - Mar 2022. We model these death counts using a time series approach in Section 3. Section 4 introduces the vine copula model for the dependence structure across different causes of death. We then analyze and discuss results from the empirical studies in Section 5. Finally, Section 6 concludes.

2 Data

Our empirical studies are conducted using the US cause-specific death counts from January 2015 to March 2022 at a weekly frequency. We collect death data from the CDC National Center for Health Statistics (NCHS) via the "COVID-19 Death Data and Resources"[1]. Weekly death counts from circulatory diseases (*circulatory*), Alzheimer disease and dementia (*dementia*), diabetes (*diabetes*), influenza and pneumonia (*influenza*), malignant neoplasms (*malignant*), and chronic lower respiratory diseases (*respiratory*) are considered. These causes are selected based on the CDC analyses of most common co-morbid conditions reported on death certificates where COVID-19 was listed. The International Classification of Diseases 10th Revision (ICD-10) is used by the NCHS. We provide detailed codifications of the six aforementioned causes of death in Table 1.

Cause of death	**ICD-10 codes**
dementia	G30, G31, F01, F03
circulatory	I00–I99
diabetes	E10–E14
influenza	J09–J18
malignant	C00–C97
respiratory	J40–J47

Table 1: Codification of the 6 major causes of death

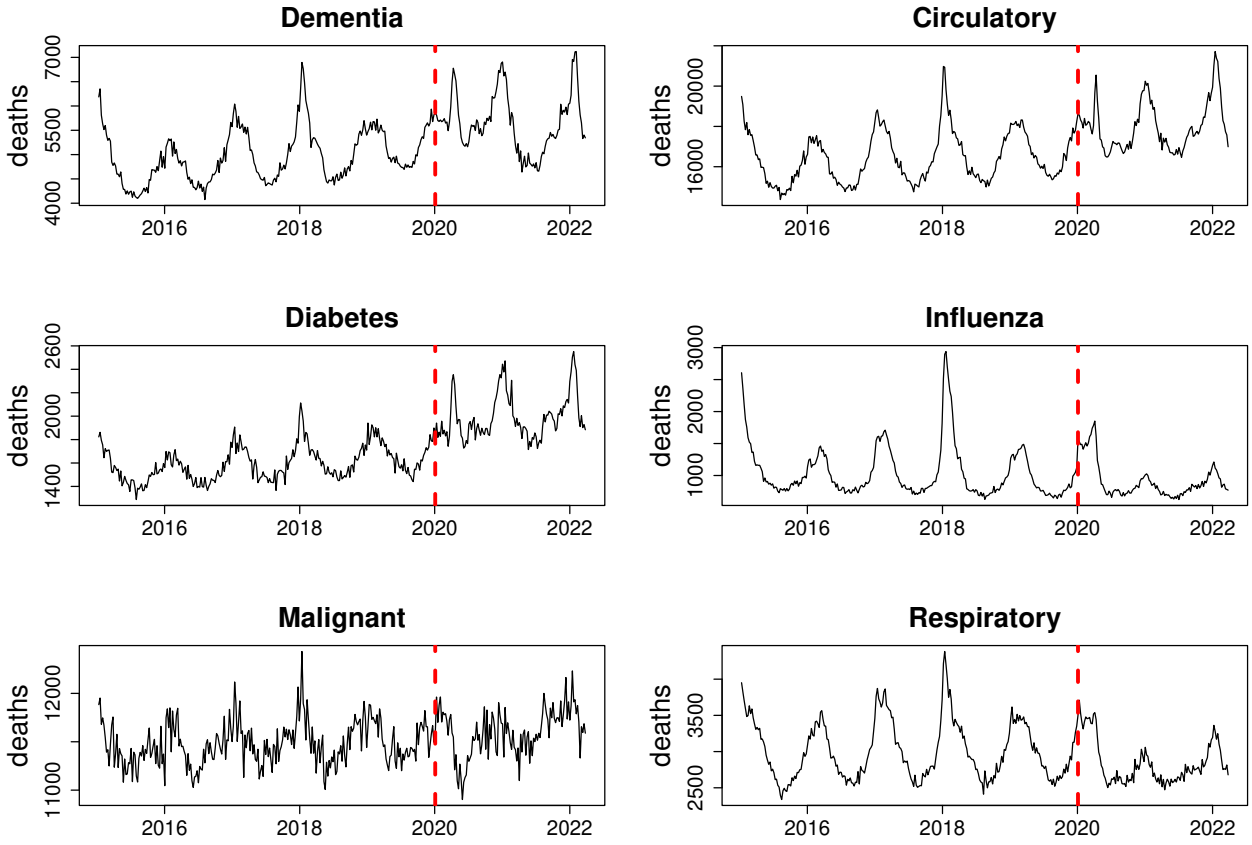


Figure 1: Weekly death counts from 6 selected causes

[1]The data can be downloaded at https://www.cdc.gov/nchs/nvss/vsrr/covid19/excess_deaths.htm.

In Figure 1, we plot the US weekly death counts during 2015–2022. It can be seen that mortality rates from the six causes of death differ considerably in sizes and seasonal patterns. Unsurprisingly, *circulatory* is the number one leading cause-of-death, followed by *malignant* and *dementia*. For all six causes, we see that peaks in death counts usually happen in winter months, particularly for *influenza* and *respiratory*. On the other hand, *malignant* shows the weakest seasonality of mortality compared to other causes. We plot a vertical line at Week 1, 2020 to illustrate the potential change in mortality experience before and after the pandemic. There has been a spike in death counts around March/April 2020 from *dementia*, *circulatory*, and *diabetes*. The death counts remain relatively high for these causes especially during major COVID waves. It is also worth noting that, since mid-2020, deaths from *influenza* and *respiratory* have been very low in the US, likely attributed to the social-distancing and lock-down measures in placed.

