Cause-of-death diversity from a multiple-cause perspective in the US

Sergi Trias-Llimós* & Iñaki Permanyer

Centre d'Estudis Demogràfics, Centres de Recerca de Catalunya (CERCA), Carrer de Ca n'Altayó, Edifici E2, Universitat Autònoma de Barcelona, 08193 Bellaterra, Spain

* Corresponding author. Email: strias@ced.uab.cat

Abstract

The dispersion of deaths across underlying causes of death has increased over the last two decades. We contribute to the debate of cause-of-death diversity by proposing and measuring it in the more general 'Multiple Causes of Death' (MCOD) settings in 2003-18. Our approach is an adaptation of popular approaches commonly applied to assess heterogeneity in human societies or animal ecosystems. Cause-of-death diversity increased in the studied period, and specially up to 2012. These trends were driven by increases in ages 65 and over. The inclusion of MCOD resulted in higher increases in cause-of-death diversity over time compared to merely using underlying causes of death, except for older men where no differences were observed. The observed increases in cause-of-death diversity should be monitored to better understand mortality dynamics in ageing populations. Our new MCOD diversity measures suggest that traditional approaches relying on single causes of death might be underestimating cause-of-death diversity dynamics.

Keywords: multiple causes of death, diversity, inequalities, health, multimorbidity

Introduction

Patterns in causes-of-death are known to be key indicators of population well-being. While much is known about leading causes of death and how they have shifted over time (e.g., from a majority of deaths caused by infections and communicable diseases to a majority of deaths caused by non-communicable diseases (Omran 1998)), much less is known about the diversity of causes from which individuals die at a given point in time and how it has evolved over time. The diversity in causes of death is unarguably an important marker of health heterogeneity, with higher values suggesting more difficulties both in diagnosis/treatment (i.e., micro-level effects) and health planning (i.e., macro-level effects). The study of cause-of-death diversity has received very little attention from academia (Bergeron-Boucher et al. 2020) despite its important practical implications – especially in the swift ageing process that is sweeping the world. In this paper, we propose a new approach to measure cause-of-death diversity that has been adapted to the specificities prevailing in low-mortality settings, and therefore that is able to account for multiple causes of death (MCOD).

Cogent indicators aiming to measure cause-of-death diversity in low-mortality settings should ideally be sensitive to (1) the increasingly high prevalence of comorbidity, and (2) the increasing differences in health outcomes across socioeconomic groups. On the one hand, as mortality shifts towards older ages, the presence of comorbidities becomes progressively more prevalent, thus complicating the assignation of a single cause of death – which is commonly referred to as 'the underlying cause of death' (Alpérovitch et al. 2009; Tinetti et al. 2012). This is nowadays more important than ever before as mortality has shifted toward older ages (e.g. Brown et al., 2012), and managing comorbidities at old age has become a major challenges for health care and public health management (Barnett et al. 2012; Parekh and Barton 2010). Indeed, some researchers have raised concerns about the oversimplification of underlying cause of death approaches given the complexity of its cause-of-death assignation (Lu et al. 2001) and its quality (Flagg and Anderson 2021), which suggests the need to go beyond single cause-of-death settings, for example by exploring MCOD for better understanding current mortality dynamics (Désesquelles et al. 2014). On the other hand, in a context of widening health inequalities across socioeconomic groups (e.g. Montez et al., 2019; Permanyer et al., 2018; Sasson, 2016), it would be desirable that cause-of-death diversity measures could be broken down in such a way that informed users about the extent of diversity occurring within and between socio-economic groups. Such decomposition could be extremely useful to potentially identify the main locus of cause-of-death diversity (i.e., unveiling whether diversity comes from differences occurring within or between groups) and pinpoint the key drivers of changes over time.

In this extended abstract, we examine cause-of-death diversity in the US by age groups and sex over the period 2003-2018 by developing and applying a new approach based on the analysis of multiple causes of death. Educational gradients will be assessed in the full version of the paper.

