Tesis Doctoral

CARACTERIZACIÓN Y TENDENCIAS DE LA MORTALIDAD GLOBAL Y POR CAUSAS DE LOS HOMBRES Y MUJERES VIH-POSITIVOS DE LAS COHORTES DE LA RED DE INVESTIGACIÓN EN SIDA EN ESPAÑA

Tesis Doctoral presentada por:

Belén Alejos Ferreras

Doctorado en Ciencias de la Salud
Departamento de Cirugía, Ciencias médicas y Sociales
Alcalá de Henares, 2015

Directores:

Julia del Amo Valero
Susana Monge Corella
Inmaculada Jarrín Vera

Caminante, son tus huellas
el camino y nada más;
caminante, no hay camino,
se hace camino al andar.

Al andar se hace camino
y al volver la vista atrás
se ve la senda que nunca
se ha de volver a pisar.

Caminante no hay camino
sino estelas en la mar.
Campos de Castilla
Antonio Machado

A mi familia

Table of contents

INSTITUTIONAL ACKNOWLEDGEMENTS

This study would not have been possible without the collaboration of all the patients, medical and nursing staff, and data mangers who have taken part in the project. This work has been (partially) funded by the RD12/0017/0018 project as part of the Plan Nacional R + D + I and cofinanced by ISCIII- Subdirección General de Evaluación y el Fondo Europeo de Desarrollo Regional (FEDER). "

We are also grateful for funds provided by the Predoctoral Training in Health Research Program (National Plan for Scientific Research, Development and Technological Innovation Nº exp: FI10/00368) and by the instruction programme of the Spanish Network of HIV/AIDS Research.

ACKNOWLEDGEMENTS

First, I would like to thank all the patients participating in CoRIS cohort altruistically, because without them none of this would be possible, to all the research teams members and in general to all those that "have put a hand in this thesis"

To Julia, for giving me a place in her team. her expertise in HIV epidemiology and her insight in statistics have been essential in my development over the years. I am grateful for all her support and capacity of motivation that draw out the best that is in each one of us and for her endless efforts that make this worthwhile.

To my co-directors, for their patience to re-re-re-review this document and for their friendship. To Inma, for her generosity, I wouldn't even have half of my knowledge if she wouldn't have shared her mind with me. And to Susana, for all the "thinking tanks" because group thinking is better (particularly with her).

To the rest of my colleagues, Cris, Sara, Carolina, Adela, Nieves, Yaiza, Debo and Vicky, for your friendship, your time and taking care of me. And especially, to Paz, for giving me the opportunity to come in this world and supporting my first steps. Julia says, sometimes, that "I have been brought up alone" but the truth is that you have made me grow.

To my girls ... Mary, Jessi, Lourdes, Tere, Marta, Maria Mérida, Mari-Carmen and Nerea because you are the best friends in the whole world, for supporting me always and taking the drama out when needed. Because it is always easy, it is always

good and always without problems with you. Also to Ana and Rocy, for being friends, in addition to family.

To Enrique, for standing by me as far as he could. We have grown up together and a part of me is yours, so this Thesis is also a bit yours.

To those who were with me during my year in London. To my friends at LSHTM (Jenna, Efi, Fahed, Chris, Jenric ....) because thanks to them I only remember the good times and laughs at 365 and to my family at ILSC (Tea, Giulia, Haris, Juan, Haydar, ...) for being my parents and brothers and not letting me feel alone. And, especially, to Andre, for showing me that London can be a wonderful city, for his unconditional support on everything I do and for motivating me always (Ich habe dich lieb).

And finally to my family. to my brother, for helping me since I can remember and insist on the idea that computers can be used for more things than "Messenger". And to my parents, for their support and love, for buying me my first “Rubio problems” workbook, and especially for teaching me that education it is more than a right. They, more than anyone, through their efforts and their sacrifice, have made this Thesis possible.

AGRADECIMIENTOS

En primer lugar quiero dar las gracias a todos los pacientes que participan en la cohorte CoRIS de forma altruista, porque sin su colaboración nada de esto sería posible, a todos los miembros de los equipos de investigación y en general a todos aquellos que “han echado una mano en esta tesis”.

A Julia, por hacerme un hueco en su equipo. Sus conocimientos sobre epidemiología y su forma de valorar la estadística han sido esenciales para mi aprendizaje y desarrollo durante estos años. Su apoyo y su capacidad de motivación hacen que saquemos lo mejor de nosotras y su esfuerzo inagotable hace que esto merezca la pena.

A mis co-directoras, por su paciencia para re-re-re-revisar este documento y por su amistad. A Inma, por su generosidad, ya que no sabría ni la mitad si ella no hubiera compartido su mente conmigo. Y a Susana, por todos los “thinking tanks”, porque en equipo (y en particular con ella) se piensa mejor.

Al resto de mis compis, las que estuvieron y las que están, Cris, Sara, Carolina, Adela, Nieves, Yaiza, Debo y Vicky, por su amistad, su tiempo y por cuidarme tanto. Y, en especial, a Paz, por darme la oportunidad de entrar en este mundo y acompañar mis primeros pasos. Julia dice, a veces, que “me he criado sola” pero lo cierto es que me habéis hecho crecer vosotras.

A mis chicas… María, Jessi, Lourdes, Tere, Marta, María Mérida, Mari-Carmen y Nerea porque sois las mejores amigas del mundo mundial, por apoyarme siempre y desdramatizar cuando es necesario. Porque con vosotras siempre es fácil, siempre es bien y siempre sin problemas. También a Ana y a Rocy, por ser amigas, además de familia.

A Enrique, por acompañarme hasta donde le fue posible. Hemos crecido juntos y una parte de mi es suya, por eso esta Tesis también es un poco suya.

A los que estuvieron conmigo durante mi año en Londres. A mis amigos de la LSHTM (Jenna, Efi, Fahed, Chris, Jenric….) porque gracias a ellos solo me acuerdo de los buenos momentos y risas en la 365 y a mi familia de la ILSC (Tea, Giulia, Haris, Juan, Haydar, …) por hacer de padres y hermanos y no dejar que me sintiera sola. Y, en especial, a Andre, por enseñarme que Londres puede ser una ciudad maravillosa, por su apoyo incondicional en todo lo que hago y por motivarme siempre (Ich habe dich lieb)..

Y por último a mi familia. A mi hermano, por ayudarme en todo desde que me alcanza la memoria y por insistir en que los ordenadores servían para más cosas que el "Messenger". Y a mis padres, por su apoyo y cariño, por comprarme mis primeros cuadernillos Rubio de problemas, y sobre todo por inculcarme que la educación es algo más que un derecho. Ellos, más que nadie, gracias a sus esfuerzos y a sus renuncias, han hecho que esta Tesis sea posible.

LIST OF ABBREVIATIONS

ADI	AIDS Defining Illness
AIDS	Acquired Immunodeficiency Syndrome
cART	combination Antiretroviral Therapy
CC	Complete-Case
CDC	Clinical Staging and Disease Classification System
CI	Confidence Interval
CoD	Cause of Death
CoRIS	Spanish AIDS Research Network adult cohort
eHR	excess Hazard Ratio
HBV	Hepatitis B Virus
HCV	Hepatitis C Virus
HIV	Human Immunodeficiency Virus
ICD	Classification of Diseases
IDU	Injection Drug Users
IM	Indicator Method
IQR	Interquartile Range
IRR	Incidence Rate Ratios
JAV	Just Another Variable
LA	Latin America
LD	Late Diagnosis
LP	Late Presentation
MAR	Missing at Random
MCAR	Missing Completely at Random
MI	Multiple Imputation
MICE	Multiple Imputation by Chained Equations
MID	Multiple Imputation, then deletion
MNAR	Missing Not at Random
MSM	Men who have Sex with Men
NBDF	National Basic Death File
NIS	National Institute of Statistics
NLRD-NARD	Non–Liver-Related Non–AIDS-Related Deaths
py	person-years
RIS	Spanish AIDS Research Network
SINIVIH	National information system on new HIV diagnosis
SMR	Standardized Morbidity or Mortality Ratios
SSA	Sub-Saharan Africa
STD	Sexually Transmitted Disease
TB	Tuberculosis
WHO	World Health Organization

1 Summary/Resumen

1.1 Summary (English version)

Since the introduction of combination Antiretroviral Therapy (cART), mortality in HIV-positive subjects has sharply decreased in countries with good access to health care including HIV testing and treatment. Consequently, the life expectancy of those individuals starting cART and attaining restoration to a normal CD4 count may approach that of the general population. However, excess mortality (mortality above what would be expected in the general population) seems to remain as duration of HIV infection lengthens.

In high income countries, including Spain, the distribution of causes of death in the HIV-positive population has changed over the last 15 years. While the proportion of AIDS-related deaths has decreased, non-AIDS-related deaths such as liver-related, non-AIDS-related malignancies or cardiovascular disease have emerged as important causes of morbidity and mortality as consequence of a more prolonged survival in HIV-positive patients. As a final point, excess mortality among HIV-positive subjects is probably attributable to a variety of factors including: aging, presentation at late stage of infection, late start of cART or treatment failure, chronic inflammation, long exposure to treatment toxicities, , active Injecting Drug Use (IDU), use of other recreational drugs, life style-related risk factors (smoking, alcohol, etc.), coinfection with oncogenic viruses such us hepatitis C or B virus (HCV or HBV) or barriers in the initiation and completion of anti-hepatitis C therapy.

Missing data are common in observational studies and need to be considered before performing statistical analyses. Conventional methods, such us Complete-Case (CC) analysis or Indicator Method (IM) have been widely used due to simplicity but might produce biased and underpowered results. Multiple Imputation by Chained Equations (MICE) has been suggested as the best approach to cope with missing values in multiple variables under Missing At Random (MAR) assumption.

The objective of this Doctoral Thesis is to explore overall and cause-specific mortality in HIV-positive men and women: to study trends, to determine how mortality rates in HIV-positive subjects are related to mortality rates in general population, as well as to assess the effect of prognostic factors. We also aimed to give an approach to how to handle missing data in cohorts of HIV-positive patients and to explore different methods to deal with missing data.

We analysed data of CoRIS-MD and CoRIS cohorts from 1997 to 2014. We used Poisson regression to model mortality rates and to assess trends over calendar period. We calculated: (i) all-cause mortality rates, (ii) Standardized Mortality Ratios (SMR) and (iii) excess mortality rates for both cohorts per 100 person-years (py) of follow-up, comparing overall and cause-specific mortality with that of the general population of similar age and sex. We used generalized linear models with Poisson error structure to model excess mortality and to assess the impact of multiple risk factors on this excess mortality.

Missing data patterns in CoRIS cohort were explored using frequency distributions. Excess mortality estimates were calculated and compared using four methods to handle missing Cause of Death (CoD) and covariates; CC: restricting to patients with complete data, IM: creating an extra category for missing values, MICE: developing an imputation model for each variable with missing values (also for CoD), MID: deleting cases with imputed CoD after performing multiple imputation.

Taking the first years of the cART era as a reference (1997-2003), we observed a decrease in overall and cause-specific mortality. The observed decreases in mortality were largely at the expense of the HCV-negative patients, as declines were not observed among HCV-positive individuals.

Patients included in CoRIS from 2004 to 2014 have shown to have overall, liver, non-AIDS malignancies and cardiovascular excess mortality compared to the general population associated with being HIV-positive. Differential short-term and long-term effect of having AIDS before cohort entry and HCV coinfection was found for overall mortality. Indeed, having AIDS before cohort entry was a strong predictor of excess mortality during the first year of follow-up nevertheless this effect was diluted during the rest of follow-up. Opposite, having a positive Hepatitis C test at entry was just related with higher long-term excess mortality.

Missing values are infrequent in CoRIS and we did not find significant differences between the estimates of excess mortality using four different methods to handle missing data, nevertheless lack of differences must be taken with caution as the

number of deaths was low. We detected that CC approach led to less precise estimations and incorrect classification of CoD (IM) or deletion of cases with missing CoD (CC, MID) seems to lead to underestimation of the excess mortality rates.

The results of this Doctoral Thesis are key to influence changes in the clinical management and treatment of HIV-positive patients in order to further diminish the mortality of HIV-positive patients, especially of those that are also coinfected with HCV.

1.2 Resumen

Desde la introducción del Tratamiento Antirretroviral combinado (TAR), la mortalidad en sujetos con VIH ha disminuido drásticamente en los países con un buen acceso a atención médica incluyendo cribado y tratamiento de VIH. En consecuencia, la esperanza de vida de las personas que inician TAR y consiguen recuperar niveles de CD4 normales puede parecerse a la de la población general. Sin embargo, el exceso de mortalidad (mortalidad por encima de lo que cabría esperar en la población general) parece mantenerse a medida que la duración de la infección aumenta.

En los países desarrollados, entre ellos España, la distribución de las causas de muerte en la población VIH-positiva ha cambiado en los últimos 15 años. Si bien la proporción de muertes por SIDA ha disminuido, las muertes no-SIDA, tales como, neoplasias no-SIDA, hepáticas o cardiovasculares, han emergido como causas importantes de morbilidad y mortalidad a consecuencia de una supervivencia más prolongada en pacientes VIH-positivo.

Como nota final, el exceso de mortalidad entre los sujetos VIH-positivo es probablemente atribuible a una variedad de factores que incluyen: la edad, la presentación tardía, el retraso en el inicio del TAR o fallo del tratamiento, la inflamación crónica, la exposición prolongada a las toxicidades del tratamiento, el consumo activo de drogas inyectadas (UDI), el uso de otras drogas recreativas, los factores de riesgo relacionados con el estilo de vida (tabaco, alcohol, etc.), la co-

infección con virus oncogénicos como el virus de la hepatitis C o B (VHC o VHB) o las barreras en la iniciación y terminación de la terapia anti-hepatitis C.

Los datos faltantes son comunes en los estudios observacionales y deben tenerse en cuenta antes de realizar análisis estadísticos. Los métodos convencionales, tales como el análisis de los Casos-Completos (CC) o Método del Indicador (MI) han sido ampliamente utilizados debido a su simplicidad, pero pueden producir resultados sesgados y poco potentes. El método de la Imputación Múltiple por Ecuaciones Encadenados (MICE) ha sido propuesto como el mejor enfoque para hacer frente a valores perdidos en múltiples variables asumiendo que el mecanismo de perdida es Missing At Random (MAR).

El objetivo de esta Tesis Doctoral es explorar la mortalidad global y causa-específica en hombres y mujeres VIH-positivo: estudiar las tendencias, determinar cómo las tasas de mortalidad en sujetos VIH-positivo se relacionan con las tasas de mortalidad en la población general, así como evaluar el efecto de los factores pronósticos. También se intentó dar una aproximación a cómo tratar los datos faltantes en cohortes de pacientes VIH-positivo y se exploraron diferentes métodos analíticos para hacer frente a los datos faltantes.

Se analizaron los datos de las cohortes CoRIS-MD y CoRIS desde 1997 a 2014. Se utilizó la regresión de Poisson para evaluar las tendencias en la mortalidad porperíodo calendario. Calculamos: (i) las tasas de mortalidad global, (ii) Razones de Mortalidad Estandarizadas (SMR) y (iii) las tasas de exceso de mortalidad para ambas

cohortes por cada 100 persona-año (pa) de seguimiento, comparando la mortalidad global y causa-específica observada con la mortalidad en la población general de la misma edad y sexo. Se utilizaron modelos lineales generalizados con estructura de error de Poisson para modelizar el exceso de mortalidad y para evaluar el impacto de múltiples factores de riesgo en este exceso de mortalidad.

Los posibles mecanismos de datos faltantes en la cohorte CoRIS se analizaron utilizando distribuciones de frecuencia. Las estimaciones de exceso mortalidad se calcularon y compararon utilizando cuatro métodos para hacer frente a los valores perdidos en la variable respuesta causa de muerte (CoD) y en las covariables; CC: restringe el análisis a los pacientes con datos completos, IM: en cada variable se crea una categoría adicional para los valores perdidos, MICE: se desarrolla un modelo de imputación para cada variable con valores perdidos (también para CoD), MID: se eliminan los casos con la CoD imputada una vez realizada la imputación múltiple.

Tomando los primeros años de la era antirretroviral como referencia (1997-2003), se observó una disminución en la mortalidad global y causa-específica. Las disminuciones observadas en la mortalidad fueron en gran medida debidas a los pacientes VHC-negativo mientras que no se observaron disminuciones entre los individuos VHC-positivo.

Los pacientes incluidos en CoRIS en el periodo 2004-2014 mostraron excesos de mortalidad global, por neoplasias no-SIDA, hepática y cardiovascular asociado a ser VIH-positivo comparado con la población general. Se encontró que el efecto de

presentar SIDA a la entrada de cohorte y de la co-infección por el VHC sobre el exceso de mortalidad global era diferente a corto plazo y a largo plazo. De hecho, presentar SIDA a la entrada en la cohorte resulto ser un fuerte predictor del exceso de mortalidad durante el primer año de seguimiento, sin embargo, este efecto se diluía durante el resto del seguimiento. Por otro lado, tener una prueba positiva de la hepatitis C en la entrada estaba relacionado con un mayor exceso de mortalidad a largo plazo.

Los valores faltantes son poco frecuentes en CoRIS y no encontraron diferencias significativas entre las estimaciones del exceso de mortalidad utilizando cuatro métodos analíticos diferentes para hacer frente a los datos faltantes, sin embargo, la falta de diferencias deben tomarse con cautela, ya que el número de muertes fue bajo. Se detectó que el método CC produce estimaciones menos precisas y que la clasificación incorrecta de CoD (MI) o la eliminación de los casos con CoD perdida (CC, MID) parece producir una subestimación de las tasas de exceso de mortalidad.

Los resultados de esta tesis doctoral son clave para influir en los cambios en el manejo clínico y el tratamiento de los pacientes VIH-positivo con el fin de reducir aún más la mortalidad de los pacientes VIH-positivo, sobre todo de los que también están coinfectados por el VHC.

2 Background

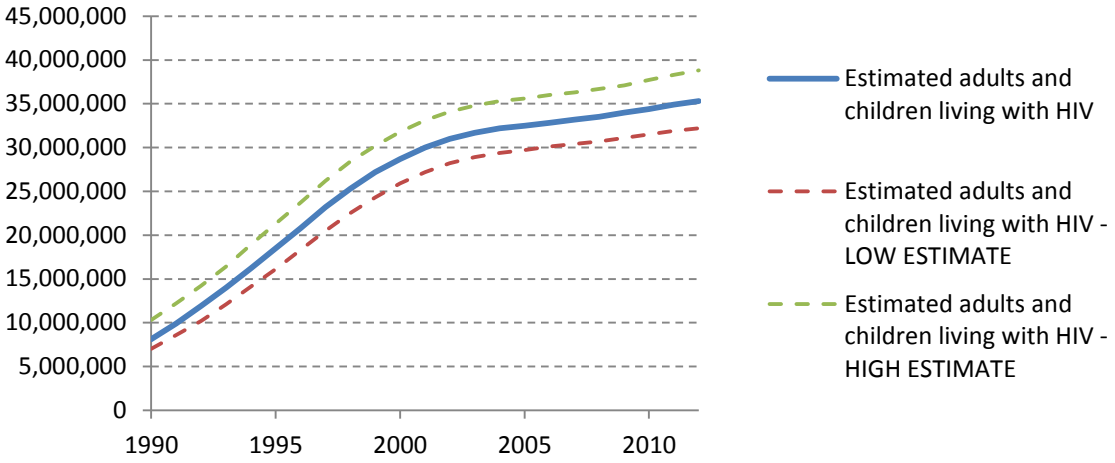
2.1 Epidemiology of HIV infection and AIDS

2.1.1 Epidemiology of HIV/AIDS in the world

Over the last 15 years, a noticeable increase in the number of people living with Human Immunodeficiency Virus (HIV) in the world has been observed due to more people receiving the life-saving combination Antiretroviral Therapy (cART). UNAIDS estimated that 20.8 million people were living with HIV in 1996, 32.2 in 2004 and it reached 35.3 million in 2012 Figure 1 (1).

The annual number of new infections in adults has fallen by 30% since 2001, from 3.4 (3.1-3.7) million to 2.3 (1.9-2.7) in 2012. The same pattern of decline has been observed for the number of AIDS deaths: there were 1.6 (1.4-1.9) million AIDS deaths in 2012, showing a decline from 2.3 (2.1-2.6) million in 2004.

Figure 1: Estimated number of adults living with HIV, 1990-2012

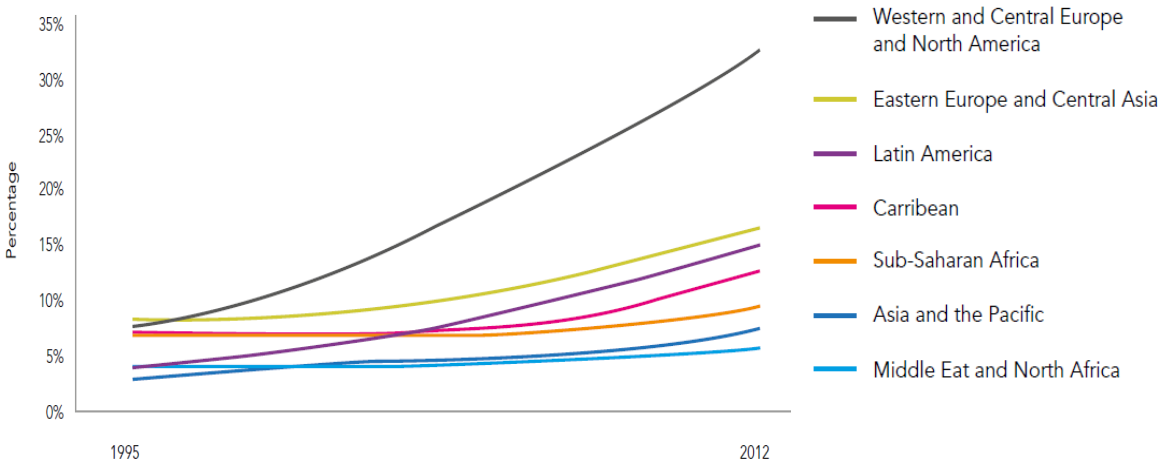


Source: UNAIDS Report on the Global AIDS Epidemic – 2013 (1).

Nevertheless, despite the efforts in scaling-up access to HIV treatment over the past few years, ensuring universal access to treatment remains still a challenge. Access to treatment needs to be expanded, especially in middle-income countries and children. In 2013, the World Health Organization (WHO) released its "Consolidated Guidelines on the use of antiretroviral drugs for treating and preventing HIV infection" (2), which recommend starting treatment in individuals with CD4 count under 500 cells/μL, pregnant women, HIV-positive partners in serodiscordant couples, children under five and people with active Tuberculosis (TB) disease or coinfected with Hepatitis B Virus (Hepatitis B Virus) with evidence of severe chronic liver disease. The estimated number of people eligible for cART under these recommendations reached 28.6 million in middle and low-income countries however only 11.7 million were receiving treatment in 2013 (3).

In its supplement *HIV and Aging*, UNAIDS reported worldwide estimates on the global HIV epidemic showing an increasing trend in the number of people aged 50 years or older who are living with (Figure 2)(4). For first time since the start of HIV epidemic, 10% of the adult population living with HIV in low- and middle-income country was aged 50 years or older and approximately 30% in high income countries.

Figure 2: People aged 50 years or older, as a percentage of all adults 15 years or older living with HIV, by region, 1995–2012.



2.1.1.1 Source: Reproduced from UNAIDS 2013 supplement: HIV and Aging (4)

The increase in the people living with HIV aged 50 years or older is a combined result of a more prolonged survival due to the success of cART, a decrease in the HIV incidence in younger adults (shifting the disease to older ages), and the fact that many of the risk behaviours common among young people are also frequent among those aged 50 years and over. CASCADE (Concerted Action on Seroconversion to AIDS and Death in Europe) is a collaboration between the investigators of 29 cohorts of persons with well-estimated dates of HIV seroconversion (seroconverters), in 14 European countries, as well as Australia, Canada, and Africa. According to CASCADE data, the median age at HIV seroconversion has increased from 26.2 years (Interquartile Range (IQR): 21.5-32.7) in those infected before 1986 to 33.2 years (IQR: 27.3-40.3) in those infected between 2006 and 2010 (5) .

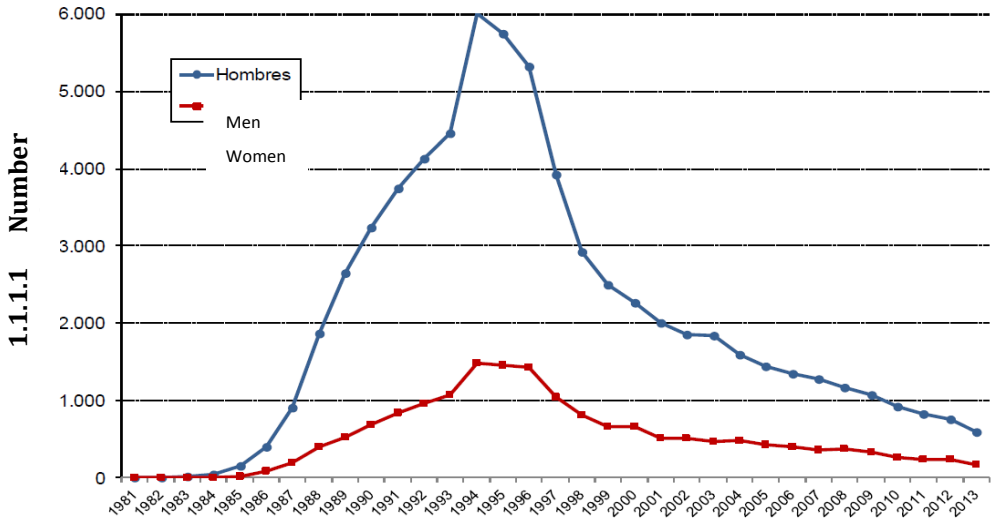
Along these lines, the aging process will continue among people living with HIV in the years ahead therefore it is essential to evaluate and to adapt the HIV testing and

treatment services to the specific characteristics and needs of the people living with HIV aged 50 years and over.

2.1.2 Epidemiology of HIV/AIDS in Spain, 1981-2013

The characteristics of the HIV epidemic in Spain have varied very much since its beginning. In the 1980s HIV spread widely, especially among Injection Drug Users (IDU) and, in a distant second position, among Men who have Sex with Men (MSM). This was in contrast with the situation in other high-income countries, where homosexual intercourse was the predominant route of transmission (6). During the late 1980s and early-mid 1990s, Spain reported the highest AIDS rates within the European Union; the epidemic peaked in 1994 with 7,489 AIDS cases diagnosed. Due to the introduction of harm reduction programs such as needle exchange and methadone maintenance treatment in the late 80s and early 90s, and the generalization of cART in 1997, AIDS incidence and AIDS-related mortality declined sharply. Currently, the new AIDS cases notified each year are around 1,000, still above European average, but similar to Western Europe rates (Figure 3) (7).

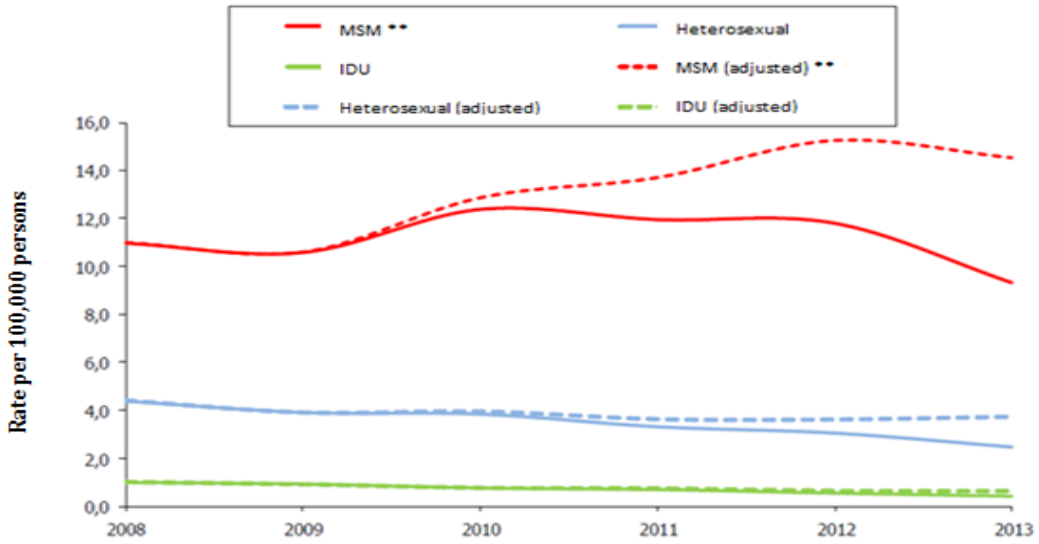
Figure 3: AIDS cases in Spain per year and sex. Data adjusted by the delayed reporting



Source: Epidemiological surveillance of HIV in Spain (7).

The national information system on new HIV diagnosis (SINIVIH) was established in 2003; the population covered has progressively increased from 14,469,101 inhabitants in 2003 to 37,863,951 in 2012 (82% the total Spanish population) and finally reaching 100% coverage in 2013. Figure 4 shows how trends over time in the rate of new diagnoses of HIV differ substantially by mode of HIV transmission; IDUs rates progressively descend during the period. In the case of heterosexual route, a downward trend in women is observed whereas the rates remain stable among men. In contrast, a steady increase among MSM has been detected (7). Reported data suggests that nowadays the Spanish HIV epidemic is being mainly fuelled by MSM transmission.

Figure 4: New diagnoses of HIV by Transmission mode and year in Spain 2008-2013. Data adjusted by the delayed reporting



**Aragón, Asturias, Baleares, Canarias, Castilla La Mancha, Cataluña, Ceuta, Extremadura, Galicia, La Rioja, Madrid, Melilla, Navarra, and País Vasco*
*** Rate per 100,000 men*
Data source: Epidemiological surveillance of HIV in Spain

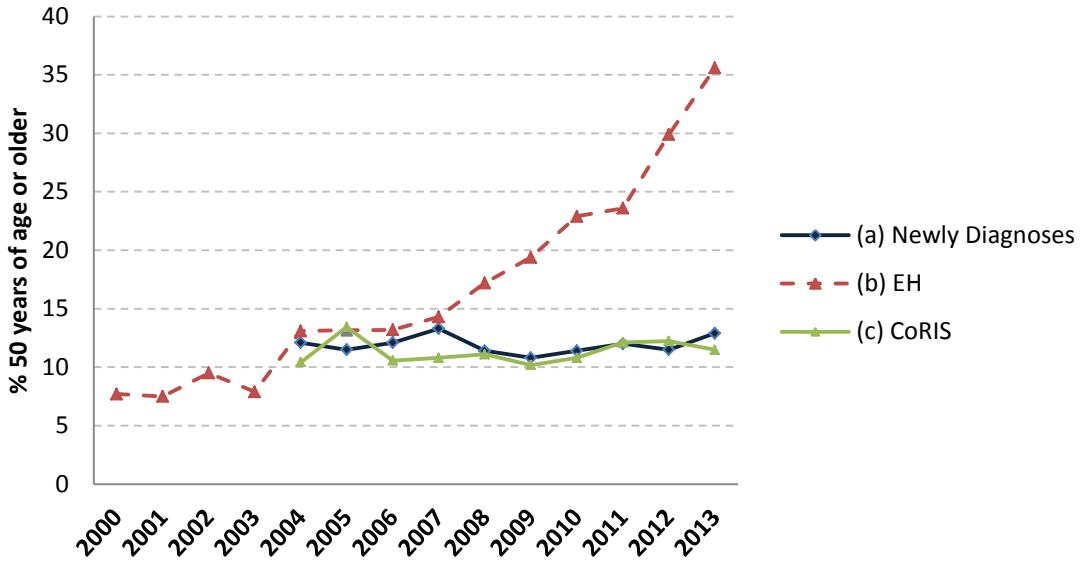
During the first half of the 2000s the proportion of people diagnosed who are from other countries different to Spain increased progressively reaching 42% in 2008. However, this proportion remained stable in the late 2000s and it started to decrease in 2010 down to 32.5%. The most common origin was Latin America (16%).

Spanish epidemic has been concentrated predominantly in men during all the period, with 85% of males; and it is relevant to point that among women, more than 50% of new diagnoses were immigrants.

The aging process that has been detected in the population living with HIV in high-income countries has been also observed in Spain. The "Encuesta Hospitalaria de pacientes infectados con VIH (EH)" (Hospital survey of patients uninfected with HIV) is an annual one-day cross sectional survey that collects information on all people living

with HIV, attending general public hospitals for HIV-related care on the day of the survey and it provides population-based information on people living with HIV in Spain (8). The EH has reported that the percentage of HIV-positive subjects aged 50 years or older has increased over the last decade; from 7.7% in 2000 to 35.5% in 2013 (Figure 5). Regarding new HIV diagnoses, in other high-income countries, such us France or USA, the age at diagnosis has increased over the time, while in Spain the proportion of newly diagnosed individuals aged 50 or older seems to be stable in Spain, (Figure 5). This is also observed in the proportion of newly enrolled patients aged 50 years or more in the cohort of the Spanish Network of HIV/AIDS research (CoRIS) (Paz Sobrino, personal communication). This might be explained by the increase in the number of new diagnoses among young MSM in Spain (9).

Figure 5: Proportion of individuals aged 50 years or older in Spain by year (2000-2013)



(a) Newly Diagnoses. Proportion of newly diagnosed individuals aged 50 years or older in Spain from SINHIV report
(b) EH. Proportion of patients aged 50 years or older from all people living with HIV attending general public hospitals for HIV-related care on the day of the survey
(c) CoRIS. Proportion of newly enrolled patients aged 50 or older in the cohort of the Spanish Network of HIV/AIDS Research

The aging process of people living with HIV in Spain has risen concerns on the risk of chronic morbidities, non-AIDS-defining malignancies, cardiovascular disease, renal disease, liver failure and possibly neurocognitive decline (5,10,11). Furthermore, these comorbidities may also worsen HIV disease progression (12) and several authors have found than being 50 years or older clearly impacts the immunological response and the survival of HIV-positive subjects after starting cART (13-15) with no effect in the virological response (13,14) .

Late Presentation (LP) has been defined, by consensus, as presenting for HIV care with a CD4 count less than 350 cells/mm^3 or presenting with an AIDS defining event, regardless the CD4 cell count (16). Additionally, we can also distinguish Late

Diagnosis (LD) that considers only subjects with a new HIV diagnosis. Late diagnosis is defined as having a CD4 count <350 cells/ mm^3 and/or AIDS-defining illness at first HIV diagnosis.

Although a reduction in the percentage of HIV-positive subjects diagnosed late has been observed since 2007, approximately 47% of the new diagnoses registered in 2013 were Late Diagnoses (7). Besides, it has been found that those subjects aged 50 years or more are more likely to present late (17) and Caro et al found in CoRIS a proportion of late diagnosis of 53% in patients aged 50 years or older versus 21.5% in those under 50 years (18). Finally, individuals with a late presentation may not get the full benefit of antiretroviral therapy (19), may experience increased morbidity as a result (20) and may contribute to the on-going transmission of HIV (21).

2.2 Mortality in HIV-positive subjects

2.2.1 Mortality in the era of cART

Since the introduction of cART, mortality in HIV-positive subjects has sharply decreased in countries with good access to health care, including HIV testing and treatment (22). Palella et al showed that death rates decreased from 29.4 per 100 person-years (py) in 1995 to 16.7 in 1996 and to 8.8 by the second quarter of 1997, in patients who were seen at nine clinics specialised in the treatment of HIV infection in eight cities in the United States.

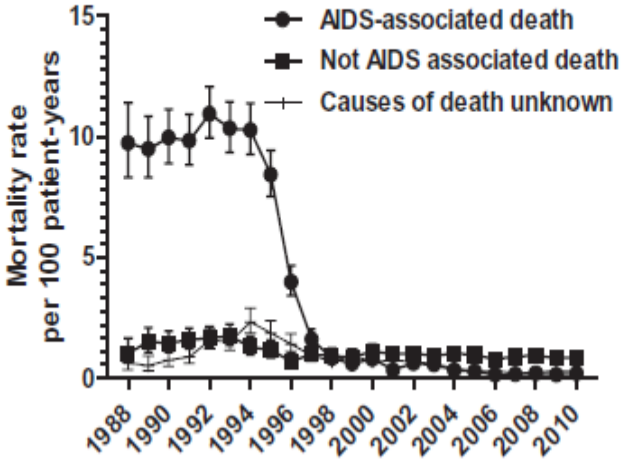
Consequently, the life expectancy of those individuals starting cART in the course of HIV infection and attaining restoration to a normal CD4 count may approach that of the general population (23). However, excess mortality due to AIDS and non-AIDS defining conditions (mortality above what would be expected in the general population) remains as duration of HIV infection lengthens (24-28). Interestingly, CASCADE found that in high income countries, persons infected with HIV through sexual contact now appear to have a risk of death similar to that of the general population in the first 5 years following HIV infection, with the excess mortality becoming evident only later in the course of infection (29).

In high income countries, including Spain, the distribution of causes of death has changed after the introduction of cART. While the proportion of AIDS-related deaths has decreased, the non-AIDS-related deaths such as liver-related, non-AIDS

malignancies or cardiovascular disease, have emerged in the last few years as important causes of morbidity and mortality as a consequence of a more prolonged survival in HIV-positive patients (Figure 6) (25,28,30,31). Higher prevalence of traditional risk factors, such us hepatitis C or B coinfection, tobacco or drugs use, together with long exposure to cART and HIV replication leading to inflammation may further contribute to the diversification of morbidity and mortality among HIV-positive patients (32-34).

Besides, several studies have suggested that non-AIDS malignancies and cardiovascular disease incidence is increasing in parallel with the aging of HIV population (25,31).

Figure 6: Mortality in the Swiss HIV Cohort Study (SHCS), 1988–2010



AIDS-associated mortality was highest in 1992 [11.0/100 patient-years (PY)] and lowest in 2006 (0.144/100 PY); non-AIDS-associated mortality was highest in 1993 (1.74/100 PY) and lowest in 2006 (0.776/100 PY); and mortality of unknown cause was highest in 1994 (2.33/100 PY) and lowest in 2007 (0.207/100 PY). From 1988 to 1998, the SHCS classified causes of death into broad categories (HIV-related, overdose of narcotics, suicide, accident, homicide, other and unknown).