To better visualize the underlying trends in the death data, we plot the de-seasonalized death counts in Figure 2.[2] We can see that before the pandemic, there has been slight upward trends in death counts from *dementia*, *circulatory*, and *diabetes*. Since the number of deaths is determined by both population exposure and mortality rates, with population growth and rapid aging over the last decade, death counts can be increasing over time even with decreasing/unchanged mortality rates in the US. After 2020, there has been a apparent structure break in the death count from all causes except for *malignant*.

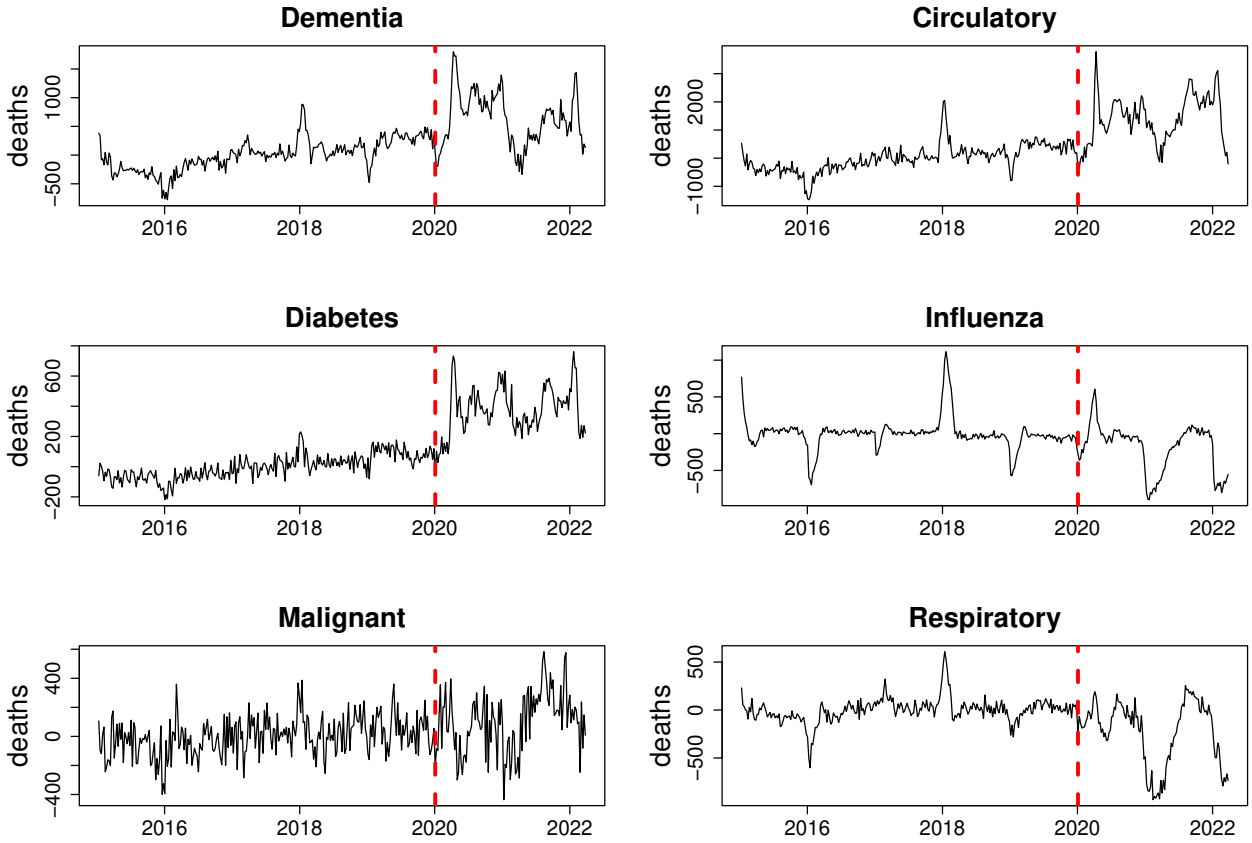


Figure 2: De-seasonalized weekly death counts from 6 selected causes

[2]The de-seasonalization was performed by subtracting weekly average death counts during 2015–2019 for each cause of death.

3 Time series models

To capture the serial correlations, in this section we adopt ARIMA time series models to fit the de-seasonalized death data described in Section 2. First, we consider the pre-pandemic data only and use the selected ARIMA model to predict cause-specific death counts assuming there had been no pandemic. We then propose a new ARIMA model incorporating COVID-19 deaths as an exogeneous regressor in the model. The obvious structural breaks shown in Figures 1 and 2 indicate that, to fit the full data from 2015 to 2022, the standard ARIMA models are no longer adequate and we need to control for the impact of COVID deaths on other related deaths.

The general form of ARIMA models can be expressed as follows:

$$(1 - \phi_1 B - \ldots - \phi_p B^p)(1 - B)^d y_t = c + (1 + \theta_1 B + \ldots + \theta_q B^q)\epsilon_t, \tag{1}$$

where B is the backward operator such that $(1 - B)y_t = y_t - y_{t-1}$. p, d, and q denote the order of the AR model, the order of differencing, and the order of the MA model, respectively. ϵ_t is the error term and $\{c, \phi, \theta\}$ are coefficients to be estimated.