Data and methods

We use individual level multiple causes of death mortality data for the period 2003-18 US from the Centers for Disease Control and Prevention covering the population ("CDC US Mortality" n.d.). Cause-of-death mortality data are classified in 13 groups, following the standards of the International Classification of Diseases revision 10 (ICD-10) main (groups of) chapters. We have selected the more commonly used chapters as underlying causes of death, namely: Infectious (ICD-10 codes: A00-B99), Neoplasms (C00-D48), Metabolic (E00-E88), Mental (F01-F99), Nervous (G00-G98), Cardiovascular (I00-I99), Respiratory (J00-J98), Digestive (K00-K92), Musculoskeletal (M00-M99), Genitourinary (N00-N98), Ill-defined (R00-R99), we have grouped all External causes within one category (S00-Y89), and we have grouped the remaining chapters under the label of Other Causes (blood (D50-D89), eye and ear (H00-H93), skin (L00-L98), pregnancy (O00-O99), perinatal (P00-P96), and congenital (Q00-Q99)).

In the MCOD setting, we use the notation $(\boldsymbol{v}; i)$, to formally indicate that a given death has been caused by an underlying *and* potentially other causes of death, where $\boldsymbol{v} = (v_1, \cdots, v_{13}) \in \{0,1\}^{13} \setminus \mathbf{0}$ is a 13-dimensional vector of zeroes and ones (with a '1' in position 'j' to indicate that cause-of-death 'j' has contributed to that specific death, and '0' for the opposite), and $i \in \{1, \cdots, 13\}$ indicates which of the 13 elements of $\mathcal{C}$ is *the* underlying cause of death. To illustrate: the vector $(\boldsymbol{a}; 2) = ((1,1,1,0, \cdots, 0); 2)$ will be used for those death certificates indicating that C_2 is the underlying cause of death and C_1, C_3 are contributing causes of death. The generic elements $(\boldsymbol{v}; i)$ will be referred to as 'sets of causes of death'.

Our attempts at measuring cause-of-death diversity in a MCOD setting require making assessments of the extent of similarity or dissimilarity among any two sets of causes of death (e.g., $(\boldsymbol{a}; i)$ vs $(\boldsymbol{b}; j)$; see below). To make this comparison, one must first decide the degree to which underlying causes of death are more important or relevant than the other causes listed in the death certificate when determining the death of individuals. Since this is a highly context-specific issue that is likely to vary considerably across individual deaths, we have decided to make some simplifying assumptions. First, we assign a weight $W > 0$ to the underlying cause of death and another weight $\omega \geq 0$ to all the remaining causes of death listed in the death certificate, with the restriction that $W \geq \omega$. Second, we further assume that in a hypothetical death certificate listing the 13 groups considered in this paper as contributing causes, the sum of the corresponding 'importance weights' must equal one (a standard normalization procedure). Since there is only one underlying cause of death and potentially 12 other contributory causes of death, the following restriction must hold:

$$\omega = \frac{1 - W}{12}$$

Imposing these mild restrictions, we conclude that the values of W are allowed to move between 1/13 and 1. While the choice of W is arbitrary, the meaning of its extreme values are clear. If we choose $W = 1$, then $\omega = 0$ and we are in the standard single-cause-of-death setting where the causes in the death certificate other than the underlying cause of death are ignored. At the other extreme, when $W = 1/13$ then $\omega = 1/13$, so we are in a setting where the underlying cause of death is only equally important, and not more important, than the other causes included in the death certificate. For all other intermediate values of W, the underlying cause of death is more important than the other causes of death included in the list in varying degrees. In the empirical section of the paper, we present all our findings for different values of W (1/13, 0.2, 0.5 and 1).

Measuring diversity in a MCOD setting

The approach we follow to measure diversity is simple: take two death certificates at random and measure how similar or dissimilar the corresponding sets of causes of death are. In this way, we obtain an estimate of the extent of (dis)similarity among deaths occurring in the population. The suggested approach has two steps. First, we measure the extent of (dis)similarity between any two sets of causes of death (i.e., distances). Second, we average such measures across all possible pairs.

Given any two sets of causes of death $(\boldsymbol{a}; i) = ((a_1, \cdots, a_{13}); i)$ and $(\boldsymbol{b}; j) = ((b_1, \cdots, b_{13}); j)$, we define the distance between them as follows:

$$d((\boldsymbol{a}; i), (\boldsymbol{b}; j)) = \sum_{k=1}^{13} \left| w_{i,k} a_k - w_{j,k} b_k \right|$$

where $w_{i,k}$ (respectively $w_{j,k}$) equals W whenever $i = k$ (respectively $j = k$) and equals ω otherwise. This distance function compares the similarity between the two causes-of-death vectors. The function is sensitive to the degree of overlap between the vectors $\boldsymbol{a}$ and $\boldsymbol{b}$ (i.e., it increases when the different causes of death in the two vectors do not coincide) and to the corresponding underlying causes of death (i, j). Implicitly, it assumes that the distance between any two single causes of death is the same.