Source: Weber et al 2013 (31)

As Bhaskaran et al stated: "It is likely that mortality of HIV-positive individuals is influenced by mortality which would have occurred regardless of HIV infection, mortality in uninfected general population is a natural reference point to take this aspect into consideration" (29). Further, it is also relevant to account for the mortality

patterns in the general population, such as the differences in mortality rates and causes of death between men and women in HIV-negative population; with a higher mortality in males for almost all causes and almost all ages. In the countries of the European Union in 2010, the difference in the age-adjusted mortality rate between men and women of the general population varied from 45% to 110% and in the case of Spain, it was 78% (35).

Previous studies have compared mortality in HIV-positive patients with mortality in general population. ART-CC, a collaborative study involving cohorts in Europe and North America, found that in individuals on cART and with high CD4 count, HIV-related excess mortality seems to be similar to other chronic conditions. However, mortality continues to be higher than in the general population in those HIV-positive subjects who start treatment late with a history of AIDS-defining illnesses or low CD4 count (36). Another study conducted in COHERE (Collaboration of Observational HIV Epidemiological Research in Europe) reported higher mortality rates among treated HIV-positive subjects as a whole compared to the general population although they did not found an excess mortality among Non-IDUs who attained a CD4 count over 500 cells/mm^3 (37).

In Spain, excess mortality in HIV-positive subjects has also been reported (30). According to a study published in CoRIS and CoRIS-MD cohorts between 1997 and 2008, all-cause mortality was approximately 6 times higher compared with the general population (38).

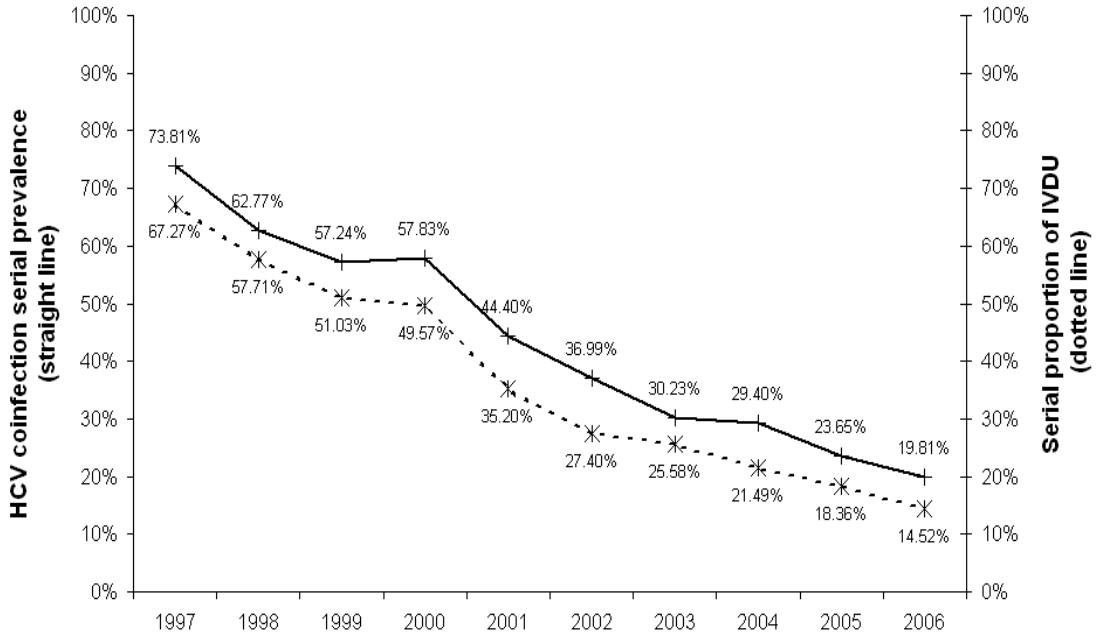
As a final point, excess mortality among HIV-positive subjects is likely attributable to a variety of factors including: aging, presentation at late stage of infection, late start of cART or treatment failure, chronic inflammation despite suppression of HIV replication, long exposure to treatment toxicities, active IDU, use of other recreational drugs, life style-related risk factors (smoking, alcohol, etc.), coinfection with oncogenic viruses such hepatitis C or B virus or barriers in the initiation and completion of anti-hepatitis C therapy (31,33,34,39-41).

2.2.2 Hepatitis C Virus related mortality in HIV-positive subjects

Coinfection by Hepatitis C Virus (HCV) is by far one of the most common comorbidities in HIV-positive patients with a prevalence ranging from 25-30% in HIV-positive persons overall; 72-95% in IDUs, 1-12% in MSM and 9-27% in heterosexuals (42).

Regarding Spain, where the HIV epidemic has been traditionally associated to IDU and high rates of coinfection with HBV and HCV, it was detected a parallel decline in the serial prevalence of IDUs and HCV coinfection during a 10-year period in HIV-positive patients (Figure 7) (43).

Figure 7: Parallel decrease in the prevalence of injection drug use (IDU) and coinfection with human immunodeficiency virus (HIV) and hepatitis C virus (HCV) in the cohorts of the Spanish AIDS Research Network



Source: Cachafeiro S, et al. Clinical Infectious Diseases 2009; 48: 1467-70 (43)

It is well known that HIV infection modifies the natural history of chronic hepatitis C, promoting more rapid progression to fibrosis and development of cirrhosis and end-stage liver disease (44). In areas with a high prevalence of coinfection of HIV and HCV, end-stage liver disease has emerged as a leading Cause of Death (CoD) in HIV-positive patients (45-49). In Spain, liver disease is also one of the major causes of death in HIV-positive persons (30).

With respect to the effect of HCV on the progression of HIV infection, data are inconsistent (50,51). Some studies found no evidence that HCV infection increases the risk of HIV disease progression or death, nor that it affects the immune response to cART (51,52). However, other studies found that HCV infection is associated with increased risk of developing AIDS-defining conditions (53,54) and impaired CD4

recovery, even after taking into account the years of exposure to cART (55). HCV infection may also contribute to AIDS-related death and Non-Liver-Related Non–AIDS-Related deaths (NLRD-NARD) through other associated factors such as lifestyle (substance abuse, accidents, suicide, poorer access to care), immune activation and defective immunity (56), systemic inflammation (57), microbial translocation (58), and liver disease itself (59). Nonetheless, the contribution of HCV to an increased incidence of different NLRD-NARD unrelated to lifestyles remains to be proven. The notion that HCV infection has a negative impact on HIV infection is further supported by a study which found that eradication of HCV in HIV/HCV-coinfected patients was associated not only with a reduction in liver-related complications and liver-related deaths, but also with a reduction in HIV progression and mortality not related to liver disease (60).

In the last decade, several factors could have influenced mortality in the HIV/HCV coinfected population including the use of more efficacious and safe cART regimens for all clinical scenarios, and the availability of anti-HCV treatment based on interferon plus ribavirin (61,62); the response to which is associated with a reduction of liver-related complications and mortality irrespectively of the stage of fibrosis (63). Hereof, COHERE has shown that the effect of HCV treatment on mortality was in the direction of benefit, probably at the expense of a lower risk of liver-related death, though differences were not statistically significant (64).

Several studies have reported an important excess mortality in HIV/HCV coinfected patients (65-68). A meta-analysis that estimated the effect of HCV infection

on HIV disease progression and overall mortality found that, in the cART era, HCV coinfection, compared with HIV infection alone, increased the risk of mortality (67). And a recent study conducted in CASCADE collaboration has shown that since 1997, when cART became widely available, HIV/HCV-coinfected patients were found to have a higher risk of death from HIV and/or AIDS, and from hepatitis or liver disease, than patients infected with only HIV (68).

2.2.3 Ascertainment and classification of deaths

It has been previously described that accurate information on CoD in HIV-positive subjects is not easy to obtain and is subject to selection and information biases (69). Most high-income countries have national death indexes that collect information from death certificates and code the causes of death according to international classification rules.

In Spain, the National Statistics Institute collects data on deaths and CoD and classifies them according to the International Classification of Diseases, 10th revision (ICD-10). ICD-10 system is defined as 'the standard diagnostic tool for epidemiology, health management and clinical purposes and it is used to monitor the incidence and prevalence of diseases and other health problems' (70). This coding system provides a set of rules to assign a code to each death using the information provided by the death certificate, generally based on a single CoD. ICD-10 was implemented in 1999 and introduced specific codes for deaths in persons with AIDS which were not present in ICD-9.

On the other hand, cohort studies of HIV-positive subjects that collect data on CoD usually use other algorithms such us the Coding Causes of Death in HIV (CoDe) Protocol (71), rather than the ICD-10. CoDe Project is a uniform coding system that can be applied to studies of individuals with HIV infection. It includes a detailed data collection on the CoD and contributing factors, as well as a centralized review process of the data collected developed by the Copenhagen HIV Programme (CHIP) (71). 'Revised CoDe' is a simplified version of CoDe selection rules that has been proposed by the Antiretroviral Therapy Cohort Collaboration (ART-CC) (36). It relies on the CD4 cell-count closest to the time of death (up to 3 months before death), AIDS-defining conditions, coinfection with hepatitis B and C, and data on antiretroviral therapy close to death. The CoD is assigned to one of the 30 conditions proposed (69).

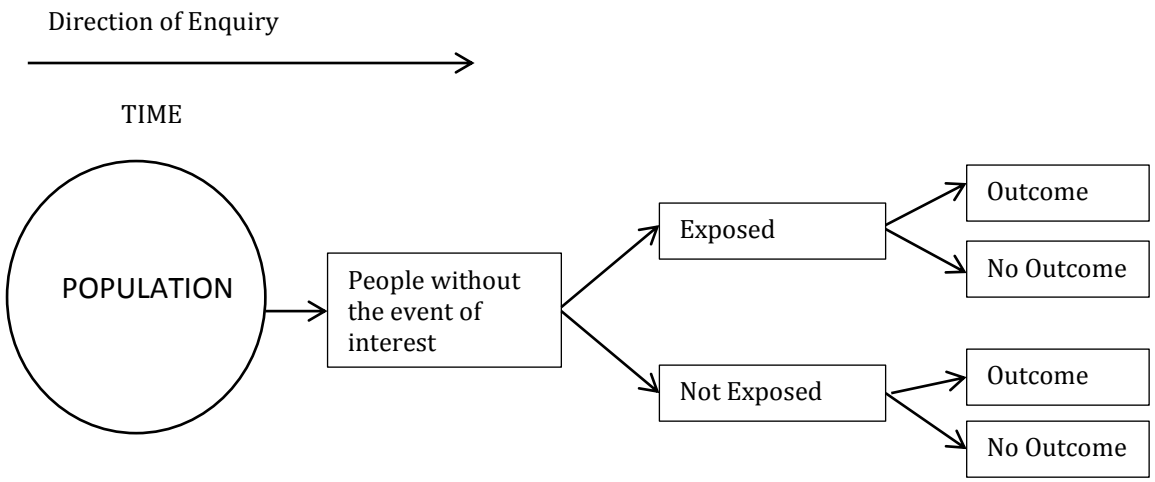
Hernando et al compared deaths reported in CoRIS cohort using two different coding algorithms: the ICD-10 and the revised CoDe. They found that applying ICD-10 overestimated HIV/AIDS-associated deaths largely at the expense of labelling as HIV/AIDS-associated causes those coded as "ill-defined causes" by the revised CoDe which converted into B24 codes (HIV disease) anyone who is known to be HIV-infected. We found that ICD-10 also overestimates AIDS-related deaths due to misclassification of "liver diseases" because deaths from cirrhosis of viral cause or of unknown aetiology in people known to be HIV-infected are assigned to HIV/AIDS-related causes (69).

2.3 Cohort studies of HIV-positive subjects

2.3.1 Overview

A cohort is generally defined as "any designated group of individuals who are followed or traced over a period of time" (72). Cohorts are longitudinal studies which are started in a group of individuals free of an event of interest who are classified by their exposure or exposures and followed-up to observe one or more potential outcomes (Figure 8).

Figure 8: Design of a cohort study



Source: Adapted from Beaglehole et al., Basic Epidemiology (73)

An individual's period of observation, period at risk or follow-up time starts when the individual is enrolled in the cohort and stops when they experience the outcome, are lost to follow-up or observation period ends (administrative censoring), whichever happens first (74). It is often measured in years and it is called person-years at risk.

Cohort studies resemble intervention studies in that people are selected on the basis of their exposure status and then followed up in time, however the allocation to the study groups is not under control of the investigator and is not at random, thus they reflect real-life conditions.

Cohort studies can be classified according to a number of characteristics. In the case of HIV cohorts, an important distinction is into seroconverter and seroprevalent cohorts (75). Seroconverter cohorts include HIV-positive subjects with known or well estimated dates of HIV seroconversion which is the natural, and thus best, origin. Seroprevalent cohorts include subjects whose date of HIV infection cannot be well ascertained as they do not have a previous HIV negative test. HIV seroconversion date is unknown in these subjects, in other words, the duration of the time they have been seropositive prior to the beginning of the enrolment is unknown and other proxies of duration, such as CD4 cell counts, need to be used.

2.3.2 The Spanish Network of HIV/AIDS Research and its cohorts of HIV-positive adults

In 2003, The Institute of Health Carlos III, currently dependent of the Ministry of Economy and Competitiveness of Spain, set up a limited number of Research networks of excellence in health, co-financed by European Regional Development Fund (ERDF). In this context the Spanish Network of HIV Research (RIS) was created (76).

The RIS is a structure coordinated and organized for the HIV/AIDS research which incorporates basic scientists, immunologists, virologists, clinicians, epidemiologists, and statisticians. The objective of The RIS is to ensure a high level of quality in HIV/AIDS research and to encourage constant improvement of the results in the National Health System (NIS), through cooperation of groups from different institutions.

The philosophy of this research network is to address new problems through innovative technical approaches that make use of all human resources and the different network platforms. Participating research groups have defined a large number of research questions.

As a part of the RIS, the Cohorts of the Spanish HIV/AIDS Research –CoRIS-MD and CoRIS- were created. Both are cohorts of largely seroprevalent subjects, who are enrolled irrespective of the availability of information to estimate the date of HIV-seroconversion.

CoRIS-MD is a retrospectively assembled multicentre cohort of HIV-positive patients over 18 years of age, with at least 6 months of follow-up during the period of January 1, 1997 to December 31, 2003.

CoRIS is an open, multicentre, prospective cohort of HIV-positive naïve adults, linked to a BioBank (BBRIS) where biological samples (blood, plasma and DNA), from

subjects included in the cohort, are preserved (77), started in January 2004 to current date.

The objectives of CoRIS are to support high-quality research in the field of HIV/AIDS, to provide a platform for RIS members to perform research projects about the epidemiology and natural history of the infection, and clinical or basic sciences questions.

2.3.3 Analytical approaches in HIV/AIDS cohorts

The main variables of interest in HIV/AIDS cohort studies are: events in person-time data, time to events of interest and markers measured repeatedly over time. The statistical methods used for theses analyses have been extensively described elsewhere (78).

The comparison of outcome rates in a cohort with rates in the general population via Standardized Morbidity or Mortality Ratios (SMR) or via the estimation of excess mortality (Relative Survival) is included within events in person-time analytical techniques.

The SMR uses the indirect method of adjustment and it is expressed as a ratio that quantifies the increase or decrease in mortality of a given group of subjects with respect to the general population.

Excess mortality is defined as the difference between the observed deaths in a given group (i.e. cohort) with a specific disease and the expected deaths in a comparable group from the general population, assumed free from the disease of interest. In cancer research settings, relative survival (analog of excess mortality) models have been widely used to estimate the mortality associated with a diagnosis of cancer (79,80).

In HIV cohort studies, cause-specific mortality cannot be estimated or might be biased when the CoD is missing or is unreliable, nevertheless net mortality associated with being HIV-positive in terms of excess mortality can be calculated. Thus, relative survival method captures both direct and indirect mortality. Direct mortality is the mortality experienced by HIV-positive subjects caused by the pathogenic effect of HIV and traditionally has been represented by AIDS deaths. Parallel, indirect mortality captures mortality not caused directly by the virus but associated with being HIV-positive; mortality "caused" by factors associated with being HIV-positive such as being heavier smoker or having higher risk of social exclusion.

Often HIV cohort studies are based on clinical routine data and some pieces of information are incomplete. Missing data are an important and common issue that has to be considered before performing a statistical analysis and certain assumptions have to be made in order to make inferences to the population of interest. The subject of missing data will be presented in the following section of this document.

2.4 Missing data

2.4.1 Overview

Missing data are data we intended to collect on observations but that due to different reasons were not collected. One of the inconveniences of using observational studies based on clinical routine data is that some pieces of information are not available in spite of large efforts to achieve high completeness. Anyhow, missing data are also an issue in randomized clinical trials.

In longitudinal studies, specific information may be incomplete at baseline, for example transmission category, educational level, age, etc.; or during follow-up, for example, weight, Viral Load (VL) or CD4 count measurements in clinical visit, etc. It can also occur that a patient is lost to follow-up and therefore information is missing after a certain point of time. One last possible situation is that a patient is lost to follow-up for a period of time but returns at a later stage; in this situation, there is missing information in a gap of follow-up.

In addition, it is also important to differentiate two broad situations: a) missing data in exposure or adjustment variables b) missing data in the outcome. These situations will be discussed further in this document.

2.4.2 Importance and consequences of missing data

The Statistical Methodology guideline published by the International Council on Harmonisation says that despite missing data, a clinical trial may be valid provided statistical methods used are sensible (81). This statement applies also for cohort studies. STROBE guidelines recommend to explain how missing data are addressed in the study and to report the number of subjects with missing data in each variable of interest (82).

A revision published in 2012 on reporting and handling of missing data in cohort studies found that despite the fact that some studies use advanced statistical approaches to handle missing data, the majority of the studies removed individuals with missing information and performed the analysis only those with complete information (83). The main problem of this approach arises when those subjects with missing values differ from those with complete information (Complete-Cases analysis); in this case, the estimates are biased and the inference to the population of interest is not valid. Besides, in studies with large proportion of missing data, the precision of the estimates would be reduced due to the reduction in the number of patients considered for the analysis.

Suppose a cohort study of HIV-positive patients whose aim is to assess the effect of educational level (Low, High) on Delayed Diagnosis (DD). We observe in Figure 9 that in both situations, missing data only in patients with high education or missing

data only in those with low education, the estimation of the association (OR and p-value) between level of education and Delayed Diagnosis is biased.

Figure 9: Example of Complete-Case analysis with different missing data patterns in a cohort of HIV-positive patients

SITUATION A (no missing data)

Education /DD	No	Yes
Low	48	34
High	44	74

N=200
OR=2.37
P=0.003

SITUATION B (missing data only on those with low education)

Education /DD	No	Yes
Low	17	18
High	44	74

N=153
OR=1.59
P=0.233

SITUATION C (missing data only on those with high education)

Education /DD	No	yes
Low	48	34
High	31	40

N=153
OR=1.82
P=0.067

There is no universal rule to indicate the number or the proportion of missing data that would lead to bias or to invalid results. Carpenter and Kenward (84) stated the impact of missing data is determined by:

- the question;
- the information in the observed data, and
- the reason of missing data

When a statistical analysis is performed in the presence of missing data, three aims are targeted: minimize bias, maximize the use of available information and get

good estimates of uncertainty (85). However, before deciding what is the best technique to deal with missing information it is essential to consider the missing data mechanisms and patterns.

The success of a statistical analysis in the presence of missing data will depend on the extent we can identify plausible missing data patterns and mechanisms and the degree to which conclusions are robust to the different patterns and mechanisms (84). This is to say, the success will depend on the reasons why data are missing.

2.4.3 Missing data mechanisms

Let be,
$X = (X_o, X_m)$ = covariates we intend to collect ; where X_o is observed data and X_m is missing
$Y = (Y_o, Y_m)$ = outcomes we intend to collect ; where Y_o is observed data and Y_m is missing
$R = Complete\ value\ Indicator = \begin{cases} 1 & If\ X\ and\ Y\ observed \\ 0 & If\ X\ or\ Y\ missing \end{cases}$

In words, the covariates we intend to collect for certain analysis are denominated X; X_o represents the observed data and X_m the missing values in the covariates. The outcome variable is denominated Y; Y_o are the observed values in the outcome and Y_m the missing values in the outcome. The indicator variable R takes the value 1 if the covariates X and the outcomes Y are complete and the value 0 if the covariates or the outcomes are missing.

Different mechanisms that can generate missing data were classified by Little and Rubin (86) into:

Missing Completely at Random (MCAR): The probability of a variable being missing (so called missingness) is independent to the value of this variable and to the other observed or unobserved variables. In this ideal situation, the subjects with missing data are similar to those with complete information. We can write this as:

$$P(R|\mathrm{Y}_o, \mathrm{Y}_m, X_o, X_m) = P(R)$$

MCAR means that missingness does not depend neither on observed nor unobserved data.

We note that MCAR is always an assumption that cannot be proven to hold, exploring the observed data. We can explore the relationship of missingness with observed data by fitting models [R~X,Y] and if we find any association between missingness and observed data we can discard MCAR mechanism. However, even if R is not related with all other X and Y, we cannot guarantee data are MCAR because we cannot explore whether missingness depends on the underlying (not observed) values of the variable.

Example

A certain page of the questionnaire is accidentally missing and this fact is unrelated to the specific information collected in the missing page and to any other information.

Implications

MCAR is usually considered an ignorable mechanism. When data are MCAR, the analysis including only subjects with complete data (Complete-Case analysis) gives unbiased estimates and valid inferences because the subjects included in the analysis can be considered as a random sub-sample of the total database. However, results are less precise than in the data full observed and this mechanism is not reasonable in the majority of the cases.

Missing at Random (MAR): The probability of a variable being missing is independent of the underlying value of this variable given the observed data (either the outcome or any covariates). In other words, missingness is unrelated to the missing values but related to other observed variables.

$$(R|\mathrm{Y}_o, \mathrm{Y}_m, X_o, X_m) = P(R|\mathrm{Y}_o, \mathrm{X}_o)$$

To declare data are MAR we need to assume that missing data depend uniquely on observed data. We can explore whether this assumption is plausible, however we could never asses if missing values depend on unobserved data. Therefore, as with MCAR, we could never guarantee that the mechanism is MAR based on the observed data.

Example

At cohort entry, short-term survivors have poorer conditions thus the probability of collecting socio-demographic information is lower since they cannot be asked about their educational level or country of birth. The probability of having

missing values in the variable level of education is associated with short-term mortality but it is independent to the value of level of education itself.

Figure 10: Data set Missing At Random (MAR)

ID	Short-term death	Educacional Level
1	1	.
2	1	.
3	0	High
4	0	Low

A special case of MAR is uniform non-response within classes (87). For example, suppose we try to collect data on income and property tax band. Normally, those with higher incomes may be less willing to reveal them. This would be a case of Missing Not At Random (MNAR) (see below) and a simple average of incomes from respondents will be downwardly biased.

Figure 11: Data set with uniform non-response within classes

id	Income	Tax band
1	12000	Low
2	30000	Medium
3	.	High
4	.	High
5	100000	High

However, if we have complete information on property tax band which is highly associated with income then missing data on income is missing at random given property tax band. Given property band, missingness does not depend on income itself, so it cannot be considered MNAR .

Implications

Briefly, the main idea of MAR patterns is that observed values predicts solely missingness. Hence, as missing data depends on other observed values, Complete-Case approach might provide biased estimates when the missing data mechanism is MAR (e.g. we would be selecting those with better clinical prognosis, in the example in Figure 10 page 39).

Cases with complete information are not any longer a random sub-sample and therefore not representative of the complete sample. If the probability of being a complete case is associated with the outcome, the statistics calculated (mean, median Confidence Interval (CI), coefficients from generalize lineal models, estimating equations, etc.) would not be valid.

Nevertheless, the effects can be ignorable controlling adequately by the observed variables that predict missingness. Besides, MAR seems to be more plausible and likely to occur than MCAR.

Missing Not at Random (MNAR): The probability of a variable being missing is dependent of the underlying value of that variable. Sometimes it is also known as Not Missing At Random (NMAR).

MNAR mechanism cannot be definitely discarded by the observed data. Specifically, it is not possible to distinguish between MAR and MNAR mechanism just

based on observed data and therefore additional assumptions are needed based on expert opinions.

Example

In a cohort of HIV-positive subjects those patients with missing HCV antibody test result at baseline are more likely to be HCV negative.

Figure 12: Data set Not Missing At Random (NMAR)

ID	HCV antibody (with missing values)	HCV antibody ((complete)
1	Postive	Postive
2	Negative	Negative
3	.	Negative
4	.	Negative
5	Positive	Positive

Implications

This missing data mechanism is considered non-ignorable. Under MNAR we need to specify both a valid model for the outcome and a valid model for the missingness mechanism, which will be unknown in the majority of the cases. Sensitivity models are essential, as well as an informed expert may be very helpful to explain departures and the different results obtained.

Rubin et all (88) suggest that even under MNAR, it quite often happens that after accounting for the information about missingness mechanism in the observed data, there is relatively little information remaining in the unseen data and therefore MAR models may provide accurate answers (84).

In conclusion, although assumptions about missingness mechanisms must be consistent with the data, they can rarely be verified and therefore it is necessary to formulate appropriate analysis and to perform exhaustive sensitivity analyses; to explore the robustness of our inferences and conclusions to different assumptions about the missingness mechanisms.

2.4.4 Missing data patterns

Arbitrary or random pattern

It is said that the pattern of missingness is arbitrary or random when missing values are present in any subject and any variable. Missingness does not occur in a pre-specified way. This pattern is also known as non-monotone.

Figure 13: Arbitrary or random pattern of missingness

	Variables			
ID	**X1**	**X2**	**X3**	**X4**
1	1	.	A	0
2	.	38	.	1
3	1	.	B	.
4	.	15	A	0

Monotone pattern

A pattern of missingness is said to be monotone when variables are ordered by the number of missing values and it is observed that if a subject has missing values in one variable k this subject also will have missing values in all subsequent variables (ordered). Monotone missingness pattern is common in longitudinal data if the data are ordered by observation time (i.e. subjects that are lost to follow-up).

Monotone pattern is relevant because some methods to deal with missing data such as Multiple Imputation or Inverse Probability Weighting can be simplified.

Figure 14: Arbitrary or random pattern of missingness

	Variables					
ID	**X1**	**X2**	**X3**	**X4**	**X5**	**X6**
1	1	33	A	0	5	89
2	2	48	A	1	4	.
3	1	54	B	.	.	.
4	2	15	A	.	.	.
5	2	.	.	.	.	.

2.4.5 Methods to deal missing data

During the last few decades, ad-hoc or conventional methods, such us Complete-Case (CC) or Indicator Method (IM), have attempted to cope with missing data. However, it has been previously shown that these methods might produce biased results, loss of power and they are not based on statistical principles (89-91). On the other hand, the main advantage of conventional methods is their simplicity. They are easy to implement in most of the scenarios, even when specific software for missing data is not available. Some of these methods are:

Complete-Case

CC or likewise deletion consists of restricting the statistical analyses to the cases with complete information for all the variables in the model. The strength of this method is that it is easy to implement, it can be used for all statistical methods and if data are MCAR, it does not introduce any bias in the parameter estimates. The final

data set (after deleting the incomplete cases) is a random subsample of the initial sample (Figure 15). Consequently, if the missing data mechanism is MAR or MNAR the results might be biased. Besides, if the proportion of cases with missing data is large, this method would lead to an important lack of statistical power (89).

Figure 15: Complete-Case analysis data set

Original data set			
ID	**Y**	**X**	**R**
1	5	4	1
2	4	.	0
3	.	2	0
4	3	.	0
5	4	5	1

Complete-Case data set			
ID	**Y**	**X**	**R**
1	5	4	1
5	4	5	1

Y: outcome; X:Covariate; R:Complete-Case

In the literature, it has been often suggested that CC approach is only valid under MCAR mechanism (92) nevertheless this method is also valid under certain Non-MCAR (MNAR and MAR) mechanisms. CC approach validity has been extensively discussed by White & Carlin (93).

Briefly, considering the situation that missing values are present in both the outcome (Y) and the covariates (X) and that missingness can be solely explained by the covariates (the probability of being a complete case is independent of the outcome given the covariates $\mathrm{p}[\mathrm{R}=1|\mathrm{X},\mathrm{Y}]=\mathrm{p}[\mathrm{R}=1|\mathrm{X}]$), then:

$$p[\mathrm{Y}|\mathrm{X},\mathrm{R}=1]=\frac{\mathrm{p}[\mathrm{Y},\mathrm{X},\mathrm{R}=1]}{\mathrm{p}[\mathrm{X},\mathrm{R}=1]}=\frac{\mathrm{p}[\mathrm{R}=1|\mathrm{X},\mathrm{Y}]\mathrm{p}[\mathrm{Y},\mathrm{X}]}{\mathrm{p}[\mathrm{R}=1|\mathrm{X}]\mathrm{p}[\mathrm{X}]}=\frac{\mathrm{p}[\mathrm{R}=1|\mathrm{X}]\mathrm{p}[\mathrm{Y},\mathrm{X}]}{\mathrm{p}[\mathrm{R}=1|\mathrm{X}]\mathrm{p}[\mathrm{X}]}=\mathrm{p}[\mathrm{Y}|\mathrm{X}]$$

Hence, Complete Case approach is unbiased.

When missing values are present in the outcome, the condition for validity is the same than under the MAR mechanism. Besides, with missing data in the covariates, several situations and conditions for validity are possible. For example, in the restricted situation of just one covariate, CC approach is unbiased and provide the maximum likelihood analysis (94), if the data are MCAR or follow the MNAR mechanism $p(R|X,Y)=f(X)$. However, under the MNAR mechanism $p(R|X,Y)=f(X+Y)$ or the MAR mechanism $p(R|X,Y)=f(Y)$ maximum likelihood analysis is not guaranteed. In conclusion, CC analysis can be the maximum likelihood analysis even under some MNAR mechanisms.

Nevertheless, even under this particular situation, CC approach does not make full use of the observed information, since the information from incomplete cases is not included.

Finally, frequency estimates (such as prevalence, rates) from CC approach are always biased if data are MAR (rather MCAR) whereas conditional associations may not be (if we condition on the MAR mechanism variables).

Indicator Method

IM creates an extra category for missing values in each incomplete, independent and categorical variable (Figure 16).

Figure 16: Indicator method analysis data set

Original data set			
D	Y	X	R
1	5	0	1
2	4	.	0
3	4	1	1
4	3	.	0
5	4	1	1

Indicator Method data set			
D	Y	X	R
1	5	0	1
2	4	9	0
3	4	1	1
4	3	9	0
5	4	1	1

Y: outcome; X=Covariate; R=Complete-Case

The name of IM was taken because it is first used in linear regression analyses where an indicator variable was created for missing values in each covariate X. Suppose we want to explore the effect of Educational Level X (Low, High, missing values) on Delayed Diagnosis Y. Then the regression formula applying Indicator Method to deal with missing data would be:

$$Delayed\ Diagnosis = \alpha + \beta_1\ Indicator\ high\ education + \beta_2\ Indicator\ Missing,$$

$$where\ Indicator\ Missing = \begin{cases} 1\ missing\ value \\ 0\ other \end{cases}$$

$$where\ Indicator\ high\ education = \begin{cases} 1\ high\ education \\ 0\ other \end{cases}$$

This method has been widely used in epidemiological studies due to its simplicity. Nevertheless it lacks of statistical principles and provides biased results in most situations (90,95). When all the missing values of certain variables are grouped into just one category, it might occur that very different classes (unseen values) are lumped into just one group and subsequently, significant bias might arise. In addition, this method is not valid to adjust for confounding factors.

Imputation methods

Imputation methods are based on the practice of "filling in" the missing values. The information collected in the sample is used to assign one or several values to those variables with missing values. The main advantage of imputation methods is that a complete data set is generated and afterwards standard statistical methods can be applied.

Simple mean or regression mean imputation

The base of both these methods is replacing each missing observation by a single value and subsequently producing the originally intended complete data set. The statistical analysis is performed on the complete dataset.

Simple mean imputation replaces each missing observation in X by the completers mean for X. This method is not appropriate for categorical variables where the average of the categories is not a plausible value. In addition, the variance of the estimates is underestimated because all missing values are imputed with the same value. Therefore, these methods would lead to biased results.

Regression mean imputation replaces each missing observation in X with the predicted value calculated by fitting a regression model where X is the outcome and the observed variables are the explanatory variables. For continuous variables, linear regression is used. Under MAR mechanism the estimates of the means are unbiased

nevertheless variances are still underestimated and therefore the associations would be biased.

Both methods do not account correctly for the uncertainty due to the presence of missing data, variances are underestimated and therefore results are biased.

Random or stochastic regression imputation

Random or stochastic regression imputation seeks to acknowledge the uncertainty due to the presence of missing data (missing values cannot be perfectly predicted by the other observed variables).

The missing values are imputed drawing random values from the conditional distribution of the variable being imputed, given the other observed variables. In other words, to create an imputed value, an appropriate random residual is added to the value predicted using regression mean imputation.

For example, in normal linear regression to create an imputed value, a random draw from a normal distribution with mean equal to zero and standard deviation equal to the regression equation residual error, is added to the predicted value.

Estimates using this method to deal with missing data are unbiased if the data mechanism is MAR and the regression model to create the random or stochastic imputations is correctly specified.

Multiple Imputation methods

The disadvantage of imputation methods with a unique stochastic imputation is that they do not consider the variability introduced by the imputation process. Multiple Imputation (MI) is an imputation method developed by Rubin (96) that provides appropriate measures of precision, under MAR assumption. Multiple Imputation techniques assign several imputed values to each missing value. Secondly, standard statistical methods are applied to each completed data set and the average across imputations is calculated. Variances are estimated using Rubin rules taking into consideration the variability between the imputed values.

Multiple imputation method includes all the parametric and non- parametric techniques that assign several imputed to each missing value and use Rubin rules to combine the results.

Other advanced methods

Other advanced methods can be also used to deal with missing data: maximum likelihood estimation, Bayesian models or weighting methods.

Maximum likelihood estimation method models simultaneously the outcome and the reason why data are missing. This method selects as parameter estimates those values which, if true, would maximize the probability of observing what has, in fact, been observed (97). An example of this would be the use of random effects models that incorporate information on partially observed variables from intermediate

time points; mixed models correct automatically for missing data when there is not strong evidence against assuming data are MAR.

Bayesian methods estimate a statistical model for full data (including missingness mechanism and the outcome). This approach considers missing data as additional unknown parameters for which a posterior distribution can be estimated; it is essential to build appropriate joint model for the observed and missing data and model parameters (98).

Weighting methods calculate the predicted probability for certain variable to be observed of each patient and use these weights in the outcome model to account for missing data.

In general, these advanced principled methods produce more precise estimates than CC approach however they also depend on missing data mechanism assumptions and in some cases on the correct definition of the missing data model. Besides, these methods are not usually implemented in most statistical software packages and therefore require programming expertise.

2.4.6 Multiple Imputation by Chained Equations

Multiple Imputation by Chained Equations (MICE) is a particular multiple imputation technique that allows to impute missing values in multiple variables when the missing pattern is not monotone (random or arbitrary) and under MAR

assumptions. Besides logistic, multinomial or ordered regression can be used instead linear regression for non-normal variables (99,100).

MICE approach is also known as fully conditional specification (99) and sequential regression multivariable imputation (101). This method assumes that a multivariable joint of the data exists, however it does not specifies it, then samples from Gibbs sampling of the conditional distributions are generated.

MICE procedure

MICE methods follow these general steps:

1) Initially all missing values in each variable are filled in at random with an initial value;
2) Then from one of the incomplete variables, *x*, missing values are discharged (back to missing)
3) And *x* is regressed (using appropriate regression model in each case) on all the other variables (complete and incomplete). In other words, x is the dependent variable of the regression models and all other variables are the independent variables. This regression model must satisfy same assumptions than usual regression models.
4) After this, missing values in *x* are imputed using simulated draws from the posterior predictive distribution of *x (random regression imputation)*. This process is repeated for all the variables with missing data; when any x is subsequently used in the regression models of the rest of the variables, both the observed and the imputed values will be used.
5) Once each variable has been imputed the first *cycle* has been finished. Then, to generate an imputed dataset, the process is repeated to stabilize the results.

6) The complete process is repeated *m* times, *m* datasets are produced and then the convenient statistical analysis is performed in each of these datasets.
7) Finally, the results from each dataset are combined using Rubin rules (96,100) as follows:
    a. The parameter of interest, β it is calculated as the mean of the parameters obtained in each completed dataset

$$\bar{\beta} = (1/m) \sum_{j=1}^{m} \hat{\beta}_j$$

    b. The variance has two components:
        i. Within-imputation variance (U): the mean of the variances in each dataset m

$$\bar{U} = (1/m) \sum_{j=1}^{m} \hat{s}_j^2$$

        ii. Between imputation variance: additional variance due to the uncertainty about the missing values

$$\mathrm{B} = (1/(m-1)) \sum_{j=1}^{m} (\hat{\beta}_j - \bar{\beta})^2$$

        iii. The total variance is the sum of U and B corrected by for a finite number of imputations:

$$T = \bar{U} + (1 + m^{-1})B$$

If the missing data pattern is monotone, imputation process becomes much simpler as imputation can be performed sequentially: initially the variable with least missing values is imputed, then the second variable with least missing values, etc.

MICE modelling

Variable selection

To avoid bias in the analysis of interest, the MICE model for missing values must include: all model covariates (complete and incomplete), the analysis outcome and if possible auxiliary variables.

Auxiliary variables are those variables (usually without missing values) that are not included in the analysis of interest. However, they can improve the imputation models. The inclusion of auxiliary variables that are strongly related to the variables being imputed will always increase efficiency because we maximize the information used (102). On the other hand, the inclusion of variables that are not related with the variable being imputed (even if related with missingness) may slightly decrease efficiency but would not produce biased estimates (103).

The imputation model usually requires a large number of variables because it is used for several statistical analyses and it should include all the variables considered in these analyses. However, complex regression models, such us multinomial regression, might fail to converge when the number of categories is high and/or a high number of categorical covariates are included. One possible solution is adding one by one the covariates to the model, to identify the variable/s that are generating the lack of converge. Afterwards, the problematic variable/s must be excluded from the model or re-categorized in order to create a feasible imputation model.

Non-normally distributed variables

Regression models for missing data must satisfy all other regression models assumptions. The consequence of imputing non-normally distributed variables, by

assuming normality, is that the distribution of imputation values might not resemble the one of the observed values and therefore the imputation model will not be valid.

Categorized variables

Categorized variables which originally are continuous (for example age) must be imputed as continuous and then categorized in each imputed data set.

Interactions and non-linearities: different approaches

In general, the imputation models should reflect the structure of the analysis of interest, and should be at least as rich as the analysis model (including non-linear terms and interactions). Ignoring interactions and non-linearities at the imputation stage leads to attenuated (closer to zero) estimates of the corresponding parameters.