The Akaike information criterion (AIC) is used to select the optimal ARIMA model for each de-seasonalized death series. Based on these models, we project death count levels for Jan 2020–Mar 2022, assuming no COVID and the trend in 2015–2019 continues into the future. We plot the cause-specific forecasts in Figure 3, where the blue lines represent the predictions based on the 2015–2019 experience. Note that the forecasts are presented on the original scale after adding the seasonal components for each week.

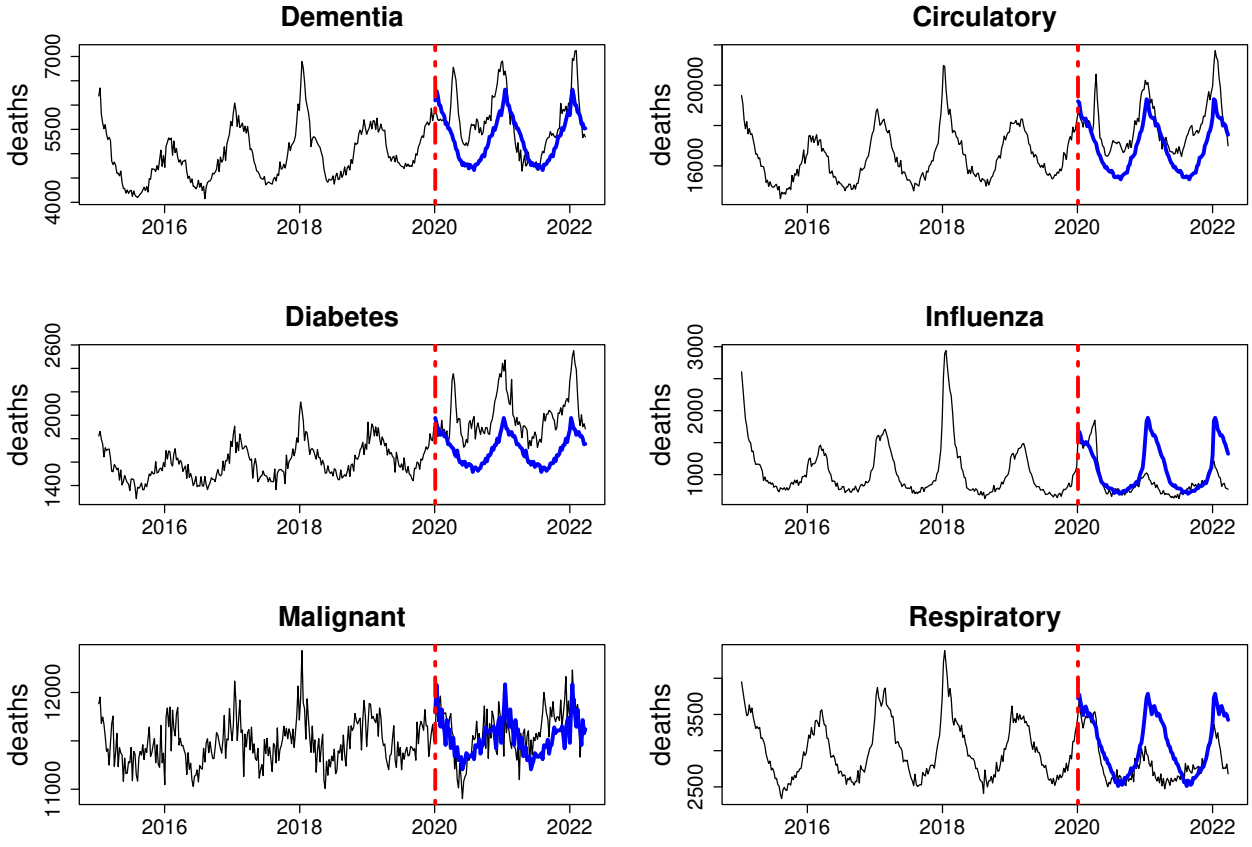


Figure 3: Predicted weekly death counts from 6 selected causes from Jan 2020 to Mar 2022

From Figure 3, again we see that mortality experience from all 6 causes have been affected by the pandemic to different extents. Excess deaths are observed from *dementia*, *circulatory*, and *diabetes*. Regarding the increase in *diabetes* deaths, research has found diabetes to be one of the most common comorbidities present in patients with COVID-19. This is likely to be a contributing factor that has led to higher deaths counts from diabetes during the pandemic. On the other hand, lower-than-expected number of deaths is observed from *influenza* and *respiratory*. It should be noted that during the COVID-19 pandemic, the fall in respiratory-related deaths has not only been observed in the US, but also been observed in other countries such as Australia. COVID may have acted as a competing cause for respiratory diseases, and this is a potential explanation behind this phenomenon. Consistent with the previous discussions, *malignant* is the least affected cause among the six causes, where forecasts based on 2015–2019 experience are quite close to the observed values in 2020–2022.

To model the full dataset from 2015 to 2022, we propose a new ARIMA model as follows:

$$(1 - \phi_1 B - \ldots - \phi_p B^p)(1 - B)^d y_t = c + \beta \times d_{\text{COVID}} + (1 + \theta_1 B + \ldots + \theta_q B^q)\epsilon_t, \qquad (2)$$

where d_{COVID} is the number of COVID deaths and β is the associated coefficient.
We present the estimated values of β in Table 2, as well as the standardized errors and their corresponding p-values. Based on the table, it can be seen that *dementia*, *circulatory*, and *diabetes* are positively correlated with COVID deaths, while *influenza* and *respiratory* are negatively correlated with COVID deaths. The impact of COVID on *malignant* deaths is very small and not significant with a p-value of 0.501.

