Analysis of multiple causes of death: a review of methods and practices

Karen Bishop¹, Saliu Balogun¹, James Eynstone-Hinkins², Lauren Moran², Melonie Martin¹, Emily Banks¹, Chalapati Rao¹, Grace Joshy¹*

¹National Centre for Epidemiology and Population Health, Australian National University
²Australian Bureau of Statistics, Canberra, Australia

*Corresponding Author: National Centre for Epidemiology and Population Health, ANU College of Health and Medicine, Australian National University, 62 Mills Road, Acton ACT 2601, Australia. Email: grace.joshy@anu.edu.au
Abstract

Background

Research and reporting of mortality indicators typically focus on a single underlying cause of death selected from multiple causes recorded on a death certificate. The need to incorporate the multiple causes in mortality statistics - reflecting increasing multimorbidity and complex causation patterns - is recognised internationally. This review aims to identify and appraise relevant multiple cause analytical methods and practices.

Methods

We searched Medline, PubMed, Scopus and Web of Science from inception to December 2020 without language restrictions, supplemented by consultation with international experts. Eligible articles included those analysing multiple causes of death from death certificates. The process identified 4,080 articles; after screening, 434 full texts were reviewed.

Results

Most reviewed articles (77%, n=332) were published since 2001. The majority examined mortality by “any-mention” of a cause of death (87%, n=377) and assessed pairwise combinations of causes (56%, n=245). Recently emerging (since 2001) were applications of methods to group deaths based on common cause patterns using, for example, cluster analysis (2%, n=9), and the application of multiple cause weights to re-evaluate mortality burden (1%, n=5). Multiple cause methods applied to specific research objectives are described for recently emerging approaches.

Conclusion

This review confirms rapidly increasing international interest in the analysis of multiple causes of death and provides the most comprehensive overview of methods and practices to date. Available multiple cause methods are diverse but suit a range of research objectives, that with greater data availability and technology could be further developed and applied across a range of settings.

Keywords: multiple causes of death, mortality, comorbidity, death certificates, population health
**Introduction**

Mortality statistics are crucial to population health as they provide fundamental information about health status, disease aetiology, trends, and patterns of diseases in different populations. They inform health services, health policy development and planning, as well as research and can be used to evaluate the impact of health intervention programs.\(^1\)\(^,\)\(^2\) Therefore, it is critical that accurate and reliable information about the diseases and health conditions that cause death are appropriately analysed.

Death typically results from the interplay between multiple health conditions. The standard international format of the death certificate (Figure 1) facilitates recording the certifying doctor’s medical opinion of all diseases and conditions involved in the death including the underlying and non-underlying (intermediate and immediate) causes in Part I, and significant other contributing causes in Part II. If the certificate is completed correctly, the underlying cause (UC) reflects the initiating condition, that is, one that could be avoided by some preventative mechanism to interrupt the sequence leading to death. The medical certification process reflects the multifaceted pathological processes leading to death. However, the recording of a single disease as the UC can be complex and misclassification of the UC can occur when several causal pathways are involved.\(^3\)\(^-\)\(^8\) Following the medical certification of the cause of death, an international coding standard is applied to all causes reported on the death certificate to endorse the reported UC or select a more appropriate alternative to be used for statistical reporting (including international comparisons) and epidemiological studies.\(^9\) Deaths data thus contain the standardised UC and all other causes that were involved in the death (associated causes).

Despite the vast amount of information that is collected about the causes of deaths, mortality statistics typically use only the (single) UC. Researchers have long recognised that the UC alone does not adequately describe the pathological processes responsible for most deaths,\(^10\)\(^-\)\(^12\) and potentially understates the importance of other significant contributing causes of death.\(^1\)\(^,\)\(^13\) International support for the need to assess the multiple causes of death (MC) to complement statistics based on the UC approach is well established but methods used are diverse. To adequately inform population health initiatives, it is essential that all diseases and conditions contributing to death receive appropriate attention.

Previous studies\(^10\)\(^,\)\(^14\)\(^-\)\(^16\) have attempted to appraise and summarise the assortment of methods used to measure the involvement of all causes on the death certificate, however they are
limited in scope and were performed over a decade ago thus not capturing recent methodological advances in the analysis of MC. This review, supplemented by consultations with subject experts, aims to identify and appraise the methods used in analyses of MC data, providing a comprehensive and up-to-date account of methods and practices that are used to describe and measure the involvement of multiple health conditions in causing death.