All relationships investigated in the substantive model must be included in the imputation model. For example, if we want to explore if the effect of calendar period on mortality differs by HCV status; then, we need to include the interaction between calendar period and HCV status in the imputation model. All interactions that want to be explored in the outcome analyses should be included in the imputation models to avoid bias; however this may be computationally non feasible. In this case it has been suggested to set up an expert committee to decide which interactions must be included (104).

Note that there is no universally agreed method for imputing non-linear terms (105,106), and it is a very controversial topic. The most common approaches to deal with interactions and non-linearities are:

JAV (Just Another Variable): the interaction or non-linear term is created and added to the imputation model as another variable to impute. The main disadvantage of this method is the deterministic relationship with the source variables (variables which interaction are non-linear terms are created from) is not kept.

Passive approach: the variables that form the interaction term are imputed in the usual way and the interaction term is then imputed passively (interaction is created based on imputed values). The disadvantage of this method is that the interaction term is not considered in the imputation model and this might lead to lack of power to detect such an interaction in the model of interest; however it keeps the deterministic relationship between the variables and it does not add new variables to impute.

Passive improve approach: this method is based in congenial imputation models (105). Congeniality means that although the models are not fully correctly specified they are compatible with some larger data. It is necessary to meditate deep about all the relationships within the data and to include most of these in the imputation model.

For example, let us suppose that the association between calendar period and mortality differs by HCV status; then congeniality implies that the association between HCV status and mortality differs by calendar period and therefore the interaction between calendar period and mortality has to be also included in the imputation model for HCV (not only the interaction between HCV status and calendar period).

Separate imputation: if one of the variables of the interaction-term is a categorical variable and it is fully observed, then another possible approach is to impute separately in each group of the categorical variable. Following with the previous example, this would be to build and perform the imputation model separately by calendar-period status (when calendar period is fully observed).

Imputation in a subgroup

MICE approach allows to impute separately (i.e. by gender), in different subgroups (i.e. number of cigarettes only imputed for smokers), conditional on a specific range of values (i.e. age possible value from 20-45).

Outcome imputation

If missing data are present in both the outcome and the independent variables we need to impute the outcome since this imputed outcome must be included in the imputation models for the independent variables (23).

The inclusion and imputation of missing outcomes has been largely discussed (100,106). Shortly, it has been shown that cases with imputed outcomes contain no

information (log-likelihood) on the relationship between the outcome and the covariates if auxiliary variables (complete variables that are not considered in the substantive model) are not included in the imputation model. In this latest case, when the auxiliary variables are correlated with the outcome results are more precise due to standard error reduction and, when the auxiliary variable are both correlated with the outcome and with the probability of the outcome being missing, biased is reduced (100).

After imputation of the missing outcome, two different approaches can be considered:

- Multiple Imputation by Chained Equations: Cases with imputed outcome are included in the analysis of interest.

- Multiple Imputation, then deletion (MID) (106): cases with missing outcome are included and imputed in the imputation models. However, after imputation is performed and m datasets are created, cases with imputed outcome are deleted from each dataset and therefore the analysis of interest is only performed on those with initial complete information on the outcome.

  The justification for using MID is that despite deleting cases with imputed outcomes increases standard error within each imputation, when the influence of randomly imputed values is reduced (by deleting the imputed cases) then the difference in the estimates from one imputed dataset to another is also reduced and subsequently the variance between imputed data sets its also reduced.

Number of imputations and proportion of missing data

Standard guides about Multiple Imputation suggest that a small number of imputed datasets (3 to 10) yield adequate results (96). However, White et al suggest a rule of thumb based on Bodner's approximation: the number of imputed datasets (m) should be at least equal to the percentage of incomplete cases (100).

On the other hand, there is not established a cut off from the literature regarding an acceptable percentage of missing data that still produces valid inferences. In principle, MICE should handle large amounts of missing data. The error terms of variables with a high proportion of missing values will be larger than those with fewer missing values and consequently the capacity to detect significant associations will be limited. Nevertheless, multiple imputation is computationally intense hence when the number of incomplete cases is large (i.e. 50%) the imputation process might be unfeasible.

Guidelines of the number of imputations needed, depending on the number of incomplete cases and the relative efficiency desired have been published (96,107). However, the size of the data set and the computational resources available must be also considered.

3 Justification

Few studies have compared overall and cause-specific mortality of HIV-positive subjects with that of the uninfected general population nor have they examined trends in excess mortality of HIV-positive patients; and in Spain, this has not been done. CoRIS cohort includes data of HIV-positive persons recruited during a period where highly effective antiretroviral treatment was available and all patients are naïve to treatment at study entry. However, we hypothesize that even though these patients may be in a better starting point than patients in other studies, the risk of mortality compared with the general population is still higher.

Moreover, relative survival method captures both direct and indirect mortality and provides a measure of the excess mortality experienced by HIV-positive patients. Comparing with cause-specific mortality usual analysis (HIV/AIDS mortality), where only deaths directly attributable to HIV are considered, relative survival method also considers causes of deaths such as HCV and HBV coinfection related, or suicide, indirectly associated with being HIV-positive.

In addition, liver and non-AIDS defining malignancies related mortality are the second and third more frequent causes of death among HIV-positive subjects and cardiovascular disease has raised concerns due to the aging process of the HIV-positive population (5,25); therefore, it is relevant to estimate the liver, non-AIDS malignancies and cardiovascular excess mortality (cause-specific excess mortality) associated with being HIV-positive.

Overall and cause specific excess mortality must be calculated to determine how mortality rates in HIV-positive subjects are related to mortality rates in general population and to assess the effect of prognostic factors. Determining which modifiable factors may contribute to excess mortality is essential in decision-making and policy planning needs.

Limited data are available on the impact of HCV coinfection on the mortality in HIV-positive patients, in Spain. Despite the fact that the prevalence of HCV infection in new HIV-positive patients has decreased sharply in parallel with a decrease in IDU as mechanism of HIV transmission (108), HCV coinfection in HIV-positive subjects remains high, and mortality due to liver diseases in HIV populations partly replaces mortality from AIDS prior to cART (109).

In addition, the contribution of HCV to an increased incidence of different CoD, others than liver or AIDS-related, and unrelated to lifestyles remains to be proven. These data could also contribute to disentangle the pathogenesis of these co-morbidities which could then be used to develop preventive and curative strategies. Better estimates of the effect of HCV coinfection on cause-specific excess mortality are also needed to influence changes in the clinical management of HIV-positive patients in order to guide the care and diminish further the mortality of HIV-positive subjects that are also coinfected with HCV.

Missing or incomplete data are a common issue in cohort studies. Information is usually collected by clinicians during the clinical routine and missing values can arise due to diverse reasons.

Before performing the statistical analysis selected to investigate our objective, it is essential to explore and understand why observations in our study might be missing and to gather information, from both the data and the expert knowledge that might help to clarify this occurrence.

There is no consensus on which is the best method to deal with missing data in cohorts and it is unavoidable making additional assumptions. However it has been shown that incomplete information has been usually inadequately handled in cohort studies (83). An inappropriate approach, without accounting for the missing data patterns or with an incorrect model specification, could lead to biased results. For that reason, it is essential to give an approach on how to handle missing data in cohorts of HIV-positive patients.

4 Objectives

4.1 Objective 1

To assess trends in overall mortality and cause-specific mortality stratified by HCV serostatus in a cohort of naïve HIV-positive patients in Spain.

Objective 1.1

To evaluate overall and cause-specific mortality rates observed in HIV-positive subjects followed-up in the cohorts of the Spanish Network on HIV/AIDS Research -CoRIS and CoRIS-MD-, stratified by HCV serostatus.

Objective 1.2

To assess changes over calendar time in overall and cause specific mortality of HIV-positive individuals stratified by HCV serostatus.

4.2 Objective 2

To calculate the overall mortality rates, standardized mortality ratios and excess mortality rates in the cohorts of the Spanish AIDS Research Network -CoRIS-MD and CoRIS-, comparing the overall mortality rates observed in HIV-positive subjects in both cohorts with the mortality rates of the general population of similar age and sex.

Objective 2.1

To calculate the overall mortality rates, standardized mortality ratios and excess mortality rates in the cohorts of the Spanish AIDS Research Network, from 1997 to 2008, compared to the general population of similar age and sex in Spain.

Objective 2.2

To explore differences in the overall mortality rates, standardized mortality ratios, and excess mortality rates by socio-demographic, epidemiological and clinical characteristics.

4.3 Objective 3

To evaluate the overall and cause-specific excess mortality observed in HIV-positive subjects followed-up in the cohort of the Spanish Network on HIV/AIDS Research (CoRIS), compared with the mortality of the general population of similar age and sex, and to identify prognostic factors of excess mortality.

Objective 3.1

To calculate the overall, non-AIDS malignancies, liver and cardiovascular- related excess mortality observed in HIV-positive subjects followed-up in the Spanish Network on HIV/AIDS Research (CoRIS), from 2004 to 2014, compared with the mortality in the general population of the same age and sex in Spain.

Objective 3.2

To identify prognostic factors for overall, non-AIDS malignancies, liver and cardiovascular-related excess mortality.

Objective 3.3

To assess changes over calendar time in the overall, non-AIDS malignancies, liver and cardiovascular-related excess mortality of HIV-positive individuals compared with expected mortality in the general uninfected population.

4.4 Objective 4

To describe the distribution of incomplete data, to build an imputation model and to compare the estimated excess mortality in the cohort of the Spanish Network of HIV Research (CoRIS) using different methods to handle missing data.

Objective 4.1

To describe the distribution of incomplete data and to explore predictors of missingness in CoRIS.

Objective 4.2

To build an appropriate imputation model for missing data in CoRIS with a focus on model specification and to assess the imputation procedure.

Objective 4.3

To estimate and to compare the overall and non-AIDS malignancies-related excess mortality in CoRIS using four different methods to handle missing data.

5 Methods

5.1 Sources of Information /Study population

5.1.1 CoRIS-MD

This cohort has been previously described elsewhere (110).

Study design

CoRIS-MD is a multicentre and retrospectively assembled cohort of HIV-positive patients.

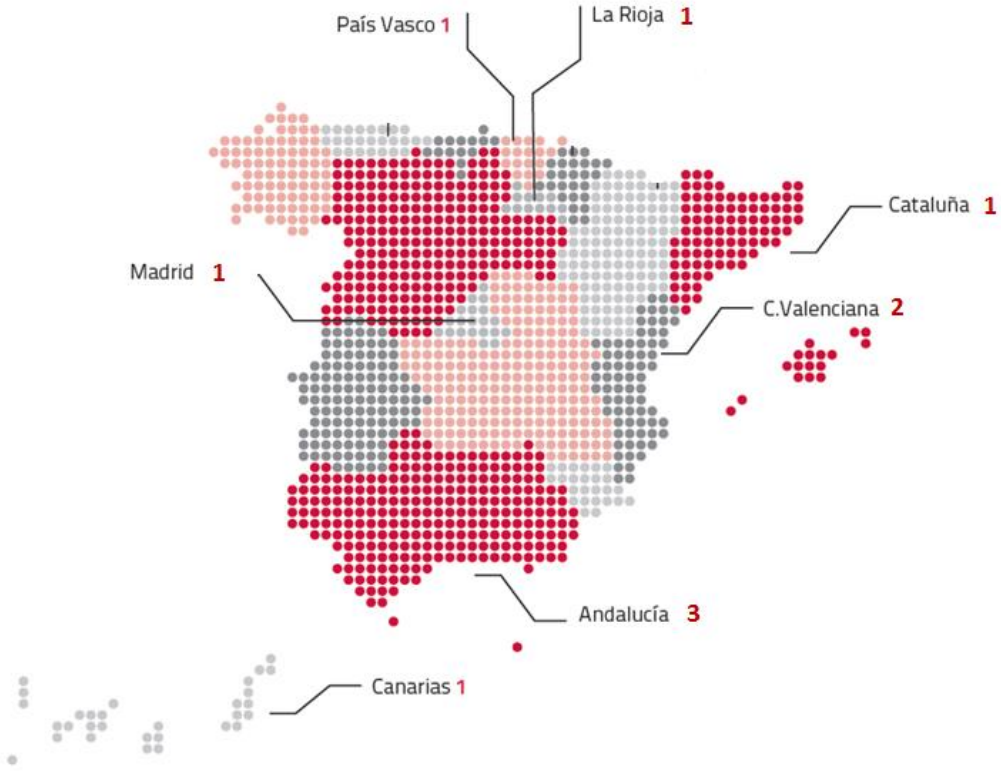
Time period

1997-2003

Setting, participating Centres, subjects and inclusion criteria

The total number of research groups that participated was 10 located in 7 of the 17 Autonomous Regions of Spain: Hospital Universitario Virgen del Rocio (I), Hospital Universitario Virgen del Rocio (II), Hospital Clinico Universitario de Granada, Hospital Universitario de Canarias, Hospital Universitari de Tarragona Joan XXIII, Hospital General de Elche, Hospital La Fe, Hospital General de la Rioja, Hospital de Donostia, and Hospital Ramón y Cajal. The total number of patients included in this study was 4643.

Figure 17: Map of participating centres in CoRIS-MD



Each participating centre included in the cohort all patients that fulfilled the following criteria:

- HIV-positive subjects attending any of these centres between 1997 and 2003
- Older than 18 years old
- To have more than 6 months of follow-up

Data collection and information collected

Each participating centre provided both a set of predefined variables and sent it in an exchangeable electronic format. Collected variables included: sex, date of birth, probable category of HIV transmission, date of first confirmed HIV-positive test, date of first HIV-positive test in the centre, AIDS defining diseases and dates, antiretroviral

therapy prescribed and starting and ending dates, CD4 counts and VL measurements at every visit, drop out dates and/or date of death.

The information collected from each centre was merged and a single data set was generated for statistical analyses.

Information of cause of death

A cross-check with the National Death Index was performed. Causes of death for all deceased subjects were obtained from the National Basic Death File (NBDF) provided by the National Institute of Statistics (NIS) in 2010 for the period 1997–2003. For 1997 and 1998, CoD was coded according to the 9th revision of the International Classification of Diseases (ICD 9) and was converted to ICD 10 codes.

The NBDF provides information on the date and CoD of all persons dying in Spain coded by the underlying CoD – defined as "the disease or injury which initiated the train of events directly leading to death or the circumstances of the accident or violence which produced the fatal injury" – in accordance with the ICD-10. The data are obtained from the civil registries and from the National Institute of Statistics itself through the Statistical Bulletin on Deaths which is compiled from death certificates. The underlying CoD is coded as the initial or basic cause as noted on the death certificate.

5.1.2 CoRIS

This cohort has been previously described elsewhere (111).

Study design

CoRIS is an open, multicentre and prospective cohort of HIV-positive patients who are naïve to antiretroviral therapy at study entry. It is linked to a BioBank (BBRIS) where biological samples (blood, plasma and DNA), from subjects included in the cohort, are preserved (112).

Study scope and time period

CoRIS is a national cohort that is integrated by all the health care centres that stated their interest in participating in the project (most of them are part of the Spanish Network of HIV/AIDS Research).

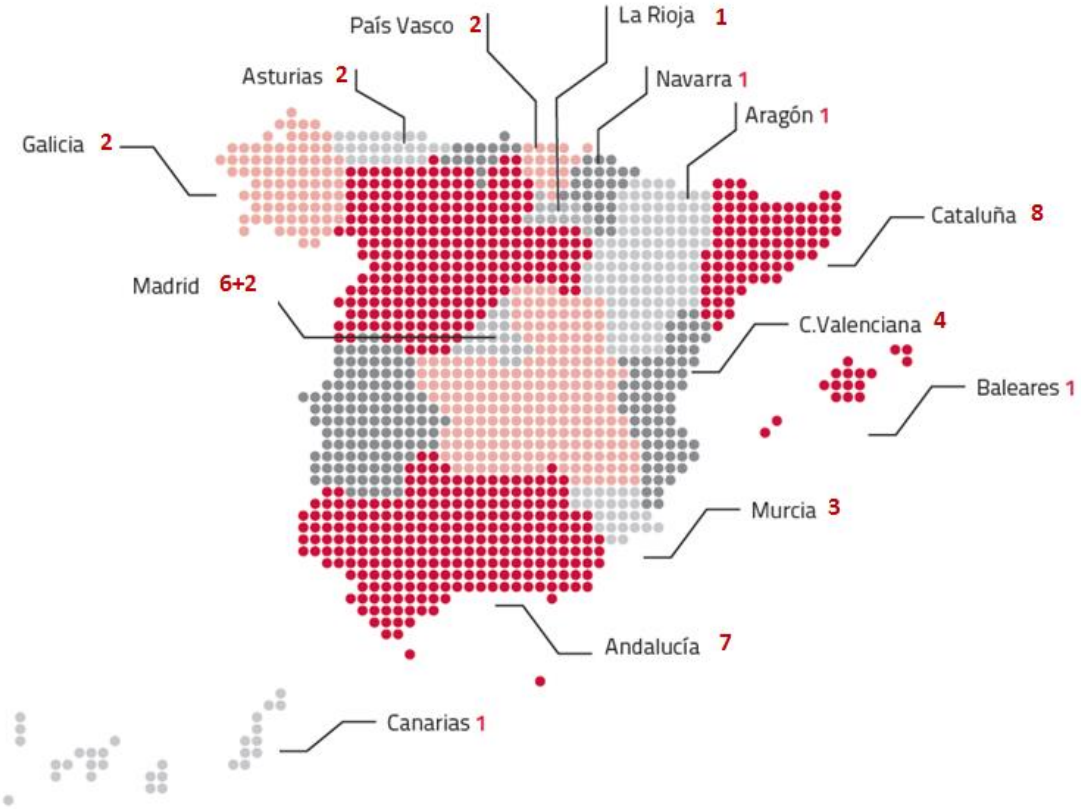
2004 – Onward. The recruitment and follow-up of patients in CoRIS cohort is still ongoing.

Participating Centres, subjects and inclusion criteria

At a first stage of this cohort, between 2004 and 2006, 19 public hospitals located in 9 of the 17 Autonomous Regions of Spain participated. From 2007 to 2011, 13 of the 17 Autonomous Regions of Spain participated and subjects were recruited in 28 health care centres (27 hospitals and one Sexually Transmitted Disease (STD) clinic).

From 2014 and onwards, 41 centres participate in CoRIS; 39 health care centres recruit and follow-up subjects and 2 centres provide only follow-up for those subjects that have been previously recruited in other centre participating in CoRIS.

Figure 18: Map of the participating centres in CoRIS



Each participating centre includes in the cohort all patients that fulfil the following criteria:

- Seen for the first time in the recruiting health care centre, after its incorporation in the cohort;
- To be older than 13 years-old;
- To have a confirmed HIV-positive diagnosis;
- Not having received antiretroviral therapy before cohort enrolment (cART naïve);

- To sign the informed consent to participate in the study. In the case of patients that fulfil the inclusion criteria but die before signing the informed consent, a minimum amount of information is anonymously collected.

Data collection, Quality Control and external audit

The frequency of patient visits is established by the responsible clinician and the patients themselves, usually every 4-6 months. The follow-up of the patients finishes when the patient dies, changes to another health care centre that does not participate in this cohort and therefore is not capable of providing his/her information or is lost to follow-up.

The coordinating team designed an electronic tool to collect the patient's information. Participating centres send, to the coordinating team, all the information collected since the inclusion of the patient in the cohort until his/her last visit. The coordinating team performs exhaustive data quality control and data checking; a report with the results of the quality control is sent to each participating centre a and then each participating centre sends, again, all the corrected or confirmed information to the coordinating centre.

All patients included in the cohort are given an univocal identifier, formed by their two first initials of both surnames, their date of birth and their sex. This procedure codes the dataset and allows to identify patients included in several participating centres. A patient can be included in one of the centres participating in the cohort and can continue his/her follow-up in a different one. Therefore, this method helps to minimize loses to follow-up in the cohort and participants duplicates.

Subsequently, the coordinating team merges the information from all participating centres and generates a single data set that is used to perform statistical analysis and that is sent for international data collaborations (CASCADE, COHERE, ART-CC, and HIV CAUSAL collaboration). The data update is performed annually.

Biennially, a specialized agency develops an external audit of 10% of the patients included in the cohort; it is verified that the information collected in the data base is concordant with the information in the medical records.

Collected Information

Data from patients are collected at enrolment and during follow-up. The variables collected include socio-demographic information, epidemiological, clinical, treatment and mortality data.

Biological samples (blood, plasma and DNA) from subjects included in the cohort are collected using the Biobank protocol (112).

Governance and Ethical Issues

In 2003, a scientific committee was established to elaborate the cohort operational protocol, which was approved by all the participating centres at that moment. CoRIS project was approved by several Institutional Ethics Review Boards.

All subjects included in the cohort had to sign an Informed Consent to participate in the study. This consent complies with the Spanish law, RD 1716/2011 of

18 November, that establishes the basic requirements for authorization and operation of biobanks for biomedical research and treatment of biological samples of human origin.

The corresponding personal data are incorporated into a computerized confidential file registered in the Spanish Data Protection Agency, under the terms established in Law 15/1999, of December 13, Protection of Personal Data.

Information of cause of death

The variables collected on mortality are: vital status, date of death, underlying CoD and contributing causes as free text variables.

A cross-check with the National Death Index was performed. Causes of death for all deceased subjects were obtained from the NBDF provided by the NIS in 2010 for the period 2004–2008. CoD was coded according to the 10th revision of the ICD 10.

In 2010, CoRIS established a cause-of-death committee formed by statisticians, clinicians, epidemiologists and coding experts. As detailed information was not always available, the cause-of-death committee decided to apply a simplified version of CoDe to CoRIS (‘Revised CoDe’).

For the purposes of this thesis, the CoD was coded according to two different coding algorithms: ICD-10 and the “Revised CoDe” (71).

5.1.3 Spanish National Institute of Statistics (NIS)

Death rates and number of deaths in the general population for all causes and for cause-specific deaths, from 1/01/2004 to 31/12/2013 and stratified by sex and age at 5 years intervals, were obtained from the NIS webpage (36). Information on yearly age death rates was not available therefore we assumed a constant overall and cause-specific death rate within each 5 years age interval.

Deaths in the general population are coded according the ICD-10.

5.2 Participants, follow-up and variables

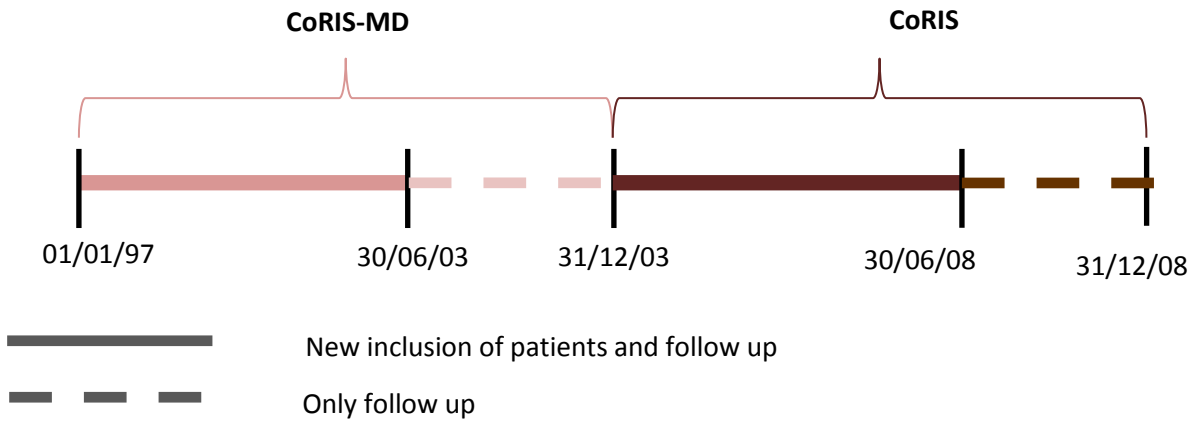
5.2.1 Objective 1

Participants and follow-up

For this objective, we included subjects recruited in CoRIS-MD at any time from 01/01/1997 to 30/06/2003; and in CoRIS from 01/01/2004 to 30/06/2008. To be eligible for these analyses subjects had to be naïve to cART at cohort entry, older than 19 years, had a follow-up of more than 6 months and had at least one diagnostic test for HCV.

Individuals were followed-up from study entry to death, last study contact or the administrative censoring date (31/12/2003 in CoRIS-MD and 31/12/2008 in CoRIS) whichever arose first.

Figure 19: Diagram of inclusion and follow-up of patients in CoRIS-MD and CoRIS



Variables

We considered the following variables: age at cohort entry (<30; 30-40; >40); sex (male, female); HIV transmission category (IDU, MSM, heterosexual contact, others); Clinical Staging and Disease Classification System (CDC) stage (Primo-infection, Asymptomatic, Symptomatic, AIDS-Indicator); CD4 cells/mm^3 count at entry[1] (<200, 200-350,>350); HIV VL copies/ml at entry[1] (<100000, ≥100000); cART initiation during follow-up; HCV serological status (positive or negative antibodies); vital status; cause and date of death.

The cut-offs used to categorize CD4 cell counts, VL and age have been widely used in the literature (113).

Cause of death

Deaths and date of death were ascertained through cohort reporting and/or a cross-match with the Spanish National Death Index. Causes of deaths for deceased subjects were obtained from NBDF. Data were cross-matched in the first quarter of 2010 for subjects who had died between January 2004 and December 2008, date of the last available update of the NBDF at date of request, according to the IDC-9 and IDC-10 revision.

[1] *Measured over a period of six months from the inclusion in the cohort (and prior to cART initiation for CD4 counts and Viral Load).*

[2] *Measured over a period of six months from the inclusion in the cohort (and prior to cART*

The following cause-specific deaths groups were created:

AIDS-related:

Codes: A02-A029, A073, A15-A19, A30-A31, A812, B00-B009, B20-B24, B250-B259, B371, B383-B389, B393-B399, B451-B459, B582, C53, C83

Liver-related:

Codes: B15-B19, C22-C229, K70-K77, K922, K65, R18

Non–liver-related non–AIDS-related (NLRD-NARD)

Rest of codes

5.2.2 Objective 2

Participants and follow-up

For this objective, we included subjects recruited in CoRIS-MD at any time from 01/01/1997 to 30/06/2003; and in CoRIS from 01/01/2004 to 30/06/2010. To be eligible for these analyses subjects had to be naïve to cART at cohort entry, older than 19 years, had a follow-up of more than 6 months and had at least one diagnostic test for HCV.

Individuals were followed up from study entry to death, last study contact or the administrative censoring date (31/12/2003 in CoRIS-MD and 31/12/2010 in CoRIS) whichever arose first.

Variables

We considered the following variables: age at cohort entry (20-29; 30-39; 40-49; >=50); sex (male, female); year of cohort entry; HIV transmission category (IDU, MSM, heterosexual contact, others); AIDS before entry and changes in AIDS status during follow-up; CD4 cells/mm^3 count at entry[2] (<200, 200-349,≥350); HIV VL copies/ml. at entry[2] (<20000, 20000-100000, ≥100000); cART initiation during follow-up; HCV serological status (positive or negative antibodies); vital status and date of death.

The cut-offs used to categorize CD4 cell counts, viral load and age has been widely used in the literature (113).

To calculate mortality rates, AIDS variable was classified as "Yes" when the person had AIDS before entering the cohort, AIDS at cohort entry or AIDS during follow-up and "No" when the person did not develop AIDS at any moment during the study.

[2] *Measured over a period of six months from the inclusion in the cohort (and prior to cART initiation for CD4 counts and Viral Load).*

5.2.3 Objective 3 & 4

Participants and follow-up

For this objective, we included patients who were recruited in CoRIS from 01/01/2004 to 31/05/2014 (administrative censoring date), cART naïve and older than 19 years old.

Individuals were followed up from study entry to death or last study contact, whichever arose first.

Variables

We considered the following variables: age at entry (20-49; >=50); sex (male, female); HIV transmission category (IDU, MSM, heterosexual contact, others); Level of education (no education or compulsory, Upper Secondary, University, others (no possible to classify according this system)); country of origin (Spain, Latin-America; Sub-Saharan Africa, others); AIDS status before entry; CD4 cells/mm^3 count at entry[3] (<200, 200-350,>350); HIV VL at entry[3] copies/ml (<100000, ≥100000); cART initiation at entry[3]; HCV serological status at entry[3] (positive or negative antibodies); vital status; cause and date of death.

The cut-offs used to categorize CD4 cell counts, VL and age have been widely used in the literature (113).

[3] *Measured over a period of six months from the inclusion in the cohort (and prior to cART initiation for CD4 counts, Viral Load and HCV test).*

Classification of deaths in the cohort and in the general population

Cause-specific death groups were created. We used two different processes to classify deaths in the cohort and in the general population, respectively. Cause-specific groups were created based on revised CoDe codes in CoRIS and ICD-10 codes were applied in the general population. Table 1 shows causes of deaths grouped in each category.

Table 1: International Classification of Diseases 10th Revision (ICD-10) and Revised Coding causes of death in HIV (CoDe) included in each cause of death group

Cause death group	ICD-10 code	CoDe code
AIDS-associated diseases	A02, A073, A15-A19, A30, A31, A812, B00, B20-B24, B25, B371, B383-B389, B393-B399, B451-B459, B582, B588-B589, C46, C53, C83, C857, C859, R75	01 (AIDS)
Liver disease	B15-B19, K65, K70-K77, K922, HCV or HBV with C22-C229	03(Chronic viral hepatitis), 14 (Liver failure)
Non-AIDS Malignancies	C00-C52 , C54-C82, C84-D48 (excluding HCV or HBV with C22-C229)	04 (Malignancy)
Infectious diseases	A00-A019, A03-A072, A078-A09, A20-A28, A32-A563, A568-A811, A813-A99, B01-B09, B25-B370, B372-B382, B39-B392, B40-B450, B46-B581, B583, B99, G00-G02, J12-J18	02 (Infection)
Cardiovascular disease	I00-I45, I47-I99	08 (Ischemic heart disease) 09 (Stroke), 24 (Heart or vascular disease)
Diseases of the blood	D50-D89	20 (Haematological disease)
Pulmonary diseases	J40-J99	11(Primary pulmonary hypertension), 12 (Lung embolus), 13 (Chronic Obstructive lung disease), 25 (Respiratory disease)
Central nervous system diseases	G048-G99	23 (CNS disease)
Drug abuse	F192, X40-X44, T36-T50, Y10-Y15	19 (Substance abuse)
External causes	X00-X30, X45-X49, X50-X59, X85-X90, V01-Y98	16 (Accident or violent death)
Suicide	X60-X84	17 (Suicide)
Other diseases	Any other code	Any other code
Il defined and unknown causes	R092, R95, R960-R961, R98, R99	91(Unclassifiable cause), 92 (Unknown)

CNS, central nervous system

Adapted from Hernando V. et al. Differences in the causes of death of HIV-positive patients in a cohort study by data sources and coding algorithms. AIDS. 2012 Sep 10;26(14):1829-34. Table 1 p 1831. (69)

For the aim of this thesis we analysed the following causes of death: liver disease, non-AIDS-related malignancies and cardiovascular disease.

5.3 Statistical analyses

5.3.1 Objective 1

To assess trends in overall and cause-specific mortality, follow-up was divided into two calendar periods (1997-2003 and 2004-2008), which correspond to period of follow-up of each cohort CoRIS-MD and CoRIS.

The descriptive analysis of baseline characteristics for individuals in the two calendar periods was performed using frequency distributions. Differences by calendar period were analysed using the $\chi 2$ test.

Overall and cause-specific crude mortality rates per 1,000 py of follow-up stratified by calendar period and HCV-status were calculated. We used Poisson regression to calculate crude and adjusted Incidence Rate Ratios (IRR) of death in the second period taking the first period as a reference and stratified by HCV-status. Multivariable models were adjusted for the following potential confounders: CD4 count at entry (to account for duration of infection prior to recruitment), age at entry, HIV-transmission category, and active cART as time-updated variable. cART was modelled as time-dependent covariate with an intention-to-treat analysis (once started on cART patient was assumed to remain on it). Robust methods were used to estimate 95% confidence intervals (95% CI).

Complete-Case method was used to deal with missing data in HCV test, including only those with at least one diagnostic test. Indicator method was used to

deal with missing data in the rest of covariates; an extra category for missing values was created.

All statistical analyses were performed in Stata version 11 (Collage Station, TX, USA)

5.3.2 Objective 2

To assess trends in mortality, SMR and excess mortality, follow-up was divided into two calendar periods (1997-2003 and 2004-2010), which correspond to period of follow-up of each cohort CoRIS-MD and CoRIS.

Descriptive analysis of patients' characteristics was carried out using frequency distribution for categorical variables and median (IQR) for continuous variables.

We calculated mortality rates, overall and according to socio-demographic and clinical characteristics, as the number of deaths per 100 py of follow-up with 95% CI calculated using the exact Poisson method.

SMR were estimated for all-cause mortality in CoRIS-MD and CoRIS, comparing with the overall mortality rates of the general population standardized by sex and age. SMR were estimated as the ratio of observed deaths to expected deaths, had our patients had the same distribution of mortality as the general population. SMR were calculated through Poisson models offsetting expected mortality rates, and adjusted for sex, age, category of transmission and HCV test result.

$$SMR = \frac{Observed\ number\ of\ deaths}{Expected\ number\ of\ deaths} \quad (1)$$

The expected number of deaths was calculated by applying the probability of death in the general population to the py distribution of the cohort. Life-tables from general population were matched to the cohort patients by sex, calendar period (year) and age in 5 years bands. Matching by other variables such us socio-economic level, transmission category or country of origin might have been valuable, however this information was not available in the general population life-tables.

A sensitivity analysis was performed to assess a possible selection bias due to the inclusion of more severely ill individuals. To avoid overestimation of SMR, when comparing with mortality in the general population, we may need to exclude a time period of 6 or 12 months after the cohort entry, excluding both observation time and patients with a studied outcome during a lag-time period. To determinate whether it is necessary to include a lag time, the SMR was calculated separately for the first 12 months after cohort entry and for all patients together.

Excess mortality rate or excess hazard rate at time t, *λ(t),* associated with being HIV-positive, was calculated as the difference between the death rate (hazard) observed in the cohort *h(t)* and the expected death rate of the background population *h*(t) (baseline hazard) (79,80).*

$$\lambda(t) = h(t) - h^*(t) \quad (2)$$

Excess mortality rates per 100 py of follow-up by covariates were calculated and confidence intervals were estimated using Poisson's exact method.

In terms of survival, the analogue of the excess mortality rate is the relative survival ratio *R(t)*, and it is calculated dividing the observed survival in the cohort ,*S(t),* by the expected survival $S^*(t)$.

To calculate mortality rates, SMR and excess mortality rates, AIDS, and cART were treated as time-dependent variables.

Complete-Case method was used to deal with missing data in HCV test, including only those with at least one diagnostic test. Indicator method was used to deal with missing data in the rest of covariates; an extra category for missing values was created.

All statistical analyses were performed by using Stata software (Version 11.0, College Station, Texas).

5.3.3 Objective 3

Descriptive analysis of patients' characteristics was performed using frequency distributions for categorical variables and median (IQR) for continuous variables.

We used death rates and number of deaths in the general population in 2013 to calculate excess mortality in 2014, as information on death rates in 2014 was not available when this analysis was performed.

Excess mortality rate or excess hazard rate at time t, *λ(t),* associated with being HIV-positive was calculated as in the objective 2:

$$\lambda(t) = h(t) - h^*(t) \qquad (2)$$

Relative Survival: modelling excess mortality

In objective 3, we took one step ahead to model the excess mortality observed in HIV-positive subjects using Relative Survival models.

Dickman et al revised in 2004 different methods to model excess mortality (analogously relative survival). Concurrently, it was shown that the survival model can also be estimated in the framework of generalized linear models using a Poison distribution assumption for the observed number of deaths (79).

Equation (2) can be re-written as

$$h(t) = \lambda(t) + h^*(t) \Leftrightarrow {}^{\mu_j}/_{y_j} = {}^{d_j^*}\Big/_{y_j} + \exp(\boldsymbol{x}\beta) \qquad (3)$$

We assume that the number of deaths, d_j, for observation *j* can be described by a Poisson distribution, $d_j \sim Poisson(\mu_j)$, where $\mu_j = \lambda_j y_j$ and y_j is person-time at risk for the observation. And through applying logarithms:

$$\ln(\mu_j - d_j^*) = \ln(y_j) + \boldsymbol{x}\beta \quad (4)$$

Where, $\mu_j = E[d_j]$ and d_j^* is the expected number of deaths.

Dickman et al 2004 explained that the equations presented above involve a generalized linear model with endpoint d_j , Poisson error structure, link function $(\mu_j - d_j^*)$ and offset $\ln(y_j)$. These models have been also previously described (114,115).

The exponential of β is known as excess Hazard Ratio (eHR) and this is interpreted as a common HR but in terms of excess mortality.

Excess mortality rates per 100 py of follow-up by covariates and confidence intervals were estimated based on the crude fitted regression models.

We used generalized linear models with Poisson error structure to estimate the excess mortality in HIV-positive patients compared to the general population and to assess the impact of multiple risk factors (education level, HIV transmission, country of origin, sex, CD4, VL and HCV, AIDS and age at entry). The selection of potential risk factors was based in the underlying conceptual framework.

These models assume piecewise constant hazards in each follow-up time interval and that the excess hazards for two groups of patients are proportional throughout all the follow-up time. We divided follow-up time (time since entry in the cohort to last contact or death) in first year after the inclusion in the cohort and from the second year to tenth year; it has been previously shown that mortality in the first year after inclusion is usually higher than in the following years (116). In this way, we assumed piecewise constant hazards in each interval. We checked the sensitivity of this assumption using one-year intervals of follow-up. To check the proportional excess hazards assumption, we included follow-up time by covariates interaction terms in the model and tested if the excess hazards ratios for these covariates varied by follow-up time.

Excess mortality by specific causes of death

Methods

We applied relative survival models to estimate non-AIDS malignancies, liver, and cardiovascular related excess mortality. Cause-specific excess mortality rates were calculated as the difference between the observed cause-specific mortality rates and the expected cause-specific mortality rate in the general population.

To fit these models, a competing risk framework was taken into account. The outcomes of interest were non-AIDS malignancies, liver or cardiovascular deaths hence the rest of causes of death were considered as competing events as they prevent the endpoint of interest from occurring (117).

A competing risk analogue of the generalized linear model with Poisson error structure to estimate the excess mortality in HIV-positive patients can be expressed as:

$$\ln\left(\mu_{ji} - d_{ji}^{*}\right) = \ln\left(y_{ji}\right) + \boldsymbol{x}\boldsymbol{\beta_i} \;\; let\; i$$

$$\in (liver, Non - AIDS\; malignancies, cardiovascular \;\; others)$$

where β_i represents the effect of the covariate *x* on excess of cause *i* mortality.