	Estimate	Std. Error	p-value
dementia	0.054	0.006	0.000
circulatory	0.113	0.013	0.000
diabetes	0.021	0.002	0.000
influenza	-0.016	0.005	0.001
malignant	-0.002	0.004	0.501
respiratory	-0.020	0.008	0.010

Table 2: Estimated β values for each cause-of-death

For each selected ARIMA model, we obtain an approximate *i.i.d.* residual time series ϵ_t, which will then be used to fit the vine copulas to model the cross-sectional dependence. We conduct the Ljung–Box test on ϵ_t to detect any auto-correlation in the residuals. For all causes of death, the p-value of the test is found to be greater than 0.05, which means that the null hypothesis of zero auto-correlation cannot be rejected at 5% level of significance.

Before fitting the vine copulas, we adopt a kernel smoothing approach to transform ϵ_t from the z-scale to the uniform interval $[0, 1]$ giving rise to copula data.[3] Kernel density estimation is one of the most widely used nonparametric methods which does not require any assumption on the distribution of the data. Let $\{x_1, x_2, \ldots, x_n\}$ be a sample of *i.i.d.* observations from a univariate distribution $F_X(x)$, the conventional kernel density estimator is defined as

$$\hat{f}_X(x) = \frac{1}{nh} \sum_{i=1}^{n} K(\frac{x - x_i}{h}), \qquad (3)$$

[3]The kernel density estimation is implemented by R package `kde1d` (Nagler *et al.*, 2019).

where K is the kernel function, and h is the bandwidth parameter which controls the amount of smoothness imposed in the estimation.

4 A vine copula-based approach

The key idea behind vine copulas is to construct a flexible dependence structure across variables using pair-copulas as bivariate building blocks (Aas *et al.*, 2009; Joe, 2014). Unlike multivariate Archimedean copulas, vine copulas allow for a different direction and magnitude of dependence for each pair of variables, which is especially suitable for cause-of-death modeling. Unlike Gaussian and t copulas, vine copulas can assign a different type of dependence in each pair of causes and allow for asymmetric tail dependence between causes. Under the vine copula framework, we aim to quantify and visualize the interaction and joint dynamics across different causes, both before and after the pandemic.

We illustrate the construction of vine copulas using a simplified example with three variables. Let $\boldsymbol{X} = \{X_A, X_B, X_C\}$ be a random vector of dimension 3 with joint probability density $f(x_A, x_B, x_C)$, which can be factorized as follows

$$f(x_A, x_B, x_C) = f_{C|AB}(x_C|x_A, x_B) f_{B|A}(x_B|x_A) f_A(x_A). \tag{4}$$

According to Sklar's Theorem (Sklar, 1959), $f_{C|AB}(x_C|x_A, x_B)$ can be re-expressed as

$$\begin{aligned} f_{C|AB}(x_C|x_A, x_B) &= \frac{f_{AC|B}(x_A, x_C|x_B)}{f_{A|B}(x_A, x_B)} = c_{AC|B}\big(F_{A|B}(x_A|x_B), F_{C|B}(x_C|x_B)\big) f_{C|B}(x_C|x_B) && (5)\\ &= c_{AC|B}\big(F_{A|B}(x_A|x_B), F_{C|B}(x_C|x_B)\big) c_{BC}\big(F_B(x_B), F_C(x_C)\big) f_C(x_C), && (6) \end{aligned}$$

where c represents the copula function. Similarly, $f_{B|A}(x_B|x_A)$ can be re-written as

$$f_{B|A}(x_B|x_A) = c_{AB}\big(F_A(x_A), F_B(x_B)\big) f_B(x_B). \tag{7}$$

Therefore, we obtain a pair copula decomposition in three dimensions as follows

$$\begin{aligned} f(x_A, x_B, x_C) =& c_{AC|B}\big(F_{A|B}(x_A|x_B), F_{C|B}(x_C|x_B)\big) c_{BC}\big(F_B(x_B) F_C(x_C) \times && (8)\\ & c_{AB}\big(F_A(x_A), F_B(x_B)\big) f_C(x_C) f_B(x_B) f_A(x_A). && (9) \end{aligned}$$

As such, the joint density is broken down into a product of bivariate copulas and marginal densities. This paper will consider the most general class of vines that provides a valid density without further restrictions on the vine tree sequence, namely the R-vine.

Figure 4 shows an example of an R-vine tree sequence on three variables. The nodes in the first tree represent the three causes A, B, and C. The edges are identified with bivariate pair-copulas, which describe the dependence between each pair of variables. In the second tree, the nodes are the edges of the first tree. The edges describe the dependence between node AB and BC conditional on C. The full flexibility of vine allows for a different direction and magnitude of dependence in each pair of variables.

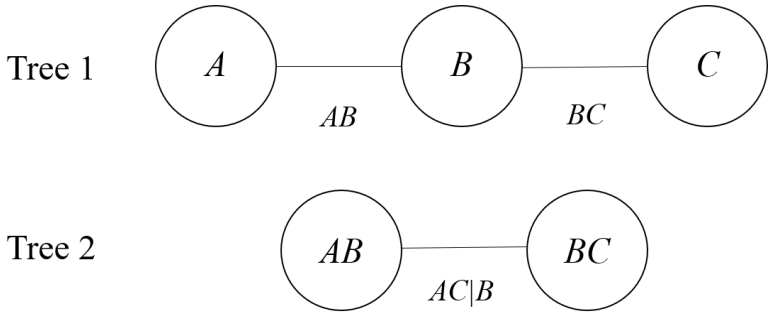


Figure 4: Example of an R-vine tree sequence

Following the same factorization procedure, we can construct pair-copulas in d dimensions. The joint probability distribution in dimension n can be expressed as:

