



**INTERNATIONAL WORKSHOP
ON THE MULTIPLE CAUSE-OF-DEATH
ANALYSIS**

INED, Salle Sauvy, 1st floor

22&23 November, 2012

Thursday November, 22nd

☪ 9:30 – 10:00 am Welcome coffee and Tour de table

10:00-11:00 am OPENING SESSION - Chair: Viviana Egidi

International workshop on the multiple cause-of-death analysis: why and what for?
Aline Désesquelles (INED)

Causes of death statistics: needs for multiple causes data in an ageing society
Monica Pace, Elodie Cayotte & Hartmut Buchow (Eurostat, Directorate F – Unit F5)

☪ 11:00-11:15 am Coffee Break

11:15-12:45 pm How to use the data? - Chair: France Meslé

Using the multiple cause-of-death approach in a cross-country perspective: a review of methodological issues and answers
Viviana Egidi (La Sapienza Università) et le groupe de recherche franco-italien sur les causes multiples

Analyzing mortality differentials by multiple causes of death
Isaac W. Eberstein, Janet Weeks & Charles B. Nam (Florida State University)

Selection in high-dimensional binary graphical models. Application to the search of associations among causes of death in French death certificates
Vivian Viallon (Université Lyon 1)

🕒 12:45 - 2:00 pm Lunch break

2:00 - 3:30 pm Data collection - Chair: Monica Pace

The revision of the ICD10 and the multiple causes of death
Robert Jakob (WHO) (Video presentation with SKYPE)

Coding multiple cause of death: some basic concepts for a better understanding of the data.
Francesco Grippo (Istat) & Lars Age Johansson (National Board of Health and Welfare, Sweden & Nordic Centre for Classifications in Health Care)

The Iris international coding system of the causes of death
Gérard Pavillon (CépiDc, Inserm) & Lars Age Johansson (National Board of Health and Welfare, Sweden & Nordic Centre for Classifications in Health Care)

☪ 3:30 - 3:45 pm Break

3:45 - 5:15 pm First outlooks into the data - Chair: Aline Désesquelles

Multiple causes in the Czech Republic: registration, availability, first outlooks
Marketa Pechholdova (University of Economics, Prague)

Multiple cause-of-death data in the United States
Magali Barbieri (INED/University of Berkeley) & Irma Elo (University of Pennsylvania)

Potential sources for the analysis of the Spanish mortality using the multiple cause-of-death method
Rosa Gómez-Redondo (UNED), Oscar Zurriaga (CSISP), Immaculada Melchor (Alicante University), Joaquín Moncho (Alicante University)

Friday November, 23rd

9:30-11:00 am More results (1)! – Chair: Marilena Pappagallo

Exploring multiple causes for infant deaths in Italy

Enrico Grande, Francesco Grippo & Monica Pace (Istat) - *Monica Pace is currently at Eurostat*

The distribution of garbage codes among the underlying cause and other diagnoses and conditions mentioned the death certificate – Norway 1996-2010

Stein Emil Vollset et al. (Norwegian Institute of Public Health and University of Bergen)

Use of death certificates to assess survival and causes of death among people with AIDS in Italy, 1999-2009

Luisa Frova, Francesco Grippo, Monica Pace (Istat), Stefano Boros, Laura Camoni, Barbara Suligo (ISS, Rome), Luigino Dal Maso, Diego Serraino & Antonella Zucchetto (CRO, Aviano) - *Monica Pace is currently at Eurostat*

☪ 11:00-11:15 am Coffee Break

11:15-12:45 More results (2)! – Chair: Luisa Frova

Co-morbidity and the burden of disease in England and Wales: what can be learned from multiple causes of death analysis?

Emily Grundy (Cambridge University) & Christopher Marshall (London School of Hygiene and Tropical Medicine)

Analysis of multiple causes of death in the elderly population of Québec (Canada), 2000-2005

Robert Bourbeau & Allison Blagrove (Department of Demography, Université de Montréal)¹

A comparison of the mortality due to Parkinson's disease, Alzheimer's disease, and other senile dementias of France and Italy using the multiple cause-of-death approach

Aline Désesquelles, France Meslé (INED), Elena Demuru, Viviana Egidi (La Sapienza Università), Luisa Frova, Monica Pace, Marilena Pappagallo & Michele Antonio Salvatore (Istat) - *Monica Pace is currently at Eurostat*

☪ 12:45-2:00 pm Lunch break

2:00-5:00 pm Debriefing and perspectives

¹ Robert Bourbeau cannot attend the meeting. The presentation will be made by Jacques Légaré (Université de Montréal).