We fitted separate models for liver, non-AIDS malignancies and cardiovascular hence we allowed each cause-specific baseline hazards of excess mortality to be different (117). We did not allow common parameters between outcomes therefore the statistical analysis was standard; individuals who developed a competing event before the event of interest were censored on the occurrence date of the competing event.

For the purpose of these analyses death rates in the general population were stratified by CoD (non-AIDS malignancies, liver and cardiovascular related).

Crude and multivariable generalized linear models with Poisson error structure were performed to evaluate the association between overall, non-AIDS malignancies, liver and cardiovascular excess mortality and potential prognostic factors. CD4 count at entry (to account for the duration of infection at enrolment) and follow-up time intervals (to assume piecewise constant hazards in each interval) were included in the multivariable models regardless of statistical significance. For all other covariates, only

those with p-values <0.1 in the crude analysis, were included in the multivariable models.

Due to the small number of events and to avoid overfitting we did not consider the variable "AIDS before entry" for cause-specific multivariable models.

We also used these models to investigate changes in the overall and cause-specific excess mortality over calendar period of inclusion in the cohort adjusting for potential confounders. Year of inclusion in the cohort was divided into two calendar periods (2004-2007 and 2008-2014), each one corresponds to half inclusion period of the CoRIS cohort. Interaction terms between covariates and calendar period were added to the model to assess modifications in the effect of prognostic factors on excess mortality over calendar period of inclusion. The potential confounders considered were: follow-up interval, sex, age at entry, HIV transmission category, CD4 counts and HCV test result at entry.

Multiple Imputation by Chained Equations was used to deal with missing data. The results from the 12 imputed datasets were combined using Rubin's rules (96). Missing data exploration and multiple imputation model specification are detailed further in this document as part of objective 4.

All statistical analyses were performed using STATA (Version 13.0, College Station, Texas). **strs** module (118) was used to perform the models described above to CoRIS database.

5.3.4 Objective 4

Missing data exploration and description

Data checking in CoRIS was performed to explore the existence of missing values. Missing data mechanism and distribution by covariates were explored to identify possible reasons for incomplete data and to decide the best analytical approach to cope with these missing data. Percentage of missing values per covariate was calculated.

Multiple imputation methods are only valid under MAR assumption. Although this assumption cannot be definitely determined, incomplete data can be MAR conditional on certain covariates. We used frequency distributions and χ^2 test and built logistic models for missingness (equation 5) to explore which variables were predictive of missingness in the different variables.

$$logit\, \mathrm{p}(R_i = 1) = \beta_0 + \beta_1 X_i\,, \qquad i = 1, \ldots n \;\; and \;\; X = covariates \qquad (5)$$

$$R = missing\ value\ indicator$$

Multiple Imputation model for missing data

We built a Multiple Imputation by Chained Equations model to deal with missing data in CoRIS. We developed an imputation model for each variable with missing values including all covariates (complete and incomplete) considered for the analysis of interest, the analysis outcome and auxiliary variables.

We used transformation towards normality to handle departures from normality before imputation and then back-transformed to the original scale. We explored different power transformations and selected the one that made the data more normal distribution-like.

Categorized variables which originally were continuous were imputed as continuous and then categorized in each imputed data set.

Several approaches to handle interactions were explored. Finally, we decided to use passive approach in the imputation model.

Cases with missing CoD (outcome) were included and imputed in the imputation models because all variables involved in the model of interest have to be included in the imputation models.

The number of imputations performed, in these analyses, was 12 and the burn-in period was 50 iterations.

Assessing imputation procedures

Validity of imputation models was extensively checked and discussed. We examined residual plots of models fitted to observed data to check the distributional assumptions made by the imputation models (i.e. normality, heteroscedasticity, etc.)

The distribution of imputed variables was explored after performing MICE models to detect potential problematic variables or misspecifications of the imputation model. Summary statistics providing information of the observed values, the imputed values and the combined values for each variable were compared. Graphical tools such as histograms, quantile-quantile plots, and density plots were also used to assess differences between imputed and observed data and to check if imputed values are plausible.

Finally, MICE is an iterative method hence checking of convergence was evaluated using trace plots of means and standard deviations of imputed values from multiple chains.

Excess mortality

Generalized linear models with Poisson error structure (explained in detail in objective 3– page 93) were used to estimate the eHR of overall and non-AIDS malignancies mortality in HIV-positive patients using four approaches to deal with missing data and results were compared:

- Complete-Case: restricting to patients with complete data (full description page 43).
- Indicator Method: creating an extra category for missing values (full description page 45)
- Multiple Imputation by Chained Equations: developing an imputation model for each variable with missing values (also for CoD) (full description page 50)

- Multiple Imputation, then deletion (MID): deleting cases with imputed CoD after performing multiple imputation (full description page 56).

Sensitivity Analyses

We performed exhaustive sensitivity analyses using only complete cases and excluding individuals with imputed outcomes and the results were compared to those obtained using imputed values.

We checked whether the imputed values were reasonable under the assumptions that values were MAR. MAR and MNAR mechanisms cannot be distinguished based on the observed data, therefore we also tested robustness of key inferences to possible departures from the MAR assumption using extreme cases approach. We imposed a MNAR mechanism by replacing missing observations by extreme values of the variables and checking how the results varied. This sensitivity analysis is most convenient for categorical variables however it can be also applied to continuous variables by replacing the missing values by the bounds of possible values of the continuous variable. For example, repeating the analysis of interest assuming that all the missing values of HCV status are negative or assuming that all age values missing are 85.

Software

All imputation models were run with STATA 13. *MI impute chained package* was used for the imputation step. However, *Strs* module (118) to estimate relative

survival is not supported by *MI estimate* command, hence the pooling step was performed applying *micombine (ice package (119))* to the previously imputed data.

6 Results

6.1 Objective 1

6.1.1 Patients

We analysed 5,974 HIV-positive cART-naïve patients: 2,471 in the 1997-2003 (CoRIS-MD) period with a median follow-up of 4.3 years (IQR: 2.3-6.2) and 3,503 in the 2004-2008 (CoRIS) period with a median follow-up of 2.4 years (IQR: 1.3-3.6). The characteristics of the patients at inclusion are shown in Table 2.

In comparison with patients in the first period, patients in the second period were more frequently males, were less likely to have acquired HIV through IDU, were more frequently on CDC category C at inclusion in the cohort, and were less likely to be coinfected with HCV.

Table 2: Baseline characteristics of 5,974 HIV-positive cART-naïve patients at the inclusion in the study

	1997-03 (n = 2,471)	2004-08 (n = 3,503)	P-value
	n (%)	n (%)	
Male sex	1791 (72.5)	2737 (78.1)	<0.001
Age at entry-years			
<30	563 (22.8)	876 (25.0)	<0.001
30-40	1432 (58.0)	1425 (40.7)	
>40	476 (19.3)	1202 (34.3)	
HIV-transmission category-			
MSM	361 (14.6)	1583 (45.2)	<0.001
IDU	1329 (53.8)	530 (15.1)	
Heterosexual	575 (23.3)	1259 (35.9)	
Other/unknown	206 (8.3)	131 (3.7)	
CDC disease category C	282 (11.4)	490 (14.0)	0.003
CD4+ cells/mm^3			
<=200	608 (24.6)	1047 (29.9)	<0.001
201-350	372 (15.1)	713 (20.4)	
>350	881 (35.7)	1662 (47.4)	
Unknown	610 (24.7)	81 (2.3)	
HIV VL copies/mL			
<100,000 copies/mL	1076 (43.6)	2242 (64.0)	<0.001
≥100,000 copies/mL	434 (17.6)	1178 (33.6)	
Unknown	961 (39.0)	83 (2.4)	
HCV serostatus			
Negative	974 (39.4)	2814 (80.3)	<0.001
Positive	1497 (60.6)	689 (19.7)	

cART, combination Antiretroviral therapy; MSM, Men who have Sex with Men; IDU, Injection Drug Use; CDC, Centres for Disease Control and Prevention; HIV, Human Immunodeficiency Virus; HCV, Hepatitis C Virus.

6.1.2 Causes of death

During follow-up, 232 deaths were identified, 158 during the first period (85 AIDS-related deaths, 54 NLRD-NARD, and 19 liver-related), and 74 during the second period (42 AIDS-related deaths, 26 NLRD-NARD, and 6 liver-related). There were 58 NLRD-NARD among HCV-positive patients (cardiovascular disease n=14; accidental

poisoning n=9; accidental/traumatic events n =7; suicide n=6; non-AIDS-defining malignancy n=4; non-AIDS-defining infections n=4; and other causes n=14). There were 22 NLRD-NARD among HCV-negative patients (non-AIDS-defining malignancy n=7; cardiovascular disease n=5; accidental/traumatic events n=1; non-AIDS-defining infections n=1; and other causes n=8).

In both periods—and in both HCV-positive and HCV-negative patients—HIV/AIDS was the leading cause of mortality, followed by NLRD-NARD, and finally by liver disease (Table 3).

Table 3: Absolute number of deaths stratified by HCV serostatus and calendar period

	All patients		HCV-negative		HCV-positive	
	1997-2003	2004-08	1997-2003	2004-08	1997-2003	2004-08
Follow-up (person-years)	10,372.2	8,744.8	3,619.6	6,965.6	6,752.5	1,779.2
Total deaths	158	74	32	32	126	42
HIV/AIDS-related	85	42	16	22	69	20
Non-liver/ non-AIDS -related	54	26	13	9	41	17
Liver-related	19	6	3	1	16	5

HCV, Hepatitis C Virus; HIV, Human Immunodeficiency Virus.

6.1.3 Mortality rates

During 19,117 py of follow-up, 232 patients died (mortality rate: 12.1 per 1,000 py).

The crude overall mortality rate declined significantly (44%) from 1997-2003 to 2004-08. When we stratified by HCV serostatus, a significant 48% (IRR: 0.56 95% CI: 0.42; 0.75) decrease in the overall mortality rate was observed among HCV-negative

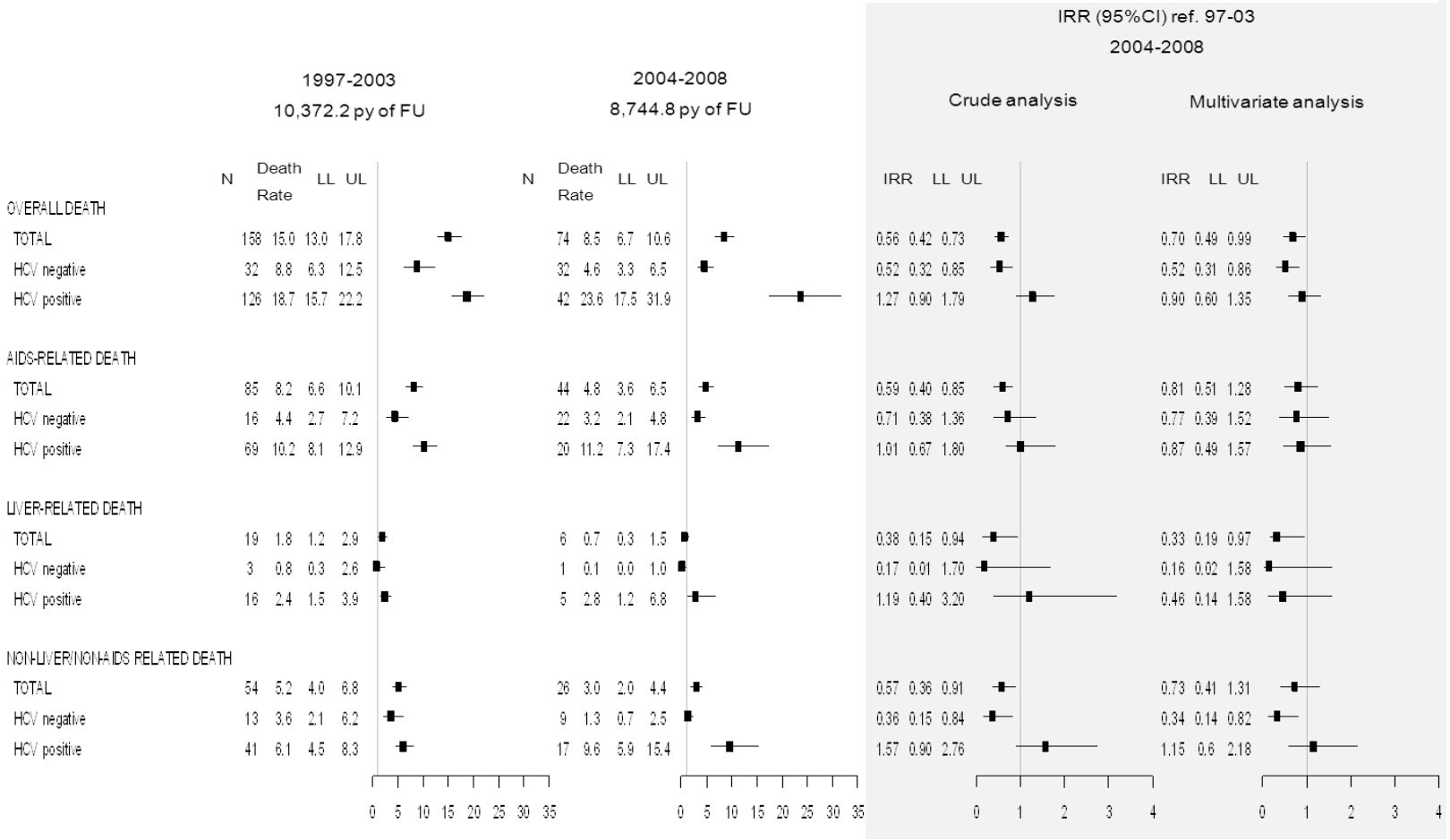
patients, whereas a non-significant 27% (IRR: 1.27 95% CI: 0.90; 1.79) increase was observed among HCV-positive patients (Figure 20).

The crude AIDS-related mortality rate declined significantly by 41% from 1997-2003 to 2004-2008 (IRR: 0.59 95% CI: 0.40; 0.85). When we stratified by HCV serostatus, the AIDS-related mortality rate declined from 1997-2003 to 2004-2008 among HCV-negative patients, although the difference did not reach statistical significance (IRR: 0.71 95% CI: 0.58; 1.36). The AIDS-related mortality rate remained virtually unchanged among HCV-positive patients (IRR: 1.01 95% CI: 0.67; 1.80) (Figure 20).

The crude Liver-related mortality rate declined significantly by 62% from 1997-2003 to 2004-2008 (IRR: 0.38 95% CI: 0.15; 0.94). When we stratified by HCV-serostatus, the liver-related mortality rate declined non-significantly from 1997-2003 to 2004-2008 among HCV-negative patients (IRR 0.17 95% CI: 0.01; 1.70) and remained unchanged among HCV-positive patients (IRR 1.19 95% CI: 0.40; 3.20) (Figure 20).

The crude non-liver-related non-AIDS-related mortality rate declined significantly (43%) from 1997-2003 to 2004-2008. When we stratified by HCV serostatus, non-liver-Related non-AIDS-related mortality rates decreased significantly by 54% from 1997-2003 to 2004-2008 (IRR: 0.36 95% CI: 0.15; 0.84) among HCV-negative patients but non-significant increases were observed among HCV-positive patients (IRR: 1.57 95% CI: 0.90; 2.76) (Figure 20).

Figure 20: Effect of calendar period (2004-08 vs 1997-03) on cause-specific mortality and mortality rates (and 95% confidence intervals) for each HCV serostatus



For each death category and HCV serostatus, the IRRs of the calendar period were calculated considering the 1997-03 as the reference category. One separate multivariate Poisson model for each cause of death was performed, adjusted for baseline CD4+ cell count, age at enrolment, HIV transmission category, and active cART as a time-updated variable. IRR, Incidence Rate Ratio; CI, Confidence Interval; py, Person-Years; FU, Follow-Up; LL, Lower Limit of 95% CI; UL, Upper Limit of 95% CI; HCV, hepatitis C Virus.

The effects seen in the crude analyses are maintained in the multivariable models presented in Figure 20. A statistically significant 48% (IRR: 0.52 IC 95%: 0.31; 0.86) reduction in all-cause mortality was observed for HCV-negative patients, while no reductions were seen for HCV-positive individuals.

6.2 Objective 2

6.2.1 Baseline characteristics of the study population

A total of 8,214 subjects were included in the study, 2,453 (29.9%) in the period 1997-2003 (CoRIS-MD) and 5,761 (70.1%) in the period 2004-2010 (CoRIS), adding up to 28,743 persons-year of follow-up, and 294 deaths.

Men represented 78.0% (n=6,412) of the sample, and median age at the cohort entry was 35.0 years (IQR: 30.2–41.0), 35.5 years (IQR: 30.2-41.7) for men and 34.2 years (IQR: 29.1-40.1) for women. Regarding transmission categories, the sample was distributed between IDUs or ex-users, 25.0% (n=2,050), MSM, 39.6% (n=3,255), and heterosexuals, 30.7% (n=2,524). A 20.4% of the subjects had a history of an AIDS defining illness (ADI), although for 59.4% (n=994) of them the ADI diagnosis was previous to cohort entry. A 30.9% (n=2541) were coinfected with HCV. Median CD4 count at cohort entry was 350 cell/mm^3 (IQR: 170–552), and median viral load was 39,811 copies/ml (IQR: 7,520–135,988) (Table 4).

Among the 294 deceased subjects, 80.6% (n=237) were men, and median age was 37.7 years (IQR 33.3 – 44.5). Some 60.2% (n=177) were IDU or ex-IDU, 51.0% (n=150) had an AIDS diagnosis and 67.4% (n=198) were coinfected with HCV. Median CD4 count at entry was 154 cell/mm^3 (IQR: 66–390) and median HIV viral load was 78,200 copies/ml (IQR: 17,335–230,000) (Table 4).

Table 4: Socio demographics and clinical characteristics at cohort entry for total of analysed subjects and deceased subjects

		Total		Deaths	
	py	n	%	n	%
Total	28,743	8,214	100	294	100
Gender					
Males	21,903	6,412	78.0	237	80.6
Females	6,839	1,802	22.0	57	19.4
Age at cohort entry (years)					
20-29	6,946	2,064	25.1	34	11.6
30-39	13,778	3,722	45.3	145	49.3
40-49	5,584	1,075	13.1	71	24.1
>=50	2,436	723	8.8	44	15.0
Median age (IQR)		35.0 (30.2-41.0)		37.7(33.5-44.5)	
Category of transmission					
IDUs	8,515	2,050	25.0	177	60.2
MSM	9,994	3,255	39.6	41	14.0
Heterosexual	8,909	2,524	30.7	67	22.8
Others/Unknown	1,325	385	4.7	9	3.0
AIDS					
No	22,255	6,542	79.6	144	49.0
AIDS before entry	3,667	994	12.1	72	24.5
AIDS after entry	2,821	678	8.3	78	26.5
CD4 count at entry (cells/mm^3)					
<200	7,525	2,217	27.0	141	48.0
200-349	5,191	1,567	19.1	39	13.3
>=350	12,366	3,744	45.6	64	21.8
Unknown	3,660	686	8.4	50	17.0
Median (IQR)		350 (170-552)		154 (66-390)	
HIV viral load (copies/ml)					
<20.000	8,748	2,769	33.7	61	20.7
20.000-100.000	7,141	2,202	26.8	65	22.1
>100.000	7,181	2,196	26.7	89	30.3
Unknown	5,673	1,047	12.8	79	26.9
Median (IQR)		39,810 (7,520-135,988)		78,200 (17,335-230,000)	
Cohorts					
CoRIS (2004-2008)	18,447	5,761	70.1	137	46.6
CoRIS-MD (1997-2003)	10,296	2,453	29.9	157	53.4
HCV test					
Negative	18,332	5,673	69.1	96	32.6
Positive	10,411	2,541	30.9	198	67.4
Antiretroviral treatment during follow-up					
No	9,993	1,948	23.7	63	21.4
Yes	18,751	6,266	76.3	231	78.6

IDUs: Injecting Drugs Users; MSM: Men have Sex with Men; HCV: Hepatitis C virus; IQR: Interquartile Range

6.2.2 Mortality Rates, Standardized Mortality Ratios and Excess Mortality Rates

Figure 21 shows mortality rates per 100 py of follow-up, standardized mortality ratios and excess mortality rates for 100 py in both RIS cohorts.

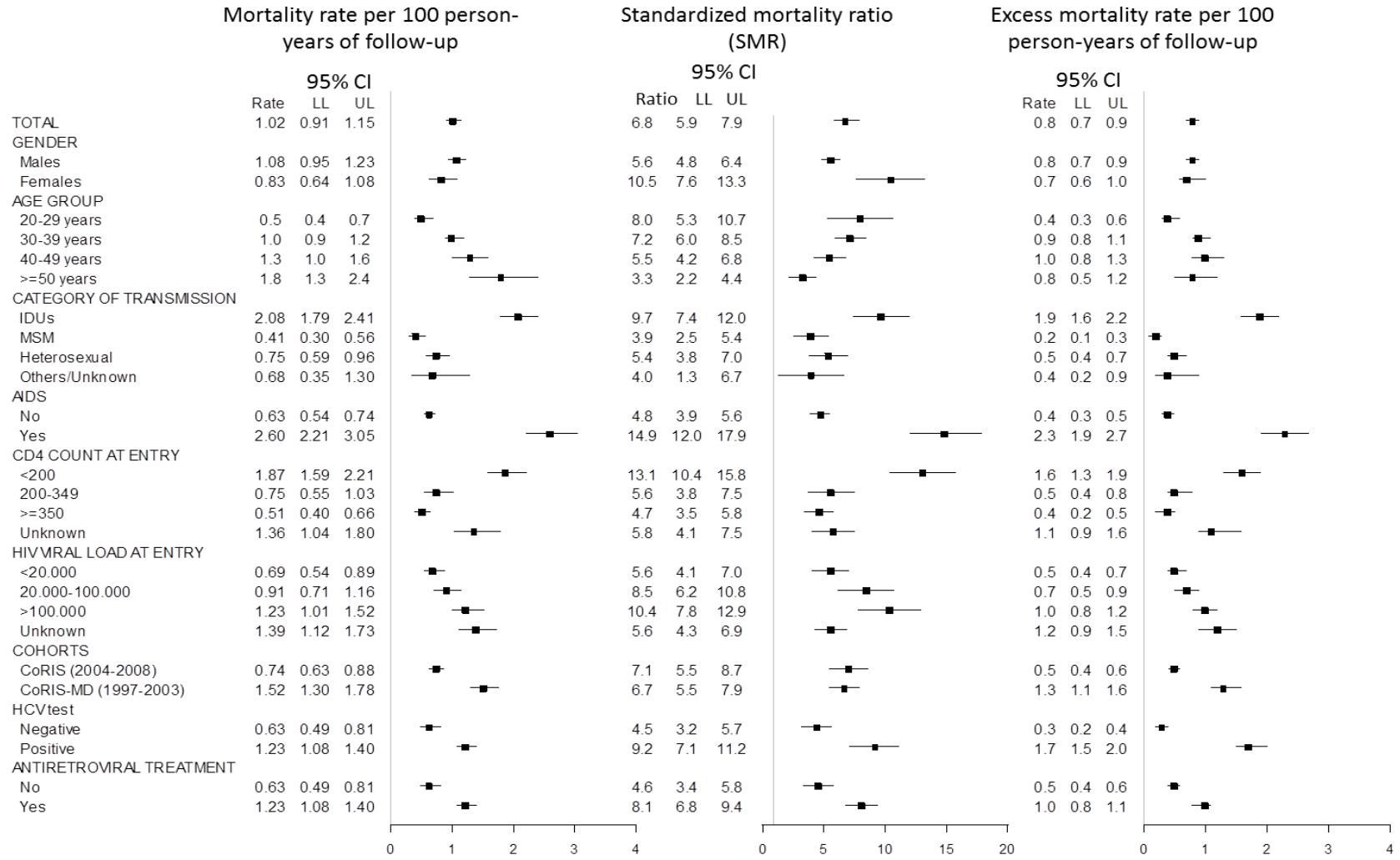
Overall mortality rate was 1.02 (95% CI: 0.91; 1.15) deaths per 100 py of follow-up, higher for men (1.08; 95% CI: 0.95; 1.23), for subjects over 50 years-old (1.81; 95% CI: 1.34; 2.42), for IDU (2.08; 95% CI: 1.79; 2.41) compared to both MSM (0.41; 95% CI: 0.30;0.56) and heterosexuals (0.75; 95% CI: 0.59; 0.96) and for patients included in CoRIS-MD (1.52; 95% CI: 1.30; 1.78). For patients who had an AIDS diagnosis, mortality rate was 2.06 (95% CI: 2.21; 3.05), compared to 0.63 (95% CI: 0.54; 0.74) for those who were AIDS-free. For HCV coinfected patients mortality rate rose up to 1.90 (95% CI: 1.65; 2.19) in contrast with 0.52 (95% CI: 0.42; 0.64) for those not coinfected.

Overall mortality in both CoRIS cohorts was 6.8 (95% CI: 5.9; 7.9) times higher than mortality of the general population of same age and sex. As opposed to the crude mortality rates, standardized mortality ratios were higher in women (10.5; 95% CI: 7.6; 13.3) compared to men (5.6; 95% CI: 4.8; 6.4). Still, a higher SMR was found for IDUs (9.7; 95% CI: 7.4; 12.0), persons with an AIDS diagnosis (14.9; 95% CI: 12.0; 17.9), persons coinfected with HCV (9.2; 95% CI: 7.1; 11.2) and those whom receiving antiretroviral treatment (8.1; 95% CI: 6.8; 9.4).

In the sensitivity analysis, considering only the first 12 month of follow-up, overall SMR was lower than in the complete analysis (4.0; 95% CI: 2.4; 5.6).

Finally, regarding excess mortality rate, as an absolute estimator, groups with the higher excess mortality were similar to those observed for crude mortality rates (Figure 21).

Figure 21: Mortality rates per 100 person-years of follow-up, Standardized Mortality Ratio (SMRs) and excess mortality rates per 100 person-years of follow-up according to socio-demographic, epidemiological and clinical characteristics



IDUs: Injecting Drugs Users; MSM: Men who have Sex with Men; 95% CI LL: Lower Limit; 95% CI UL: Upper Limit of the Confidence Interval
SMR adjusted by gender, age, category of transmission and HCV

6.3 Objective 3

6.3.1 Participants

The total number of patients was 10,340. The median follow-up was 3.20 years (IQR: 1.01- 5.66) and during 36,984.18 py of follow-up, 368 deaths were observed.

Socio-demographic and clinical baseline characteristics for total analyzed subjects and deceased subjects after imputation are presented in Table 5. After imputation of missing values, 83.5% were men, 39.0% had compulsory or no education. Regarding HIV transmission mode, 9.9% were infected IDU, 58.8% were MSM and 30.3% were heterosexuals, Spain was the country of origin of 69.3% of the patients. The median age at cohort entry was 35 (IQR: 29-43) years old. With respect to HIV disease markers at entry, 55.2% presented baseline CD4 cell count greater than 350 cells/mm^3 and 33.7% of the patients had HIV VL lower than 20,000 copies/ml, 1,019 (9.9%) patients had a previous history of AIDS defining disease. Overall, 13.5% had a positive HCV test at entry or within the first six months. The number of patients included during the period 2004-2007 was 4,541 (43.9%) and it was 5,799 (56.1%) during the period 2008-2014.

Among deceased patients we observed that the most common level of education was also no or compulsory education (64.2%) but the percentage was higher compared with the total of patients included in the analysis. The median age at entry was 43 (IRQ: 36-51) years-old, 41.1% were heterosexuals, 25.2% MSM and 31.3% IDUs.

Baseline clinical characteristics were poorer in the group of deceased patients compared with the total sample; 61.2 % had CD4 lower than 200 cells/mm^3 and 51.4% presented VL greater than 100,000 copies/ml at entry. The percentages of patients with a positive result in HCV test and/or AIDS defining disease at entry were also higher with values of 40.0% and 35.1%, respectively. The 72.0% of the deaths occurred in this sample happened in the first calendar period (2004-2007).

Socio-demographic and clinical characteristics of patients who died because of liver disease, non-AIDS malignancies or cardiovascular were also detailed in Table 5. After imputation, the number of liver-related deaths was 62 (95% CI: 55; 69), of non-AIDS malignancies-related deaths was 40 (95% CI: 35; 46) and of cardiovascular was 13 (95% CI: 9; 17).

Table 5: Distribution of patients included by vital status, Socio-demographic and clinical characteristics, after imputation

	ALL PATIENTS		DEATHS			
	py	N (%)	All causes	Non-AM[1]	Liver[1]	Cardio[1]
TOTAL	36984.18	10340 (100.0)	368 (100)	62	40	13
Education						
No/ compulsory	14824.27	4032 (39.0)	236 (64.2)	30	29	9
Upper Secondary	9814.982	2685 (26.0)	58 (15.9)	18	4	1
University	9054.512	2800 (27.1)	42 (11.4)	8	3	0
Other	3290.417	823 (8.0)	32 (8.6)	7	3	3
HIV transmission mode						
UDI	4252.226	1020 (9.9)	115 (31.3)	7	27	2
MSM	19729.57	6083 (58.8)	93 (25.2)	10	2	3
Heterosexual	12524.94	3131 (30.3)	151 (41.1)	23	11	6
Others	477.4456	106 (1.0)	9 (2.4)	0	0	1
Origin country						
Spain	26824	7167 (69.3)	305 (82.8)	53	37	9
SSA	1908.95	568 (5.5)	15 (4.1)	3	2	0
LA	5983.749	1845 (17.8)	32 (8.7)	4	0	2
Others	2267.484	759 (7.3)	16 (4.4)	2	1	1
CD4 at entry (cel/mm3)						
<=200	10015.65	2569 (24.8)	225 (61.2)	29	22	6
201-350	7496.354	2060 (19.9)	63 (17.1)	14	10	4
>350	19472.18	5711 (55.2)	80 (21.7)	19	8	2
VL at entry (copies/ml)						
<20000	12618.47	3481 (33.7)	79 (21.5)	17	13	6
20000-100000	12288.07	3461 (33.5)	100 (27.1)	16	14	4
100000	12077.64	3398 (32.9)	189 (51.4)	29	14	3
HCV at entry						
Negative	31238.85	8947 (86.5)	221 (60.0)	42	7	7
Positive	5745.326	1393 (13.5)	147 (40.0)	20	33	6
AIDS before entry						
No	32952.18	9321 (90.1)	239 (64.9)	46	33	11
Yes	4031.997	1019 (9.9)	129 (35.1)	16	7	1
Age at entry						
20-49	33221.24	9305 (90.0)	263 (71.5)	34	33	9
>=50	3762.938	1035 (10.0)	105 (28.5)	28	7	4
Median (IQR)	35 (29-43)		43 (36-51)			
Sex						
Male	29908.44	8633 (83.5)	298 (81.0)	50	30	11
Female	7075.743	1707 (16.5)	70 (19.0)	11	11	2
Period of inclusion						
2004-2007	25409.05	4541 (43.9)	265 (72.0)	48	34	11
2008-2014	11575.13	5799 (56.1)	103 (28.0)	14	7	2

Results after imputation are based on rounded mean values of the 12 imputed datasets; results may not always count up exactly to the total value

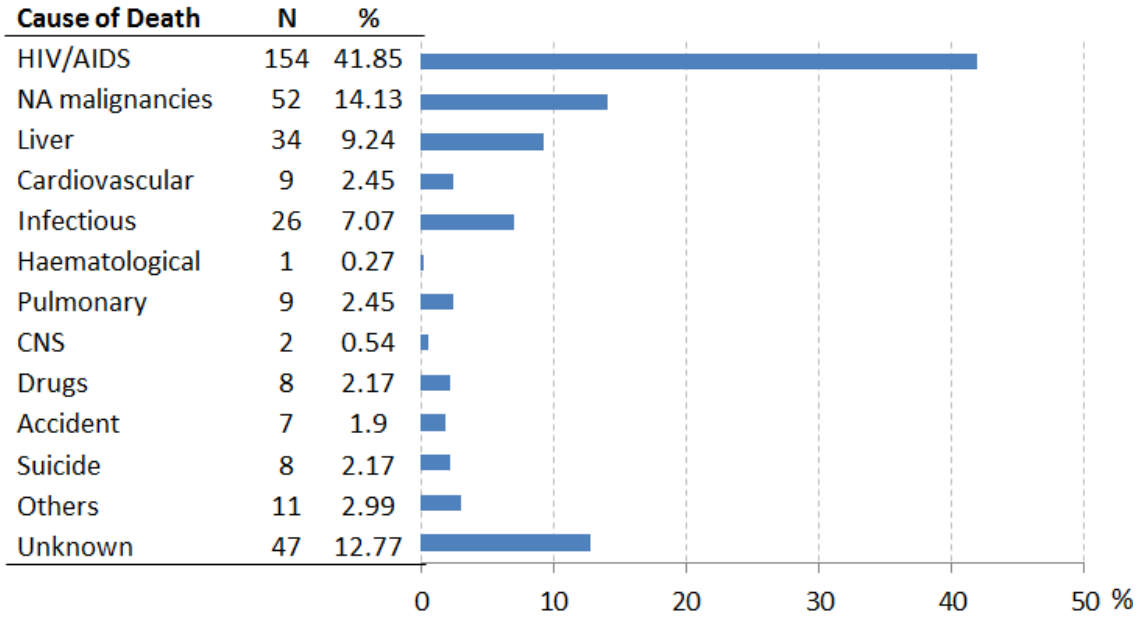
[1] *Number of events*

LA: Latin America; SSA: Sub-Saharan Africa

Non-AM: Non-AIDS Malignancies mortality; Cardio: Cardiovascular mortality

Figure 22 outlines the distribution of causes of death in CoRIS before Multiple Imputation. Nearly half of the deaths observed during the period 2004-2014 were HIV/AIDS-related (41.85%). The second most common CoD was non-AIDS malignancies (14.13%), closely followed by liver-related diseases (9.24%). Infectious diseases were responsible for 7.07% of the total number of deaths and cardiovascular diseases for 2.45 %. The rest of the causes (pulmonary, Central nervous system, drugs, accident, suicide or others) represented less than 3% of the deaths. CoD was unknown in 12.77 % of the deceased patients.

Figure 22: Distribution of causes of death before Multiple Imputation

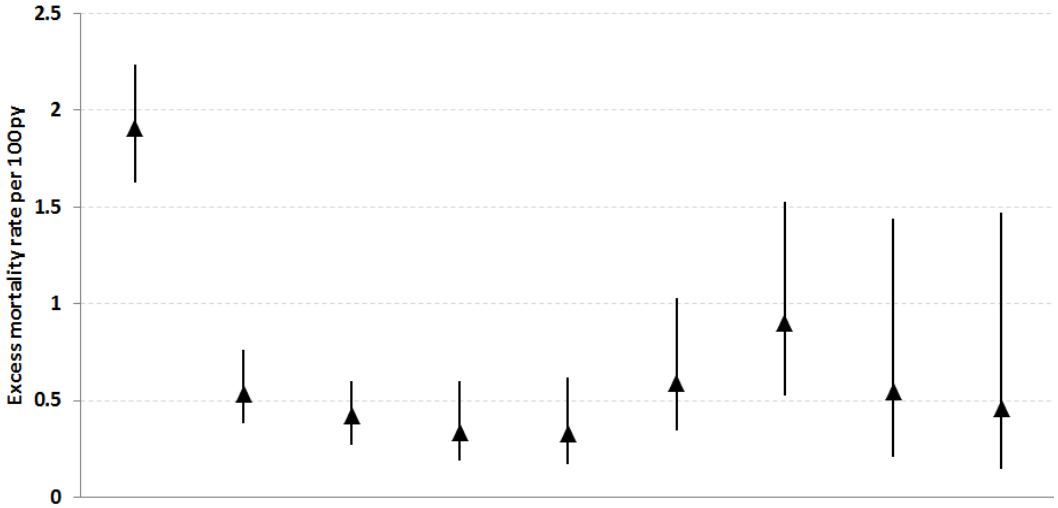


NA: Non-AIDS
CNS: Central Nervous System

6.3.2 Overall excess mortality

The overall excess mortality observed in the cohort was 0.82 deaths per 100 py of follow-up (95 CI%: 0.73; 0.93), that is, the mortality rate was 0.82 deaths per 100 py higher in CoRIS than in the Spanish population of the same age and sex for all causes of death. A total of 182 patients died during the first year of follow-up, compared with an estimated 9.1 that would have been expected in a matched general population cohort (Figure 23). The excess mortality rate was 1.91 per 100 py (95% CI: 1.63; 2.24) during the first year of the follow-up, decreasing in the subsequent follow-up intervals to excess mortality rates under 1 excess deaths per 100 py of follow-up. We observed a negligible increase of excess mortality for those subjects with longer follow-up; 0.59 (95% CI: 0.34; 1.03) and 0.90 (95% CI: 0.53; 1.53), for those with 6 and 7 years of follow-up, respectively.

Figure 23: Overall excess Mortality Rate by follow-up time interval



Follow-up interval (year)	1	2	3	4	5	6	7	8	9-10
Persons-year	8667.9	7119.2	5972.1	4857.7	3829.8	2751.5	1794.4	1079.3	900.5
Observed deaths	182	48	33	26	19	23	20	9	8
Expected deaths	9.1	8.3	7.6	6.7	5.7	4.5	3.1	2.1	1.9
eMR per 100 py	1.91	0.54	0.42	0.34	0.33	0.59	0.90	0.55	0.46
95 % CI	(1.63; 2.24)	(0.38; 0.76)	(0.27; 0.60)	(0.19; 0.60)	(0.17; 0.62)	(0.34; 1.03)	(0.53; 1.53)	(0.21; 1.44)	(0.14; 1.47)

eMR: excess Mortality Rate
py: Persons-year
CI: Confidence Interval

Crude excess Mortality Rates (eMR) and excess Hazard Ratio (eHR) by potential prognostic factors are shown in Table 6. We found that crude overall excess mortality decreases as the level of education and the CD4 count at entry increases and that it was higher in IDUs (eMR: 2.59, 95% CI: 2.13; 3.13), patients of Spanish origin (eMR: 0.94, 95% CI:10.82;1.08), over 50 years of age (eMR: 2.53, 95% CI: 2.00; 3.20), HCV positive (eMR: 2.41, 95% CI: 2.03; 2.88), individuals with previous AIDS diagnosis (eMR: 2.98, 95% CI: 2.48; 3.58), and those with VL greater than 100000 copies/ml (eMR: 1.38, 95% CI: 1.17; 1.63). It was detected that short-term excess mortality was higher than long-term excess mortality; during the first year of follow-up the excess mortality rate was 1.91 deaths per 100 py of follow-up, approximately 4 (eHR: 0.26, 95% CI: 0.20; 0.33) times higher compared with the rest of follow-up.