$$f(x_1, \ldots, x_d) = \big(\prod_{k=2}^{d} f(x_k|x_1, \ldots, x_{k-1})\big) f_1(x_1) \tag{10}$$

$$= \prod_{i=1}^{d-1} \prod_{e \in E_i} c_{j_e,k_e|D_e}(F_{j_e|D_e}(x_{j_e}|\boldsymbol{x}_{D_e}), F_{k_e|D_e}(x_{k_e}|\boldsymbol{x}_{D_e})) \prod_{k=1}^{d} f_k(x_k), \tag{11}$$

where E_i is the set of all possible edges in Tree i, and $e = \{j_e, k_e\} \in E_i$. D_e is the conditioning set associated with edge e, and $\boldsymbol{x}_{D_e} = \{x_i \,|\, i \in D_e\}$.

There are three key components in a vine copula-based model, namely the vine structure (*i.e.* trees), copula families, and copula parameters. It is important to note that the construction of pair copula densities in Equation (11) is not unique, which could lead to a large number of possible vine tree sequences. As the dimension of the vine distribution increases, the number of possible tree sequences could be enormous. In this paper, the automated selection algorithm proposed by Dissmann *et al.* (2013) will be employed which jointly searches for the vine tree sequence, the pair-copula families and the parameter values of the chosen pair-copula families. The method aims to maximize dependence on higher order trees and is also referred to as the "sequential treewise approach". The pair-copula families are selected based on the Bayesian Information Criterion (BIC), and the corresponding parameter(s) are estimated using the sequential Maximum Likelihood Estimation method.[4]

5 Empirical study

We present the empirical results based on the US weekly cause-specific data described in Section 2. Vine copula models are fitted to the residual terms for both pre and post pandemic periods, based on the ARIMA models in Section 2. We first show pairs plot of the copula data, together with estimated Kendall's τ (valued between 0 and 1) to illustrate the level of ordinal association between two causes. The vine tree structure, parameter estimates, as well as the tree sequences will then be presented and discussed.

[4]R packages "`VineCopula`" (Schepsmeier *et al.*, 2015) and "`rvinecopulib`" (Nagler and Vatter, 2018) are used to perform model selection and estimation.

5.1 Dependence structure during 2015–2019

Figure 5 illustrates the pairwise dependence structure across 6 causes of death for the pre-pandemic period during 2015–2019. The histograms presented in the diagonal show that the copula data is reasonably close to uniform distribution on [0, 1].

From the upper triangle, we can see that the highest Kendall's τ (0.32) lies between *circulatory* and *malignant*. *circulatory* and *dementia* also have a relatively strong ordinal association with a Kendall's τ of 0.26. The third highest Kendall's τ is found between *circulatory* and *respiratory*. Overall, *circulatory* exhibits relatively strong dependence with all other causes except for *influenza*. On the other hand, in many cases the estimated Kendall's τ is close to 0 which means that the dependence between those paired causes is very weak. This result is not surprising as not all causes are exposed to common risk factors or external events. It is also worth noting that all of the pairwise Kendall's τ are positive. In the lower triangle, we plot the empirical contour plots of the normalized copula data which is the marginal transformed copula data using the inverse of the standard normal distribution function. Elliptical shapes indicate that the bivariate Gaussian copula might give a good fit.