Abstracts

International workshop on the multiple cause-of-death analysis: why and what for?

Aline Désesquelles (INED)

In 1940 Theodore Janssen, the then Chief of the Nosology Section at the Division of Vital Statistics of the US Bureau of the Census wrote: “statistics showing combinations of causes come nearer the truth than do those based on the single cause principle because the majority of deaths actually result from a combination of causes”. As we enter the 21st century, cause-specific mortality analysis is still primarily achieved on the basis of the underlying cause of death. Yet, multiple cause-of-death (MCOD) data are increasingly produced and disseminated by statistical offices. These developments encouraged us to engage in a comparative research using the French and the Italian MCODE data. In the light of this experience, our presentation aims at explaining why we believe that the cause-specific mortality analysis should not account for the underlying cause of death only but also for all contributing causes. But the method is still in its infancy. Methodological choices regarding the use of the MCODE data must be discussed largely within the scientific community, and there is room for new developments. Apart from data availability issues, one of the reasons why MCODE data has not been analysed as extensively as it could be is the quite widely shared scepticism about the quality of the data. Efforts towards better data quality (e.g. harmonization of the death certificates, better account taken in the physicians’ training of the specific problems related to the reporting of the multiple causes) are needed.

Causes of death statistics: needs for multiple causes data in an ageing society

Monica Pace, Elodie Cayotte & Hartmut Buchow (Eurostat, Directorate F – Unit F5)

Background: Causes of death statistics are currently based on the major disease or condition as reported on the death certificate (DC). However, since the mid-eighties we know that the DC “hides” a much higher quantity of information because it usually contains several diagnoses reported by the certifying professional (multiple causes of death, MCODE), which are lost for analyses due to current coding and/or tabulation schemes.

State of art: Over recent years Eurostat has promoted the methodological improvement of public health statistics through the "Partnership Health" with EU-27 Member States, EFTA and candidate countries. The MCODE approach has constantly been on the road map and information on MCODE data availability in EU member states is currently updated and available on Circa.

In times of an ageing society the lack of knowledge on MCODE is an imminent problem in view of analysing the major cause of death among a multiple set of diseases that contribute to the death of the elderly. In addition, for analysing productive years of life lost or avoidable deaths internationally comparable statistics on MCODE would be a major step forward.

Future perspectives: The improvement of internationally comparable EU statistics by MCODE will be linked to several aspects:

1. the national implementation of automated coding systems (ACS) that had been financially supported by Eurostat from 2009 to 2011;
2. the establishment of a set of internationally agreed rules and methodologies on coding and data processing, and
3. an agreed list of indicators for dissemination and the appropriate level of data aggregation for research purposes.

Analyzing mortality differentials by multiple causes of death

Isaac W. Eberstein, Janet Weeks & Charles B. Nam (Florida State University)

This paper reports illustrative results from an examination of cause-of-death specific adult mortality with a particular emphasis on the problems and potential of using information on multiple causes of death (MCODE). Data are taken from the public use version of the US National Health Interview Survey for 1997-2004, linked to the National Death Index through 2006 for persons ages 45 and older, which includes 29,029 death records corresponding to a survey population of 222,727. The public use files contain complete information on underlying cause of death (UCODE) but only limited information on three associated causes due to concerns over confidentiality: diabetes, hypertension, and hip fractures. We make use of the MCODE information on diabetes and hypertension and have three specific objectives: 1/ Compare unadjusted and age-adjusted mortality rates by UCODE and MCODE, according to sex and race/ ethnicity. In this portion of the analysis we arrange the MCODE information to identify both the total mentions of diabetes and hypertension and two alternative strategies of combining these causes, within and across reported underlying cause, and we compare the pattern of mortality differentials indicated by UCODE and MCODE across these social categories. 2/ Compare these mortality differentials in a multivariate context after adjusting for a range