Table 6: Overall crude excess Mortality Rates (eMR) and crude excess Hazard Ratios (eHR)

	eMR (95% CI)[a]	Crude eHR (95% CI)
TOTAL	**0.82 (0.73;0.93)**	
Education		
No/Compulsory	1.40 (1.20;1.64)	1
U. Secondary	0.44 (0.29;0.66)	0.31 (0.20;0.49)
University	0.33 (0.21;0.52)	0.24 (0.15;0.38)
Others	0.80 (0.52;1.25)	0.57 (0.35;0.00)
HIV tramission mode		
IDU	2.59 (2.13;3.13)	1
MSM	0.35 (0.27;0.46)	0.14 (0.10;0.19)
Heterosexual	0.96 (0.79;1.17)	0.37 (0.28;0.49)
Others	1.71 (0.81;3.62)	0.66 (0.30;1.44)
Origin		
Spain	0.94 (0.82;1.08)	1
SSA	0.69 (0.39;1.22)	0.73 (0.40;1.32)
LA	0.46 (0.31;0.69)	0.49 (0.32;0.75)
Others	0.58 (0.32;1.05)	0.61 (0.33;1.13)
CD4 at entry (cells/mm3)		
<200	2.04 (1.76;2.37)	1
201-350	0.67 (0.48;0.93)	0.33 (0.23;0.47)
>350	0.29 (0.21;0.39)	0.14 (0.10;0.20)
VL at entry (copies/ml)		
<20000	0.47 (0.34;0.64)	1
20000-100000	0.65 (0.50;0.83)	1.39 (0.92;2.09)
> 100000	1.38 (1.17;1.63)	2.96 (2.07;4.25)
Sex		
Male	0.83 (0.73;0.95)	1
Female	0.77 (0.57;1.04)	0.93 (0.67;1.28)
Age at entry		
20-49	0.70 (0.61;0.80)	1
>=50	2.53 (2.00;3.20)	3.25 (2.49;4.25)
HCV at entry		
Negative	0.53 (0.44;0.63)	1
Positive	2.41 (2.03;2.88)	4.59 (3.58;5.89)
AIDS at entry		
No	0.57 (0.48;0.66)	
Yes	2.98 (2.48;3.58)	5.26 (4.13;6.71)
Follow-up		
1 year	1.91 (1.63;2.24)	1
2-10 years	0.50 (0.41;0.60)	0.26 (0.20;0.33)

[a] *Excess mortality rate per 100 person-years of follow-up*

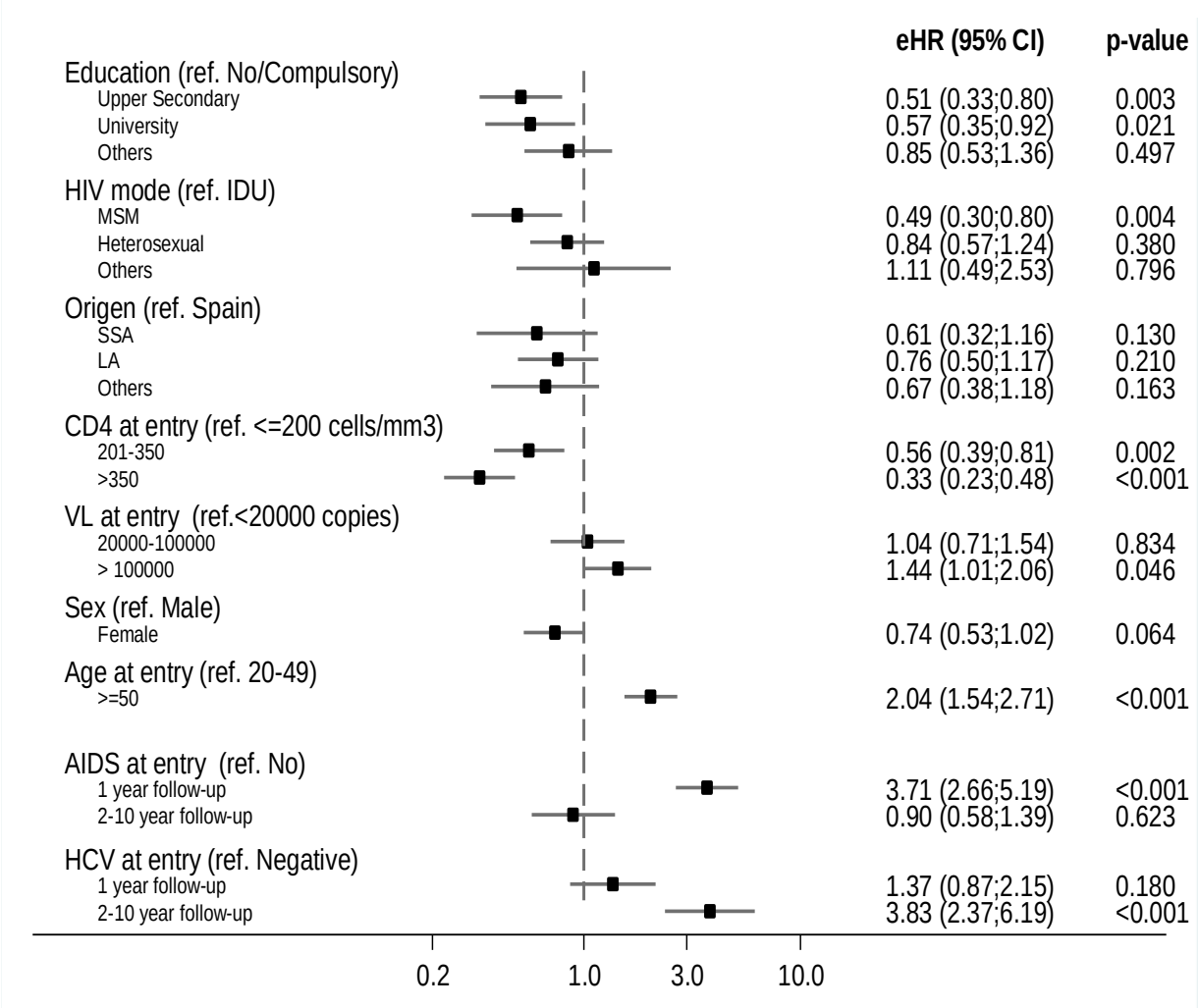
CI: Confidence Interval SSA=Sub-Saharan Africa, LA=Latin-America

The effect of potential predictive factors detected in the crude analysis was in general attenuated in the multivariable analysis (Figure 24). As expected, the level of CD4 count and the age at cohort enrolment were strong predictors for excess mortality adjusted for all other risk predictors. Excess mortality was higher in those with higher VL at entry (eHR: 1.44, 95% CI: 1.01; 2.06). There was a clear inverse association between excess mortality and level of education; excess mortality in those with upper secondary or university was 0.51 (95% CI: 0.33; 0.80) and 0.57 (95% CI: 0.35; 0.92) times lower respectively than in the group of no or compulsory education and excess mortality was lower among MSMs compared to IDUs. There was borderline evidence that being female was associated with lower risk of excess mortality (eHR: 0.74, 95% CI: 0.53; 1.02) and differences by country of origin were not statistically significant after controlling for the rest of predictors.

There was enough evidence that short term and long term effect of AIDS (p-value <0.001) and HCV (p-value <0.001) on excess mortality were different, hence results were presented separately for the first year after the inclusion in the cohort and from the second year onwards. There was strong evidence that having AIDS before cohort entry was a strong predictor of excess mortality during the first year of follow-up (eHR: 3.71, 95% CI: 2.66; 5.19) nevertheless this effect was attenuated (eHR: 0.90, 95% CI: 0.58; 1.39) thereafter. Opposite, having a positive Hepatitis C test result at entry predicted higher long term excess mortality (eHR: 3.83, 95% CI: 2.37;6.19) while there was no evidence of an effect during the first year of follow-up (eHR: 1.37, 95% CI: 0.87; 2.15).

Interactions between follow-up interval and all other covariates were checked but we failed to find any significant differences.

Figure 24: Overall adjusted excess Hazard Ratio (eHR) obtained from multivariable model



eHR adjusted for all other potential predictors
P-value from Wald test
CI: Confidence Interval ,SSA=Sub-Saharan Africa, LA=Latin-America

6.3.3 Non-AIDS malignancies-related excess mortality

The non-AIDS malignancies-related excess mortality rate observed in the cohort was 0.08 deaths (95 % CI: 0.05; 0.14) per 100 py of follow-up. Crude excess mortality rates and eHRs by potential prognostic factors are shown in Table 7. Excess mortality

rate was 0.09 per 100 py (95% CI: 0.03; 0.24) in those patients with no or compulsory education and 0.02 deaths per 100 py (95% CI: 0.00; 0.26 in those with university education. Higher crude non-AIDS malignancies excess mortality rates were observed in older patients (eHR: 7.76, 95% CI: 3.45; 17.46), with a previous diagnosis of AIDS (eHR: 4.56, 95% CI: 1.78; 11.70) and in subjects with a positive HCV result at entry (eHR: 4.68, 95% CI: 1.77; 12.38). Moreover, crude non-AIDS malignancies-related excess mortality rate in the first year of follow-up was 0.10 deaths per 100 py (95% CI: 0.04; 0.24) and it was 0.08 deaths per 100 py (CI 95% 0.04; 0.14) from 2^{nd} to 10^{th} year of follow-up.

Table 7: Crude non-AIDS malignancies excess Mortality Rates (eMR) and crude excess Hazard Ratios (eHR)

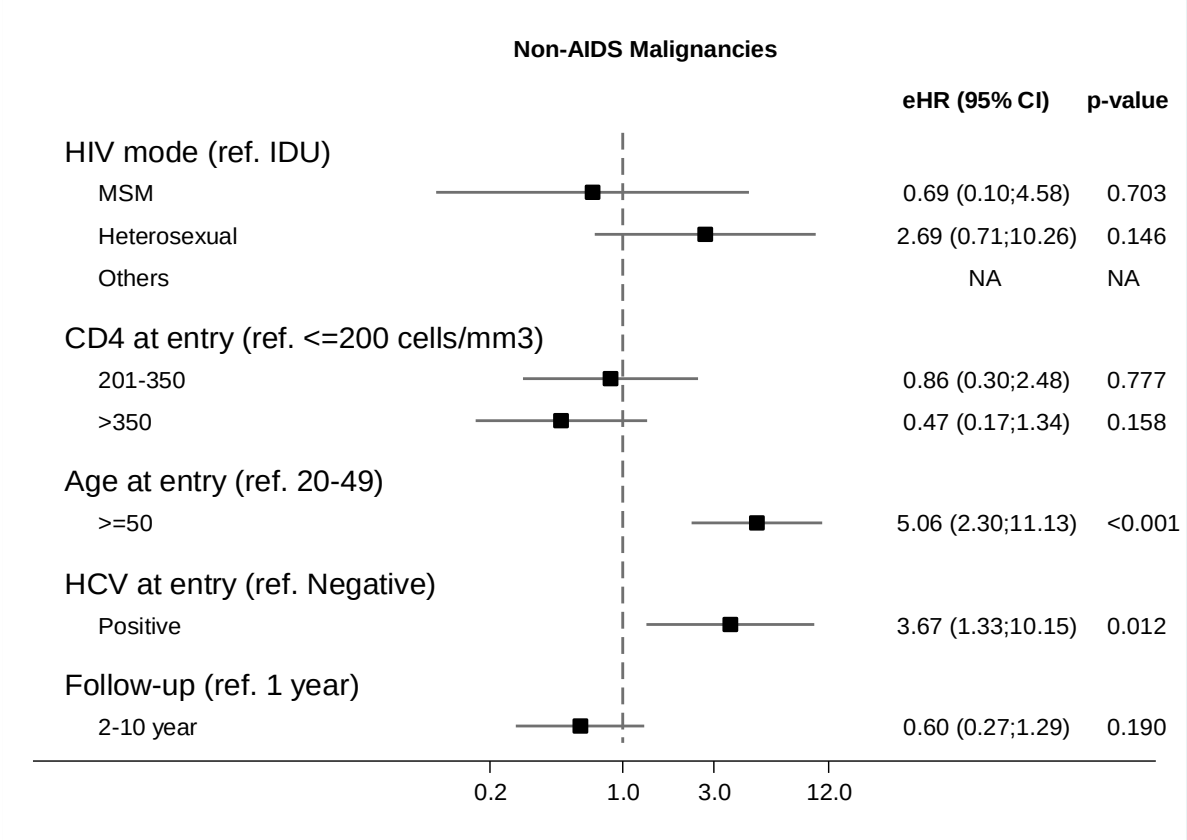
	eMR (95% CI)[a]	Crude eHR (95% CI)
TOTAL	0.08 (0.05;0.14)	
Education		
No/compulsory	0.09 (0.03;0.24)	1
Upper Secondary	0.09 (0.04;0.23)	1.03 (0.29;3.62)
University	0.02 (0.00;0.26)	0.25 (0.02;3.99)
Others	0.16 (0.06;0.42)	1.81 (0.47;0.00)
HIV transmission mode		
IDU	0.17 (0.07;0.45)	1
MSM	0.03 (0.01;0.10)	0.15 (0.03;0.86)
Heterosexual	0.17 (0.10;0.29)	0.97 (0.32;2.94)
Others	NA	NA
Origin		
Spain	0.10 (0.06;0.17)	1
Others	0.05 (0.02;0.17)	0.54 (0.15;1.94)
CD4 at entry (cel/mm3)		
<=200	0.18 (0.09;0.35)	1
201-350	0.09 (0.03;0.28)	0.48 (0.11;2.02)
>350	0.04 (0.01;0.12)	0.23 (0.07;0.78)
VL at entry (copies/ml)		
<20000	0.06 (0.02;0.19)	1
20000-100000	0.05 (0.02;0.16)	0.92 (0.17;4.89)
> 100000	0.14 (0.08;0.27)	2.53 (0.65;9.88)
Sex		
Male	0.07 (0.04;0.14)	1
Female	0.11 (0.05;0.26)	1.55 (0.57;4.25)
Age at entry		
20-49	0.06 (0.03;0.11)	1
>=50	0.49 (0.27;0.86)	7.76 (3.45;17.46)
HCV at entry		
Negative	0.05 (0.03;0.11)	1
Positive	0.25 (0.13;0.48)	4.68 (1.77;12.38)
AIDS before entry		
No	0.06 (0.03;0.12)	1
Yes	0.28 (0.14;0.55)	4.56 (1.78;11.70)
Follow-up		
1 year	0.10 (0.04;0.24)	1
2-10 years	0.08 (0.04;0.14)	0.73 (0.27;2.02)

[a] *Excess mortality rate per 100 person-years of follow-up*

CI: Confidence Interval; NA: Not a Number

Multivariable generalized linear model for non-AIDS malignancies excess mortality included HIV transmission mode, CD4 count, follow-up interval, HCV and age (Figure 25). There was strong evidence age and HCV test result at entry were independently associated with a higher non-AIDS malignancies excess mortality; patients of 50 years old or older at entry had 5.06 times higher risk (95% CI: 2.30; 11.13) compared with those younger than 50 years old at entry and the adjusted eHR for those subjects HCV coinfected at entry was 3.67 (95% CI: 1.33; 10.15).

Figure 25: Adjusted non-AIDS malignancies excess Hazard Ratio (eHR) obtained from multivariable model



eHR adjusted for all other potential predictors
P-value from Wald test
CI: Confidence Interval, NA: Not a Number, SSA=Sub-Saharan Africa, LA=Latin-America

6.3.4 Liver-related excess mortality

The liver-related excess mortality observed in the cohort from 2004 to 2014 in CoRIS was 0.11 deaths per 100 py of follow-up. Higher liver-related excess mortality rates were observed in patients with lower level of education (Table 8): 0.20 deaths per 100 py in patients with no or compulsory education vs. 0.03 deaths per 100 py in those with university education (eHR: 0.15, 95% CI: 0.03; 0.73). The liver-related excess mortality was higher among IDUs: 0.64 deaths per 100 py.

Regarding HIV disease markers, there was evidence that crude liver excess mortality rate was lower in patients with CD4 counts greater than 350 cells/mm^3 compared with those with levels under 200 cells/mm^3at entry (eHR: 0.16, 95% CI: 0.05; 0.57). Values greater than 100,000 copies in VL at entry had a non-significant effect compared to those with VL lower than 20,000 (eHR: 1.17, 95% CI: 0.43;3.16). Higher liver excess mortality rates were observed in patients with a positive HCV test result (eHR: 26.76, 95% CI: 10.26; 69.75). Females appeared to have similar risk compared to males (eHR: 1.44, 95% CI: 0.64; 3.24). A reduction in the risk of liver mortality excess was detected from 2^{nd} to 10^{th} year of follow-up compared to first year of follow-up (eHR: 0.44, 95% CI: 0.21; 0.89).

Interactions between follow-up interval and all covariates were tested in multivariable models for non-AIDS malignancies and liver excess mortality; however, there were not found significant differences. As the sample size was small, short-term and long-term effect of covariates on cause-specific excess mortality was explored

separately but not relevant differences were found, hence the results were presented for the complete period.

Table 8: Crude liver excess Mortality Rates (eMR) and crude excess Hazard Ratios (eHR)

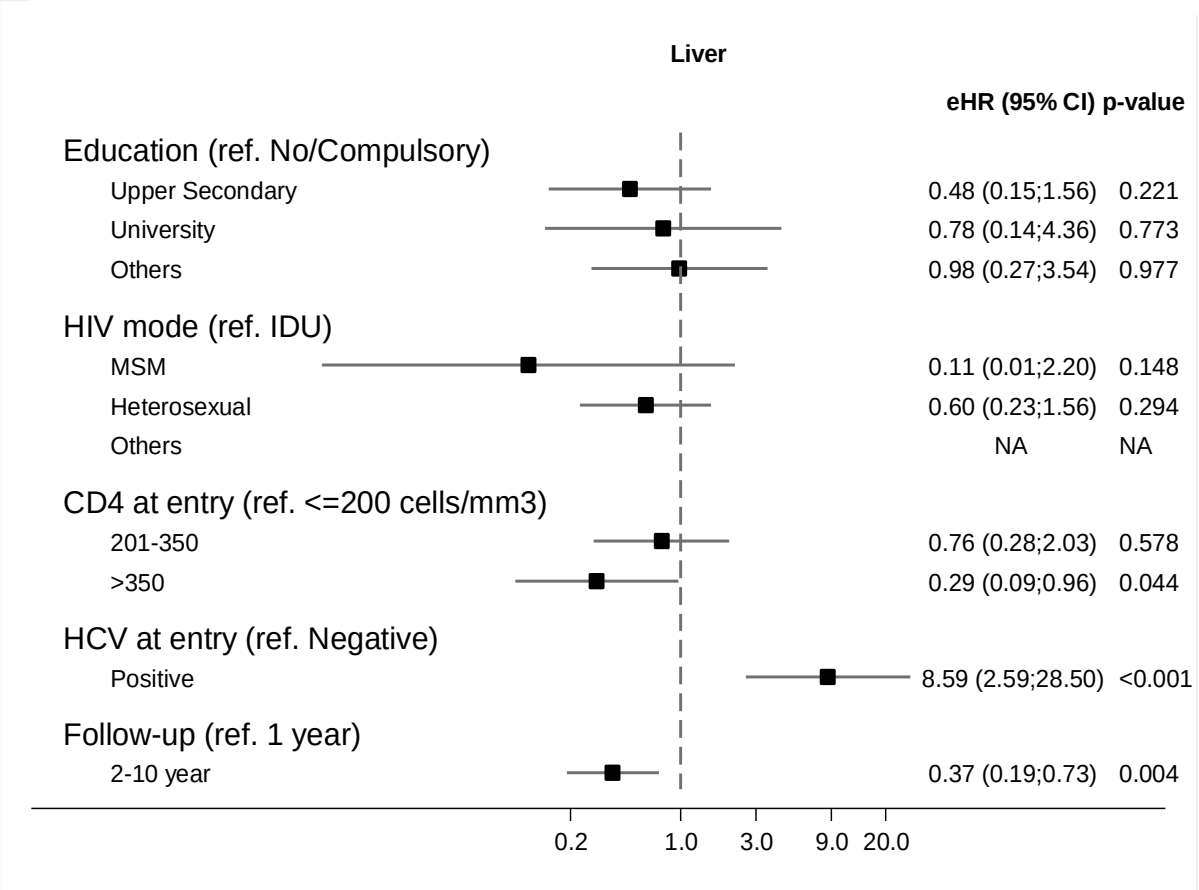
	eMR[a] (95% CI)	Crude eHR (95% CI)
TOTAL	0.11 (0.08;0.16)	
Education		
No/Primary	0.20 (0.13;0.30)	1
Upper Secondary	0.03 (0.01;0.14)	0.17 (0.04;0.79)
University	0.03 (0.01;0.13)	0.15 (0.03;0.73)
Others	0.08 (0.02;0.32)	0.43 (0.11;0.00)
HIV transmission		
IDU	0.64 (0.43;0.95)	1
MSM	0.01 (0.00;0.08)	0.01 (0.00;0.12)
Heterosexual	0.08 (0.04;0.16)	0.13 (0.06;0.27)
Others	NA	NA
Origin		
Spain	0.13 (0.09;0.19)	1
Others	0.03 (0.01;0.12)	0.25 (0.07;0.95)
CD4 at entry (cel/mm3)		
<=200	0.21 (0.13;0.34)	1
201-350	0.13 (0.05;0.30)	0.60 (0.22;1.66)
>350	0.03 (0.01;0.11)	0.16 (0.05;0.57)
VL at entry (copies/ml)		
<20000	0.09 (0.05;0.20)	1
20000-100000	0.10 (0.05;0.19)	1.06 (0.37;3.00)
> 100000	0.11 (0.06;0.21)	1.17 (0.43;3.16)
Sex		
Male	0.10 (0.06;0.14)	1
Female	0.14 (0.07;0.28)	1.44 (0.64;3.24)
Age at entry		
20-49	0.10 (0.06;0.14)	1
>=50	0.17 (0.08;0.38)	1.77 (0.72;4.37)
HCV at entry		
Negative	0.02 (0.01;0.05)	1
Positive	0.56 (0.39;0.81)	26.76 (10.26;69.75)
AIDS before entry		
No	0.10 (0.06;0.14)	1
Yes	0.16 (0.07;0.37)	1.70 (0.69;4.19)
Follow-up		
1 year	0.19 (0.11;0.31)	1
2-10 years	0.08 (0.05;0.13)	0.44 (0.21;0.89)

[a] *Excess mortality rate per 100 person-years of follow-up*

CI: Confidence Interval; NA: Not a Number

Multivariable generalized linear model for liver-related excess mortality included level of education, HIV transmission mode, country of origin, CD4 counts, HCV at entry and follow-up interval (Figure 26). Through this model there was evidence that CD4 count, HCV at entry and follow-up interval were associated with higher risk of liver excess mortality. There was evidence that a CD4 count at entry greater than 350 cells/mm^3 was associated with approximately 71% reduction on liver excess mortality, compared with CD4 counts lower than 200 cells/mm^3 (eHR: 0.29, 95% CI: 0.09; 0.96). As expected, those with a positive HCV result at entry had 8.59 times (95% CI: 2.59; 28.50) higher risk of liver excess mortality. Besides, there was strong evidence that long-term adjusted eHR was lower compared with short-term (eHR: 0.37, 95% CI: 0.19; 0.73).

Figure 26: Adjusted liver-related excess Hazard Ratio (eHR) obtained from multivariable model



eHR adjusted for all other potential predictors
P-value from Wald test
CI: Confidence Interval, NA: Not a Number, SSA=Sub-Saharan Africa, LA=Latin-America

6.3.5 Cardiovascular-related excess mortality

The cardiovascular-related excess mortality rate observed in the cohort was 0.02 deaths (95% CI: 0.01; 0.06) per 100 py of follow-up. Crude excess mortality rates and eHRs by potential prognostic factors are shown in Table 9. Excess mortality rate was 0.03 per 100 py (95% CI: 0.03; 0.24) in those patients with no or compulsory education but no cardiovascular-related excess mortality was found among those with university education. We also observed that the cardiovascular mortality rate among females and among those patients with CD4 > 350 cel/mm^3 did not differ from the rate

in the Spanish population of the same age and sex. Regarding HCV coinfection, the estimated cardiovascular excess mortality was 0.01 deaths py of follow-up (95% CI: 0.00; 0.05) among those with an HCV negative test at entry and 0.07 (95% CI: 0.01; 0.31) deaths per 100 py among those patients with a HCV positive test at entry (eHR: 4.96, 95% CI: 0.58; 42.46).

Nevertheless, no significant predictor for the crude cardiovascular-related excess mortality was found hence multivariable generalized linear model was not performed.

Table 9: Crude cardiovascular excess Mortality Rates (eMR) and crude excess Hazard Ratios (eHR)

	eMR[a] (95% CI)	Crude eHR (95% CI)
TOTAL	0.02 (0.01;0.06)	
Education		
No/Primary/ Upper Secondary	0.03 (0.01;0.09)	1
University	0.00 (0.00;0.00)	0.00 (0.00;NA)
Others	0.05 (0.01;0.39)	2.06 (0.19;22.22)
HIV transmission		
MSM	0.02 (0.00;2.02)	1
IDU	0.01 (0.00;0.06)	0.34 (0.00;186.85)
Heterosexual	0.03 (0.00;0.17)	1.69 (0.01;436.97)
Others	0.29 (0.04;2.02)	13.68 (0.04;5264.95)
Origin		
Spain	0.01 (0.00;0.10)	1
Others	0.03 (0.01;0.14)	2.07 (0.15;28.56)
CD4 at entry (cel/mm3)		
<=200	0.04 (0.01;0.14)	1
201-350	0.04 (0.01;0.18)	1.05 (0.18;6.16)
>350	0.00 (0.00;NA)	NA
VL at entry (copies/ml)		
<20000	0.02 (0.00;0.15)	1
20000-100000	0.02 (0.00;0.14)	0.99 (0.06;16.19)
> 100000	0.01 (0.00;0.15)	0.66 (0.03;14.99)
Sex		
Male	0.02 (0.01;0.06)	1
Female	0.00 (0.00;NA)	NA
Age at entry		
20-49	0.02 (0.01;0.06)	1
>=50	0.03 (0.00;0.54)	1.66 (0.08;34.58)
HCV at entry		
Negative	0.01 (0.00;0.05)	1
Positive	0.07 (0.01;0.31)	4.96 (0.58;42.46)
AIDS before entry		
No	0.02 (0.01;0.07)	1
Yes	0.01 (0.00;16.94)	0.45 (0.00;1070.98)
Follow-up		
1 year	0.03 (0.01;0.16)	1
2-10 years	0.02 (0.00;0.06)	0.54 (0.06;5.19)

[a] *Excess mortality rate per 100 person-years of follow*
[b] *Adjusted for all other potential predictors*
[c] *P-value from Wald test*
CI: Confidence Interval; NA: Not a Number

6.3.6 Changes in Overall and Cause-specific excess mortality by calendar period of inclusion

Changes on the excess mortality over the period of inclusion are presented in Table 10. The estimated overall excess mortality was 0.86 deaths per 100 py of follow-up in the period of inclusion 2004-2007 and it was 0.73 per 100 py in the period of inclusion 2008-2014. After controlling for all other factors, no significant differences were found in overall excess mortality by period of inclusion (eHR: 0.91, 95% CI: 0.70; 1.19).

Non-AIDS malignancies and liver excess mortality appeared to be lower in the second period of inclusion compared with the first one. However there was no statistical evidence for this result.

The cardiovascular excess mortality appeared to be concentrated in the first period of inclusion. Adjusted analyses were not performed due to the small sample size.

Table 10: Changes in overall and cause-specific excess Hazard Ratios (eHR) over period of inclusion

Period	Excess Mortality (95% CI)[a]	Crude eHR (95% CI)	Adjusted[b] eHR (95% CI)
Overall mortality			
2004-2007	0.86 (0.75;1.00)		
2008-2014	0.73 (0.58;0.92)	0.85 (0.65;1.11)	0.91 (0.70; 1.19)
Non-AIDS malignancies mortality			
2004-2007	0.10 (0.06;0.18)		
2008-2014	0.05 (0.01;0.16)	0.46 (0.11;1.89)	0.75 (0.30; 1.88)
Liver mortality			
2004-2007	0.13 (0.09;0.19)		
2008-2014	0.05 (0.02;0.13)	0.41 (0.15;1.15)	0.73 (0.26; 2.07)
Cardiovascular mortality			
2004-2007	0.03 (0.01;0.07)		
2008-2014	0.00 (0.00;0.00)	0.01 (0.00;NA)	

[a] *Excess mortality rate per 100 person-years of follow-up*
[b] *for follow-up interval, sex, age at entry, HIV transmission category, CD4 counts and HCV test result at entry*
CI: Confidence Interval; NA: Not a Number

6.3.7 Sensitivity Analyses

Sensitivity analyses in which follow-up time was divided in one-year intervals provided similar estimations.

CC results from final multivariable models for overall and non-AIDS malignancies excess mortality are later described in objective 4 and presented in Table 15 (page 155) and Table 16 (page 159). CC results from liver and cardiovascular related models are not shown in this document but no significant differences between eHRs were detected.

CC results from final multivariable models for changes over period of inclusion in overall, non-AIDS malignancies, liver and cardiovascular excess mortality are

presented in Table 11. There was no evidence of significant differences between results from CC analysis and results from models performed after Multiple Imputation.

Table 11: Complete-Case analysis for changes in overall and cause-specific excess Hazard Ratios (eHR) over period of inclusion

Period	Excess Mortality (95% CI)[a]	Crude eHR (95% CI)	Adjusted[b] eHR (95% CI)
Overall mortality			
2004-2007	0.69 (0.57;0.83)		
2008-2014	0.54 (0.39;0.74)	0.78 (0.54;1.14)	0.98 (0.68;1.42)
Non-AIDS malignancies mortality			
2004-2007	0.04 (0.01;0.12)		
2008-2014	0.04 (0.01;0.19)	0.91 (0.13;6.22)	1.05 (0.31;3.62)
Liver mortality			
2004-2007	0.09 (0.06;0.15)		
2008-2014	0.03 (0.01;0.11)	0.33 (0.08;1.34)	0.71 (0.19;2.62)
Cardiovascular mortality			
2004-2007	0.00 (0.00;0.12)		
2008-2014	0.00 (0.00;54.75)	0.57 (0.00;NA)	

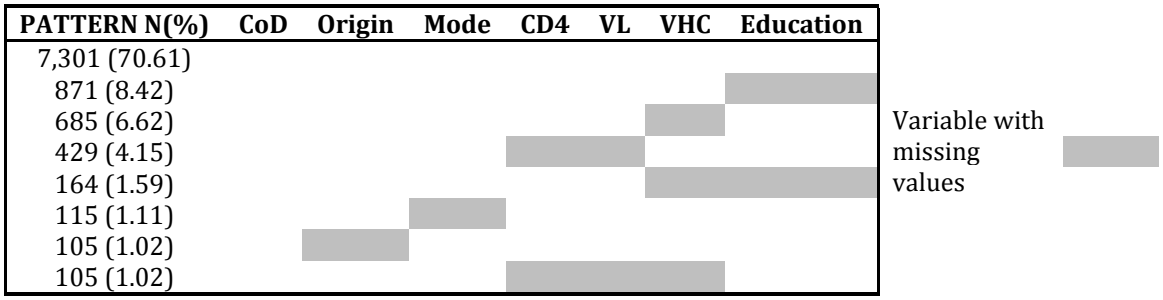
[a] *Excess mortality rate per 100 person-years of follow-up*
CI: Confidence Interval
[b] *for follow-up interval, sex, age at entry, HIV transmission category, CD4 counts and HCV test result at entry*

6.4 Objective 4

6.4.1 Missing data in CoRIS

A total of 10,340 HIV-positive patients were included. Figure 27 shows the most frequent patterns of missing data: 7,301 (70.61%) subjects had complete information in all the variables (outcome and covariates), 871 (8.42%) had incomplete data only in level of education, 685 (6.62%) only in HCV coinfection variable and 429 (4.15%) had incomplete information on both CD4 count and VL measurement at enrolment. There were 44 additional missingness patterns other than those shown, representing each less that 1% of the cases. We finally checked that the data had a non-monotone pattern.

Figure 27: Eight most common missing data patterns of missingness



Code of Death, Country of Origin, Mode HIV transmission, CD4 at entry, VL at entry, HCV at entry, Level of education

Table 12 shows the distribution of the patients before imputation and the number of missing values per variable. Data were missing on level of education (1,349, 13.05 %), HIV transmission category (242, 2.34%), Country of Origin (218, 2.11 %), CD4 count (775, 7.50%), VL (811, 7.84%), HCV status at entry (1,081, 10.45%) and CoD (47, 12.77% among death subjects).

The right side of Table 12 shows the percentage of missingness for each variable with missing data among covariate category (for each incomplete variable, the percentage of missing values per covariate). The probability of missing data on education was positively related with being IDU, lower CD4 count, HCV positive, AIDS and initiating cART at entry. There was evidence that missing data on HIV transmission mode was also associated with country of origin, CD4 count, age, CoD, AIDS and initiating cART. Missing data on country of origin was higher among those with upper secondary education, IDU HIV transmission mode, higher CD4 count, lower Viral Load, without cART and without AIDS. For example, the percentage of patients with missing data on country of origin was 2.3% on those free of AIDS and it was 1.0% on those with AIDS. We detected that VL and CD4 count shared the same predictors of missingness: lower education, being IDU, from Sub-Saharan Africa, HCV positive, having AIDS and cART at entry, older than 50 years and female. We also observed that in case of incomplete data, the probability of being missing was higher for those who died during follow-up.

Table 12: Distribution of patients and percentage of missing values among covariate category

	N	%	Percentage with missing data (%)						
Total	**10340**	**100**	**Education**	**Mode**	**Origin**	**CD4**	**VL**	**HCV**	**CoD**
Education									
No compulsory	3,442	33.29	-	2.1	1.3	10.2	10.7	10.5	0.8
Upper Secondary	2,339	22.62	-	1.5	2.7	4.8	5.2	7.0	0.3
University	2,477	23.96	-	1.5	1.2	4.1	4.3	9.6	0.2
Other	733	7.09	-	2.0	0.1	8.9	9.0	12.6	0.1
Missing	1,349	13.05	-	6.2	6.1	10.9	10.9	16.9	0.7
Mode									
IDU	998	9.65	17.0	-	3.0	13.3	13.8	12.7	1.8
MSM	5,963	57.67	10.4	-	2.3	4.5	4.8	9.4	0.2
Heterosexual	3,035	29.35	15.1	-	1.6	10.9	11.5	11.2	0.5
Others	102	0.99	15.7	-	0.0	8.8	9.8	7.8	2.0
Missing	242	2.34	34.7	-	2.1	12.4	11.6	18.2	0.4
Country of Origin									
Spain	7,008	67.78	11.9	2.6	-	7.4	7.6	11.3	0.5
Sub-Saharan Afri	558	5.40	17.7	3.9	-	14.3	15.1	10.6	0.5
Latin American	1,812	17.52	12.6	1.0	-	5.8	6.2	8.3	0.3
Others	744	7.20	14.4	2.0	-	6.3	7.1	9.0	0.3
Missing	218	2.11	37.6	2.3	-	11.9	13.8	6.9	0.5
CD4 at entry (cel/mm3)									
<=200	2,219	21.46	14.9	3.9	0.8	-	2.3	9.5	0.7
201-350	1,876	18.14	12.6	2.6	1.6	-	0.9	8.6	0.6
>350	5,470	52.9	11.6	1.4	2.6	-	0.7	9.8	0.3
Missing	775	7.50	19.0	3.9	3.4	-	90.8	21.9	0.6
VL at entry									
<20000	3,228	31.22	12.9	1.7	2.7	0.4	-	10.7	0.4
20000-100000	3,227	31.21	12.1	2.3	1.7	0.7	-	9.0	0.5
100000	3,074	29.73	12.8	2.8	1.5	1.1	-	8.8	0.5
Missing	811	7.84	18.1	3.5	3.7	86.8	-	21.2	0.6
HCV at entry (copies/ml)									
Negative	8,039	77.75	11.8	2.2	2.1	5.9	6.3	-	0.3
Positive	1,220	11.80	14.3	1.6	2.8	10.8	11.1	-	1.6
Missing	1,081	10.45	21.1	4.1	1.4	15.7	15.9	-	0.3
CoD									
Alive	9,972	96.44	12.7	2.2	2.1	7.3	7.6	10.3	-
HIV/AIDS	154	1.49	33.1	12.3	0.6	13.0	14.3	18.2	-
Liver	52	0.50	13.5	1.9	1.9	17.3	15.4	13.5	-
Non-AIDS Maligna	34	0.33	17.6	2.9	0.0	17.6	20.6	5.9	-
Cardiovascular	9	0.09	22.2	0.0	0.0	0.0	0.0	44.4	-
Others	72	0.70	11.1	2.8	1.4	16.7	13.9	12.5	-
Missing	47	0.45	21.3	2.1	2.1	10.6	10.6	6.4	-
AIDS at entry									
No	8,973	86.78	12.7	2.0	2.3	6.5	6.8	10.4	0.4
Yes	1,367	13.22	15.4	4.5	1.0	13.9	14.8	11.0	1.0
cART at entry									
No	5,364	51.88	12.5	1.6	2.7	2.3	2.6	11.9	0.4
Yes	4,976	48.12	13.6	3.2	1.4	13.1	13.4	8.9	0.5
Age at entry									
20-49	9,305	89.99	13.0	2.0	2.1	7.2	7.6	10.3	0.4
>=50	1,035	10.01	13.8	5.8	2.6	10.1	10.3	11.9	0.9
Sex									
Male	8,633	83.49	12.8	2.2	2.2	6.7	7.0	10.2	0.5
Female	1,707	16.51	14.1	2.9	1.4	11.5	12.1	11.5	0.6
Survival time									
< 6 months	1,748	16.91	14.8	3.4	1.5	9.1	9.8	10.4	0.2
6-12 months	828	8.01	14.9	3.6	3.4	6.5	6.2	8.7	0.7
2 years	1,207	11.67	13.7	2.9	3.1	7.4	8.0	9.5	0.5
3 years	1,182	11.43	12.4	1.4	3.1	7.9	8.1	10.6	1.0
4 years	1,051	10.16	14.0	2.7	2.4	8.2	8.9	11.7	0.6
>4 years	4,324	41.82	11.8	1.7	1.5	6.8	7.0	10.7	0.3

Table 13 summarizes, for each incomplete variable, the predictors of a variable being missing (missingness) and the predictors of values of variables with missing data. To investigate which variables were related with missingness, we estimated mutually adjusted associations between the covariates and being missing on level of education, mode of HIV transmission, country of origin, CD4 count, VL measurement, HCV coinfection and CoD.