Having defined how to measure the extent of (dis)similarity between any two sets of causes of death, we now average the results across all possible pairs. Following this approach, we obtain the following measure of MCOD diversity:

$$D = \sum_{(\boldsymbol{a};i)\in\Omega} \sum_{(\boldsymbol{b};j)\in\Omega} p_{(\boldsymbol{a},i)}\, p_{(\boldsymbol{b},j)} d((\boldsymbol{a}; i), (\boldsymbol{b}; j))$$

where $\Omega = (\{0,1\}^{13} \backslash \mathbf{0}) \times \{1, \dots, 13\}$ is the set of all possible sets of causes of death, and $p_{(\boldsymbol{a};i)}$ is the share of death certificates having $(\boldsymbol{a}; i)$ as the corresponding set of causes of death. The diversity indicator D can be interpreted as the mean distance between sets of causes of death, or in other words as the expected distance between two randomly chosen deaths. In a hypothetical society where everyone died exactly from the same causes, the distance function would always be 0, so D would also be equal to 0. This would be the lowest level of diversity. At the other extreme, when individuals tend to die from a more variegated set of causes, the distance function tends to increase, and so does the corresponding diversity indicator D. In order to render our results comparable across different W values and over time we normalized the results using 2003 as baseline (D_{2003}=1) for any given W.

Results

Cause-of-Death Diversity

Cause-of-death diversity trends are presented in **Figure 1**. The different values of W represent the importance of the underlying versus other causes of death listed in the death certificate. When W equals to 1 we are in the underlying cause of death setting. Oppositely, when W equals to 1/13 we are in a MCOD setting in which each group of cause listed in the death certificate has the same importance, irrespective of its position. W between 1/13 and 1 represent scenarios in between, where the underlying cause has more importance than the other causes, but that are also included. For any given W we see increases in cause-of-death diversity trends. These increases are more pronounced in the first half of the analysed period (2003-2010) and are more remarkable as W reduces (e.g. growing importance of MCOD). In the second half of the period (2010-2018) increases in cause of death diversity slowed down, particularly in the scenario when uniquely underlying causes of death are considered (W=1). Trends in cause-of-death diversity were increasing across all age groups (**Figure 2**).

Figure 1. Trends in normalized (D_{2003}=1) cause of death diversity according to the importance of the underlying versus other causes of death (W), 2003-18, US