of important demographic, socioeconomic, behavioral, and health characteristics. In this portion of the analysis we utilize a framework of discrete-time hazards models with multi-category outcomes corresponding to the various combinations from the MCODE data. In this way we initially identify the relative risk of MCODE-specific mortality relative to survival, adjusted for important controls. We also rearrange the analysis to directly assess the pattern and significance of observed differences in relative risks across the various MCODE categories, net of compositional differences in other variables examined. 3/ Compare medical conditions self-reported by the respondents in the baseline survey with causes of death as reported in the MCODE records. This portion of the analysis is descriptive and points to significant concerns about the precision, reliability, and validity of the MCODE records. Overall, we highlight what are mainly consistent findings across UCOD and MCODE and, further, across the various MCODE approaches. However, we do identify differences across these strategies for measuring cause of death in terms of the magnitude and, occasionally, the direction of observed mortality differentials, whether represented by sex- and race/ethnicity-specific mortality rates (from aim 1) or from adjusted relative risk ratios (from aim 2), and it is clear that these differences may be important for specific substantive applications or analytical comparisons. In addition, we use the descriptive data (from aim 3) to raise questions of data quality that must be addressed before MCODE-based analysis can reach its full potential for helping to understand health disparities among older adults in high income countries.

Selection in high-dimensional binary graphical models. Application to the search of associations among causes of death in French death certificates

Vivian Viallon (Université Lyon 1)

Looking for associations among multiple variables is a topical issue in statistics due to the increasing amount of data encountered in biology, medicine and many other domains involving statistical applications. Graphical models have recently gained popularity for this purpose in the statistical literature. Following the ideas of the LASSO procedure designed for the linear regression framework, recent developments dealing with graphical model selection have been based on l_1 -penalization. In the binary case, however, exact inference is generally very slow or even intractable because of the form of the so-called log-partition function. Various approximate methods have recently been proposed in the literature. We compare them through an extensive simulation study. We apply the retained method on a real application example, in which we search for associations among causes of death recorded on French death certificates.

Coding multiple cause of death: some basic concepts for a better understanding of the data

Francesco Grippo (Istat) & Lars Age Johansson (National Board of Health and Welfare, Sweden & Nordic Centre for Classifications in Health Care)

Multiple cause analysis is an innovative approach to the study of mortality, which allow to take account of all the conditions causing death and their relationship. Nevertheless, while the process of producing information is well documented for the underlying cause, little is known on how multiple cause coding is performed. In fact, multiple cause coding is not a simple translation of medical text into Icd- codes, but the application of rules is needed in order to provide a consistent information and allow the selection of the right underlying cause.

In the presentation, some basic rules of multiple cause coding will be discussed and their impact in the production data will be shown. Some actual cases will be discussed, also with reference to the behaviour of traditional and innovative systems for automatic coding.

Multiple cause-of-death data in the United States

Magali Barbieri (INED/University of Berkeley) & Irma Elo (University of Pennsylvania)

This presentation will give a brief overview of the availability and use of multiple cause-of-death data in the United States. We will give select examples of types of analyses undertaken and the methods used in studies conducted in the last 30 some years. In addition, we present some preliminary results from analyses based on multiple cause of death data from NCHS vital statistics data for 2009.

Potential sources for the analysis of the Spanish mortality using the multiple cause-of-death method

Rosa Gómez-Redondo (UNED), Oscar Zurriaga (CSISP), Immaculada Melchor (Alicante University), Joaquín Moncho (Alicante University)

We present a preliminary work about the potential basic sources and their limits to analyse the mortality profile of Spain, regarding the multiple-cause-of-death approach during the last decade. Following the case of one Autonomous Community (Valencia), our presentation shows the analytical possibilities of the multiple-cause-of-death method to properly follow a substantial change observed in hypertension and diabetes at the beginning of the 2000's. This example offers information about the dissimilarities found in the results using this approach respect to the generalized basic cause of death, which supports the reliability of the multiple-cause-of-death method.

Exploring multiple causes for infant deaths in Italy

Enrico Grande, Francesco Grippo & Monica Pace (Istat) - *Monica Pace is currently at Eurostat*

Background: Infant mortality in Italy is among the lowest in Europe, especially for infants aged 28 days or more. Although the trend for this major public health indicator is decreasing, some difference are still observed as for area of residence, citizenship and other demographic variables. These differences are well documented for the overall mortality, but little is known about causes of death.

The objective of this study is to explore the patterns of mortality by cause, according to some demographic variables. A multiple cause approach has been used in order to evaluate the performance of this kind of analysis compared to the use of the underlying cause only in defining cause-specific mortality profiles.

Data and methods: Data refer to infant deaths (within 1 year of age) in Italy in the period 2007-2009. Underlying and multiple causes of death, coded according to the Icd-10, have been grouped into major categories. Multiple Correspondence Analysis (MCA) has been performed in order to explore spatial association among underlying causes, other causes reported on the death certificates and demographic variables. In order to take account of different specificity in causes of death, separate analysis has been carried out on two different age groups: neonatal deaths (less than 28 days of age) and infant deaths (from 28 up to one year of age), based on different groups of causes.