**Methods**

**Search strategy**

To identify articles that analyse MC data, we searched the Medline, PubMed, Scopus and Web of Science databases, each from inception to 31 December 2020, for original research without any restriction on language or country of study. To maximise the number of relevant articles related to multiple causes of death, we included search terms covering ‘multiple*’ AND ‘cause*’ AND (‘death’ OR ‘mortality’) and other variations such as (‘associated’ OR ‘contributory’ OR ‘underlying’) AND ‘cause*’ (Supplementary File 1).

**Selection of articles**

Two authors (KB and SB) each independently screened 50% of titles and abstracts of the identified articles, beginning with a random sample of 5% articles in duplicate. Disagreements were solved by consensus.

**Eligibility criteria**

Articles were eligible if they: reported using death registration or death certificate data and applied a multiple cause method to calculate a multiple cause indicator; or derived a measure of mortality based on the multiple causes of death. We considered all study designs, except case reports, case series and forensic reports. Research that used only the UC for analysis, or that used MC data but did not apply or report a measure based on multiple causes were excluded. Articles based on verbal autopsy, narrative, reviews, and non-peer reviewed literature were also excluded.

**Data extraction**

For the included articles, study characteristics (authors’ names, journal name, year published, study design, study period, country, source of multiple cause data and main cause of interest) and decedent characteristics (age, sex and number of deaths evaluated) were extracted using a full text review. Each article was categorised into one or more categories according to the
objectives of the application of MC methods as articles that: described cause-related mortality based on ‘any-mention’ of a cause; assessed the joint involvement of causes according to pairwise disease occurrence on death certificates; described mortality for clusters of >2 commonly co-occurring causes; and measured cause-related mortality burden by weighting multiple causes.

Audit of experts

To identify unpublished methods in practice, the search strategy was supplemented by consultation with subject experts. Contacts were identified from affiliations of relevant papers and recommendations from experts in the field, and included representatives from agencies such as the Multiple Causes-of-Death Network (https://mcod.web.ined.fr), the World Health Organization Family of International Classification collaborating centres, and national statistical offices of countries, including the United States, England, Canada, New Zealand, Italy and Australia. During December 2020, emails were sent to 261 contacts asking regarding their awareness of relevant studies, particularly recently accepted or unpublished papers or reports that used multiple cause methodology. We received 41 responses resulting in a response rate of 15.7%.

Statistical methods

Articles included in full text review were classified based on a priori selected mutually exclusive categories of statistical methods used to analyse multiple causes of death as: methods based on any-mention; methods to assess pairwise occurrence of causes; methods based on groups of >2 co-contributing causes; and methods based on weighting of multiple causes. Following full text review, included articles published from 2015 onwards were classified based on the main research objective into four categories: describe cause-related mortality; identify co-contributing causes; assess relationships between co-contributing causes; assess impact of risk factors; and other objectives. For articles published between 2015 and 2020, we mapped the research objectives against the MC methods that were used to achieve them. We did not evaluate publication bias as this review focussed on methodological practices applied in each article rather than on results.

Results

Overall 8,070 articles were identified: 8,002 from the database search and 68 from responses to the consultations. After removing duplicates, 4080 articles were selected for title and
abstract screening (Figure 2). Due to the large number of articles and limited resources, 101 articles that appeared to only use methods based on any-mention were excluded without full text review. Four review articles were identified in the screening, and a manual search of these identified a further 25 potentially eligible articles. A total of 602 articles were selected for full text review, including 25 from the consultations (Figure 2). From these, 434 articles were included, and the multiple cause methods applied in each were assessed. The remaining 168 articles did not meet the inclusion criteria: four were reviews, 61 were not in scope, 53 were irretrievable (including 13 in a foreign language), 28 were communications or conference abstracts and 22 were non-peer reviewed articles (none of them employed methods other than those a priori identified) and were excluded from full text review.

A summary of the characteristics of included articles is presented in Table 1. The number of articles using MC methods increased over time; more than three-quarters (n=332, 76.5%) were published after 2001 (Table 1). Most articles assessed deaths registered in the United States (n=222, 12.9%), Brazil (n=47, 2.7%) and the United Kingdom (n=43, 2.5%) (Table 1). In most countries, the application of MC methods rose over time, with notable increases during 2001–2020 (Figure 3). Infectious diseases (largely HIV/AIDS) were the most common cause of interest in the application of MC methods (n=76, 17.5% articles), followed by external causes (n=69, 15.9% articles) of which most assessed drug-related deaths. Population-level analysis of all causes of death using multiple cause methods were found in 44 (10.1%) articles. Cross-sectional evaluation of deaths was the most common study design (n= 353, 81.3% articles). Most articles reviewed were in English (n=416, 95.9%) and Portuguese (n=13, 3.0%).