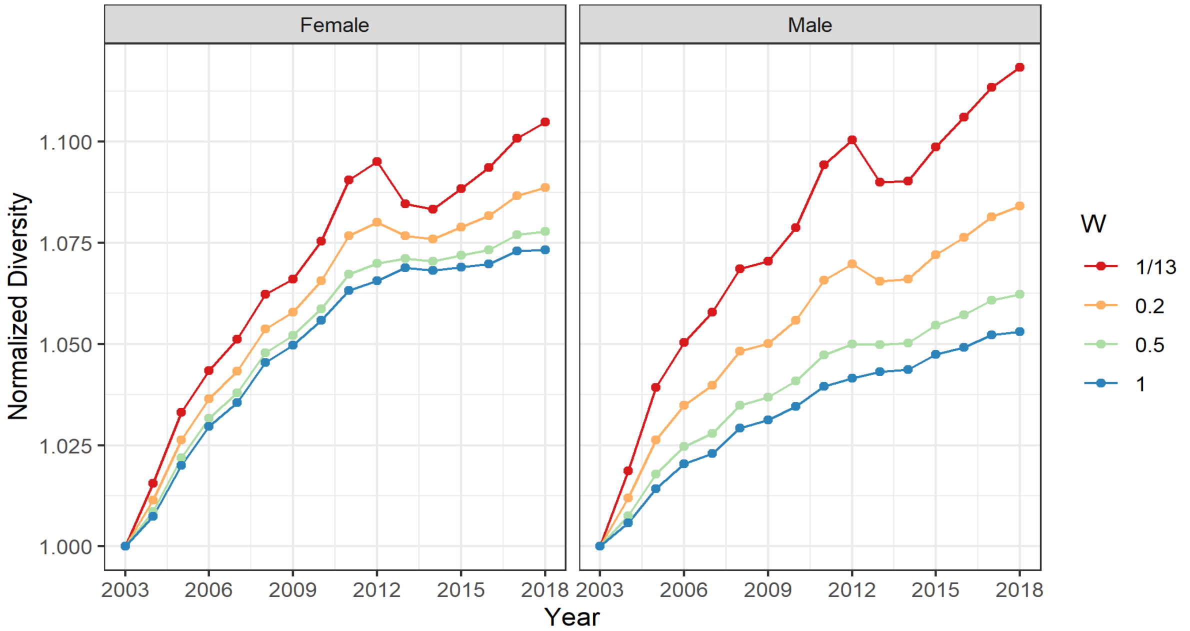
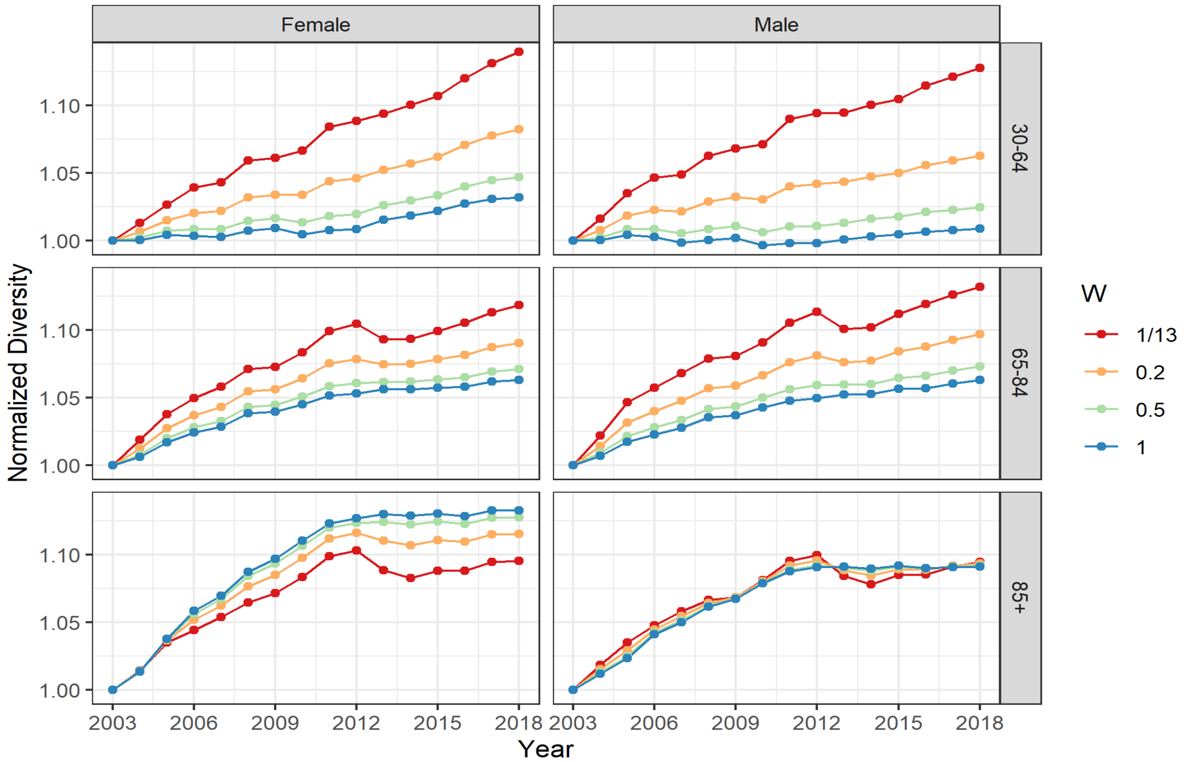


Figure 2. Trends in normalized (D_{2003}=1) cause-of-death diversity according to the importance of the underlying versus other causes of death (W) by age groups, 2003-18, US



Preliminary discussion

Cause of death diversity has been increasing from 2003 onwards both analysing deaths from the underlying cause and from the MCOD. In the full version of the paper we will provide results by educational attainment, including decompositions to control for the changing educational composition of deaths.