Results: Some relevant associations of causes of death with both age of decedent and mother's area of residence have been noted. In particular, decedent aged less than 28 days are associated with maternal and perinatal causes; older ages with congenital malformations of circulatory system and other causes not originating in the perinatal period.

Discussion: Results from the MCA suggest that the inclusion of multiple causes provides a major contribution for the characterization of profiles for specific demographic subgroups. Moreover, this approach allows to study the associations among underlying and multiple causes.

The distribution of garbage codes among the underlying cause and other diagnoses and conditions mentioned the death certificate – Norway 1996-2010

Stein Emil Vollset et al. (Norwegian Institute of Public Health and University of Bergen)

Following the typology of garbage codes for the underlying cause of death described by Naghavi et al (Pop Health Metrics 2010), we study the distribution of garbage codes over calendar time, by age and gender and other relevant characteristics of the examination of the death, including number and type of other diagnoses and conditions mentioned on the death certificate. The most common "garbage code" underlying cause of death in Norway during the period 1996-2010 was heart failure (I50). Using a redistribution approach we will investigate the occurrence of other diagnoses on the death certificate that could provide a more informative underlying cause of death. The presentation will include a description of how multiple causes of death have been registered, coded and made available to researchers and other users of death certificates for statistical purposes. In Norway this practice has varied over time and changed with the introduction of ACME in 2005.

Use of death certificates to assess survival and causes of death among people with AIDS in Italy, 1999-2009

Luisa Frova, Francesco Grippo, Monica Pace (Istat), Stefano Boros, Laura Camoni, Barbara Suligo (ISS, Rome), Luigino Dal Maso, Diego Serraino & Antonella Zucchetto (CRO, Aviano) - *Monica Pace is currently at Eurostat*

The introduction in the mid '90s of highly active antiretroviral therapies (HAART) has substantially prolonged the survival of people with HIV/AIDS. However, people with AIDS (PWA) are still at higher risk of death, as compared to the age-matched general population, for several concomitant conditions, particularly for infectious diseases and cancer. In the specialized medical literature, investigations on the causes of death of PWA span from the description of conditions listed in death certificates (i.e., perimortal conditions) to the estimates of cause specific excess risks of death (i.e., standardized mortality ratios – SMR).

In this proposed presentation, we intend: i) to illustrate the research work already carried out, in Italy, on the causes of death of PWA who died between 1999 and 2006; and ii) to discuss new areas of epidemiologic research based on the use of death certificates up to 2009.

With regard to studies already published, all death certificates provided by ISTAT on files of people who died, in Italy between 1999 and 2006, were linked (using an automated standardized record linkage procedure and manual quality control) with the National AIDS Registry (RNAIDS) to identify PWA deceased in those years. This record linkage procedure allowed us to identify 3209 deaths among 10 392 PWA diagnosed between 1999 and 2006 and to address topics like:

- 1) the survival of PWA after the diagnosis of AIDS, and the comparison of the role of AIDS-associated cancers on survival (Serraino D et al, *J Acquir Immune Defic Syndr* 2009, 52:99-105);
- 2) proportions of death of PWA due to Kaposi sarcoma and non-Hodgkin lymphomas (Serraino D et al, *Cancer Epidemiol* 2010, 34:257-261) or to chronic conditions like diabetes mellitus and cardiovascular diseases (Serraino D et al, *AIDS Res Ther* 2010, 24:7-11)
- 3) the excess mortality (i.e., estimates of SMRs) among PWA for cancers not included in the AIDS-definition (Zucchetto A et al, *Clin Infect Dis* 2010, 51:1099-1101).

New areas of research that this collaborative group intends to develop include the update -as of 2009- of the archives of death certificates and of AIDS cases (RNAIDS) for all of Italy; and the use of multiple causes of death to better describe the mortality patterns of PWA.

Co-morbidity and the burden of disease in England and Wales: what can be learned from multiple causes of death analysis?

