Differences in the causes of death of HIV-positive patients in a cohort study by data sources and coding algorithms

Victoria Hernando a,b, Paz Sobrino-Vegas a,b, M. Carmen Burriel c, Juan Berenguer d, Gemma Navarro e, Ignacio Santos f, Jesús Reparaz g, M. Angeles Martínez h, Antonio Antela i, Félix Gutiérrez j, Julia del Amo a,b, for CoRIS cohort

Objectives: To compare causes of death (CoDs) from two independent sources: National Basic Death File (NBDF) and deaths reported to the Spanish HIV Research cohort [Cohorte de adultos con infección por VIH de la Red de Investigación en SIDA CoRIS] and compare the two coding algorithms: International Classification of Diseases, 10th revision (ICD-10) and revised version of Coding Causes of Death in HIV (revised CoDe).

Methods: Between 2004 and 2008, CoDs were obtained from the cohort records (free text, multiple causes) and also from NBDF (ICD-10). CoDs from CoRIS were coded according to ICD-10 and revised CoDe by a panel. Deaths were compared by 13 disease groups: HIV/AIDS, liver diseases, malignancies, infections, cardiovascular, blood disorders, pulmonary, central nervous system, drug use, external, suicide, other causes and ill defined.

Results: There were 160 deaths. Concordance for the 13 groups was observed in 111 (69%) cases for the two sources and in 115 (72%) cases for the two coding algorithms. According to revised CoDe, the commonest CoDs were HIV/AIDS (53%), non-AIDS malignancies (11%) and liver related (9%), these percentages were similar, 57, 10 and 8%, respectively, for NBDF (coded as ICD-10). When using ICD-10 to code deaths in CoRIS, wherein HIV infection was known in everyone, the proportion of non-AIDS malignancies was 13%, liver-related accounted for 3%, while HIV/AIDS reached 70% due to liver-related, infections and ill-defined causes being coded as HIV/AIDS.

Conclusion: There is substantial variation in CoDs in HIV-infected persons according to sources and algorithms. ICD-10 in patients known to be HIV-positive overestimates HIV/AIDS-related deaths at the expense of underestimating liver-related diseases, infections and ill defined causes. CoDe seems as the best option for cohort studies.

Keywords: AIDS, causes of death, hepatitis, HIV, malignancies, mortality

AIDS 2012, 26:1829–1834
Introduction

Accurate information on causes of death (CoDs) is hard to obtain and is subject to selection and information biases. Most countries have national death registries that compile information from death certificates and code CoD by international classification rules. In Spain, the National Statistics Institute collects data on CoD according to the International Classification of Diseases, 10th revision (ICD-10). ICD-10 provides a set of rules to extract the maximum information from the death certificate, usually based on a single CoD [1]. ICD-10 was implemented in 1999 and introduced specific codes for HIV/AIDS which were not present in ICD-9 [2]. Cohort studies of HIV-positive patients have not traditionally used the ICD-10. Other algorithms, such as the Coding Causes of Death in HIV (CoDe) protocol are used instead [3,4]. CoDe was developed in 2004 by the Copenhagen HIV Programme (CHIP) (http://www.cphiv.dk) and collects extensive information on CoD which is examined in a centralized review process. CoDe is used by the Data Collection on Adverse events of Anti-HIV Drugs (D:A:D) and EuroSIDA, among others [5–7]. Unfortunately, as it is not always possible to collect detailed information, a simplified version of CoDe selection rules has been used by the Antiretroviral Therapy Cohort Collaboration (ART-CC) in what we call the ‘revised CoDe’ (http://www.art-cohort-collaboration.org).

CoRIS, the Cohort of the Spanish Network of Excellence on HIV/AIDS Research, collects information on multiple CoDs from reporting physicians. In 2008, we obtained CoDs for deceased cohort members from the National Statistics Institute, which provides single CoD coded with ICD-10. We hypothesize that CoD identified from each of source will vary substantially due either to sources and/or the coding rules. Our objectives are to compare the CoD obtained from two independent data sources using ICD-10: the National Basic Death File (NBDF) and reports to CoRIS provided directly by the physicians between 2004 and 2008. We also compare deaths reported in CoRIS using two different coding algorithms: the ICD-10 and the revised CoDe.

Methods

Sources of information

CoRIS

CoRIS is an open multicentre cohort of HIV-positive patients naïve to antiretroviral therapy from 32 centres from 12 of the 17 autonomous regions of Spain. The project was approved by the Institutional Ethics Review Boards and each participant is required to sign an informed consent form [8]. The variables collected include socio-demographic, epidemiological, clinical, treatment and mortality data; vital status, date of death and underlying CoD, as well as contributing causes as free text variables. For the purpose of this study, we coded CoDs according to two coding systems:

1. ICD-10 – we assigned a code to each death with the assistance of an ICD-10 coding expert familiar with coding rules.
2. Revised CoDe – we assigned a code to each death following CoDe protocol assisted by ART-CC.