Articles were broadly categorised by the methods applied in analysing MC as those describing mortality based on any-mention of a cause; examination of pairwise occurrence of diseases on death certificates; assessment of mortality based on groupings of >2 co-contributing causes; and calculation of cause-related mortality burden based on weighted multiple causes. The results for each category are described in detail below.

**Descriptive measures of multiple causes based on any-mention**

The full text review identified 86.9 % (n=377) articles that assessed cause-related mortality based on any-mention of the cause (Table 2, Supplementary File 2). The methods applied in these articles included basic summary (or univariate)\(^1\) statistics applied to a specific index cause. Examples of specific methods include multiple cause indicators that describe the
number of causes \((n)\) involved in each death, the average number of causes per death and frequency or percentage distributions of \(n\), each illustrating the extent to which multiple causes occur in the deaths data. Also included here were articles that aimed to evaluate cause-related mortality using rates based on any-mention of a cause; that is, by counting each death that mentions the cause of interest anywhere on the death certificate.\(^\text{18-21}\) Some evaluated the leading causes of death using any-mention\(^\text{22-27}\) (to understand the most common causes involved in deaths. Assessment of temporal trends in any-mention rates against rates based on the UC were used to highlight changes in certification or coding practices and changing patterns of disease contribution to death.\(^\text{28-30}\) Comparisons between countries in multiple cause indicators can be used to emphasise differences in certification practices between countries.\(^\text{3,31-33}\)

Rate ratios were commonly applied to indicate the extent to which a cause occurs as the UC versus non-UC. A common approach compared the occurrence of the cause as any-mention to its occurrences as the UC (the approach varies according to whether counts or rates are used in the ratio and whether the any-mention versus non-underlying mentions are compared to the UC.\(^\text{10}\) One form of this measure, the standardised ratio of multiple to underlying causes (SRMU),\(^\text{13,34}\) encourages a harmonised approach to calculating this indicator; here the rate ratio is calculated as the age-standardised rate for any-mention of the cause compared to the age-standardised rate when the cause is the UC. This indicator describes the extent to which the cause is selected as non-UC relative to UC (with 1 indicating the cause is always the UC, 2 indicating equal representation as UC and non-UC and >2 indicating the cause is more often a non-UC). Country comparisons of this method can be used to assess variation in certification.\(^\text{32}\)

Articles that used counts or rates of any-mention of causes to derive other summary measures of mortality, for example potential years of life lost\(^\text{35}\) and life expectancy\(^\text{36-41}\) were also categorised here. Further examples of articles that applied these methods are in Table 2.

**Assessing pairwise contribution to mortality**

Articles that investigated the relationships between two causes of death reported on the same death certificate were categorised here; we found 56.5\% \((n=245)\) articles in this review (Table 2, Supplementary File 2). A distinguishing feature of these methods is that the joint frequencies of just two causes are the focus of the analysis. To investigate relationships between two causes, researchers assessed the involvement of a specific UC with one or more
associated causes or vice versa (e.g. the nature and spread of an associated or immediate cause for an UC of interest). Typically, the objective is to understand which comorbid conditions commonly occur with a specific cause using frequency or percentage distributions of the most common associated causes for a specific UC and vice versa.42-45

More comprehensive applications of odds ratios (mortality odds ratios) and relative rates was found in articles that aimed to measure the associations between comorbid medical conditions involved in the death for a specific cause of interest. Mortality odds ratios were used to assess the odds of any-mention of a specific cause with other comorbidities at death. 9,35,46-57 Rate ratios were used to assess mortality burden according to whether another specific cause was present or absent, or to deaths in general.58,59 In most cases the application of mortality odds ratios disregards the role of the cause, thereby enabling relationships between two non-UCs to be included in the assessment. Previous reviews of measures of association for multiple cause of death discuss the applicability of several approaches including: matched mortality odds ratios, matched exposure odds ratios,15 and Yules Q, Positive Matching Index, Forbes’ coefficient and, Wise and Sorvillo ratio),14 concluding the most suitable to be those that do not consider non-matches.

A more recently introduced measure, the ‘cause of death association indicator’ (CDAI) compares the standardised rate of involvement of a cause of interest reported together with a specific UC to its involvement in death overall (that is with any UC).13,34,60 The CDAI aids understanding whether a non-UC of interest is more common with a specific UC than with all UCs combined. In this indicator, the role of the causes is fixed thereby requiring reverse comparisons for each UC and non-UC of interest.