Emily Grundy (Cambridge University) & Christopher Marshall (London School of Hygiene and Tropical Medicine)

Background: It is increasingly recognised that in ageing populations in which multiple morbidity is common, analyses of cause specific mortality based on one identified underlying or major cause may give a partial or misleading indication of the overall burden of disease. Some diseases, such as diabetes, may contribute significantly to morbidity levels and to death from a range of causes but be less often identified as the immediate or primary cause than, for example, myocardial infarction. Additionally, changes between revisions of the ICD or the rules accompanying them may result in marked changes in identification of 'leading' causes of death in the absence of any real change. Such changes have, for example, led to fluctuations in the proportions of deaths attributed to pneumonia and other respiratory diseases. This has led to increasing interest in use of multiple cause of death coding information as a way of providing broader picture of important causes of morbidity and mortality and the extent of co-morbidities. Previous studies using this approach have been undertaken in a number of countries. Comparisons between countries using this approach have revealed differences, some of which may be due to variations in coding practices and others of which may reflect real differences in morbidity patterns. Extending this comparative work to include a larger number of countries may provide some insight both this balance and into variation in patterns of disease. In this study we present results from analyses of data from the England and Wales Longitudinal Study which have been modelled on previous analyses conducted for France and Italy.

Data and Methods: We use data from the ONS Longitudinal Study, a record linkage study including individual level data from population censuses and vital registration records. The sample was drawn on the basis of birthday from those included in the 1971 Census of England & Wales with a sampling fraction of just over 1%. Records are linked to the National Health Service Central Register so that vital events can be captured. The study has been maintained by the addition of 1% of new births and immigrants and now

includes information from the 1971, 1981, 1991, and 2001 Censuses, with 2011 data being linked at the moment.

The study population for this analysis consisted of all LS members who were present at the 2001 Census, excluding students not at their term time address and those with multiple enumerations.

We examine deaths among this group between 1st January 2002 and the end of 2008. Those whose age was recorded as 100 or more at 2001 and for whom no date of death has since been recorded were excluded from the analysis.

The underlying cause of death was recoded into specific categories of disease using the ICD10 code within the database. Where the original underlying cause of death had been superseded by a new code (possibly following an inquest or full post mortem) then the underlying cause of death was recorded as being the latter cause. The ICD10 codes assigned to 15 underlying causes of death can be seen in Table 1. The same groupings were used when checking the recorded ‘contributory causes of death’ of which up to eight per LS member could be recorded.

Analysis: Analysis was undertaken using STATA. Cause specific mortality rates were derived and directly standardised using the European Standard population.

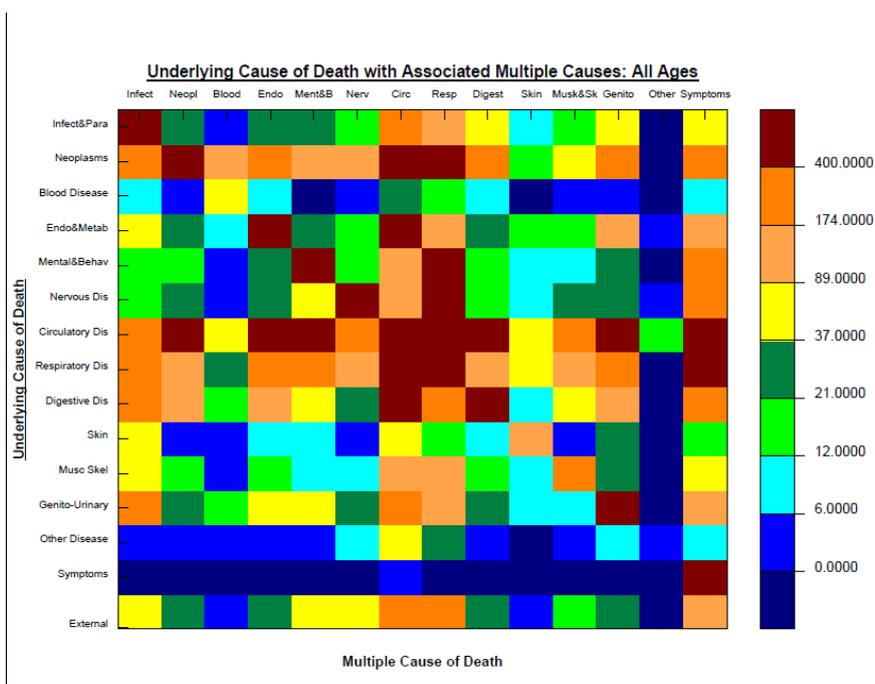
The above procedure was then repeated three times for the contributory causes of death. In the first instance the contributory cause group was only counted if the cause under consideration was not also the underlying cause of death group. In the second instance the contributory group cause was counted once only whether or not it was also the underlying cause of death group. In the third instance each occurrence of the contributory cause group was counted, again whether or not it was also the underlying cause of death group.