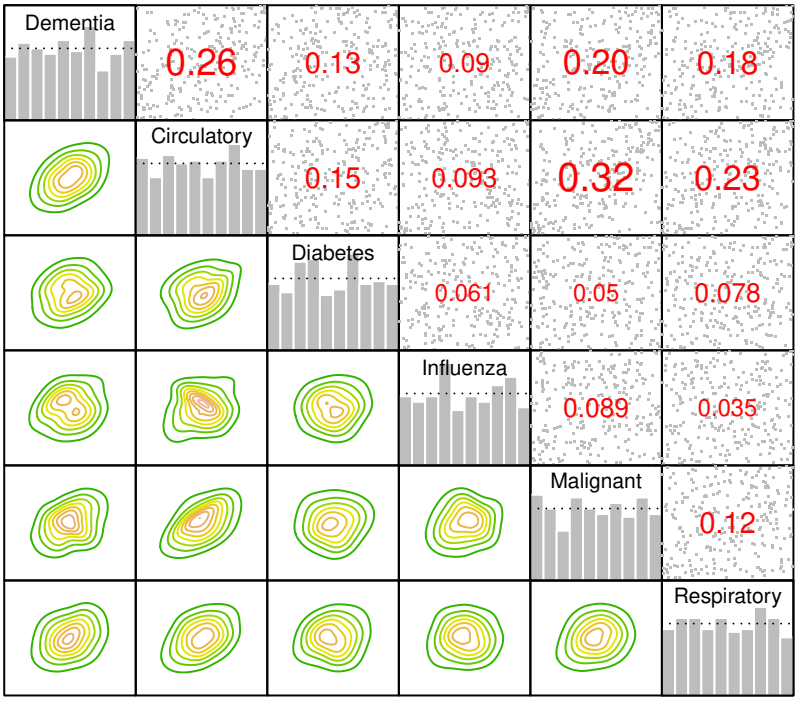


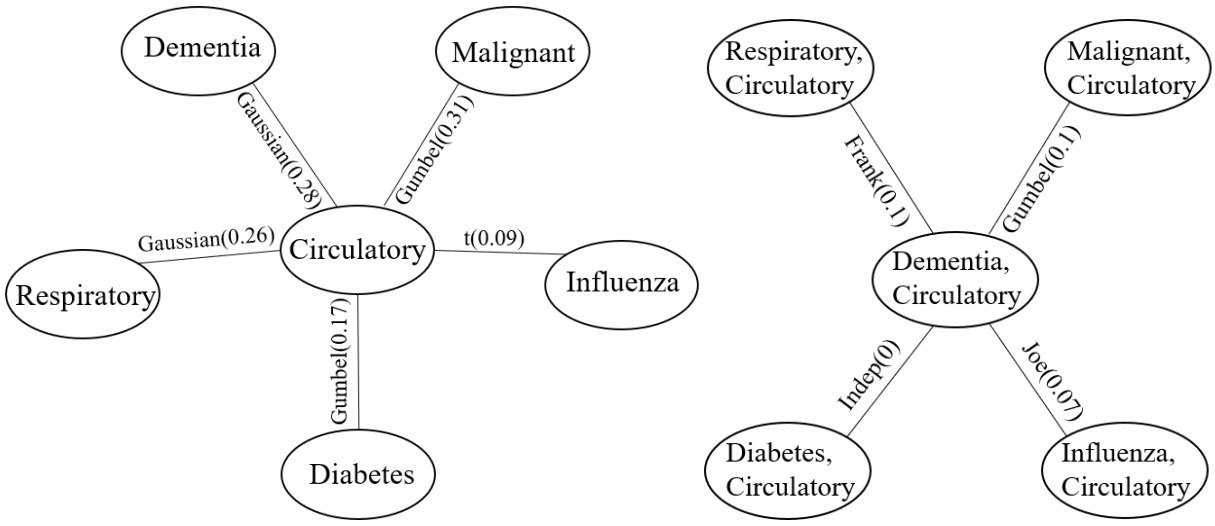
Figure 5: Pairwise copula data for 2015–2019. Upper triangle: scatter plots of copula data with estimated Kendall's τ; Diagonal: marginal histograms of copula data; Lower triangle: empirical contour plots of normalized copula data.

Table 3 presents the selected R-vine tree structure and sequential parameter estimates. In fact, the selected structure is a C-vine which has one node in each tree having a maximum number of edges attaching to it. In this case, *circulatory* is the hub connecting all other causes. The tree structure is truncated at level 2 as the first two trees have fully captured the dependence across different causes. Higher order trees have 0 Kendall's τ and involve independence copulas only.

Tree	Edge	Conditioning	Family	Kendall's τ	Tail dependence
1	*diabetes*, *circulatory*	-	Gumbel	0.166	Upper tail
1	*influenza*, *circulatory*	-	t	0.094	Upper & Lower tail
1	*malignant*, *circulatory*	-	Gumbel	0.308	Upper tail
1	*dementia*, *circulatory*	-	Gaussian	0.276	-
1	*respiratory*, *circulatory*	-	Gaussian	0.258	-
2	*diabetes*, *dementia*	*circulatory*	Independence	0	-
2	*influenza*, *dementia*	*circulatory*	Joe	0.070	Upper tail
2	*malignant*, *dementia*	*circulatory*	Gumbel	0.098	Upper tail
2	*respiratory*, *dementia*	*circulatory*	Frank	0.102	-

Table 3: Vine tree structure and sequential parameter estimates for 2015–2019

During the vine construction, we consider bivariate pair copulas including the Gaussian and t copulas, as well as Archimedean copulas (including Gumbel, Clayton, Frank, and Joe copulas).[5] On the first tree, it is worth noting that *circulatory* exhibits upper tail dependence with *malignant* and *diabetes*, which means that large excess deaths from these causes are likely to occur simultaneously. The dependence captured on the second tree is much weaker compared to the first tree, due to the sequential treewise selection algorithm (Dissmann *et al.*, 2013) adopted in this analysis. Followed the work by Bedford and Cooke (2001), we present a graphical view of Tree 1 and Tree 2 sequence in Figure 6.



(a) Tree 1 (b) Tree 2

Figure 6: Graphical view of vine copula tree sequence: 2015–2019

5.2 Dependence structure during 2020–2022

In this section, we present the vine copula model based on cause-of-death data during 2020–2022. We also compare the selected copula structure with the pre-pandemic structure shown

[5]For detailed specifications of the copulas shown in the table, refer to Brechmann and Schepsmeier (2013).

in the previous section. Figure 7 illustrates the pairwise dependence structure for the post-pandemic period. The histograms of copula data show that the transformed copula data based on the marginal models is still close to Uniform(0,1), in view of a much smaller sample size compared to 2015–2019.

From the upper triangles, we can see that overall, there has been an increase in the dependence level across 6 causes. The increase is most noticeable for *respiratory*, which now has much higher values of τ with other causes compared to the pre-pandemic structure. Stronger pairwise association means that unexpected large excess deaths are more likely to occur simultaneously across multiple causes, which is consistent with the mortality experience during 2020–2022. The lower triangles in Figure 7 show more variations in shapes, indicating potential deviations from elliptical copulas, as well as tail dependence behaviour.