Assessing mortality patterns for grouped causes of deaths

Articles categorised here aimed to assess mortality from either a cluster of decedents where the grouping is based on patterns in the multiple causes or by known disease associations, or by a cluster of specific causes grouped according to some measure of ‘similarity’ or closeness. These approaches were applied in 2.1% (n=9) articles and were considered as newly emerging methods, with all but one article published since 2009 (Table 2, Supplementary File 2). Methods for grouping multiple causes of death were largely used to understand more complex relationships between multiple causes of death or to highlight patterns of disease that commonly co-contribute to death with more than two causes of interest. In some articles, the groupings were user-specified combinations of causes11,61-63
while others used data-driven methods such as cluster analysis, social network analysis and more exploratory methods of data mining. Social network analysis identified links (and their strengths) between causes of death, cluster analysis methods grouped decedents on the basis of similarity between causes, that is, based on the causes of death that commonly co-contribute to death, and data mining techniques were applied to identify complex patterns in mortality data and assess temporal evolution of the leading clusters of conditions that cause death.

**Multiple-cause weighting methods**

We identified five (1.1%) articles that aimed to calculate cause-related mortality by ascribing weights to each cause in the death record (Table 2, Supplementary File 2). The multiple cause weighting methods identified in this review, assigned weights to each cause such that within each death, the weights summed to 1.0. With this approach, the counting unit (deaths) is preserved enabling measures of cause-related mortality (rates, years of life lost, etc) to be recalculated based on the weighted counts of deaths. The included articles illustrated their proposed method by estimating socioeconomic inequalities in mortality, proportional mortality and the relative risk of cause-specific mortality among individuals with human immunodeficiency virus versus those without.

Various strategies were noted for ascribing weights, for example, weighting all causes equally as $1/n$ where $n$ is the number of causes involved in the death, weighting the UC a fixed amount (e.g. 50%) with the remaining amount (in this case 50%) apportioned equally across the other causes, or weighting the UC twice that of other causes. The strategies also varied according to the causes included for weighting – all causes versus the UC plus contributing causes from Part II of the death certificate.

Among included texts, 45.7% (n=198) applied >1 MC method; most frequently (n=187) an any-mention method and a pairwise assessment of causes of death (Table 3). Four of five applications of weighting methods also reported indicators using any-mention.

Across all applications of MC methods, the techniques varied in regard to 1) whether the study included a comparison of multiple to underlying causes; 2) whether the study used all non-UCs, or a specified set based, for example, on the location on the death certificate (Part I or Part II); 3) whether ill-defined or external causes of death were considered; and 4) whether duplicate mentions of the cause of interest were excluded.
**Research Objectives**

The objectives for applying multiple causes analyses varied. Some broad themes were identified in the aims: to describe cause-related mortality using MC; to identify co-contributing causes (that is, the associated causes for a specific UC and vice-versa); to assess relationships between causes using a measure of association; or assess the impact of risk factors on mortality. A residual category of ‘other’ objectives captured articles that intended to measure the contribution of all causes of death in a population using MC methods and those with aims that centred around ‘exploring’ the MC data. In this subset of articles published between 2015 and 2020 (n=133), most (54.9%, n=73) aimed to describe the cause of interest using multiple causes of death and applied methods based on any-mention in 66 articles and pairwise assessment in 39 (Table 4). Of the 28 articles whose main aim was to identify co-contributing causes for a specific cause of interest, nearly all applied methods based on any-mention (n=24) and pairwise assessment of causes on the death certificate (n=25) with only 1 that grouped more than two causes for analysis. Most notable, were the applications of emerging approaches to analysing multiple causes. Foremost were whole of population analyses to quantify the contribution of all causes that contributed to death by weighting each one as described above.\(^1,69\) Though weights were arbitrary, these novel methods lend to capturing all causal contribution at a population level, which by restricting the selection of causes can convey the mortality contributed by antecedent causes\(^1,69\) or if required, only the complications of diseases for targeted prevention efforts.\(^73\) Other examples of recently emerging methods assessed the impact of risk factors by quantifying the contributing causes of death (using weighting) among known AIDS cases for comparison of causes of death by exposure to injecting drug use,\(^71\) assessing the impact on mortality trends of deaths reporting the presence of conditions known to be associated with obesity,\(^61\) and by using cluster analysis techniques to assess mortality differentials between clusters of discrete groupings of causes associated with obesity.\(^64\)

**Discussion**

Our review and audit confirm rapidly increasing international interest in the use of multiple causes of death data in mortality research and provides the most comprehensive overview of methods and practices to date. Most articles identified were published since 2011 and were from countries with health information management systems enabling collection and recording of multiple causes of death; the United States, Brazil, the United Kingdom, Italy, and France being the most common. Articles analysing multiple causes of death were highly...
skewed towards those using descriptive measures applied to any-mention of a cause of death and those assessing pairwise contribution of causes to mortality. Cluster analysis techniques and weighting multiple causes were found to be newly emerging applications, applied only since 2016 and used in less than 3% of articles. Irrespective of the methodological approach, all the included articles demonstrated that multiple cause analysis complements the single UC approach by increasing the descriptive utility of the data and improving the quantification of causal attribution to mortality at both individual and population level.