National Statistics Institute: National Basic Death File

The NBDF provides information on the date and CoDs of all persons dying in Spain coded by the underlying CoDs in accordance with the ICD-10. The data are obtained from civil registries and from the National Statistics Institute itself through the Statistical Bulletin on Deaths which is compiled from death certificates.

Period of analysis

Data from the two sources were cross-matched in the first quarter of 2010 for individuals who had died between January 2004, when CoRIS was created, and December 2008, date of the last available update of the NBDF.

Data analysis

Thirteen clinical categories were created for the CoD in HIV-positive persons. These are described in Table 1, which also shows the individual ICD-10 and revised CoDe codes included in each category.

The CoD from each source – those reported to CoRIS and those included in the NBDF; and from each algorithms – ICD-10 rules and revised CoDe, for CoD reported to CoRIS– were compared according categories of diseases.

Results

Overall, 4687 patients, 11 491.28 person-years follow-up and 160 deaths were observed from January 2004 to December 2008, yielding a crude mortality rate of 1.39 per 100 person-years [95% confidence interval (CI) 1.19–1.62]. Of these 160 deaths, 87% (n = 139) were men, median age was 44 years [interquartile range (IQR) 37–51], 36% (n = 57) were injecting drug users, 38% (n = 60) heterosexuals, 18% (n = 29) MSM and 9% (n = 14) had others or unknown categories. The median CD4 cell count within 6 months of death for the 55% patients with data available was 117 cells/μl (IQR 39–259) and 67% received antiretroviral treatment.

Comparing cause of death in National Basic Death File and CoRIS using International Classification of Diseases, 10th revision

Comparing the CoD for the 13 categories from each source yielded concordance in 111 (69%) cases ($\kappa = 0.48$;
However, only 29 of the 82 HIV/AIDS-associated deaths (35%) were classified with exactly the same ICD-10 code in each source, so concordance for individual codes was very low, although moderate for the 13 categories. Of the 49 (31%) discordant CoD, 19 (39%) had a completely different CoD, but 14 (29%) would have been concordant if the information about HIV-positive status had been recorded in the Bulletin on Death. For example, if CoD is cirrhosis and HIV status is unknown, according to ICD-10, it gets coded as K74.6 (other and unspecified cirrhosis of liver), but if HIV status is positive, it gets coded as B23.8 (HIV disease resulting in other specified conditions).

Comparing cause of death in CoRIS using International Classification of Diseases, 10th revision and revised Coding Causes of Death in HIV

We compared the CoD according to ICD-10 rules and the revised CoDe algorithm and found concordance in 115 (72%) cases for the 13 categories ($\kappa = 0.54$; CI 95% 0.43–0.64). Of the 45 discordant causes, 33 (73%) were classified as ‘HIV/AIDS-associated causes’ by ICD-10, whereas 15 of these 33 (45%) are coded as poorly defined by revised CoDe, 10 (30%) are classified as liver-related and eight (24%) as infectious diseases. This is due to an ICD-10 coding rule that converts into B24 codes (HIV disease) all CoD which are unknown or ill defined in anyone who is known to be HIV infected. When the CoD is pneumonia in an HIV-positive person, the assigned ICD-10 code is B20, B20.9 or B24 (HIV/AIDS related), whereas revised CoDe considers CD4 cell counts prior to death: if CD4 cell count is below 100 cells/$\mu$m, it gets coded as 01 (AIDS), but if it is unknown or above 100 cells/$\mu$m, it gets coded as 02 (infection).

Distribution of the cause of death using different data sources and coding algorithms

‘HIV/AIDS-associated causes’ was the most frequent CoD irrespective of the source and coding algorithm, accounting for 91 (57%) of the causes in the NBDF, 112 (70%) in CoRIS_ICD-10 and 85 (53%) in CoRIS_CoDe. The second most common category was ‘non-AIDS-defining tumours’: 16 (10%) in the NBDF, 21 (13%) in CoRIS_ICD-10 and 18 (11%) in CoRIS_CoDe. ‘Liver diseases’ was the third most common cause in CoRIS_CoDe, accounting for 15 (9%) cases, as well as in the NBDF, accounting for 12 (8%) cases. In contrast, ‘liver diseases’ accounted for only five (3%) cases in CoRIS_ICD-10 (Fig. 1).