Assessing cause-of-death diversity trends from a MCOD perspective provides value information on the variability of groups of causes to be tackled in order to contribute to all-cause mortality reduction. As individuals tend to die from increasingly heterogeneous sets of causes, the effort towards reducing mortality should be divided within a larger group of causes. This implies that mortality improvements may have a higher cost as mortality from more causes of death would need to be tackling and improved in order to further reduce in all-cause mortality. In other words, a higher diversification of causes of death might require a wider variety of specific treatments, and therefore potentially reducing the effectiveness of health policies.

References

Alpérovitch, A., Bertrand, M., Jougla, E., Vidal, J.-S., Ducimetière, P., Helmer, C., et al. (2009). Do we really know the cause of death of the very old? Comparison between official mortality statistics and cohort study classification. *European Journal of Epidemiology*, *24*(11), 669–675. https://doi.org/10.1007/s10654-009-9383-2

Barnett, K., Mercer, S. W., Norbury, M., Watt, G., Wyke, S., & Guthrie, B. (2012). Epidemiology of multimorbidity and implications for health care, research, and medical education: a cross-sectional study. *The Lancet*, *380*(9836), 37–43. https://doi.org/10.1016/S0140-6736(12)60240-2

Bergeron-Boucher, M.-P., Aburto, J. M., & Raalte, A. van. (2020). Diversification in causes of death in low-mortality countries: emerging patterns and implications. *BMJ Global Health*, *5*(7), e002414. https://doi.org/10.1136/bmjgh-2020-002414

Brown, D. C., Hayward, M. D., Montez, J. K., Hummer, R. A., Chiu, C.-T., & Hidajat, M. M. (2012). The significance of education for mortality compression in the United States. *Demography*, *49*(3), 819–840.

CDC US Multiple Causes of Death Data. CDC Wonder. Available at: https://wonder.cdc.gov/mcd.html.

Désesquelles, A., Demuru, E., Egidi, V., Frova, L., Meslé, F., Pappagallo, M., & Salvatore, M. A. (2014). Cause-specific mortality analysis: Is the underlying cause of death sufficient? *Revue Quetelet/Quetelet Journal*, *2*(1), 119–135. https://doi.org/10.14428/rqj2014.02.01.05

Flagg, L. A., & Anderson, R. N. (2021). Unsuitable underlying causes of death for assessing the quality of cause-of-death reporting.

Lu, T.-H., Shih, T.-P., Lee, M.-C., Chou, M.-C., & Lin, C.-K. (2001). Diversity in death certification: a case vignette approach. *Journal of clinical epidemiology*, *54*(11), 1086–1093.

Montez, J. K., Zajacova, A., Hayward, M. D., Woolf, S. H., Chapman, D., & Beckfield, J. (2019). Educational disparities in adult mortality across US states: how do they differ, and have they changed since the mid-1980s? *Demography*, *56*(2), 621–644.

Omran, A. R. (1998). The epidemiologic transition theory revisited thirty years later. *World health statistics quarterly*, *53*(2, 3, 4), 99–119.

Parekh, A. K., & Barton, M. B. (2010). The challenge of multiple comorbidity for the US health care system. *Jama*, *303*(13), 1303–1304.

Permanyer, I., Spijker, J., Blanes, A., & Renteria, E. (2018). Longevity and Lifespan Variation by Educational Attainment in Spain: 1960–2015. *Demography*, *55*(6), 2045–2070. https://doi.org/10.1007/s13524-018-0718-z

Sasson, I. (2016). Trends in Life Expectancy and Lifespan Variation by Educational Attainment: United States, 1990–2010. *Demography*, *53*(2), 269–293. https://doi.org/10.1007/s13524-015-0453-7

Tinetti, M. E., McAvay, G. J., Murphy, T. E., Gross, C. P., Lin, H., & Allore, H. G. (2012). Contribution of Individual Diseases to Death in Older Adults with Multiple Diseases. *Journal of the American Geriatrics Society*, *60*(8), 1448–1456. https://doi.org/10.1111/j.1532-5415.2012.04077.x