Analogously, to investigate which covariates where predictors of the values of the variables with missing data, we estimated mutually adjusted associations between the covariates and incomplete variables. We found several predictors for each incomplete variable.

Table 13: Predictors of variables with missing data and of missingness

	Variables with missing values						
Predictors	**Education**	**Mode**	**Origin**	**CD4**	**VL**	**HCV**	**CoD**
Education		**V**	**M V**	**V**	**V**	**M V**	
Mode	**M V**		**M V**	**V**	**V**	**V**	**V**
Origin country	**V**	**MV**		**M V**	**V**	**M V**	
CD4 at entry	**V**	**V**	**M V**		**M V**		
VL at entry		**V**	**V**	**M V**			
HCV at entry	**V**	**V**	**V**	**V**			**V**
CoD	**M**			**M**		**M V**	
AIDS at entry	**V**	**V**	**V**	**V**	**V**		**V**
cART at entry		**V**	**V**	**V**	**V**	**M**	
Age at entry	**M V**	**V**	**M V**	**V**	**V**	**V**	**M**
Sex	**V**		**V**	**V**	**V**	**V**	
Survival time	**M V**	**V**	**V**		**V**		**M V**

***M**: Covariate associated with missingness (variable being missing). P-value < 0.05 from logistic regression between missingness and the covariates, based on analyses with complete cases.*
***V:** Covariate associated with the values of the variable with missing data. P-value < 0.05 from multinomial regression between the variable with missing values and the covariate, based on analyses with complete cases.*

In summary, we found a reasonable number of variables that predicted the probability of being missing in each incomplete variable, therefore it seems implausible than data were MCAR. As well, we detected a high number of predictors for the values of variables with missing data that contributed to generate accurate imputations. Finally, we did not found evidence against assuming data were MAR, neither expert knowledge nor descriptive analyses of missing values suggested that data were MNAR.

6.4.2 Multiple imputation by chained equations in CoRIS

We developed a multiple imputation model for each variable with missing values. Resulting models included the other incomplete variables (education, HIV transmission, country of origin, CD4, VL, HCV and CoD), the complete variables (AIDS at entry, age and sex), the outcome (survival time and CoD) and one auxiliary variable "Receiving cART within 6 months of entry". Linear regression was used for VL and CD4 counts; logistic regression for HCV coinfection, ordinal regression for education level; and multinomial regression for HIV transmission category, country of origin and CoD.

CD4 count and VL were considered as continuous variables in imputation models, although for the purpose of the analysis of interest, these variables were categorized.

Log transformation was used to improve the normality of VL measurements and 4th root was applied to CD4 counts. The inverse of the squared root was used to normalize the baseline age (Figure 28 and Figure 29).

Figure 28: Histogram and quantile-quantile normal plot of CD4 count variable before and after 4th root transformation

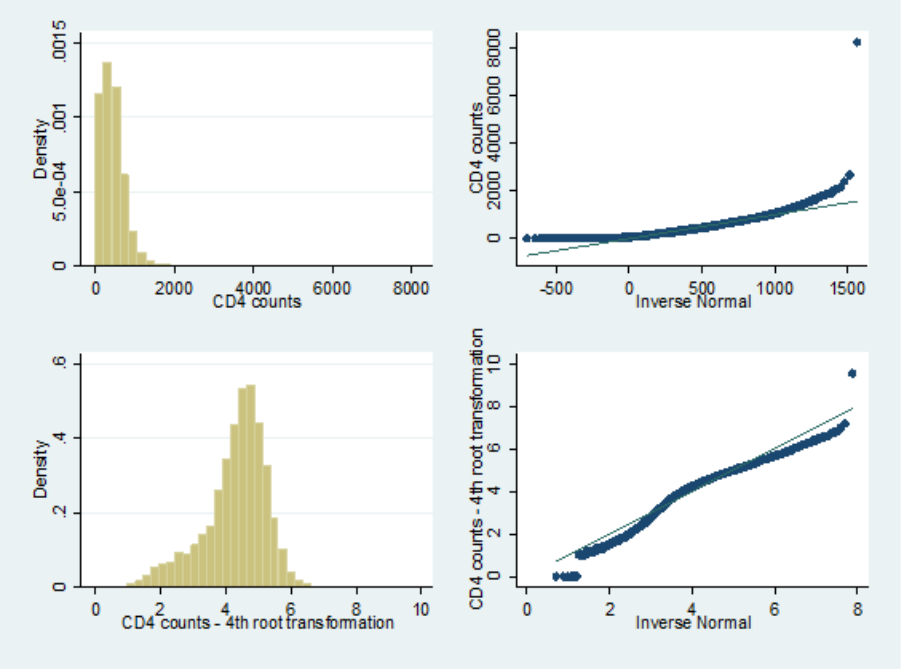
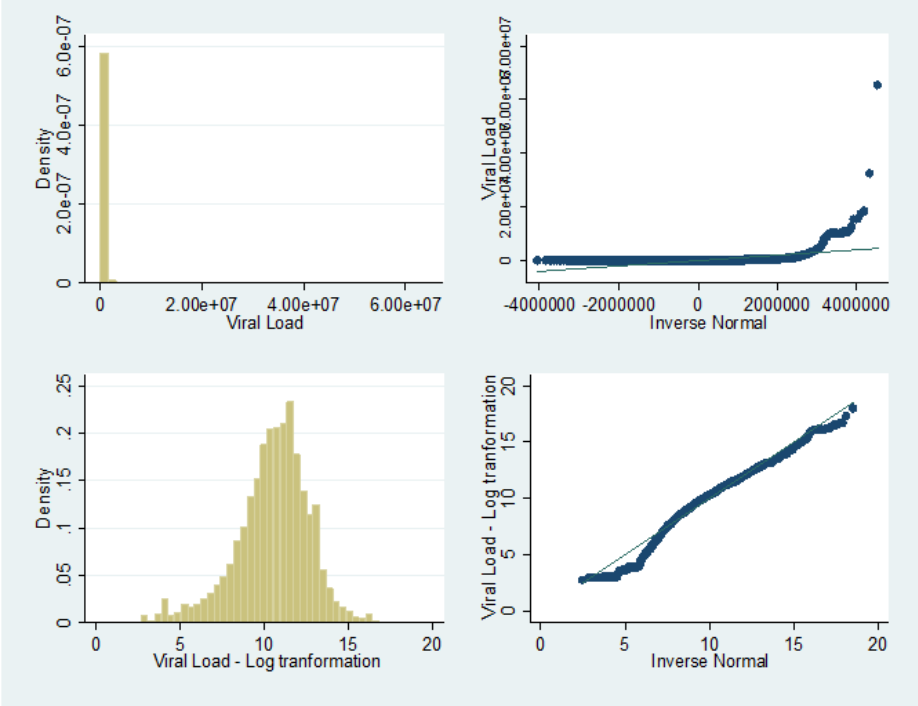
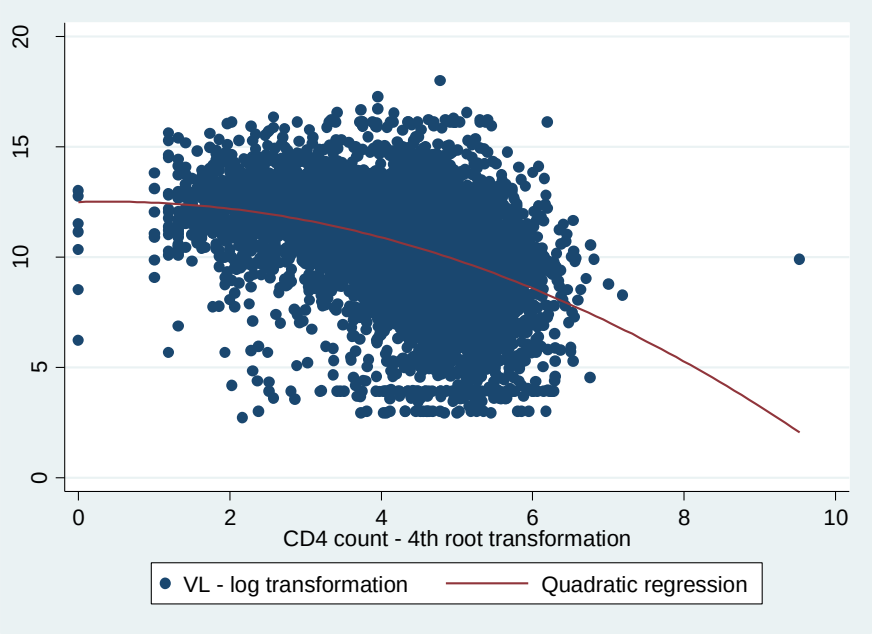


Figure 29: Histogram and quantile-quantile normal plot of Viral Load variable before and after log transformation



Passive imputation was applied to add non-linearities in the chained equations model. Firstly, we examined quadratic associations based on analyses of patients with complete information. Subsequently, significant squared terms of transformed CD4 counts, VL and age were added to each imputation model (Figure 30).

Figure 30: Scatter plot and quadratic regression curve of Log Viral Load on 4th root CD4 count



The auxiliary variable receiving cART was included in the imputation models to correct potential bias and to make missingness assumptions more plausible although this variable was not considered for analytical purposes.

Regarding the inclusion of interactions, separation approach was not convenient because HCV variable had missing values and passive improved imputation and JAV implied the estimation of additional parameters and the imputation model failed to converge. Our imputation model was really complex and a lot of parameters had to be estimated therefore interaction terms of AIDS at entry, HCV and period of inclusion by follow-up time were passively imputed as they were included in the analyses of interest.

The outcomes, survival time and CoD, were also included in the imputation models. It was not plausible to assume a linear effect of survival time on each of the

variables imputed hence the logarithm of follow-up time was considered. CoD was imputed conditional on death having occurred.

Finally, variables considered for the purpose of the analysis were passively or deterministically imputed. CD4 counts and VL were back-transformed and then categorized as pre-specified.

6.4.3 Multiple imputation model assessment

Once the imputation model was specified and the imputations were created, the model was assessed. The proportions presented in Table 14 show the distribution among the observed, imputed and combined values separately.

We observed that the distribution of both the imputed and combined values for the variable level of education seemed plausible. Regarding HIV transmission mode, the percentage of MSM in those subjects with imputed values (49.48%) was slightly lower than in the groups of observed (59.05%) or combined values (58.83%). However the distribution in all the groups seemed reasonable. Just 218 values were found to be incomplete for country of origin and approximately 73.13% were imputed in the Spain group. Regarding VL at entry, distributions of observed, imputed and combined values were reasonable. For CD4 count, we detected that approximately 50% of CD4 count missing values were imputed in the group of <=200 cells/mm^3; however the combined and the observed distribution show high agreement pointing no evident problem with imputation model. The distribution between observed,

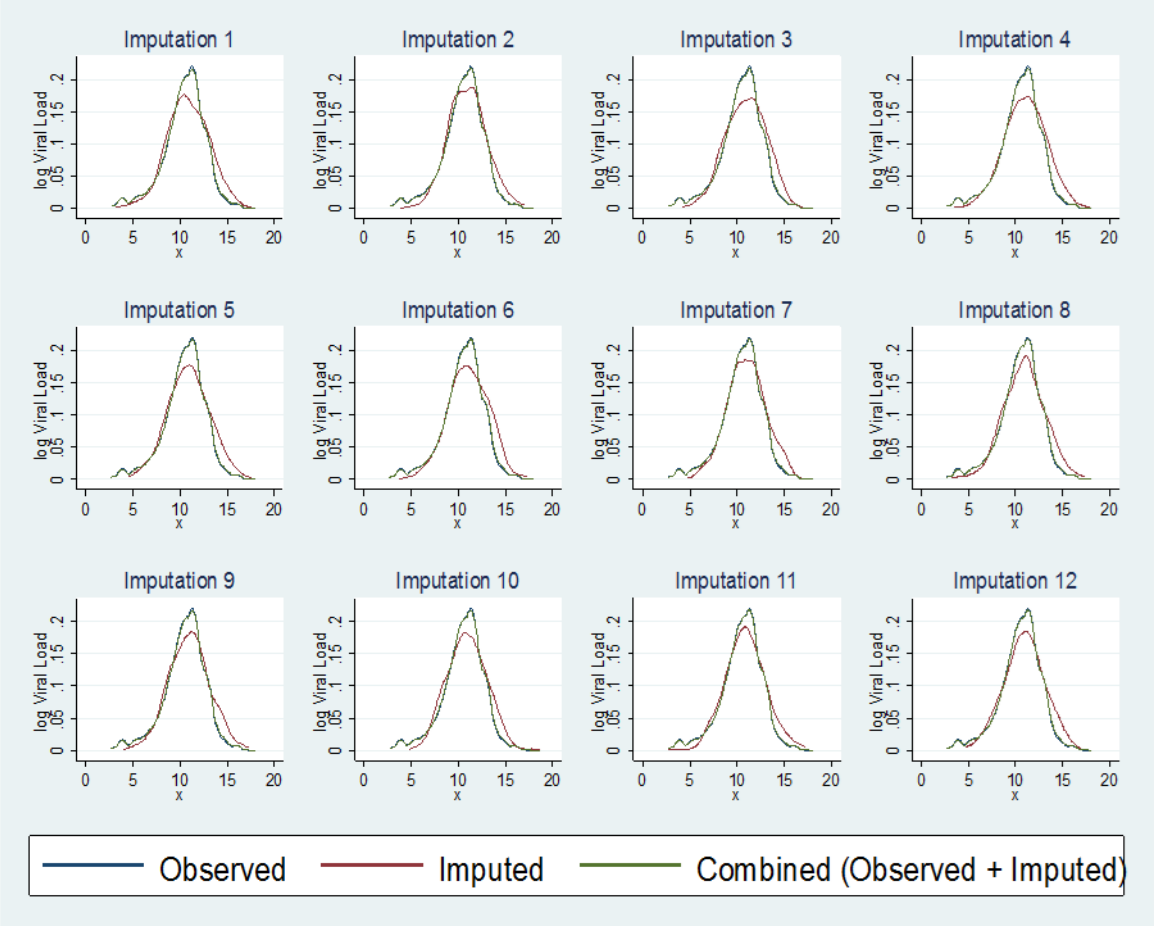
imputed, and complete values of HCV at entry was very similar. For CoD, the imputation was restricted to those individuals that were found to be death during the period of study and 47 values were missing. We detected slight differences between observed and imputed values; however the combined and the observed distribution were similar.

In general, we did not observe gross discrepancies between the observed and imputed values neither combined distributions that seem unreasonable.

6.4.3.1 Table 14: Distribution of variables with missing data for observed, imputed and combined values

VARIABLE WITH MISSING DATA	OBSERVED	IMPUTED	COMBINED (OBSERVED +IMPUTED)
	%	%	% (CI 95%)
Education			
No or compulsory	38.28	43.74	38.99 (37.98;40.01)
Upper Secondary	26.01	25.65	25.97 (25.04;26.89)
University	27.55	23.94	27.08 (26.19;27.97)
Other	8.15	6.68	7.96 (7.40;8.52)
HIV transmission			
IDU	9.88	9.26	9.87 (9.29;10.45)
MSM	59.05	49.48	58.83 (57.87;59.78)
Heterosexual	30.06	39.74	30.28 (29.39;31.17)
Others	1.01	1.52	1.02 (0.83;1.22)
Origin country			
Spain	69.24	73.13	69.32 (68.42;70.22)
Sub-saharan African	5.51	4.47	5.49 (5.05;5.93)
Latin American	17.90	15.33	17.85 (17.10;18.59)
Others	7.35	7.07	7.34 (6.84;7.85)
CD4 at entry (cel/mm3)			
<=200	22.96	48.18	24.85 (23.99;25.71)
201-350	19.60	23.81	19.92 (19.12;20.72)
>350	57.44	28.01	55.23 (54.26;56.20)
VL at entry (copies/ml)			
<20000	33.85	31.39	33.67 (32.74;34.60)
20000-100000	33.88	28.72	33.47 (32.53;34.40)
100000	32.27	39.89	32.86 (31.92;33.81)
HCV at entry			
Negative	86.82	84.02	86.53 (85.86;87.20)
Positive	13.18	15.98	13.47 (12.80;14.14)
Cuase of Death			
HIV/AIDS	47.98	25.71	45.13 (39.69;50.57)
Liver	16.20	21.10	16.83 (12.55;21.10)
Non-AIDS Malignancies	10.59	13.48	10.96 (7.39;14.53)
Cardiovascular	2.80	7.98	3.46 (1.25;5.68)
Others	22.43	31.74	23.62 (18.98;28.26)

Figure 31: Distributions of the observed, imputed, and combined values for log of the Viral Load variable (kernel density estimator plot)

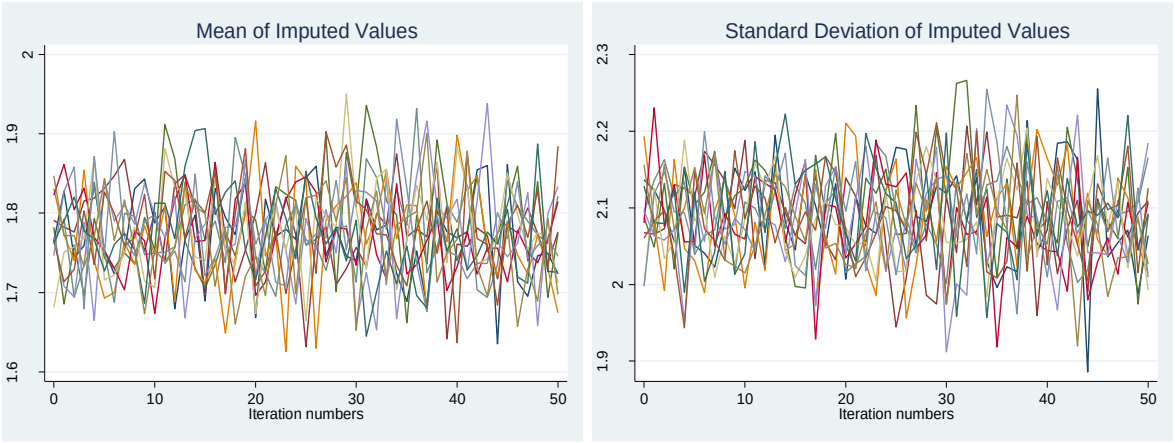


For the continuous variables Log VL and 4th root of CD4 count kernel density estimator was plotted among the observed, imputed and combined values in each imputation, separately (Figure 31). The density plots for both variables seemed plausible.

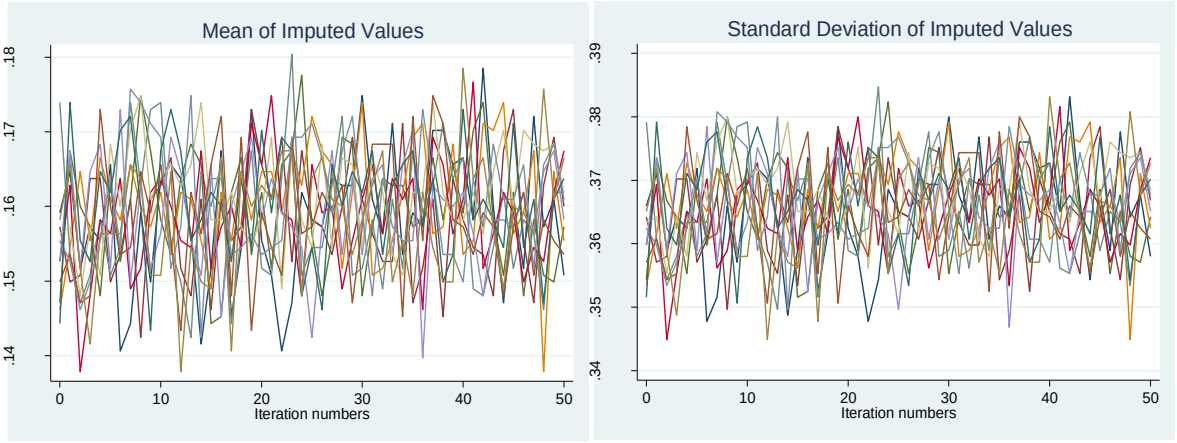
To assess the convergence of the multiple imputation model, we calculated the mean and standard deviation in each iteration of each imputed variable. Figure 32 shows the results from 12 chains (imputations) along the burn-in period of 50 iterations in the variables level of education, HCV at entry and Log VL. We observed that the mean and standard deviation varied randomly without any trend indicating that models converged successfully.

Figure 32: Trace plots of summaries of imputed values from 12 chains of 7 variables (level of education, HCV at entry, Log Viral Load)

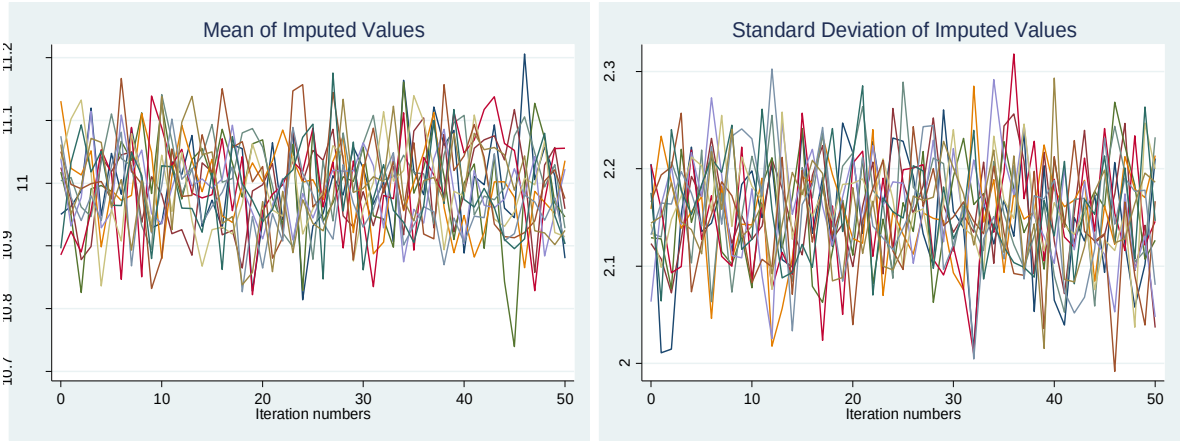
Level of education



VHC at entry



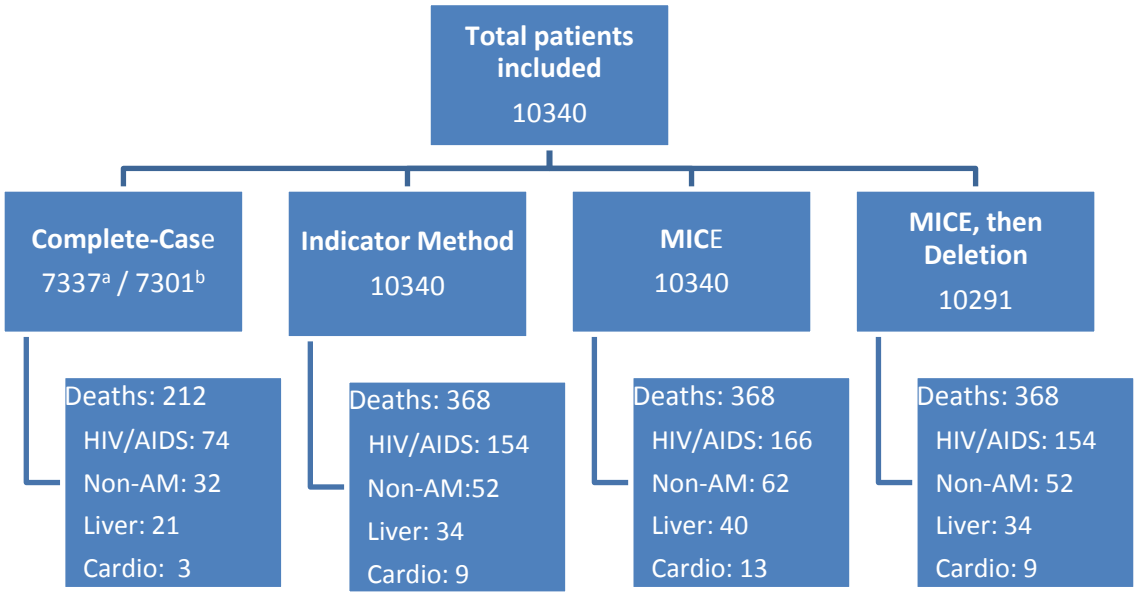
Log Viral Load



6.4.4 Excess mortality using different methods to deal with missing data

The number of patients and events included in the analyses varied depending on the method used to deal with missing data (Figure 33). When we used CC approach to assess overall excess mortality, the number of patients included was 7337 with 212 events, while to assess non-AIDS malignancies excess mortality 7301 patients could be considered and only 32 events. When we adopted the IM approach, all the patients and events could be included in the overall excess mortality analyses, however only 52 events were considered to assess the non-AIDS malignancies excess. MICE approach included all the patients and all the potential events as the unknown causes of death were also imputed. Finally, when we adopted the approach MID, 10291 patients and 52 deaths were included in the non-AIDS malignancies excess mortality analyses as the subjects with unknown CoD were deleted after imputation.

Figure 33: Number of patients included in the analysis based on Complete-Case approach, Indicator Method, MICE and MICE, then Deletion



[a] Overall mortality analysis. [b] Cause-specific mortality analysis
Non-AM: Non- AIDS malignancies. Cardio: Cardiovascular

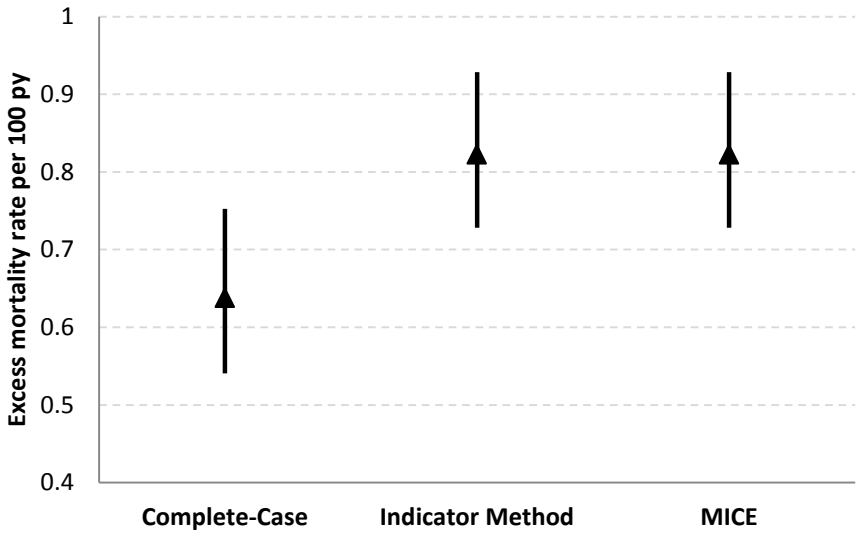
6.4.5 Overall excess mortality

Figure 34 shows the estimates of the overall excess mortality rate based on the CC approach, on the IM and on MICE approach.

We observed that the estimated overall mortality excess rate is higher when MICE or IM approaches are adopted compared with CC as the number of events included in the analysis is reduced using the later.

Figure 34: Overall crude excess mortality rates using three methods to deal with missing values

6.4.5.1



	Complete-Case	Indicator Method	MICE
eMR (CI 95%)	0.64 (0.54;0.75)	0.82 (0.73;0.93)	0.82 (0.73;0.93)
SE of Log rate	0.084	0.062	0.062

eMR: excess Mortality Rate per 1000 py (persons-year)
CI: Confidence Interval

Results from multivariable models to assess association between overall excess mortality and covariates using three different methods to deal with incomplete data are compared in Table 15. There was little difference between overall adjusted excess Hazard Ratios obtained from multivariable model using the three methods to deal with missing values. Adjusted eHR from MICE, IM and CC analyses were contained in each other 95% Confidence Intervals, respectively.

Further, we observed that the Standard Errors of the log of excess Hazard Ratios were higher on analyses based on CC approach.

Table 15: Overall adjusted excess Hazard Ratio (eHR) obtained from multivariable model using three methods to deal with missing values

	Complete-Case analysis		Indicator method		Multiple Imputation	
	eHR[a] (95% CI)	SE[b] log eHR	eHR[a] (95% CI)	SE[b] log eHR	eHR[a] (95% CI)	SE[b] log eHR
Education						
No/compulsory	1		1		1	
Upper Secondary	0.53 (0.34;0.83)	0.230	0.47 (0.31;0.71)	0.216	0.51 (0.33;0.80)	0.224
University	0.52 (0.29;0.93)	0.294	0.52 (0.32;0.86)	0.254	0.57 (0.35;0.92)	0.245
Others	0.86 (0.50;1.48)	0.278	0.92 (0.59;1.46)	0.232	0.85 (0.53;1.36)	0.239
Unknown	-	-	1.62 (1.22;2.15)	0.145	-	-
HIV tramission mode						
IDU	1		1		1	
MSM	0.41 (0.22;0.77)	0.317	0.46 (0.30;0.72)	0.227	0.49 (0.30;0.80)	0.246
Heterosexual	0.70 (0.42;1.16)	0.263	0.75 (0.52;1.10)	0.192	0.84 (0.57;1.24)	0.201
Others	1.04 (0.38;2.87)	0.519	1.02 (0.47;2.22)	0.261	1.11 (0.49;2.53)	0.418
Unknown	-	-	1.65 (0.99;2.76)	0.399	-	-
Origin						
Spain	1		1		1	
SSA	0.49 (0.15;1.54)	0.587	0.63 (0.34;1.20)	0.324	0.61 (0.32;1.16)	0.327
LA	0.79 (0.44;1.42)	0.296	0.82 (0.53;1.25)	0.217	0.76 (0.50;1.17)	0.218
Others	0.79 (0.40;1.57)	0.351	0.70 (0.39;1.23)	0.292	0.67 (0.38;1.18)	0.292
Unknown	-	-	0.66 (0.23;1.86)	0.530	-	-
CD4 at entry (cel/mm3)						
<=200	1		1		1	
201-350	0.59 (0.38;0.93)	0.238	0.56 (0.38;0.80)	0.188	0.56 (0.39;0.81)	0.190
>350	0.42 (0.27;0.66)	0.222	0.33 (0.22;0.49)	0.200	0.33 (0.23;0.48)	0.191
Unknown	-	-	0.99 (0.49;2.02)	0.363	-	-
VL at entry (copies/ml)						
<20000	1		1		1	
20000-100000	1.05 (0.66;1.67)	0.238	1.06 (0.71;1.57)	0.201	1.04 (0.71;1.54)	0.199
> 100000	1.60 (1.04;2.47)	0.222	1.50 (1.04;2.17)	0.187	1.44 (1.01;2.06)	0.182
Unknown	-	-	1.15 (0.53;2.46)	0.389	-	-
Sex						
Male	1		1		1	
Female	0.66 (0.42;1.04)	0.232	0.78 (0.57;1.08)	0.165	0.74 (0.53;1.02)	0.166
Age at entry						
20-49	1		1		1	
>=50	2.40 (1.65;3.51)	0.193	2.04 (1.53;2.72)	0.146	2.04 (1.54;2.71)	0.144

6.4.5.2 Table 15: Continued

	Complete-Case analysis				Indicator method				Multiple Imputation by Chained Equations			
	Follow-up interval				Follow-up interval				Follow-up interval			
	1st year		2nd -10th year		1st year		2nd -10th year		1st year		2nd -10th year	
	eHR[a] (95% CI)	SE[b]	eHR[a] (95% CI)	SE[b]	eHR[a] (95% CI)	SE[b]	eHR[a] (95% CI)	SE[b]	eHR[a] (95% CI)	SE[b]	eHR[a] (95% CI)	SE[b]
HCV at entry												
Negative	1		1									
Positive	1.82 (1.01;3.28)	0.300	3.94 (2.18;7.10)	0.301	1.52 (0.99;2.33)	0.220	4.04 (2.52;6.48)	0.270	1.37 (0.87;2.15)	0.233	3.83 (2.37;6.19)	0.245
Unknown	-	-	-	-	1.45 (0.93;2.25)	0.225	1.86 (0.84;4.12)	0.370				
AIDS entry												
No	1		1									
Yes	2.56 (1.58;4.16)	0.247	0.80 (0.46;1.37)	0.277	3.82 (2.75;5.30)	0.167	1.03 (0.67;1.59)	0.267	3.71 (2.66;5.19)	0.170	0.90 (0.58;1.39)	0.226

[a] *Adjusted for all other variables in the table*

[b] *SE: Standard Error of log eHR*

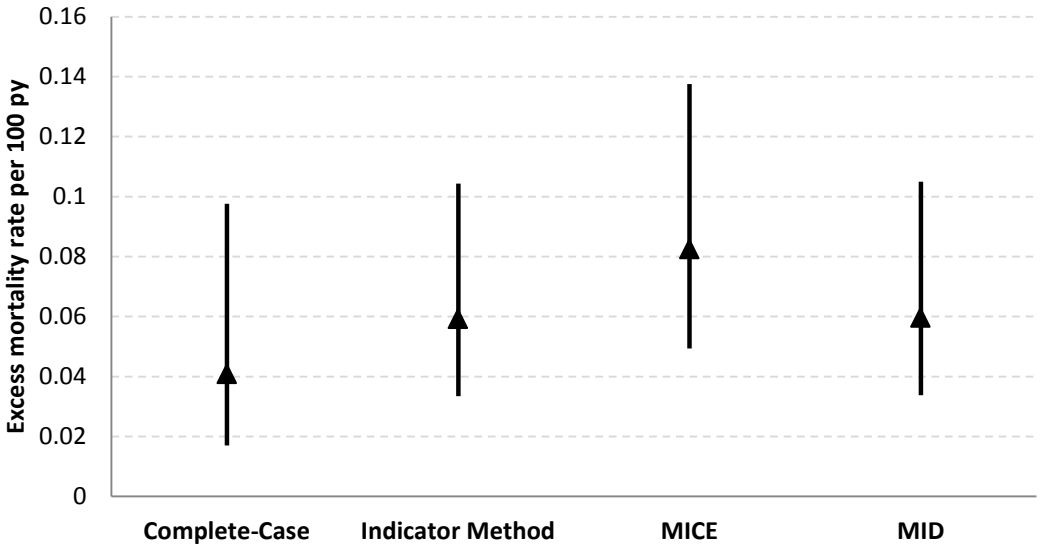
CI: Confidence Interval

SSA=Sub-Saharan Africa, LA=Latin-America

6.4.6 Non-AIDS malignancies excess mortality

Figure 35 shows the estimates of the non-AIDS malignancies excess mortality rate based on four different methods to deal with missing data including missing data in the outcome. Incorrect classification of CoD (IM) or deletion of cases with missing CoD (CC, MID) seems to lead to underestimation of the non-AIDS malignancies excess mortality rate. Besides, MICE approach provides the most precise estimations as the Standard Error of log excess rate is reduced.

Figure 35: Non-AIDS Malignancies crude excess mortality rates using four methods to deal with missing values



	Complete-Case	Indicator Method	MICE	MID
eMR (CI 95%):	0.04 (0.02;0.10)	0.06 (0.03;0.10)	0.08 (0.05;0.14)	0.06 (0.04;0.11)
SE of Log rate	0.446	0.290	0.261	0.273

eMR: excess Mortality Rate per 100 py (persons-year)
CI: Confidence Interval
MID: Multiple Imputation, then Deletion

Results from multivariable models to assess association between non-AIDS malignancies excess mortality and covariates using four different methods to deal with incomplete data (including CoD) are compared in Table 16. Significant differences between estimates were not found. Adjusted eHR from CC, IM, MICE and MID, were contained in each other 95% Confidence Intervals, respectively. Nevertheless, the estimates from IM, MICE and MID show a difference in a consistent direction, with eHRs being smaller than those from the analysis based on complete cases. Differences in point estimates between IM, MICE and MID are immaterial, however we observed that the Standard Errors of log HRs from IM and MID were higher than those from MICE models due to the reduction in the number of non-AIDS malignancies-related deaths included. MICE approach provides the most precise estimates compared to the other methods considered in this study.

Table 16: Non-AIDS Malignancies adjusted excess Hazard Ratio (eHR) obtained from multivariable model using four methods to deal with missing values

	Complete-Case analysis		Indicator method		MICE		*MID*	
	eHR[a] (95% CI)	**SE log eHR**	**eHR[a] (95% CI)**	**SE log eHR**	**eHR[a] (95% CI)**	**SE log eHR**	**eHR[a] (95% CI)**	**SE log eHR**
HIV tramission								
IDU	1		1		1		1	
MSM	0.31 (0.02;5.13)	1.437	0.69 (0.08;5.84)	1.087	0.69 (0.10;4.58)	0.964	0.54 (0.06;4.98)	1.130
Heterosexual	1.86 (0.46;7.54)	0.713	4.23 (0.92;19.37)	0.777	2.69 (0.71;10.26)	0.682	3.56 (0.84;15.08)	0.737
Others	NA	NA	NA	NA	NA	NA	NA	NA
Unknown			1.88 (0.10;34.34)	1.483				
CD4 at entry (cel/mm3)								
<=200	1		1		1		1	
201-350	1.05 (0.27;4.04)	0.689	1.16 (0.39;3.44)	0.557	0.86 (0.30;2.48)	0.540	0.87 (0.29;2.58)	0.557
>350	0.68 (0.16;2.93)	0.744	0.40 (0.09;1.77)	0.762	0.47 (0.17;1.34)	0.530	0.31 (0.08;1.28)	0.722
Unknown			2.36 (0.83;6.72)	0.533				
Age at entry								
20-49	1		1		1		1	
>=50	8.74 (2.75;27.78)	0.590	5.55 (2.35;13.07)	0.437	5.06 (2.30;11.13)	0.402	5.32 (2.27;12.44)	0.434
HCV at entry								
Negative	1		1		1		1	
Positive	6.45 (1.67;24.84)	0.688	4.71 (1.76;12.60)	0.502	3.67 (1.33;10.15)	0.518	4.11 (1.53;11.04)	0.503
Unknown			1.57 (0.47;5.26)	0.616				
Follow-up								
1st Year	1		1		1		1	
2-10th year	0.51 (0.16;1.63)	0.589	0.48 (0.21;1.11)	0.430	0.60 (0.27;1.29)	0.396	0.46 (0.20;1.07)	0.430

[a] *Adjusted for all other variables in the table*

SE log eHR: Standard Error of log excess Hazard Ratio; CI: Confidence Interval; MID: Multiple Imputation, then Deletion

6.4.7 Sensitivity analysis: extreme case analysis

MICE approach provides unbiased results assuming data are MAR and the imputation model is correctly specified. To assess the robustness of our key inferences to possible departures from MAR assumption we repeated, our main analysis under two different MNAR mechanisms. Firstly, we replaced all missing values from the variable with highest missingness (i.e. level of education) with the lowest possible value of the variable: "No or compulsory education". Secondly, we replaced all missing values from level of education with the highest possible value of the variable: "University". We found little difference between results from the MICE imputation model and both extreme cases models.