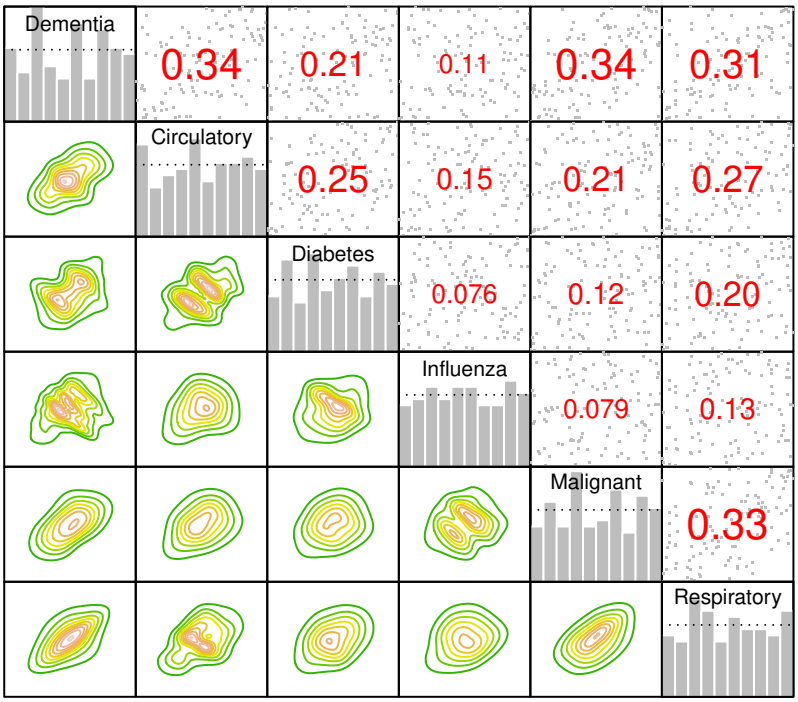


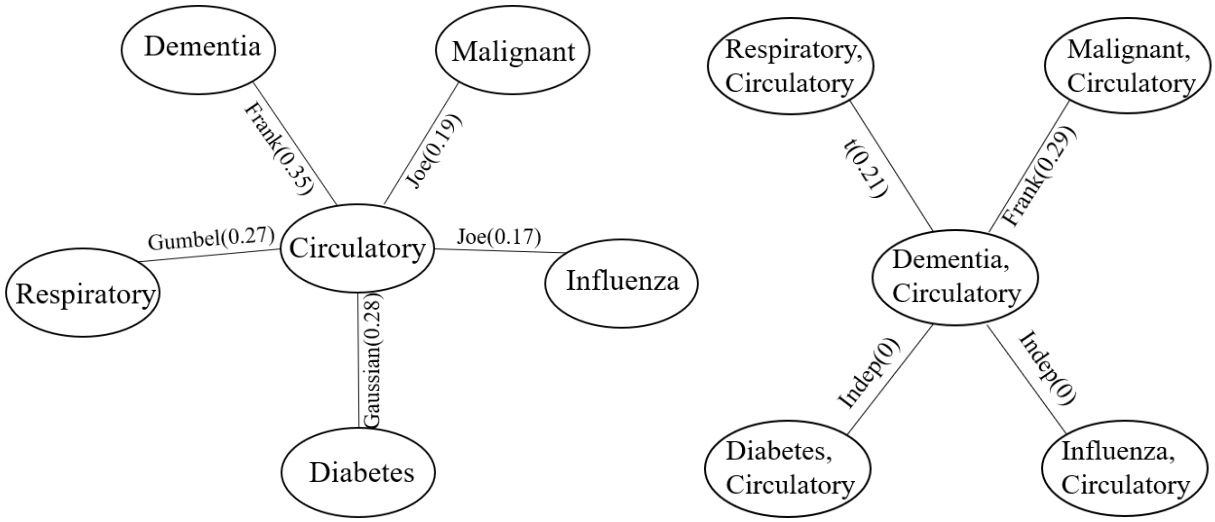
Figure 7: Pairwise copula data for 2020–2022. Upper triangle: scatter plots of copula data with estimated Kendall's τ; Diagonal: marginal histograms of copula data; Lower triangle: empirical contour plots of normalized copula data.

We present the selected vine tree structure and sequential parameter estimates in Table 4. It is important to note that the selected vine structure remains a C-Vine, where *circulatory* is the center-piece. In Tree 1, four Archimedean copulas are selected, where three of them exhibit upper tail dependence. The strongest dependence is captured between *dementia* and *circulatory* with the highest Kendall's τ (0.348), and this is a substantial increase from 0.276 shown in Table 3. The dependence between *influenza* and *circulatory* also increased noticeably from 0.094 in Table 3 to 0.168 in Table 4. On the other hand, there has been a decrease in the pairwise dependence between *circulatory* and *malignant* from 0.308 to 0.190. The dependence between *circulatory* and *respiratory* remains at a similar level before and after the pandemic.

Tree	Edge	Conditioning	Family	Kendall's τ	Tail dependence
1	*diabetes*, *circulatory*	-	Gaussian	0.278	-
1	*influenza*, *circulatory*	-	Joe	0.168	Upper tail
1	*malignant*, *circulatory*	-	Joe	0.190	Upper tail
1	*dementia*, *circulatory*	-	Frank	0.348	-
1	*respiratory*, *circulatory*	-	Gumbel	0.272	Upper tail
2	*diabetes*, *dementia*	*circulatory*	Frank	0.289	-
2	*influenza*, *dementia*	*circulatory*	Independence	0	-
2	*malignant*, *dementia*	*circulatory*	Independence	0	-
2	*respiratory*, *dementia*	*circulatory*	*t*	0.215	Upper & Lower tail

Table 4: Vine tree structure and sequential parameter estimates for 2020–2022

In Figures 8, we provide graphical representations of the selected tree sequence. We can see that the overall dependence structure remains unchanged during 2020–2022 but the dependence across causes have increased which means that extreme mortality events are more likely to happen during the pandemic.