Improved data quality and access to population level mortality data from vital registration systems may have facilitated increased use of multiple causes data over time. The review showed that for descriptive research questions, simple measures such as the numbers of deaths as well as death proportions and rates based on any-mention of a specific disease or condition may be sufficient in many cases. However, the complexity in structure of multiple causes data in terms of differentiating between causes listed in Part 1 and Part II of the death certificate, the application of rules for selection of the UC of death, and variations in death certification practices between countries and over time, pose challenges in interpreting analyses using multiple causes. The high frequency of analyses on drug-related mortality among the included articles likely arises as the ICD-10 coding mechanisms stipulate that the manner of death (intentional/unintentional poisoning) is routinely reported as the UC with the drug class represented by the associated causes; thus requiring analysis of the multiple causes.

Our audit of international experts coincided with a peak in the COVID-19 cases in Europe; the reliance on these public health experts during a period of heavy workload and uncertainty, may have contributed to the low response rate. Furthermore, analysis of multiple cause of death requires data from vital registrations, which are not available in all countries, thereby limiting the audit to high-income countries.

We could not identify a previous review capturing all contemporary methods used to assess multiple causes of death. Of the four narrative reviews identified from our search: two were published before 1990 at which time there were 47 articles (81% based on US deaths data) and the other two were published over a decade ago and neither assessed the range of analytical methods that had been used. Only methods based on any-mention and pairwise assessment of causes were captured in the existing reviews.
The four groups of methods that we identified are distinct but complementary, and each has important practical applications. For instance, descriptive analyses based on any-mention provide useful contextual information to highlight the extent to which multiple causes play a role either as part of a causal sequence or through association between various causes listed on the death certificate. More specifically, while the SRMU is a descriptive measure, it reveals the extent of the potential contribution to mortality of causes that are not deemed to be the UC.13,32,77

However, changes in coding practices over time may influence the magnitude of underlying cause mortality from death certificates. This was observed in the case of diabetes4 and other conditions29,78,79 from the transition between the 9th and 10th revisions on the ICD; hence analyses of ‘any-mention’ rates might offer a better perspective of mortality trends, than rates based on the UC. While cause-related rates based on any-mention are simple to apply, they can be misleading in the context of overall mortality as each death is counted as many times as there are mentions.69

Assessing pairwise contribution to mortality is advantageous when investigating the relationships between two causes of death. Unlike basic summary statistics that consider one cause at a time, these methods consider two causes simultaneously; they were used frequently to evaluate external cause epidemiology, specifically to assess the nature of poisoning and injury related to exposure to drugs and alcohol.80-84 While the nature of the coding process mandates the use of multiple causes of death for assessing injuries sustained from external causes, descriptive pairwise analysis was regularly applied to infectious diseases85-89 and chronic diseases.3,90-96

More complex applications such as the assessment of mortality odds ratios were applied to assess the strength of relationships between disease on the death certificates. (e.g.49,59,97-100 H the disease associations based on deaths data alone do not imply causation; they evaluate whether the presence of a particular cause increases or decreases the probability of death from another cause,14,47 and statistical tests pertaining to strength of association between the causes should be interpreted with some caution13 as statistical assumptions, for example, of independence of causes, may not be valid because both causes contributed to death. Indeed, the associations between several causes mentioned on the death certificates are more frequent than would be expected by a random occurrence of the causes.13 Further, nearly all estimates of the associations between causes mentioned on the death certificates may be influenced by some level of Berkson's paradox or collider bias.101,102 This is a form of selection bias that
occurs when both the exposure and outcome variables (the two causes of death) influence the inclusion of participants in a study (death certificate data).

Methods for grouping causes of death facilitate assessment of the complex relationships between causes that may go unnoticed by pairwise analysis. The pattern of diseases and risks leading to death may not be homogenous for the whole population, and methods that cluster deaths according to patterns in contributing causes can identify groups of individuals with specific combination of causes of death providing additional insight for setting targeted preventative interventions. A challenge with cluster analysis methods for grouping causes of death is that they are exploratory in nature and different clustering algorithms may result in different outcomes. Being data driven, the characteristics of the clusters are limited to the data, making it difficult to generalise the findings to populations in other settings. Furthermore, applications such as social network analysis describe the ‘closeness’ of causes of death, but further quantification of cause-related mortality is not possible.