7 Discussion

7.1 Discussion of results

This Doctoral Thesis provides an in-depth analysis of the overall and cause-specific mortality trends in the cART era in a cohort of naïve HIV-positive patients in Spain, as well as important evidence of the role of HCV coinfection in mortality. This study gives also guidance on how to deal with missing data on the covariates and the CoD in HIV-positive cohort based analyses. These results are key to influence changes in the clinical management of HIV-positive patients in order to diminish mortality, especially excess mortality of those coinfected with HCV.

In the **first objective** of this Doctoral Thesis, **"to assess trends in overall mortality and cause-specific mortality stratified by HCV serostatus in a cohort of naïve HIV-positive patients in Spain**", that analysed mortality trends in the cART era (year 1997-2008), we found a decline in overall mortality in the whole population taking the first years of the cART era as reference. This agrees with what has been found in other cohort studies that have analysed mortality rates among HIV-positive persons in the early and late cART era (120). This trend was also found when we considered different causes of death including AIDS-related death, liver-related death, and NLRD-NARD. Of note, the mortality declines in our cohort were largely at the expense of the HCV-negative patients, as declines were not observed among HCV-positive patients. Interestingly, among HCV-positive patients in our cohort, we found a non-significant trend not only towards increase in liver-related death, but also a trend towards increase in AIDS-related death and NLRD-NARD.

We found that the liver-related death rate among HCV-positive patients did not decline between the two periods. This is of interest because anti-HCV therapy based on interferon plus ribavirin has been available in Spain since 2000, and as mentioned before, the sustained virological response reduces liver-related complications and mortality irrespectively of the stage of fibrosis (63). It is not possible from our data to find explanations to this observation, but in our view it must be due to several factors including barriers that limit the initiation and completion of anti-HCV therapy in this population including missed clinic visits, active psychiatric illness, active drug or alcohol use, decompensated liver disease, medical illness, patients' reluctance to start anti-HCV treatment, and the low tolerability of treatment (41,121). In a cross-sectional study carried out in Spain several years ago it was found that between 43% and 45% of HIV/HCV coinfected patients were good candidates to interferon plus ribavirin despite of which only 14% and 15% of patients had initiated it (122). Another factor to be taken into consideration is the low sustained virological response rates to interferon plus ribavirin therapy particularly among patients infected with HCV genotypes 1 or 4 that represent more than 60% of HIV/HCV-coinfected patients in Spain (123). The importance of end-stage liver disease as a CoD in the cART era has also been emphasized in several studies (120,124,125). Nevertheless, numbers are small and warrant larger collaborations.

In the cART era, AIDS-defining conditions have been the first CoD in absolute terms in some cohorts (32) or institutions (126) but in others cohorts, the first CoD has been occupied by non-AIDS-defining deaths (127). In absolute numbers, AIDS-related

death was the most frequent CoD in both periods in our study. Interestingly, in our cohort, AIDS-related mortality rate declined between the two periods among HCV-negative patients but remained virtually unchanged among HCV Positive patients. This finding raises again the question as whether HCV may impact negatively in the natural history of HIV infection. This is a difficult question to answer because IDU is the primary risk factor among HCV-positive patients whereas HCV-negative patients are more likely to have acquired HIV by sexual transmission; meaning that HCV seropositivity may be a marker for poorer access to care and competing problems with drug-addiction. In fact, it has been reported that IDUs begin cART at a more advanced stage of HIV disease than non-IDUs (128,129). This may explain in part differences found in studies that have addressed this issue from a clinical (51,54) or immunological perspective (50,55). However, some publications give some weight to the plausibility that HCV may negatively influence HIV infection. For example, in one prospective study, increases in the baseline HCV RNA were found to be associated with increases in the relative risk for progression to AIDS and AIDS-related mortality after controlling for CD4+ cell count and HIV-VL level (130). In another study, HIV-positive patients coinfected with HCV exhibited higher grades of immune activation (56) and had increased risk of HIV clinical progression (131).

In our cohort, non-liver-related non-AIDS-related deaths ranked second in the causes of mortality accounting for 33% of deaths in the first calendar period and 35% of deaths in the subsequent calendar period. We found that non-liver-related non-AIDS-related mortality rate declined significantly from 1997-2003 to 2004-2008;

however, when we stratified by HCV-serostatus the decrease was present among HCV-negatives only. Some causes of death in this category including substance abuse, accidental or trauma events and suicide were particularly common among HCV-positive patients and virtually inexistent among HCV-negatives. These causes of death have been frequently reported among IDUs in several studies (132-135) and consequently seem more related to lifestyle factors rather than to biological factors related to HCV infection. However, a large proportion of non-liver-related non-AIDS-related deaths among HCV-positive patients were caused by cardiovascular disease and non-AIDS malignancies, conditions in addition to end-stage liver disease that now play a major role in most of the fatal conditions classified as non AIDS-related events as a consequence of increased survival in this population group (24-27). It is conceivable that chronic infection by HCV, may also contribute to persistent inflammation and immunosenescence in the same way that other chronic viral co-pathogens such as *Cytomegalovirus* do in the HIV-positive population with a net result of more rapid progression of atherogenesis and oncogenesis (136); factors that may emerge as frequent causes of mortality in HIV-positive persons the years to come.

Regarding the **second objective** of this Doctoral Thesis, **"to calculate the overall mortality rates, standardized mortality ratios and excess mortality rates in the cohorts of the Spanish AIDS Research Network"**, we showed that all-cause mortality in CoRIS-MD and CoRIS cohorts, between 1997 and 2010, was close to five times higher than that of the general population of the same age and sex. Significant

differences have been found depending on the history of AIDS and HCV co-infection at cohort entry.

A previously published study, carried out in similar cohorts in Europe and North America, found an overall SMR of 3.36 (95% CI: 3.16; 3.56), but with a notable heterogeneity between cohorts depending on participant-specific characteristics, and being higher for cohorts with a greater representation of IDUs (137). For example, Aldaz et al. found mortality of HIV-positive persons in Navarre (Spain) to be 14 times higher than mortality in general population; 63% of this cohort had been infected through the use of injected drugs (138).

These differences could also be related to the higher prevalence of HCV-co-infection as the standardized mortality in HCV co-infected subjects in our study was 9.2 times higher than the general population's. Similar results were found by Lewden et al., where SMR for HCV co-infected persons was 13.9 compared to 4.4 for the HCV negative subjects (37). In a previous study of CoRIS-MD and CoRIS cohorts, an important increase of the risk of both all-cause mortality and liver-related mortality was observed for HIV patients coinfected with HCV (30). Chen et al in a meta-analysis found that the risk of mortality was increased in HCV/HIV coinfected patients in cART era (67).

In our study, we found a similar SMR for patients recruited in CoRIS, from 2004 onwards, and those recruited in CoRIS-MD, from 1997 to 2003, after adjustment for gender, age, transmission category and HCV infection. That is, the differences in the

subject´s characteristics along these years, the decrease in the representation of IDUs and the percentage of HCV co-infected subjects (108,139,140) were corrected after adjustment. Others studies observed a lower mortality in recent years with the improvement in antiretroviral therapies (141-144), although when specific groups were analysed, for example: IDUs, found that mortality risk remain elevated (142).

We found non-statistically significant, lower mortality rates in women compared to men. Even though, the women in our study showed a standardized mortality ratio 10.5 times higher than women of the same age from the general population. This higher relative mortality in women could be explained by the fact that women in the general population have a higher life expectancy than men, and specifically, mortality in the general population is very low in women between ages 30 to 40, where we find the majority of HIV-positive women (35). The lower excess mortality rate in women is consistent with the higher proportion of HIV-positive men in the Spanish epidemic, and in our cohorts (9).

This study was the foundation for the third objective where it was considered, along with overall mortality, excess mortality rate for specific causes of death, such as non-AIDS malignancies, liver or cardiovascular disease and potential prognostic factors for mortality excess were investigated.

Regarding the **third objective** in this Doctoral Thesis, **"to evaluate the overall and cause-specific excess mortality observed in HIV-positive subjects and to identify prognostic factors of excess mortality",** we found that HIV-positive subjects in CoRIS

experienced overall, non-AIDS malignancies as well as liver and cardiovascular related excess mortality compared to the general population of the same age and sex, largely at the expense of mortality in the first year following the cohort inclusion.

Despite the large reduction in mortality after the introduction of cART in 1996, the overall excess mortality associated with being HIV-positive observed in CoRIS between 2004 and 2014 was still 0.82 deaths per 100 person-years of follow-up, when compared to the general Spanish population of the same age and sex in Spain. The excess mortality in HIV-positive patients has been reported both in Spain and elsewhere, as previously described in the introduction of this thesis (30,138,142). Aldaz et al found that the overall excess mortality rates in the period 1999-2006 in HIV-positive subjects in Navarra was 2.74 per 100 py of follow-up, cause-specific excess mortality rates estimated were 0.49, 0.16 and 0.08 deaths per 100 py for liver, non-AIDS defining cancers and cardiovascular disease, correspondingly (138). As previously discussed in the second objective, higher rates compared to CoRIS can be explained because of the higher proportion of IDUs in this study population, together with a reduction of overall and liver-related excess mortality associated with HIV infection over time (29,30).

On the other hand, the CASCADE collaboration reported an overall excess mortality of 0.6 deaths per 100 py during the period 2004-2006 (29). CASCADE collects data from seroconverter subjects (date of HIV seroconversion reliably estimated is available) hence the discrepancies with our results can be also explained because of

different study populations characteristics. This is supported by the fact that CASCADE's overall excess mortality rate is similar to the 0.50 per 100 py found in CoRIS after excluding short-term mortality. Not surprisingly, short-term mortality is less likely to happen in seroconverter cohorts because patients included have better clinical and immunological conditions at cohort entry.

HIV associated factors such as lower CD4 count and higher levels of VL at entry and socio-economic characteristics such as lower education, being male and older than 50 years were found to be strong predictors for overall mortality excess. Besides, there was evidence that AIDS before entry and HCV coinfection were positively associated with short-term and long-term excess mortality, respectively. In the first and second objectives of this Thesis, we have also noted the importance of HCV/HIV coinfection on overall and cause-specific related mortality.

Hernando et al found a high excess mortality in HCV/HIV coinfected patients in the Cohorts of the Spanish AIDS Research Network (30). Due to the characteristics of one of the cohorts, only patients with at least 6 months of follow-up were included in the study therefore short-term mortality was not explored. We have now found that HCV coinfection has no effect on short-term excess mortality while it has large impact on long-term mortality. Long-term excess mortality in HCV coinfected patients may be explained by the impact of HCV in the natural history of HIV infection, nevertheless, as previously discussed in the first objective, data on this question are inconsistent. In addition, Berenguer et al have recently published that the eradication of HCV in

HIV/HCV coinfected patients is associated with decreases in HIV progression and lower risks of both liver and non-liver-related mortality (60). On the other hand, long-term overall excess mortality in HCV-positive patients can be also explained by indirect mortality associated with HIV/HCV coinfection, pointing to the role of other contributing factors such drugs, alcohol and tobacco whose consumption is higher than in the general population (66).

The increase of excess mortality during the first year of follow-up among those patients with AIDS before cohort enrolment can be explained because new cases arriving to hospitals with a previous diagnosis of AIDS are those who have been diagnosed late with HIV infection; implications of a delayed HIV diagnosis and late presentation have been previously documented (144). AIDS diagnosis has been related with poorer responses to cART (145), but we have found that long-term excess mortality did not vary according to AIDS status before entry, pointing to the benefit of cART initiation even accepting lower clinical and virological responses.

We observed significantly higher non-AIDS malignancies-related mortality rates among HIV-positive patients compared with the general population. Our finding is not surprising as excess cancer incidence in HIV-positive patients compared with general population has been previously documented (146,147). Non-AIDS malignancies excess mortality can be both directly and indirectly associated with HIV; although literature about influence of immunodeficiency on non-AIDS malignancies is inconsistent, some studies have suggested that immunosuppression might be associated with moderate

excesses of non-AIDS-defining cancers (148-150) and with more rapid progression of some types of cancer in predisposed individuals (151). On the other hand, non-AIDS malignancies excess mortality can be also explained by the elevated frequency of cancer risk factors in HIV-positive patients such as smoking, alcohol consumption (152,153), and coinfection with carcinogenic virus such as Hepatitis D Virus (HDV) and Human Papiloma Virus (HPV) (154) who share sexual and parental transmission routes with HIV.

We found HCV coinfection and aging to be the main risk factors for non-AIDS malignancies excess mortality in CoRIS. Several studies have previously reported increased risk of non-AIDS malignancies after excluding liver cancer (i.e. renal and prostate cancer) among HCV positive subjects (155-157); as previously mentioned in objective 1, it has been suggested that inflammation produced by HCV infection could partly explained the increased cancer incidence among these patients (158). However, it seems more likely that the increased risk is related with lifestyle factors rather than a direct oncogenic effect of HCV infection. Aging is a well-known risk factor for excess non-AIDS malignancies rates (146,147), additionally patients older than 50 years old have shown a poorer immunological response to cART in CoRIS, which might also contribute to the higher non-AIDS malignancies excess mortality observed in older patients (113).

Liver excess mortality associated with being HIV-positive can be explained because HIV has a direct impact on the natural history of HCV, promoting more rapid

to liver-related mortality and, as well as HBV, increasing the risk of hepatocarcinoma. (159,160). Besides, the toxicity of long-term antiviral treatment may also cause liver damage and, hence contribute to liver excess mortality (161,162). Finally, there is also indirect effect since hepatotoxic substances such as alcohol and legal and illegal drugs are often consumed by HIV-positive subjects (152).

Our results regarding moderate cardiovascular excess mortality detected in CoRIS were in the line with several studies that have reported increased risk of cardiovascular disease among HIV-positive patients compared with HIV-negative individuals, after appropriate adjustment for traditional risk factors (163-165). The increased risk for cardiovascular mortality among HIV-positive subjects is a combined result of the high prevalence of cardiovascular risk factors, cART-related metabolic changes, and systemic immune activation that promotes endothelial inflammation and atherosclerosis (166).

Studies in the literature (30,142) have reported a higher overall excess mortality in HIV-positive males compared with females and it has been previously discussed in the second objective. Additionally, we observed higher non-AIDS malignancies-related excess mortality in males compared to females although there was no statistically significant evidence for these findings. No differences by sex were found in liver excess mortality. This result could be somehow surprising, since previous analysis carried out in CoRIS have shown a higher risk of excess liver mortality in men compared to women. However, these discrepancies are explained by the decrement in

serial prevalence of women coinfected with HCV and injecting drug use (167). Sex stratified analyses would provide a better understanding of gender differences in overall and cause-specific excess mortality associated with being HIV-positive.

Competing explanations for lower mortality rates in migrant populations fall into two broad categories. The first posits the self-selection of healthier migrants driven by labour market conditions; migrants are healthier than general population from both countries of origin and destinations, in what is known as the "Healthy migrant effect" (168). Thus, mortality rates of migrants from the general population in Spain are likely to be lower than the Spanish rates. A second hypothesis, known as the "Salmon bias", proposes that foreign-born persons that have live in a different country for some time return to their origin country when they become severe ill (169). Therefore our analysis should have used general population mortality rates matched by sex, age and also country of origin. Unfortunately, cause-specific mortality rates by country of origin are not easily accessible at the moment in Spain, nevertheless we will work to include this information in further analyses.

We observed that the gap in overall and cause-specific mortality rates between HIV-positive individuals and Spanish general population did not vary by period of inclusion. Overall and cause-specific risk factors for excess mortality also remained unchanged over period of inclusion. These results are not surprising since remarkable changes in HIV disease management or cART regimes have not taken place in the last decade.

The **fourth objective** of this Doctoral Thesis was **"to describe the distribution of incomplete data, to build an imputation model and to compare different methods to handle missing data in CoRIS"**. Data checking in CoRIS database revealed the existence of missing values in 7 of the variables considered for the analysis. Missing data patterns and distribution by covariates were explored to identify possible reasons for incomplete data and to decide the best approach to cope with missing data.

In general, we observed that missing data were more common in those patients with worse values of disease progression, for example missing HCV at entry was more common in patients with primary/no education, IDUs, with lower values of CD4 and higher VL, in subjects with AIDS and cART at entry. Patients with these characteristics usually have worse follow-up and no stable contacts, hence missing data are explained. At cohort entry, short-term survivors have poorer conditions thus the probability of collecting socio-demographic information is lower since they cannot be asked about their educational level or country of birth. Besides, the probability of missing data can be also related with covariates as age or HIV transmission category. Contact and follow-up is more difficult among IDUs and younger people therefore collecting their information is also more challenging. In addition, data in CoRIS are collected by clinicians during the clinic routine, generally socio-demographic information is not asked in the first day visit, naturally, if the patient follow-up is short or is not stable over time, then the clinician has no further opportunities to ask for information. Sobrino et al described that lost to follow-up is more common among younger patients, IDUs, non-Spanish origin, low educational level and better clinical

conditions at cohort entry (139). Subsequently, missing data due to loss to follow-up is associated with these covariates.

Multiple Imputation models for variables with missing data must be carefully specified in order to reduce bias, improve efficiency and make the MAR assumption more plausible (100,103,170). For each incomplete variable, imputation models for CoRIS data included the other incomplete variables, the complete variables (AIDS at entry, age and sex), the outcome (survival time and CoD) and one auxiliary variable "cART at entry". The inclusion of variables that are strongly related to both the missingness and the values (observed and missing) of the variable with incomplete data enhances the plausibility of the MAR assumption to hold, thus the bias corrections in estimates of associations are more substantial (170,171). Nevertheless, there may be two other broad occurrences. First, adding predictors of the variable being imputed (even if unrelated with missingness) reduces standard errors of the estimates (maximize the use of available information), hence efficiency will usually improve and we will get more accurate estimates (170,172). Second, on the ground of making MAR assumption more plausible, the inclusion of variables not associated with the variable being imputed but predictors of missingness may slightly produce a loss of precision but should not cause bias (103,171). Spratt et al found that imputation models that added predictors of the values of the variables with most missing data had the greatest impact on the estimates and their standard errors; additionally, including variables related only to the probability of missingness had little impact on neither the estimates nor the efficiency (170). In our data the covariate with most missing data

was level of education and we found several predictors that were included in the model.

The approach called "Predictive Mean Matching", which consists in matching the missing value to the observed value with the closest predicted, has also been suggested to impute with non-normally distributed variables (105); however, we found satisfactory transformations towards normality for the non-symmetric variables.

In addition, any structure (i.e. interactions) included in the analysis models must be also considered in the imputation models. Our imputation models were really complex and a lot of parameters were estimated, hence we used passive imputation as it is the simplest approach. We were aware that some interactions from the models are ignored and the cost is bias of relevant terms of the model and loss of power to detect interactions. However, we performed several sensitivity analyses (exploring with different approaches) and the impact on the estimated parameters of the final models was immaterial. On the other hand, White et all and Von Hippel et all have performed simulations and sensitivity analyses with linear regression and logistic regression (i.e. JAV approach is unbiased under MCAR) (100). However, we have not found a strong recommendation for generalized linear models with Poisson error structure.

One of the main goals of imputation methods is to use as much information as possible to execute the imputations. In cross-sectional studies, there is only information available at one time-point; therefore this is the only information that can

be used when performing multiple imputation. On the other hand, for longitudinal studies there are variables that are measured at several times during the follow-up, therefore if one of this measurements is missing it is strongly recommended to use the information from the same patient at a different time to impute the missing value. Despite the fact that statistical software to perform imputation with multilevel data is available (173), the number of subjects with missing information on CD4 count or VL at entry but completed during follow-up was very small, hence for the sake of simplicity and due to the immaterial improvement in the results we decided to consider only variables at enrolment.

The inappropriate imputation model specification (structure and variables selection) could lead to bias and imprecise estimates therefore model specification must be carefully performed and justified. Nevertheless, MICE models, fitted using Markov Chain Monte Carlo method, are computationally intense. The required length increases where more missing data are present and more parameters must be estimated, to the extent that models might not reach convergence (174). MICE models must be appropriately specified but computationally feasible.

Data checking after Multiple Imputation revealed that all imputed values were plausible. Multiple imputation method rests on the assumption that data are Missing At Random. Although the MAR assumption cannot be definitely verified with the use of observed data, it seems plausible in this context particularly in the view of the powerful predictive covariates included in the imputation models. We also

investigated the robustness of key inferences based on MICE analysis to possible departures of MAR assumption; when we repeated the main analysis under a MNAR we found that key inferences remained unchanged.

Besides, sensitivity analysis restricted to complete cases did not reveal changes in the direction of the association between excess mortality and risk factors. Discrepancies and wider Confidence Intervals found in the Complete-Case analyses are explained because of the reduction of the number of subjects included.

Sensitivity analysis with missing data is an on-going research area; several approaches have been proposed nevertheless most of these techniques are technical and inaccessible for practitioners (105,175). We decided to carry out extreme-case and CC approaches to assess the sensitivity of our estimates to the missing data mechanism due to its ease of implementation.

Percentages of missing values are low (under 10%) for most of the incomplete variables in CoRIS, hence CC approach could be used assuming the missing pattern is MCAR, but several predictors for missing data were detected; therefore this assumption might be incorrect. Another disadvantage of CC approach is the substantial sample-size reduction which might produce lack of power to detect statistically significant associations. In CoRIS, total sample size would be reduced from 10,340 to 7,337 patients, and 156 deaths would not be included. Despite all the efforts made to improve the classification of deaths in CoRIS, information about CoD is still missing in 47 cases. Although changes in the eHR for the effect of covariates were not

substantial, we detected that CC analysis underestimated overall excess mortality rates. And no or incorrect classification of CoD (CC, MID or IM) leaded to underestimation of cause-specific excess mortality rates. As previously reported by Spratt et al, marginal means (such us rate or prevalence) are likely to show greater bias in CC analysis than estimates of association, because marginal means are bias if data are MAR (170).

Multiple imputation method has been suggested as the best approach to cope with missing data (96). Imputation by Chained Equations (MICE algorithm) (99,100) allows to impute missing values in multiple variables when the missing pattern is not monotone and under MAR assumptions. Besides logistic, multinomial or ordered regression can be used instead linear regression for non-normal variables.

It could be argued that significant differences between the four methods to deal with missing data were not found due to sample size limitations. When significant differences between approaches are detected, it is essential to explain why these differences are produced in order to increase the confidence in the conclusions from MICE analysis (174). Assuming that data are MAR and that the imputation models have been correctly specified, the differences observed between analyses based on MICE and based on Complete-Case approach can be explained by bias corrections.

Further, in order to choose the most suitable method to deal with missing values in a given dataset, first we need to study carefully the assumptions and potential pitfalls from each method and then make a decision based on which

assumptions are we able to fulfil better and the consequences of deviances from these assumptions.

7.2 Methodological discussion

The results of this Doctoral Thesis need to be interpreted in the light of its methodological limitations.

7.2.1 Internal Validity

The main biases that can affect internal validity of our analyses are those derived from: misclassification of missing data, misclassification in the event of interest, use of the general population as a proxy for background mortality, short period and losses to follow-up, joint analysis of different cohorts, lack of observation on variables that could be potential confounders, possible additional analyses and low precision in some estimates.

Through this Doctoral Thesis we have addressed one of the most common occurrences in any type of research: missing data. We started using Indicator Method and Complete-Case because its use is straightforward and widespread in epidemiological research (objectives 1 and 2). Afterwards, we built on analyses performed in objective 4 and took one step ahead to apply Multiple Imputation by Chained Equations in CoRIS cohort as it makes the maximum use of the available information (objective 3).

The possible bias caused by misclassification of missing data could also have affected our results in objective 3. As previously discussed, Multiple Imputation assumes that incomplete data are MAR, although this assumption cannot be proved

with the observed data, we included in the models several predictors for missingness as well as the outcome considered in the final model to make this assumption more plausible. Specifically, individuals might not be systematically tested for HCV as it might be possible that the ones who do not have a test result are more like to be HCV negative, hence HCV test variable would be MNAR. However, extreme cases sensitivity analysis was performed on HCV variable and we observed that our results were robust to possible departures of MAR assumption.

Besides, we decided to impute CoD that was missing in 12.77% of the deceased subjects occurred in CoRIS, which may produce misclassification in the event of interest. Inclusion of subjects with missing event values has been insufficiently studied (86,100,106). As previously discussed, auxiliary variables were highly correlated with CoD being missing and excluding these cases would have led to underestimation of cause-specific excess mortality.

Other advanced methods could have been used to impute CoD, as the one proposed Bakoyannis et al (176) . However, we decided to impute CoD using MICE because we aimed to impute simultaneously missing values in both CoD and covariates.

Regarding misclassification of the event of interest, another aspect that could be argued is the convenience of using different coding algorithms in the numerator and denominator of excess mortality estimations. However, it has been shown that revised CoDE classification is the best way to classify and group CoD in HIV-positive

cohort studies (69). Applying ICD-10 system to deaths occurring in CoRIS data would underestimate the liver excess mortality. Hernando et al detected that ICD-10 system overestimates HIV/AIDS-related deaths; cirrhosis of viral cause or unknown cause are labelled as HIV/AIDS-related deaths using ICD-10 while they are assigned to liver disease group using CoDE algorithm.

Selection bias could have been introduced by the use of the general population as a proxy for background population; we used the rate in the general population to calculate the expected deaths in the estimation of SMR and excess mortality, although this population contains also HIV-related deaths. Nevertheless, Ederer et al demonstrated that the estimation of background mortality is not affected by the HIV deaths included in the general population, as they represent a negligible proportion in our setting (79,80).

Another important consideration regarding internal validity of results from objective 1 lies in the fact that overall periods of follow-up are short, particularly in the 2004-2008 period. As HCV-related liver disease requires time to progress, this may result in underestimating its contribution to mortality.

Losses to follow-up are a usual limitation in cohort studies; CoRIS is an on-going cohort yearly up-dated hence this limitation is minimized. The proportion of patients lost to follow-up (approximately 19.8%) is similar to comparable cohorts (111). Sobrino et al also detected that loss to follow-up rate was higher in younger people, IDUs, in patients with non-Spanish origin, with no or primary education and higher CD4 count

at cohort entry. The cohort is losing both patients with major social exclusion risk and with better clinical characteristics therefore we considered our results are likely not biased.

Potential biases caused by the joint analysis of two different cohorts with slightly different inclusion criteria may be have been present in objectives 1 and 2, however we tried to minimize this limitation by restricting the analyses to patients that fulfilled the inclusion criteria for both cohorts. In the objective 2, risk of selection bias resulting in too high Standardized Mortality Ratios needs also to be discussed. One of the inclusion criteria in CoRIS-MD is that patients had at least 6 months of follow-up. For that, to merge both cohorts, we established this time for all patients: we did not consider patients with time of follow-up less than 6 months and we deleted individuals who had died in this period and had not reached six months of follow-up. The sensitivity analysis showed that when we established as inclusion criteria to have at least 6 months of follow-up, we were introducing a time window to avoid selection bias indirectly and overestimate SMRs.

Another limitation of this Doctoral Thesis is that important variables can be missing in our analysis. As it has been the case with other studies of similar characteristics, some possible confounding factors (e.g. adherence to cART, active mental illness, alcohol/tobacco use, access/adherence to IFN/RBV) were not collected and consequently their impact in the results could not be analysed. Information on smoking and alcohol consumption is necessary to estimate the excess mortality

directly attributable to HIV and/or HCV infections. Besides, no determination of HCV-RNA was performed; this could have sorted out better the contribution of active infection to mortality.

Another aspect that needs to be discussed is that several additional analyses could have been also performed in objective 4; we assumed that the hazards were constant over the first year and the rest of follow-up respectively. Sensitivity analyses were carried out fitting the final models including narrower time-bands. However, flexible parametric models (177) might have been used. These models are fitted on the continuous time on the log cumulative baseline excess hazards scale using restricted cubic splines to estimate the baseline log cumulative hazard. Then, continuous covariates and time-dependents effects can be easily added. They might be very useful to assess the variation of excess mortality over follow-up, as finely as possible, and to explore the effects of naturally metric variables such as age or CD4 counts. However, these models require high sample sizes as a large number of parameters must be estimated.

Time updated variables such as treatment initiation, AIDS or calendar period of follow-up were not considered for the purposes of objective 3 because of computational and time expenses. Nevertheless Lexis-expansion will be applied to the data to include time-updated variables in further analysis. Regarding to cART initiation, we assumed that all subjects that enrolled the cohort who fulfilled criteria to

initiate cART, were started given access to treatment in Spain can be considered universal and free.

Finally, the low number of deaths, when aiming to look at cause-specific mortality, may have introduced random error in some of our estimates and comparisons. For example, in objective 3 we failed to find significant predictors for cardiovascular excess mortality and this might be explained due to the small number of cardiovascular deaths. However, does not apply to systematic error and results are very consistent with data in the literature hence we think random error has a minimal impact on our results.

7.2.2 External Validity

It is necessary to point out that these results may not be generalized to the complete HIV-positive population in Spain because CoRIS cohort is limited to patients that are clinically followed in hospitals from some regions of Spain. Patients with unknown HIV-positive status or patients followed in primary health centres or in other regions are not considered (111). Under this situation, probabilistic sample selection cannot be used as there is no delimited HIV-positive population and therefore CoRIS cohort recruitment is performed on a convenience sampling scheme. However, CoRIS mimics extremely well the Spanish Epidemic characteristics by HIV surveillance (178).

8 Conclusions

The conclusions are:

1. We observed a decrease in overall and cause-specific mortality in the cART era. The observed decreases in mortality were largely at the expense of the HCV-negative patients, as declines were not observed among HCV-positive patients. Among HCV-positive patients, we identified a trend not only towards an increasing incidence in liver-related deaths, but also towards an increasing incidence in AIDS-related deaths and non-AIDS, non-liver-related deaths.

2. Patients included in CoRIS over the last decade have shown to have overall, liver, non-AIDS malignancies and cardiovascular excess mortality associated with being HIV-positive. However, important reductions of excess mortality were observed after the first year of follow-up.

3. AIDS before entry and HCV coinfection had a differential effect on short-term and long-term overall excess mortality. Having AIDS before cohort entry was a strong predictor of early excess mortality; nevertheless this effect was diluted during the rest of follow-up. Opposite, HCV coinfection was related with higher long-term excess mortality. Overall excess mortality was further affected by lower CD4 count, higher VL at entry, lower education, being male and over 50 years of age.

4. Liver excess mortality was higher in patients with low CD4 counts at entry and in those HCV coinfected. Patients aged over 50 years-old and HCV coinfected showed higher non-AIDS malignancies excess mortality.

5. Despite the fact that the prevalence of HCV infection in persons newly diagnosed with HIV has decreased sharply in parallel with a decrease in IDU, HCV coinfection is still an important factor determining overall and cause-specific mortality in HIV-positive persons.

6. Missing values are infrequent in CoRIS; however, incompleteness of prognostic variables (i.e. CD4 count) or the outcome (i.e. CoD) is a key issue in cohort-based HIV mortality analyses. Excess mortality estimates from four different methods to handle missing data were not significantly different. However, Complete-Case approach leads to less precise estimations and incorrect classification of CoD (Indicator Method) or deletion of cases with missing CoD (Complete-Case, Multiple Imputation then Deletion) seems to lead to underestimation of the excess mortality rate.

7. Multiple Imputation by Chained Equations is a powerful approach for analysing the association between excess mortality and prognostic factors in the presence of missing data in both the covariates and the outcome. However, it rests on the assumption that incomplete values are Missing At Random, conditional on all other covariates. Although Missing At Random assumption cannot be tested, it seems plausible in this context in view of the powerful predictive covariates included in the imputation model.

9 References

List of References

(1) Global report: UNAIDS report on the global AIDS epidemic 2013. Joint United Nations Programme on HIV/AIDS (UNAIDS) 2013.

(2) Consolidated guidelines on the use of antiretroviral drugs for treating and preventing HIV infection. Recommendations for a public health approach. World Health Organization 2013.

(3) Access to antiretroviral drugs in low- and middle-income countries. Technical report July 2014. World Health Organization 2014.

(4) Hiv and aging. A special supplement to the UNAIDS report on the global AIDS epidemic 2013. Available at: A special supplement to the UNAIDS report on the global AIDS epidemic 2013. Accessed 05/02, 2015.

(5) Costagliola D. Demographics of HIV and aging. Curr Opin HIV AIDS 2014 Jul;9(4):294-301.

(6) Castilla J, de la Fuente L. Trends in the number of human immunodeficiency virus infected persons and AIDS cases in Spain: 1980-1998. Med Clin (Barc) 2000 Jun 17;115(3):85-89.

(7) Dirección General de Salud Pública, Calidad e Innovación. Sistemas Autonómicos de Vigilancia Epidemiológica. Centro Nacional de Epidemiología. VIGILANCIA EPIDEMIOLÓGICA DEL VIH/SIDA EN ESPAÑA ACTUALIZACIÓN 30 de junio de 2014 SISTEMA DE INFORMACIÓN SOBRE NUEVOS DIAGNÓSTICOS DE VIH REGISTRO NACIONAL DE CASOS DE SIDA. Available at: http://www.msssi.gob.es/ciudadanos/enfLesiones/enfTransmisibles/sida/vigilancia/InformeVIHSida_Junio2014.pdf. Accessed 05/02, 2015.

(8) Dirección general de Salud Pública, Calidad e Innovación. Centro Nacional de Epidemiología. Encuesta Hospitalaria de pacientes con VIH/sida. Resultados 2013. Análisis de la evolución 2000-2013. 2014; Available at: http://www.isciii.es/ISCIII/es/contenidos/fd-servicios-cientifico-tecnicos/fd-vigilancias-alertas/fd-enfermedades/fd-sida/Informe_Encuesta_hospitalaria_2013.pdf;. Accessed 05/02, 2015.

(9) VIGILANCIA EPIDEMIOLÓGICA DEL VIH/SIDA EN ESPAÑA. SISTEMA DE INFORMACIÓN SOBRE NUEVOS DIAGNÓSTICOS DE VIH. REGISTRO NACIONAL DE CASOS DE SIDA . 2013; Available at: http://www.isciii.es/ISCIII/es/contenidos/fd-servicios-cientifico-tecnicos/fd-vigilancias-alertas/fd-enfermedades/fd-sida/Informe_VIH_sida_Junio_2013.pdf;. Accessed 05/02, 2015.

(10) Effros RB, Fletcher CV, Gebo K, Halter JB, Hazzard WR, Horne FM, et al. Aging and infectious diseases: workshop on HIV infection and aging: what is known and future research directions. Clin Infect Dis 2008 Aug 15;47(4):542-553.

(11) Balderson BH, Grothaus L, Harrison RG, McCoy K, Mahoney C, Catz S. Chronic illness burden and quality of life in an aging HIV population. AIDS Care 2013;25(4):451-458.

(12) Kirk JB, Goetz MB. Human immunodeficiency virus in an aging population, a complication of success. J Am Geriatr Soc 2009 Nov;57(11):2129-2138.

(13) Blanco JR, Jarrin I, Vallejo M, Berenguer J, Solera C, Rubio R, et al. Definition of advanced age in HIV infection: looking for an age cut-off. AIDS Res Hum Retroviruses 2012 Sep;28(9):1000-1006.

(14) Collaboration of Observational HIV Epidemiological Research Europe (COHERE) Study Group, Sabin CA, Smith CJ, d'Arminio Monforte A, Battegay M, Gabiano C, et al. Response to combination antiretroviral therapy: variation by age. AIDS 2008 Jul 31;22(12):1463-1473.

(15) Grabar S, Kousignian I, Sobel A, Le Bras P, Gasnault J, Enel P, et al. Immunologic and clinical responses to highly active antiretroviral therapy over 50 years of age. Results from the French Hospital Database on HIV. AIDS 2004 Oct 21;18(15):2029-2038.

(16) Antinori A, Coenen T, Costagiola D, Dedes N, Ellefson M, Gatell J, et al. Late presentation of HIV infection: a consensus definition. HIV Med 2011 Jan;12(1):61-64.

(17) Mocroft A, Lundgren JD, Sabin ML, Monforte A, Brockmeyer N, Casabona J, et al. Risk factors and outcomes for late presentation for HIV-positive persons in Europe: results from the Collaboration of Observational HIV Epidemiological Research Europe Study (COHERE). PLoS Med 2013;10(9):e1001510.

(18) Caro-Murillo AM, Gil Luciano A, Navarro Rubio G, Leal Noval M, Blanco Ramos JR, Cohorte de la Red de Investigacion en Sida (CoRIS). HIV infection in different age groups: Potential implications for prevention. CoRIS Cohort, Spain, 2004-2008. Med Clin (Barc) 2010 Apr 24;134(12):521-527.

(19) Egger M, May M, Chene G, Phillips AN, Ledergerber B, Dabis F, et al. Prognosis of HIV-1-infected patients starting highly active antiretroviral therapy: a collaborative analysis of prospective studies. Lancet 2002 Jul 13;360(9327):119-129.

(20) Lanoy E, Mary-Krause M, Tattevin P, Perbost I, Poizot-Martin I, Dupont C, et al. Frequency, determinants and consequences of delayed access to care for HIV infection in France. Antivir Ther 2007;12(1):89-96.

(21) Marks G, Crepaz N, Janssen RS. Estimating sexual transmission of HIV from persons aware and unaware that they are infected with the virus in the USA. AIDS 2006 Jun 26;20(10):1447-1450.

(22) Palella FJ,Jr, Delaney KM, Moorman AC, Loveless MO, Fuhrer J, Satten GA, et al. Declining morbidity and mortality among patients with advanced human immunodeficiency virus infection. HIV Outpatient Study Investigators. N Engl J Med 1998 Mar 26;338(13):853-860.

(23) May MT, Ingle SM. Life expectancy of HIV-positive adults: a review. Sex Health 2011 Dec;8(4):526-533.