(a) Tree 1 (b) Tree 2

Figure 8: Graphical view of vine copula tree sequence: 2020–2022

5.3 Limitations of the study

For readers of this paper, it is important to understand the potential limitations of the research. First, the death-count data collected in this study are based on the "underlying" cause-of-death definition, which does not take into account information on "multiple causes of death". For example, a death is recorded under *vascular* if vascular disease is the underlying condition that led to the death. However, information on other significant conditions that may have contributed to the death are not considered. Therefore, data on "underlying" cause-of-death alone could not provide us with a complete picture on cause-specific death experience of the population. Second, our investigation period is from January 2015 to March 2022, while the pandemic mortality experience is still unfolding, so it is difficult to perform statistical testing on

the current results. Once more data are collected and become available, further investigations can be conducted to test our hypothesis, and to see if the increased levels of dependence continue in the post-pandemic era.

6 Conclusions

In this paper, we investigate the dependence structure across 6 major causes for time periods both before and post pandemic. The vine copula model adopted in our analysis ensures great flexibility in the dependence modeling, via pair copula constructions. We apply the proposed approach to the US cause-specific weekly death counts during 2015–2022, and find that the dependence structure remains unchanged before and after the pandemic, with *circulatory* being the key hub connecting other causes. However, we also observe an increase in the general dependence level across these major causes of death. It is concluded that besides the surge in death counts due to the COVID-19 pandemic, the positive correlation between COVID deaths and other causes, together with an increase in the dependence level across major causes, are likely to trigger more frequent extreme mortality events during the pandemic.

From the perspective of insurance and health care industries, the impact of COVID-19 on population mortality and health experience should be carefully considered and taken into account during risk assessment and evaluation. Beyond this research, it is important to further investigate whether the increased level of dependency across major causes of death during the pandemic was partially due to stretched health care services, and whether different socioeconomic groups were affected differently in mortality and health experience during the pandemic. Answers to these questions will help insurance companies, health care professions, and government agencies be better prepared for future mortality risk.

References

Aas, K., Czado, C., Frigessi, A., and Bakken, H. (2009). Pair-copula constructions of multiple dependence. *Insurance: Mathematics and Economics*, **44**(2), 182–198.

Alai, D. H., Arnold, S., Bajekal, M., and Villegas, A. M. (2018). Mind the gap: a study of cause-specific mortality by socioeconomic circumstances. *North American Actuarial Journal*, **22**(2), 161–181.

Bedford, T. and Cooke, R. (2001). Monte carlo simulation of vine dependent random variables for applications in uncertainty analysis. *ESREL 2003*.

Brechmann, E. and Schepsmeier, U. (2013). CDVine: Modeling dependence with C-and D-vine copulas in R. *Journal of Statistical Software*, **52**(3), 1–27.

Diaconu, V., Ouellette, N., and Bourbeau, R. (2020). Modal lifespan and disparity at older ages by leading causes of death: a Canada-US comparison. *Journal of Population Research*, **37**(4), 323–344.

Dissmann, J., Brechmann, E. C., Czado, C., and Kurowicka, D. (2013). Selecting and estimating regular vine copulae and application to financial returns. *Computational Statistics & Data Analysis*, **59**, 52–69.

Joe, H. (2014). *Dependence modeling with copulas*. CRC press.

Li, H., Li, H., Lu, Y., and Panagiotelis, A. (2019). A forecast reconciliation approach to cause-of-death mortality modeling. *Insurance: Mathematics and Economics*, **86**, 122–133.

Nagler, T. and Vatter, T. (2018). rvinecopulib: high performance algorithms for vine copula modeling. *URL https://cran. r-project. org/package= rvinecopulib*.

Nagler, T., Vatter, T., and Nagler, M. T. (2019). Package 'kde1d'.

Schepsmeier, U., Stoeber, J., Brechmann, E. C., Graeler, B., Nagler, T., Erhardt, T., Almeida, C., Min, A., Czado, C., Hofmann, M., *et al.* (2015). Package 'vinecopula'. *R package version*, **2**(5).

Sklar, M. (1959). Fonctions de repartition an dimensions et leurs marges. *Publ. inst. statist. univ. Paris*, **8**, 229–231.

About The Society of Actuaries Research Institute

Serving as the research arm of the Society of Actuaries (SOA), the SOA Research Institute provides objective, data-driven research bringing together tried and true practices and future-focused approaches to address societal challenges and your business needs. The Institute provides trusted knowledge, extensive experience and new technologies to help effectively identify, predict and manage risks.

Representing the thousands of actuaries who help conduct critical research, the SOA Research Institute provides clarity and solutions on risks and societal challenges. The Institute connects actuaries, academics, employers, the insurance industry, regulators, research partners, foundations and research institutions, sponsors and non-governmental organizations, building an effective network which provides support, knowledge and expertise regarding the management of risk to benefit the industry and the public.

Managed by experienced actuaries and research experts from a broad range of industries, the SOA Research Institute creates, funds, develops and distributes research to elevate actuaries as leaders in measuring and managing risk. These efforts include studies, essay collections, webcasts, research papers, survey reports, and original research on topics impacting society.

Harnessing its peer-reviewed research, leading-edge technologies, new data tools and innovative practices, the Institute seeks to understand the underlying causes of risk and the possible outcomes. The Institute develops objective research spanning a variety of topics with its strategic research programs: aging and retirement; actuarial innovation and technology; mortality and longevity; diversity, equity and inclusion; health care cost trends; and catastrophe and climate risk. The Institute has a large volume of topical research available, including an expanding collection of international and market-specific research, experience studies, models and timely research.

Society of Actuaries Research Institute
475 N. Martingale Road, Suite 600
Schaumburg, Illinois 60173
www.SOA.org

Copyright © 2023 Society of Actuaries Research Institute