On the other hand, user-defined groupings of causes, for example, those based on known disease-risk associations, are especially useful for examining the joint contribution and impact on mortality trends that cannot be captured by pairwise comparison.

The release of the ICD-11 for current implementation worldwide offers additional flexibility designed to enhance the evidence for informing better health systems. Of specific significance is the ‘post-coordination’ feature which allows combining specific codes into a cluster of relevant clinical attributes. While specific applications have not yet been defined for mortality, future application of multiple cause methods that group diseases based on multiple cause patterns or according to known disease-risk associations, could provide evidence to inform meaningful clusters for application in ICD-11.

Multiple-cause weighting methods have been developed to facilitate the measurement of overall magnitude of the contribution of a specific cause to population levels of mortality. A major advantage of these methods is the preservation of the counting unit (deaths) enabling derivation of a broad range of mortality indicators based on the weighted counts of each cause (e.g. age-standardised rates, years of life lost). This method overcomes the limitation of double-counting deaths for as many causes as present when using any-mention approaches. While the arbitrary nature of weights in MC-weighting strategies poses a limitation, the incorporation of methods for considering causal pathways of diseases by weighting only the UC and causes in Part II of the death certificate, and application of multiple cause weighting
methods require careful consideration of the weighting strategy, the cause list and the handling of ill-defined causes, these methods offer a richer perspective for population health monitoring.

The choice of the methods for analysing MC is dependent on the research question of interest. Irrespective of the methodological approach, MC analysis complements the single UC approach, uses useful information that is usually ignored and offers an additional perspective of the causes that contribute to death. Though there are distinct limitations around individual multiple cause methods, the broad range of methods described here offer a toolkit, which in combination can offer a richer perspective for population health monitoring and policy development.

To the best of our knowledge, this study is the largest review to date, comprehensively capturing statistical methods used for analysing multiple cause of death data, including many papers (>430) with two independent reviewers supplemented by an audit of international experts. A further strength is the use of a systematic approach to identify relevant studies, considering papers published in multiple languages. The inclusion of articles from a range of years, countries and languages revealed the breadth and diversity of applications of MC analysis.

This review is atypical of systematic reviews in that non-significant results and publication bias were not relevant. Publications largely represent data from countries with national vital statistics collections.

The very specific methods are not always apparent in the articles included here. For example, there is often little transparency around how duplicate mentions of causes and ill-defined causes are handled. As well, the terminologies used can be inconsistent (for example, contributing causes is often used to refer to non-UCs, but can also have a specific meaning referring to the causes reported in Part II of the death certificate). Additionally, factors that are known to affect multiple causes statistics such as the size and structure of the deaths certificate are not always apparent.

Our review showed that infectious diseases commonly assessed using multiple cause methods, for example, to ascertain socio-demographic differentials, to identify associated health conditions, and to assess the impact of health interventions to inform targeted prevention strategies. With new and emerging infectious diseases, MC data is crucial for descriptive epidemiology and for providing evidence to inform prevention...
strategies. Since declaration of COVID-19 as a pandemic in March 2020, the WHO implemented rules for ascertaining when a death was due to COVID-19 (i.e. the UC) noting that under certain circumstances COVID-19 should be recorded somewhere on the death certificate.\textsuperscript{110} While we identified only one COVID-related article (due to the timing of our database search), MC methods were since applied to ascertain associations between contributing conditions and complications, assess changes in the pathological patterns, and identify significant sociodemographic variation in COVID-related deaths,\textsuperscript{111-114} leading in some circumstances to improved survival.\textsuperscript{115} Importantly, the recent emergence of COVID-19 as a leading UC\textsuperscript{116} may significantly alter proportional mortality from other UCs. As such future analyses of multiple causes will be necessary for monitoring trends in COVID-related mortality as well as the impact of COVID-19 on other causes of death. As well, future sensitivity analyses that assess the impact of removing the non-underlying condition (by varying the weight ascribed to the UC)\textsuperscript{69} may facilitate assessment of competing causes of death where COVID-19 has become a major UC.

**Conclusion**

The results from this review confirm that international interest is ongoing and increasing. This review provides the most comprehensive overview of multiple cause analytical methods and practices. The diversity of methods offers a toolkit for the analysis of these data which are becomingly increasingly important for understanding the complex involvement of multiple diseases in causing death across a range of settings including surveillance, policy, planning and research.