Discussion

There are substantial differences in the CoD of HIV-positive people according to data sources and classification algorithms. HIV/AIDS-associated deaths were the most frequent in our cohort for the 2004–2008 period,
followed by non-AIDS-defining tumours, for both data sources and both classification algorithms. The third most common CoD was liver disease according to the NBDF and revised CoDe in CoRIS but was not when cohort data was coded using ICD-10. Applying ICD-10 to cohort data overestimates HIV/AIDS-associated deaths largely at the expense of labelling as HIV/AIDS-associated ill defined causes and liver diseases because deaths from cirrhosis of viral cause or of unknown cause in people known to be HIV infected are assigned to HIV/AIDS-related causes. This explains why in CoRIS_ICD-10 liver-related causes are in fourth position, why there are virtually no deaths attributed to infectious diseases, none is labelled as ill defined, and HIV/AIDS-associated causes account for 70% of all deaths. In fact, Garcia-Fulgueiras et al. [9], in a study of hepatitis B and C mortality in Spain, included an estimation of the attributable fraction of mortality due to hepatitis virus among cases coded as AIDS.

Applying CoDe to CoRIS shows that just over half of the deaths are due to AIDS, which seems to be more in line with recent data from similar settings which show a reduction in HIV/AIDS-associated deaths compared with previous years [5,6,10,11]. In the Swiss HIV Cohort Study, HIV/AIDS-associated deaths decreased to 15% of all CoD between 2005 and 2009; 85% of deaths were due to non-AIDS-defining conditions, and non-AIDS-defining malignancies were the most important CoD. In our cohort, non-AIDS-defining malignancies were the second CoD [11]. Likewise, data from EuroSIDA show a decrease in the incidence rate of AIDS from 118.3 per 1000 person-years of follow-up in 1997 to 4.4 per 1000 person-years in 2002 [10]. The ART-CC found that 50% of deaths from 1996 to 2006 were associated with AIDS, followed by non-AIDS malignancies (12%) and non-AIDS infections (8%). Overall, a decline in HIV/AIDS-associated mortality was observed with increased duration of antiretroviral therapy [6].

In this study, data from the NBDF, which were coded by ICD-10, seem to be more concordant with those from CoRIS_CoDe than with CoRIS_ICD-10. This is because the largest source of misclassification bias we found was derived from the rule that deaths in persons known to be HIV-positive should be coded as HIV/AIDS-related deaths. As information on HIV status is more likely to be missing from death certificates than in CoRIS, a cohort of people with confirmed HIV infection, lack of information seems to result in a picture which is closer to reality. These results call for caution when comparing different studies, as this rule is often not taken into account.

One of the main limitations of NBDF – given that it depends on death certificates – is the poor quality of the latter. The importance of correct completion of death certificates has been well described, together with the common errors that can be minimized with adequate training [12,13]. Also, when deaths occur outside the hospital or the doctor certifying the death does not know the deceased, death certificates may miss important information [12,14] and HIV infection may be missing because of the stigma associated. However, for some
conditions such as suicide, which trigger legal notification, the proportion of deaths was exactly the same with all three strategies.

The main limitation of this study is the low number of deaths, which may have introduced random error in some of our estimates and comparisons. However, this does not apply to systematic error and the relatively small number of deaths permitted a more comprehensive exploration of the misclassification bias which would have been unmanageable with larger numbers. Although both coding algorithms establish clear rules, we must take into account that different reviewers can give different codes for a given death. Therefore, it is important to establish a centralized review system in order to check and correct these discrepancies.

We can conclude that the revised CoDe classification is the best way to classify CoD in a cohort study of HIV-positive patients. Linkage with external registries such as the national death index to obtain CoD in people lost to follow-up or with an unknown cause may introduce bias given that the ICD-10 overestimates HIV/AIDS-related deaths.

Acknowledgements

V.H.S. and J.d.A. conceived the study and wrote the first draft of the manuscript. All authors contributed to the final draft. V.H.S. and M.C.B. checked and coded data for all patients. V.H.S. and P.S.-V. extracted the data from CoRIS and did the analyses. F.G., J.B., G.N., I.S., J.R., M.A.M. and A.A. contributed to data interpretation.

Conflicts of interest

This study would not have been possible without the collaboration of all the patients, medical and nursing staff and data managers who have taken part in the project. The RIS Cohort (CoRIS) is funded by the Instituto de Salud Carlos III through the Red Temática de Investigación Cooperativa en Sida (RIS C03/173).

The ART-Cohort Collaboration (ART-CC), which provided a simplified version of the CoDe protocol used in this study, is funded by the UK Medical Research Council, grant number G0700820.

This work has been partially funded by grant from FIS (Spanish Networks for Research on AIDS and Public Health), 04/0900 and RIS (Spanish HIV Research Network for excellence), RD06/006. We are grateful for funds provided by CIBERESP (Ciber de Epidemiología y Salud Pública).


This work has been presented as Oral Communication at the 13th European AIDS Conference, 12–15 October 2011, Belgrade, Serbia.

References