The four sets of rates now obtained were set against the European Standard population for the corresponding age group. The rate multiplied by the standard population for the age group was accumulated over all age groups and then divided by six to get the equivalent rate per year (as our data covered a six year period). This was the directly standardised rate for all ages for each cause.

Cross tabulations of the rates for the underlying cause of death by the rates for the associated multiple cause of death (counted once only no matter how many times it appeared) were created for each of three sub populations: (a) all subjects (b) all subjects aged 80 and over (c) all subjects aged under 80 years of age. These crosstabulations were saved as comma delimited files and the data used to create Lexis maps using Lexis 1.1.

Results: Preliminary results are given in Figures 1-3. These reveal the extent to which focussing on one cause only leads to masking of the contribution of other causes to the overall morbidity and mortality burden.

Discussion: Implications of these results, and comparisons with those for France and Italy will be discussed.



Analysis of multiple causes of death in the elderly population of Québec (Canada), 2000-2005

Robert Bourbeau & Allison Blagrove (Department of Demography, Université de Montréal)

Up to now, mortality analysis has primarily focused on the underlying cause of death who represents the disease or injury which initiated the sequence of morbid events leading to the death. But since death is due to a complex process, especially at advanced ages, analysis based solely on this concept has its limitations and some causes are more likely to be identified as the underlying cause than others. Selecting only one cause per death may influence the relative importance of the various causes of death. Multiple causes of death statistics provide a more complete view of mortality patterns. Although multiple causes of death for Québec have been compiled since 2000, only one analysis using these data has been produced so far (Blagrove², 2010). This paper will explore the analytical potential for multiple causes of death in Québec at age 65 and over using the International Classification of Diseases (10th revision). The findings of this article are that the most frequent causes of death, which are tumours and cardiovascular diseases, remain unchanged. Nevertheless, the method shows the importance of causes such as Alzheimer, diabetes and hypertension, causes which frequently appear as a contributing cause of death rather than as an underlying cause.

A comparison of the mortality due to Parkinson's disease, Alzheimer's disease, and other senile dementias of France and Italy using the multiple cause-of-death approach

Aline Désesquelles, France Meslé (INED), Elena Demuru, Viviana Egidi (La Sapienza Università), Luisa Frova, Monica Pace, Marilena Pappagallo & Michele Antonio Salvatore (Istat) - *Monica Pace is currently at Eurostat*

Background: Because death at old age is often the final stage of a long process involving several conditions, the multiple cause-of-death (MCOD) approach is relevant to study the mortality of ageing populations. We use this approach for two conditions emblematic of old age mortality, Parkinson's and Alzheimer's disease, and for two countries, France and Italy.

Method: Data are for year 2008 and for deaths over the age of 65. They are based on the information reported on the death certificates. We first compute the ratio (Standardized Ratio of Multiple to Underlying Cause or SRMU) between mortality rates accounting for all causes and for the underlying cause (UC) only. We then examine in which combinations of causes the diseases under study are frequently involved. The indicator we use (Cause-of-Death Association Indicator of CDAI) has been specifically developed for that purpose.

Results: When the UC only is accounted for, standardized mortality rates for the two diseases under study are higher in France. SRMU for Alzheimer's disease is 1.6 in Italy and 1.7 in France. SRMU for Parkinson's disease is higher (2.8 vs.1.9) indicating that this condition is more frequently reported as contributing cause than Alzheimer's disease, and it is more frequently the case in Italy than in France. After accounting for all mentions, mortality rates become higher in Italy.

The conditions that frequently contribute to the deaths due to Parkinson's and Alzheimer's disease, often reflect the circumstances surrounding the end of life of the people with these diseases that is characterized by bed confinement, loss of autonomy, and frailty. When the conditions under study are not identified as the UC, death is frequently attributed to cerebrovascular diseases and to external causes (mainly: accidental falls, fractures, and obstruction of the respiratory tract). Association is also high with diabetes mellitus and septicaemia.

Conclusion: The MCO approach helps highlighting the complex pathological profile that characterizes deaths in old age, and it provides a more comprehensive assessment of the burden of these diseases, which prevalence is rapidly increasing in ageing populations.

² Blagrove, Allison (2010). Causes multiples de décès chez les personnes âgées au Québec, 2000-2004. Mémoire de maîtrise, Département de démographie, Université de Montréal. <http://hdl.handle.net/1866/5083>