Figure 1. Layout of a standard international form of Medical Certificate of Cause of Death

PART I
Disease or condition directly leading to death

Antecedent causes
The conditions, if any, giving rise to the above cause, with the underlying condition stated last

PART II
Other significant conditions contributing to the death, but not related to the disease or condition causing it

Figure 2. Study flow diagram

Articles identified (n=8,070)
Database search (n=8,002) / Audit (n=68) → Duplicates removed (n=3,990)

Articles for title and abstract screen (n=4,080)
Database search (n=4,053) / Audit (n=27) → Excluded by title and abstract screen (n=3,503)

Articles for full text assessment (n=577)
Database search (n=550) / Audit (n=27) → Full text articles excluded (n=148)
• Not MCOD study (n=52)
• Irretrievable (n=35)
• Foreign language & irretrievable (n=13)
• Conference abstract (n=17)
• Communication (n=10)
• Not peer-reviewed (n=17)
• Review article (n=4)

Review article citations for full text review (n=25) → Full text review citations excluded (n=20)
• Not MCOD study (n=9)
• Irretrievable (n=5)
• Conference abstract (n=1)
• Not peer-reviewed (n=5)

Articles included (n=434)
Database search (n=421) / Audit (n=8) / Handsearch (n=5)

† Due to high volume and limited resources, 101 articles that appeared to only apply a method based on any-mention were excluded without full text review.
Table 1. Characteristics of articles included in the study

<table>
<thead>
<tr>
<th></th>
<th>Number</th>
<th>Percent</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Total number of articles</strong></td>
<td>434</td>
<td></td>
</tr>
<tr>
<td><strong>Year published</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1980 or earlier</td>
<td>19</td>
<td>4.4</td>
</tr>
<tr>
<td>1981-1990</td>
<td>29</td>
<td>6.7</td>
</tr>
<tr>
<td>1991-2000</td>
<td>54</td>
<td>12.4</td>
</tr>
<tr>
<td>2001-2010</td>
<td>114</td>
<td>26.3</td>
</tr>
<tr>
<td>2011 to present</td>
<td>218</td>
<td>50.2</td>
</tr>
<tr>
<td><strong>Country of data</strong>¹</td>
<td></td>
<td></td>
</tr>
<tr>
<td>United States</td>
<td>222</td>
<td>12.9</td>
</tr>
<tr>
<td>Brazil</td>
<td>47</td>
<td>2.7</td>
</tr>
<tr>
<td>United Kingdom</td>
<td>43</td>
<td>2.5</td>
</tr>
<tr>
<td>Italy</td>
<td>30</td>
<td>1.7</td>
</tr>
<tr>
<td>France</td>
<td>23</td>
<td>1.3</td>
</tr>
<tr>
<td>Sweden</td>
<td>19</td>
<td>1.1</td>
</tr>
<tr>
<td>Australia</td>
<td>19</td>
<td>1.1</td>
</tr>
<tr>
<td>Canada</td>
<td>7</td>
<td>0.4</td>
</tr>
<tr>
<td>Finland</td>
<td>7</td>
<td>0.4</td>
</tr>
<tr>
<td>Norway</td>
<td>6</td>
<td>0.3</td>
</tr>
<tr>
<td>All other countries</td>
<td>50</td>
<td>2.9</td>
</tr>
<tr>
<td><strong>Research Focus</strong>¹</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Population level analysis of all causes</td>
<td>44</td>
<td>10.1</td>
</tr>
<tr>
<td>Specific causes of death</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Infectious diseases</td>
<td>76</td>
<td>17.5</td>
</tr>
<tr>
<td>HIV/AIDS</td>
<td>22</td>
<td>5.1</td>
</tr>
<tr>
<td>Hepatitis</td>
<td>9</td>
<td>2.1</td>
</tr>
<tr>
<td>Sepsis</td>
<td>8</td>
<td>1.8</td>
</tr>
<tr>
<td>External causes</td>
<td>69</td>
<td>15.9</td>
</tr>
<tr>
<td>Drug-related</td>
<td>23</td>
<td>5.3</td>
</tr>
<tr>
<td>Alcohol-related</td>
<td>15</td>
<td>3.5</td>
</tr>
<tr>
<td>Cardiovascular diseases</td>
<td>50</td>
<td>11.5</td>
</tr>
<tr>
<td>Endocrine diseases</td>
<td>43</td>
<td>9.9</td>
</tr>
<tr>
<td>Diabetes</td>
<td>37</td>
<td>8.5</td>
</tr>
<tr>
<td>Neurological diseases</td>
<td>35</td>
<td>8.1</td>
</tr>
<tr>
<td>Cancer</td>
<td>25</td>
<td>5.8</td>
</tr>
<tr>
<td>Respiratory diseases</td>
<td>23</td>
<td>5.3</td>
</tr>
<tr>
<td>Digestive diseases</td>
<td>13</td>
<td>3.0</td>
</tr>
<tr>
<td>Musculoskeletal conditions</td>
<td>12</td>
<td>2.8</td>
</tr>
<tr>
<td>Mental conditions</td>
<td>11</td>
<td>2.5</td>
</tr>
<tr>
<td>All other causes</td>
<td>33</td>
<td>7.6</td>
</tr>
<tr>
<td><strong>Study design</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cross-sectional</td>
<td>352</td>
<td>81.1</td>
</tr>
<tr>
<td>Cohort / longitudinal</td>
<td>69</td>
<td>15.9</td>
</tr>
<tr>
<td>Case control</td>
<td>7</td>
<td>1.6</td>
</tr>
<tr>
<td>Other</td>
<td>6</td>
<td>1.4</td>
</tr>
<tr>
<td><strong>Language of article</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>English</td>
<td>416</td>
<td>95.9</td>
</tr>
<tr>
<td>Portuguese</td>
<td>13</td>
<td>3.0</td>
</tr>
<tr>
<td>Other</td>
<td>5</td>
<td>1.2</td>
</tr>
</tbody>
</table>

¹ Numbers may not sum to the total as multiple categories apply.
Figure 3. Number of articles by country and year of publication
Table 2. Summary of approaches to assess multiple causes of death in included studies

<table>
<thead>
<tr>
<th>Methodological approach</th>
<th>Example articles</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>1. Methods based on any-mention (n=377, 86.9%)</strong></td>
<td></td>
</tr>
<tr>
<td>Uses: to assess the extent of multiple causes and compare measures based on UC to measures based on MC</td>
<td>Wall (2005)</td>
</tr>
<tr>
<td>Examples: average number of causes; rates by ‘any-mention’; comparison of underlying to multiple causes (e.g. SRMU)</td>
<td>Goldacre (2006)</td>
</tr>
<tr>
<td>Strengths: identifies causes less visible using UC approach</td>
<td>Desesquelles (2014)</td>
</tr>
<tr>
<td>Limitations: metrics based on any-mention inflate mortality estimates as deaths are counted more than once</td>
<td>Goldberger (2015)</td>
</tr>
<tr>
<td></td>
<td>Zoppini (2018)</td>
</tr>
<tr>
<td></td>
<td>Simmons (2019)</td>
</tr>
<tr>
<td></td>
<td>Sampaio (2020)</td>
</tr>
<tr>
<td></td>
<td>Cano (2020)</td>
</tr>
</tbody>
</table>

| **2. Methods to assess pairwise occurrence of causes (n=245, 56.5%)** | |
| Uses: to identify and assess co-contributing causes of death | Redelings (2005) |
| Examples: leading underlying causes for a specific associated cause; mortality odds ratios for the presence of specific causes | Rockett (2007) |
| Strengths: identifies most commonly occurring comorbid conditions at death; measures the strength of association between two co-contributing causes on the death certificate (e.g. odds ratio, CDAI) | Redelings (2007) |
| Limitations: does not measure causality; can overlook relationships between two non-underlying causes | Desesquelles (2012) |
| | Duncan (2014) |
| | Chazal (2018) |
| | Turner (2018) |
| | Quast (2020) |

| **3. Methods based on groups of >2 co-contributing causes (n=9, 2.1%)** | |
| Uses: to identify frequently co-occurring causes and assess mortality trends due to grouped causes | Stallard (2002) |
| Examples: combine causes based on known risk associations or cluster analysis, social network analysis and data mining according to patterns in multiple causes | Frova (2009) |
| Strengths: identifies highly correlated diseases and risk conditions among large datasets | Yoon (2011) |
| Limitations: data-driven methods are difficult to replicate or apply across different settings | Barbieri (2017) |
| | Hassanzadeh (2017) |
| | Jiang (2017) |
| | Egidi (2018) |
| | Villela (2018) |
| | Adair (2020) |

| **4. Methods based on weighting of multiple causes (n=5, 1.1%)** | |
| Uses: to re-evaluate mortality metrics based on weighted multiple causes | Piffaretti (2016) |
| Examples: ascribing fixed or variable weights to each cause such that the weights in each death sum to one | Moreno-Betancur (2017) |
| Strengths: can prescribe causal responsibility to selected diseases (e.g. antecedent conditions); incorporates a mortality contribution of all relevant conditions | González (2019) |
| Limitations: difficult to ascribe causal responsibility across cause of death data so weights are arbitrary | Breger (2020) |
| | Xie (2020) |

---

1. See Supplementary Appendix 2 for full reference
2. Articles that applied multiple cause weighting methods and data-driven methods for grouping deaths based on multiple cause patterns were all applied since 2016