(24) Cohen MH, French AL, Benning L, Kovacs A, Anastos K, Young M, et al. Causes of death among women with human immunodeficiency virus infection in the era of combination antiretroviral therapy. Am J Med 2002 Aug 1;113(2):91-98.

(25) Lewden C, Salmon D, Morlat P, Bevilacqua S, Jougla E, Bonnet F, et al. Causes of death among human immunodeficiency virus (HIV)-infected adults in the era of potent antiretroviral therapy: emerging role of hepatitis and cancers, persistent role of AIDS. Int J Epidemiol 2005 Feb;34(1):121-130.

(26) Louie JK, Hsu LC, Osmond DH, Katz MH, Schwarcz SK. Trends in causes of death among persons with acquired immunodeficiency syndrome in the era of highly active antiretroviral therapy, San Francisco, 1994-1998. J Infect Dis 2002 Oct 1;186(7):1023-1027.

(27) Marin B, Thiebaut R, Bucher HC, Rondeau V, Costagliola D, Dorrucci M, et al. Non-AIDS-defining deaths and immunodeficiency in the era of combination antiretroviral therapy. AIDS 2009 Aug 24;23(13):1743-1753.

(28) Palella FJ,Jr, Baker RK, Moorman AC, Chmiel JS, Wood KC, Brooks JT, et al. Mortality in the highly active antiretroviral therapy era: changing causes of death and disease in the HIV outpatient study. J Acquir Immune Defic Syndr 2006 Sep;43(1):27-34.

(29) Bhaskaran K, Hamouda O, Sannes M, Boufassa F, Johnson AM, Lambert PC, et al. Changes in the risk of death after HIV seroconversion compared with mortality in the general population. JAMA 2008 Jul 2;300(1):51-59.

(30) Hernando V, Perez-Cachafeiro S, Lewden C, Gonzalez J, Segura F, Oteo JA, et al. All-cause and liver-related mortality in HIV positive subjects compared to the general population: differences by HCV co-infection. J Hepatol 2012 Oct;57(4):743-751.

(31) Weber R, Ruppik M, Rickenbach M, Spoerri A, Furrer H, Battegay M, et al. Decreasing mortality and changing patterns of causes of death in the Swiss HIV Cohort Study. HIV Med 2013 Apr;14(4):195-207.

(32) Antiretroviral Therapy Cohort Collaboration. Causes of death in HIV-1-infected patients treated with antiretroviral therapy, 1996-2006: collaborative analysis of 13 HIV cohort studies. Clin Infect Dis 2010 May 15;50(10):1387-1396.

(33) Helleberg M, Afzal S, Kronborg G, Larsen CS, Pedersen G, Pedersen C, et al. Mortality attributable to smoking among HIV-1-infected individuals: a nationwide, population-based cohort study. Clin Infect Dis 2013 Mar;56(5):727-734.

(34) Aberg JA. Aging, inflammation, and HIV infection. Top Antivir Med 2012 Aug-Sep;20(3):101-105.

(35) Regidor E, Gutiérrez-Fisac J. Mortality patterns in Spain, 2010. Ministry of Health, Social Services and Equality. 2013.

(36) Antiretroviral Therapy Cohort Collaboration, Zwahlen M, Harris R, May M, Hogg R, Costagliola D, et al. Mortality of HIV-infected patients starting potent antiretroviral therapy:

comparison with the general population in nine industrialized countries. Int J Epidemiol 2009 Dec;38(6):1624-1633.

(37) Collaboration of Observational HIV Epidemiological Research Europe (COHERE) in EuroCoord, Lewden C, Bouteloup V, De Wit S, Sabin C, Mocroft A, et al. All-cause mortality in treated HIV-infected adults with CD4 >/=500/mm3 compared with the general population: evidence from a large European observational cohort collaboration. Int J Epidemiol 2012 Apr;41(2):433-445.

(38) Hernando V, Perez-Cachafeiro S, Lewden C, Gonzalez J, Segura F, Oteo JA, et al. All-cause and liver-related mortality in HIV positive subjects compared to the general population: differences by HCV co-infection. J Hepatol 2012 Oct;57(4):743-751.

(39) Medrano J, Resino S, Vispo E, Madejon A, Labarga P, Tuma P, et al. Hepatitis C virus (HCV) treatment uptake and changes in the prevalence of HCV genotypes in HIV/HCV-coinfected patients. J Viral Hepat 2011 May;18(5):325-330.

(40) Klein MB, Rollet-Kurhajec KC, Moodie EE, Yaphe S, Tyndall M, Walmsley S, et al. Mortality in HIV-hepatitis C co-infected patients in Canada compared to the general Canadian population (2003-2013). AIDS 2014 Aug 24;28(13):1957-1965.

(41) Fleming CA, Craven DE, Thornton D, Tumilty S, Nunes D. Hepatitis C virus and human immunodeficiency virus coinfection in an urban population: low eligibility for interferon treatment. Clin Infect Dis 2003 Jan 1;36(1):97-100.

(42) Alter M. Epidemiology of viral hepatitis and HIV co-infection. J Hepatol 2006;44(1 Suppl):S6-9.

(43) Perez Cachafeiro S, Del Amo J, Iribarren JA, Salavert Lleti M, Gutierrez F, Moreno A, et al. Decrease in serial prevalence of coinfection with hepatitis C virus among HIV-infected patients in Spain, 1997-2006. Clin Infect Dis 2009 May 15;48(10):1467-1470.

(44) Graham CS, Baden LR, Yu E, Mrus JM, Carnie J, Heeren T, et al. Influence of human immunodeficiency virus infection on the course of hepatitis C virus infection: a meta-analysis. Clin Infect Dis 2001 Aug 15;33(4):562-569.

(45) Puoti M, Spinetti A, Ghezzi A, Donato F, Zaltron S, Putzolu V, et al. Mortality for liver disease in patients with HIV infection: a cohort study. J Acquir Immune Defic Syndr 2000 Jul 1;24(3):211-217.

(46) Martin-Carbonero L, Soriano V, Valencia E, Garcia-Samaniego J, Lopez M, Gonzalez-Lahoz J. Increasing impact of chronic viral hepatitis on hospital admissions and mortality among HIV-infected patients. AIDS Res Hum Retroviruses 2001 Nov 1;17(16):1467-1471.

(47) Soriano V, Garcia-Samaniego J, Valencia E, Rodriguez-Rosado R, Munoz F, Gonzalez-Lahoz J. Impact of chronic liver disease due to hepatitis viruses as cause of hospital admission and death in HIV-infected drug users. Eur J Epidemiol 1999 Jan;15(1):1-4.

(48) Bica I, McGovern B, Dhar R, Stone D, McGowan K, Scheib R, et al. Increasing mortality due to end-stage liver disease in patients with human immunodeficiency virus infection. Clin Infect Dis 2001 Feb 1;32(3):492-497.

(49) Soriano V, Garcia-Samaniego J, Valencia E, Rodriguez-Rosado R, Munoz F, Gonzalez-Lahoz J. Impact of chronic liver disease due to hepatitis viruses as cause of hospital admission and death in HIV-infected drug users. Eur J Epidemiol 1999 Jan;15(1):1-4.

(50) Peters L, Mocroft A, Soriano V, Rockstroh JK, Losso M, Valerio L, et al. Hepatitis C virus coinfection does not influence the CD4 cell recovery in HIV-1-infected patients with maximum virologic suppression. J Acquir Immune Defic Syndr 2009 Apr 15;50(5):457-463.

(51) Sulkowski MS, Moore RD, Mehta SH, Chaisson RE, Thomas DL. Hepatitis C and progression of HIV disease. JAMA 2002 Jul 10;288(2):199-206.

(52) Peters L, Mocroft A, Soriano V, Rockstroh JK, Losso M, Valerio L, et al. Hepatitis C virus coinfection does not influence the CD4 cell recovery in HIV-1-infected patients with maximum virologic suppression. J Acquir Immune Defic Syndr 2009 Apr 15;50(5):457-463.

(53) d'Arminio Monforte A, Cozzi-Lepri A, Castagna A, Antinori A, De Luca A, Mussini C, et al. Risk of developing specific AIDS-defining illnesses in patients coinfected with HIV and hepatitis C virus with or without liver cirrhosis. Clin Infect Dis 2009 Aug 15;49(4):612-622.

(54) Greub G, Ledergerber B, Battegay M, Grob P, Perrin L, Furrer H, et al. Clinical progression, survival, and immune recovery during antiretroviral therapy in patients with HIV-1 and hepatitis C virus coinfection: the Swiss HIV Cohort Study. Lancet 2000 Nov 25;356(9244):1800-1805.

(55) Potter M, Odueyungbo A, Yang H, Saeed S, Klein MB, Canadian Co-infection Cohort Study Investigators. Impact of hepatitis C viral replication on CD4+ T-lymphocyte progression in HIV-HCV coinfection before and after antiretroviral therapy. AIDS 2010 Jul 31;24(12):1857-1865.

(56) Kovacs A, Al-Harthi L, Christensen S, Mack W, Cohen M, Landay A. CD8(+) T cell activation in women coinfected with human immunodeficiency virus type 1 and hepatitis C virus. J Infect Dis 2008 May 15;197(10):1402-1407.

(57) de Castro IF, Micheloud D, Berenguer J, Guzman-Fulgencio M, Catalan P, Miralles P, et al. Hepatitis C virus infection is associated with endothelial dysfunction in HIV/hepatitis C virus coinfected patients. AIDS 2010 Aug 24;24(13):2059-2067.

(58) de Oca Arjona MM, Marquez M, Soto MJ, Rodriguez-Ramos C, Terron A, Vergara A, et al. Bacterial translocation in HIV-infected patients with HCV cirrhosis: implication in hemodynamic alterations and mortality. J Acquir Immune Defic Syndr 2011 Apr 15;56(5):420-427.

(59) Propst-Graham KL, Preheim LC, Vander Top EA, Snitily MU, Gentry-Nielsen MJ. Cirrhosis-induced defects in innate pulmonary defenses against Streptococcus pneumoniae. BMC Microbiol 2007 Oct 23;7:94.

(60) Berenguer J, Rodriguez E, Miralles P, Von Wichmann MA, Lopez-Aldeguer J, Mallolas J, et al. Sustained virological response to interferon plus ribavirin reduces non-liver-related mortality in patients coinfected with HIV and Hepatitis C virus. Clin Infect Dis 2012 Sep;55(5):728-736.

(61) Carrat F, Bani-Sadr F, Pol S, Rosenthal E, Lunel-Fabiani F, Benzekri A, et al. Pegylated interferon alfa-2b vs standard interferon alfa-2b, plus ribavirin, for chronic hepatitis C in HIV-infected patients: a randomized controlled trial. JAMA 2004 Dec 15;292(23):2839-2848.

(62) Torriani FJ, Rodriguez-Torres M, Rockstroh JK, Lissen E, Gonzalez-Garcia J, Lazzarin A, et al. Peginterferon Alfa-2a plus ribavirin for chronic hepatitis C virus infection in HIV-infected patients. N Engl J Med 2004 Jul 29;351(5):438-450.

(63) Berenguer J, Alvarez-Pellicer J, Martin PM, Lopez-Aldeguer J, Von-Wichmann MA, Quereda C, et al. Sustained virological response to interferon plus ribavirin reduces liver-related complications and mortality in patients coinfected with human immunodeficiency virus and hepatitis C virus. Hepatology 2009 Aug;50(2):407-413.

(64) HCV working group of the Collaboration of Observational HIV Epidemiological Research in Europe (COHERE) in EuroCoord. Effect of hepatitis C treatment on CD4+ T-cell counts and the risk of death in HIV-HCV-coinfected patients: the COHERE collaboration. Antivir Ther 2012;17(8):1541-1550.

(65) Lewden C, Raffi F, Chene G, Sobel A, Leport C, APROCO Study Group. Mortality in a cohort of HIV-infected adults started on a protease inhibitor-containing therapy: standardization to the general population. J Acquir Immune Defic Syndr 2001 Apr 15;26(5):480-482.

(66) Backus LI, Boothroyd D, Deyton LR. HIV, hepatitis C and HIV/hepatitis C virus co-infection in vulnerable populations. AIDS 2005 Oct;19 Suppl 3:S13-9.

(67) Chen TY, Ding EL, Seage Iii GR, Kim AY. Meta-analysis: increased mortality associated with hepatitis C in HIV-infected persons is unrelated to HIV disease progression. Clin Infect Dis 2009 Nov 15;49(10):1605-1615.

(68) van der Helm J, Geskus R, Sabin C, Meyer L, Del Amo J, Chene G, et al. Effect of HCV infection on cause-specific mortality after HIV seroconversion, before and after 1997. Gastroenterology 2013 Apr;144(4):751-760.e2.

(69) Hernando V, Sobrino-Vegas P, Burriel MC, Berenguer J, Navarro G, Santos I, et al. Differences in the causes of death of HIV-positive patients in a cohort study by data sources and coding algorithms. AIDS 2012 Sep 10;26(14):1829-1834.

(70) World Health Organization. **International Classification of Diseases (ICD)**. 2013; . Accessed 08/22, 2013.

(71) *Copenhagen HIV Program. Protocol Coding Causes of Death in HIV (CoDe).* . Available at: http://www.cphiv.dk/Portals/_default/pdf_folder/code_protocol_ver_1.0.pdf.

(72) Porta M. **Dictionary of Epidemiology**. 5th ed. USA: OUP; 2008.

(73) Beaglehole R, Bonita R, Kjellstrom T. Basic epidemiology. 2nd ed. Geneva: WHO Library Cataloguing-in-Publication Data; 1993.

(74) kirkwood B, Sterne J. Essential Medical Statistics. 2nd ed.: Blacwell Science Ltd; 1988.

(75) Jarrin I, Bolumar F, del Amo J. Cohort studies and their contribution to the study of HIV infection: main characteristics and limitations. Enferm Infecc Microbiol Clin 2010 May;28(5):304-309.

(76) Spanish Network of HIV Research . 2013; Available at: http://retic-ris.net/Principal.aspx#&panel1-1. Accessed 08/26, 2013.

(77) Garcia-Merino I, de Las Cuevas N, Jimenez JL, Gallego J, Gomez C, Prieto C, et al. The Spanish HIV BioBank: a model of cooperative HIV research. Retrovirology 2009 Mar 9;6:27-4690-6-27.

(78) Jarrin I, Geskus R, Perez-Hoyos S, del Amo J. Analytical methods in cohort studies of patients with HIV infection. Enferm Infecc Microbiol Clin 2010 May;28(5):298-303.

(79) Dickman PW, Sloggett A, Hills M, Hakulinen T. Regression models for relative survival. Stat Med 2004 Jan 15;23(1):51-64.

(80) EDERER F, AXTELL LM, CUTLER SJ. The relative survival rate: a statistical methodology. Natl Cancer Inst Monogr 1961 Sep;6:101-121.

(81) ICH Harmonised Tripartite Guideline. Statistical principles for clinical trials. International Conference on Harmonisation E9 Expert Working Group. Stat Med 1999 Aug 15;18(15):1905-1942.

(82) von Elm E, Altman DG, Egger M, Pocock SJ, Gotzsche PC, Vandenbroucke JP, et al. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. J Clin Epidemiol 2008 Apr;61(4):344-349.

(83) Karahalios A, Baglietto L, Carlin JB, English DR, Simpson JA. A review of the reporting and handling of missing data in cohort studies with repeated assessment of exposure measures. BMC Med Res Methodol 2012 Jul 11;12:96-2288-12-96.

(84) Carpenter J, Kenward M. *Missing Data in Randomised Controlled Trials: A Practical Guide*
. Birmingham, AL: National Institute for Health Research 2008.

(85) Allison P. Missing Data. Statistical Horizons .

(86) Little R, Rubin D. **Statistical Analysis with Missing Data (2nd edn)**. 2nd ed. New York: Wiley; 1987.

(87) Carpenter J, Bartlett J, Kenward M. **www.missingdata.org.uk**. Available at: www.missingdata.org.uk. Accessed 10/08, 2013.

(88) Little RJ. Modeling the dropout mechanism in repeated-measures studies. Jour Amer Stat Ass 1995;90:1112-1121.

(89) White IR, Carlin JB. Bias and efficiency of multiple imputation compared with complete-case analysis for missing covariate values. Stat Med 2010 Dec 10;29(28):2920-2931.

(90) Nur U, Shack LG, Rachet B, Carpenter JR, Coleman MP. Modelling relative survival in the presence of incomplete data: a tutorial. Int J Epidemiol 2010 Feb;39(1):118-128.

(91) Jones MP. Indicator and stratification methods for missing explanatory variables in multiple linear regression. Journal of the American Statistical Association 1996;91:222-230.

(92) Molenberghs G, Kenward M. **Missing Data in Clinical Studies**. : Wiley; 2007.

(93) White IR, Carlin JB. Bias and efficiency of multiple imputation compared with complete-case analysis for missing covariate values. Stat Med 2010 Dec 10;29(28):2920-2931.

(94) Little RJ, Wang Y. Pattern-mixture models for multivariate incomplete data with covariates. Biometrics 1996 Mar;52(1):98-111.

(95) Greenland S, Finkle WD. A critical look at methods for handling missing covariates in epidemiologic regression analyses. Am J Epidemiol 1995 Dec 15;142(12):1255-1264.

(96) Rubin D. Multiple Imputation for nonresponse in surveys. New York: Wiley; 1987.

(97) Allison P. D. editor. Missing Data. : SAGE publications; 2002.

(98) Baraldi AN, Enders CK. An introduction to modern missing data analyses. J Sch Psychol 2010 Feb;48(1):5-37.

(99) van Buuren S, Boshuizen HC, Knook DL. Multiple imputation of missing blood pressure covariates in survival analysis. Stat Med 1999 Mar 30;18(6):681-694.

(100) White IR, Royston P, Wood AM. Multiple imputation using chained equations: Issues and guidance for practice. Stat Med 2011 Feb 20;30(4):377-399.

(101) Raghunathan T, Solenberger P, Hoewyk J. A Multivariate Technique for Multiply Imputing Missing Values Using a Sequence of Regression Models . Survey Methodology 2001;27(1):85-95.

(102) Collins LM, Schafer JL, Kam CM. A comparison of inclusive and restrictive strategies in modern missing data procedures. Psychol Methods 2001 Dec;6(4):330-351.

(103) Kenward MG, Carpenter J. Multiple imputation: current perspectives. Stat Methods Med Res 2007 Jun;16(3):199-218.

(104) Azur MJ, Stuart EA, Frangakis C, Leaf PJ. Multiple imputation by chained equations: what is it and how does it work? Int J Methods Psychiatr Res 2011 Mar;20(1):40-49.

(105) White IR, Royston P, Wood AM. Multiple imputation using chained equations: Issues and guidance for practice. Stat Med 2011 Feb 20;30(4):377-399.

(106) Vol Hippel P. Regression with missing Ys: an improved strategy for analyzing multiply imputed data. Sociological Methodology 2007;37(1):83-117.

(107) Graham JW, Olchowski AE, Gilreath TD. How many imputations are really needed? Some practical clarifications of multiple imputation theory. Prev Sci 2007 Sep;8(3):206-213.

(108) Perez Cachafeiro S, Del Amo J, Iribarren JA, Salavert Lleti M, Gutierrez F, Moreno A, et al. Decrease in serial prevalence of coinfection with hepatitis C virus among HIV-infected patients in Spain, 1997-2006. Clin Infect Dis 2009 May 15;48(10):1467-1470.

(109) del Amo J, Perez-Hoyos S, Moreno A, Quintana M, Ruiz I, Cisneros JM, et al. Trends in AIDS and mortality in HIV-infected subjects with hemophilia from 1985 to 2003: the competing risks for death between AIDS and liver disease. J Acquir Immune Defic Syndr 2006 Apr 15;41(5):624-631.

(110) Rodriguez-Arenas MA, Jarrin I, del Amo J, Iribarren JA, Moreno S, Viciana P, et al. Delay in the initiation of HAART, poorer virological response, and higher mortality among HIV-infected injecting drug users in Spain. AIDS Res Hum Retroviruses 2006 Aug;22(8):715-723.

(111) Sobrino-Vegas P, Gutierrez F, Berenguer J, Labarga P, Garcia F, Alejos-Ferreras B, et al. The Cohort of the Spanish HIV Research Network (CoRIS) and its associated biobank; organizational issues, main findings and losses to follow-up. Enferm Infecc Microbiol Clin 2011 Nov;29(9):645-653.

(112) Garcia-Merino I, de Las Cuevas N, Jimenez JL, Gallego J, Gomez C, Prieto C, et al. The Spanish HIV BioBank: a model of cooperative HIV research. Retrovirology 2009 Mar 9;6:27-4690-6-27.

(113) Blanco JR, Jarrin I, Vallejo M, Berenguer J, Solera C, Rubio R, et al. Definition of advanced age in HIV infection: looking for an age cut-off. AIDS Res Hum Retroviruses 2012 Sep;28(9):1000-1006.

(114) Breslow NE, Day NE. Statistical methods in cancer research. Volume II--The design and analysis of cohort studies. IARC Sci Publ 1987;(82)(82):1-406.

(115) Berry G. The analysis of mortality by the subject-years method. Biometrics 1983 Mar;39(1):173-184.

(116) Omland LH, Jepsen P, Krarup H, Christensen PB, Weis N, Nielsen L, et al. Liver cancer and non-Hodgkin lymphoma in hepatitis C virus-infected patients: results from the DANVIR cohort study. Int J Cancer 2012 May 15;130(10):2310-2317.

(117) Putter H, Fiocco M, Geskus RB. Tutorial in biostatistics: competing risks and multi-state models. Stat Med 2007 May 20;26(11):2389-2430.

(118) Dickman PW. **Estimating and modelling relative survival in SAS and Stata**. Available at: http://www.pauldickman.com/rsmodel/. Accessed 20/08, 2013.

(119) Royston P, White IR. Multiple Imputation by Chained Equations (MICE): Implementation in Stata. Journal of Statistical Software 2011;45(4).

(120) Crum NF, Riffenburgh RH, Wegner S, Agan BK, Tasker SA, Spooner KM, et al. Comparisons of causes of death and mortality rates among HIV-infected persons: analysis of the pre-, early, and late HAART (highly active antiretroviral therapy) eras. J Acquir Immune Defic Syndr 2006 Feb 1;41(2):194-200.

(121) Fleming CA, Tumilty S, Murray JE, Nunes D. Challenges in the treatment of patients coinfected with HIV and hepatitis C virus: need for team care. Clin Infect Dis 2005 Apr 15;40 Suppl 5:S349-54.

(122) Gonzalez-Garcia JJ, Mahillo B, Hernandez S, Pacheco R, Diz S, Garcia P, et al. Prevalences of hepatitis virus coinfection and indications for chronic hepatitis C virus treatment and liver transplantation in Spanish HIV-infected patients. The GESIDA 29/02 and FIPSE 12185/01 Multicenter Study. Enferm Infecc Microbiol Clin 2005 Jun-Jul;23(6):340-348.

(123) Berenguer J, Gonzalez-Garcia J, Lopez-Aldeguer J, Von-Wichmann MA, Quereda C, Hernando A, et al. Pegylated interferon {alpha}2a plus ribavirin versus pegylated interferon {alpha}2b plus ribavirin for the treatment of chronic hepatitis C in HIV-infected patients. J Antimicrob Chemother 2009 Jun;63(6):1256-1263.

(124) Palella FJ,Jr, Baker RK, Moorman AC, Chmiel JS, Wood KC, Brooks JT, et al. Mortality in the highly active antiretroviral therapy era: changing causes of death and disease in the HIV outpatient study. J Acquir Immune Defic Syndr 2006 Sep;43(1):27-34.

(125) Rosenthal E, Salmon-Ceron D, Lewden C, Bouteloup V, Pialoux G, Bonnet F, et al. Liver-related deaths in HIV-infected patients between 1995 and 2005 in the French GERMIVIC Joint Study Group Network (Mortavic 2005 study in collaboration with the Mortalite 2005 survey, ANRS EN19). HIV Med 2009 May;10(5):282-289.

(126) Martinez E, Milinkovic A, Buira E, de Lazzari E, Leon A, Larrousse M, et al. Incidence and causes of death in HIV-infected persons receiving highly active antiretroviral therapy compared with estimates for the general population of similar age and from the same geographical area. HIV Med 2007 May;8(4):251-258.

(127) Bonnet F, Morlat P, Chene G, Mercie P, Neau D, Chossat I, et al. Causes of death among HIV-infected patients in the era of highly active antiretroviral therapy, Bordeaux, France, 1998-1999. HIV Med 2002 Jul;3(3):195-199.

(128) Celentano DD, Galai N, Sethi AK, Shah NG, Strathdee SA, Vlahov D, et al. Time to initiating highly active antiretroviral therapy among HIV-infected injection drug users. AIDS 2001 Sep 7;15(13):1707-1715.

(129) van Asten LC, Boufassa F, Schiffer V, Brettle RP, Robertson JR, Hernandez Aguado I, et al. Limited effect of highly active antiretroviral therapy among HIV-positive injecting drug users on the population level. Eur J Public Health 2003 Dec;13(4):347-349.

(130) Daar ES, Lynn H, Donfield S, Gomperts E, O'Brien SJ, Hilgartner MW, et al. Hepatitis C virus load is associated with human immunodeficiency virus type 1 disease progression in hemophiliacs. J Infect Dis 2001 Feb 15;183(4):589-595.

(131) Kovacs A, Karim R, Mack WJ, Xu J, Chen Z, Operskalski E, et al. Activation of CD8 T cells predicts progression of HIV infection in women coinfected with hepatitis C virus. J Infect Dis 2010 Mar 15;201(6):823-834.

(132) Prins M, Hernandez Aguado IH, Brettle RP, Robertson JR, Broers B, Carre N, et al. Pre-AIDS mortality from natural causes associated with HIV disease progression: evidence from the European Seroconverter Study among injecting drug users. AIDS 1997 Nov 15;11(14):1747-1756.

(133) Tyndall MW, Craib KJ, Currie S, Li K, O'Shaughnessy MV, Schechter MT. Impact of HIV infection on mortality in a cohort of injection drug users. J Acquir Immune Defic Syndr 2001 Dec 1;28(4):351-357.

(134) Copeland L, Budd J, Robertson JR, Elton RA. Changing patterns in causes of death in a cohort of injecting drug users, 1980-2001. Arch Intern Med 2004 Jun 14;164(11):1214-1220.

(135) Kohli R, Lo Y, Howard AA, Buono D, Floris-Moore M, Klein RS, et al. Mortality in an urban cohort of HIV-infected and at-risk drug users in the era of highly active antiretroviral therapy. Clin Infect Dis 2005 Sep 15;41(6):864-872.

(136) Deeks SG. HIV infection, inflammation, immunosenescence, and aging. Annu Rev Med 2011;62:141-155.

(137) Antiretroviral Therapy Cohort Collaboration, Zwahlen M, Harris R, May M, Hogg R, Costagliola D, et al. Mortality of HIV-infected patients starting potent antiretroviral therapy: comparison with the general population in nine industrialized countries. Int J Epidemiol 2009 Dec;38(6):1624-1633.

(138) Aldaz P, Moreno-Iribas C, Egues N, Irisarri F, Floristan Y, Sola-Boneta J, et al. Mortality by causes in HIV-infected adults: comparison with the general population. BMC Public Health 2011 May 11;11:300-2458-11-300.

(139) Sobrino-Vegas P, Gutierrez F, Berenguer J, Labarga P, Garcia F, Alejos B, et al. **La cohorte de la red española de investigación en sida y su biobanco: organización, principales resultados y perdidas al seguimiento**. Enfermedades infecciosas y microbiologia clinica .

(140) Gutierrez F, Padilla S, Masia M, Iribarren JA, Moreno S, Viciana P, et al. Clinical outcome of HIV-infected patients with sustained virologic response to antiretroviral therapy: long-term follow-up of a multicenter cohort. PLoS One 2006 Dec 20;1:e89.

(141) HIV-CAUSAL Collaboration, Ray M, Logan R, Sterne JA, Hernandez-Diaz S, Robins JM, et al. The effect of combined antiretroviral therapy on the overall mortality of HIV-infected individuals. AIDS 2010 Jan 2;24(1):123-137.

(142) Collaboration of Observational HIV Epidemiological Research Europe (COHERE) in EuroCoord, Lewden C, Bouteloup V, De Wit S, Sabin C, Mocroft A, et al. All-cause mortality in treated HIV-infected adults with CD4 >/=500/mm3 compared with the general population: evidence from a large European observational cohort collaboration. Int J Epidemiol 2012 Apr;41(2):433-445.

(143) Smit C, Geskus R, Walker S, Sabin C, Coutinho R, Porter K, et al. Effective therapy has altered the spectrum of cause-specific mortality following HIV seroconversion. AIDS 2006 Mar 21;20(5):741-749.

(144) Sobrino-Vegas P, Garcia-San Miguel L, Caro-Murillo AM, Miro JM, Viciana P, Tural C, et al. Delayed diagnosis of HIV infection in a multicenter cohort: prevalence, risk factors, response to HAART and impact on mortality. Curr HIV Res 2009 Mar;7(2):224-230.

(145) May M, Sterne JA, Sabin C, Costagliola D, Justice AC, Thiebaut R, et al. Prognosis of HIV-1-infected patients up to 5 years after initiation of HAART: collaborative analysis of prospective studies. AIDS 2007 May 31;21(9):1185-1197.

(146) Patel P, Hanson DL, Sullivan PS, Novak RM, Moorman AC, Tong TC, et al. Incidence of types of cancer among HIV-infected persons compared with the general population in the United States, 1992-2003. Ann Intern Med 2008 May 20;148(10):728-736.

(147) Albini L, Calabresi A, Gotti D, Ferraresi A, Festa A, Donato F, et al. Burden of Non-AIDS-Defining and Non-Virus-Related Cancers Among HIV-Infected Patients in the Combined Antiretroviral Therapy Era. AIDS Res Hum Retroviruses 2013 Aug;29(8):1097-1104.

(148) Frisch M, Biggar RJ, Engels EA, Goedert JJ, AIDS-Cancer Match Registry Study Group. Association of cancer with AIDS-related immunosuppression in adults. JAMA 2001 Apr 4;285(13):1736-1745.

(149) Shiels MS, Cole SR, Kirk GD, Poole C. A meta-analysis of the incidence of non-AIDS cancers in HIV-infected individuals. J Acquir Immune Defic Syndr 2009 Dec;52(5):611-622.

(150) Silverberg MJ, Abrams DI. AIDS-defining and non-AIDS-defining malignancies: cancer occurrence in the antiretroviral therapy era. Curr Opin Oncol 2007 Sep;19(5):446-451.

(151) Patel P, Hanson DL, Sullivan PS, Novak RM, Moorman AC, Tong TC, et al. Incidence of types of cancer among HIV-infected persons compared with the general population in the United States, 1992-2003. Ann Intern Med 2008 May 20;148(10):728-736.

(152) Rodriguez-Arenas M, Gutierrez-Trujillo L, Robledano C, Rodríguez-Fortunez P, del Romero J, Rodriguez-Fernandez J, et al. Prevalencia de consumo de drogas y adherencia a TARc, por sexo, en la cohorte hospitalario CoRIS de personas con infección por VIH. XIV Congreso Nacional sobre el SIDA, June Zaragoza (Spain) 2011.

(153) Clifford GM, Polesel J, Rickenbach M, Dal Maso L, Keiser O, Kofler A, et al. Cancer risk in the Swiss HIV Cohort Study: associations with immunodeficiency, smoking, and highly active antiretroviral therapy. J Natl Cancer Inst 2005 Mar 16;97(6):425-432.

(154) Gonzalez C, Torres M, Benito A, Del Romero J, Rodriguez C, Fontillon M, et al. Anal squamous intraepithelial lesions are frequent among young HIV-infected men who have sex with men followed up at the Spanish AIDS Research Network Cohort (CoRIS-HPV). Int J Cancer 2013 Sep 1;133(5):1164-1172.

(155) Gordon SC, Moonka D, Brown KA, Rogers C, Huang MA, Bhatt N, et al. Risk for renal cell carcinoma in chronic hepatitis C infection. Cancer Epidemiol Biomarkers Prev 2010 Apr;19(4):1066-1073.

(156) Lee MH, Yang HI, Lu SN, Jen CL, You SL, Wang LY, et al. Chronic hepatitis C virus infection increases mortality from hepatic and extrahepatic diseases: a community-based long-term prospective study. J Infect Dis 2012 Aug 15;206(4):469-477.

(157) Allison RD, Tong X, Moorman AC, Ly KN, Rupp L, Xu F, et al. Increased incidence of cancer and cancer-related mortality among persons with chronic hepatitis C infection, 2006-2010. J Hepatol 2015 May 1.

(158) Burki TK. Hepatitis C virus associated with high rates of cancer. Lancet Oncol 2015 Jun;16(6):e266-2045(15)70207-X. Epub 2015 Apr 30.

(159) Moradpour D, Blum HE. Pathogenesis of hepatocellular carcinoma. Eur J Gastroenterol Hepatol 2005 May;17(5):477-483.

(160) Graham CS, Baden LR, Yu E, Mrus JM, Carnie J, Heeren T, et al. Influence of human immunodeficiency virus infection on the course of hepatitis C virus infection: a meta-analysis. Clin Infect Dis 2001 Aug 15;33(4):562-569.

(161) Rockstroh JK, Mohr R, Behrens G, Spengler U. Liver fibrosis in HIV: which role does HIV itself, long-term drug toxicities and metabolic changes play? Curr Opin HIV AIDS 2014 Jul;9(4):365-370.

(162) Akhtar MA, Mathieson K, Arey B, Post J, Prevette R, Hillier A, et al. Hepatic histopathology and clinical characteristics associated with antiretroviral therapy in HIV patients without viral hepatitis. Eur J Gastroenterol Hepatol 2008 Dec;20(12):1194-1204.

(163) Triant VA, Lee H, Hadigan C, Grinspoon SK. Increased acute myocardial infarction rates and cardiovascular risk factors among patients with human immunodeficiency virus disease. J Clin Endocrinol Metab 2007 Jul;92(7):2506-2512.

(164) Klein D, Hurley LB, Quesenberry CP,Jr, Sidney S. Do protease inhibitors increase the risk for coronary heart disease in patients with HIV-1 infection? J Acquir Immune Defic Syndr 2002 Aug 15;30(5):471-477.

(165) Durand M, Sheehy O, Baril JG, Lelorier J, Tremblay CL. Association between HIV infection, antiretroviral therapy, and risk of acute myocardial infarction: a cohort and nested case-

control study using Quebec's public health insurance database. J Acquir Immune Defic Syndr 2011 Jul 1;57(3):245-253.

(166) Hemkens LG. BH. **HIV infection and cardiovascular disease**. Eur Heart J 2014 1 Jan;35(21):1373.

(167) Serrano-Villar S, Sobrino-Vegas P, Monge S, Dronda F, Hernando A, Montero M, et al. Decreasing prevalence of HCV coinfection in all risk groups for HIV infection between 2004 and 2011 in Spain. J Viral Hepat 2015 May;22(5):496-503.

(168) Monge S, Alejos B, Dronda F, Del Romero J, Iribarren JA, Pulido F, et al. Inequalities in HIV disease management and progression in migrants from Latin America and sub-Saharan Africa living in Spain. HIV Med 2013 May;14(5):273-283.

(169) Pablos-Mendez A. Mortality among Hispanics. JAMA 1994 Apr 27;271(16):1237.

(170) Spratt M, Carpenter J, Sterne JA, Carlin JB, Heron J, Henderson J, et al. Strategies for multiple imputation in longitudinal studies. Am J Epidemiol 2010 Aug 15;172(4):478-487.

(171) van der Heijden GJ, Donders AR, Stijnen T, Moons KG. Imputation of missing values is superior to complete case analysis and the missing-indicator method in multivariable diagnostic research: a clinical example. J Clin Epidemiol 2006 Oct;59(10):1102-1109.

(172) Collins LM, Schafer JL, Kam CM. A comparison of inclusive and restrictive strategies in modern missing data procedures. Psychol Methods 2001 Dec;6(4):330-351.

(173) Carpenter J, Goldstein H., Kenward M. **REALCOM-IMPUTE Software for Multilevel Multiple Imputation with Mixed Response Types**. Journal of Statistical Software 2011;45(5).

(174) Carpenter J, Kenward M. Brief comments on computational issues with multiple imputation. 2008; Available at: http://missingdata.lshtm.ac.uk/downloads/mi_comp_issues.pdf. Accessed 03/2015, 2015.

(175) Sterne JA, White IR, Carlin JB, Spratt M, Royston P, Kenward MG, et al. Multiple imputation for missing data in epidemiological and clinical research: potential and pitfalls. BMJ 2009 Jun 29;338:b2393.

(176) Bakoyannis G, Siannis F, Touloumi G. Modelling competing risks data with missing cause of failure. Stat Med 2010 Dec 30;29(30):3172-3185.

(177) Eloranta S, Lambert PC, Andersson TM, Czene K, Hall P, Bjorkholm M, et al. Partitioning of excess mortality in population-based cancer patient survival studies using flexible parametric survival models. BMC Med Res Methodol 2012 Jun 24;12:86-2288-12-86.

(178) VIGILANCIA EPIDEMIOLÓGICA DEL VIH/SIDA EN ESPAÑA. SISTEMA DE INFORMACIÓN SOBRE NUEVOS DIAGNÓSTICOS DE VIH REGISTRO NACIONAL DE CASOS DE SIDA. ACTUALIZACIÓN 30 de junio de 2013.

10 Annexes

10.1 Annex 1: Scientific communications related to this doctoral thesis

10.1.1 Presentations at scientific meeting

J. Berenguer, **B. Alejos,** V. Hernando, P. Viciana, J.A. Oteo, J.L. Gómez Sirvent, D. Dalmau, J. Portilla, S. Moreno, J. Del Amo, and CoRIS. *Temporal Trends in Liver-Related Mortality in a Prospective Cohort of HIV-Infected cART Naïve Adults in Spain (CoRIS); 1997-2008". 6th IAS Conference on HIV Pathogenesis, Treatment and Prevention.* Rome, Italy 17-20 July 2011. (Oral Communication)

J. Berenguer**, B. Alejos,** V. Hernando, P. Viciana, J.A. Oteo, J.L. Gómez Sirvent, D. Dalmau, J. Portilla, S. Moreno, J. Del Amo, and CoRIS. *Evolución temporal de la mortalidad por causa hepática en una cohorte de pacientes VIH positive en España, 1997-2008.* III Congreso de Gesida, Sevilla 2011. (Oral Communication)

B. Alejos, V. Hernando, J. López-Aldeguer, F. Segura, J. Antonio Oteo, R. Rubio, A. Sanvisens, P. Sobrino, J. del Amo, CoRIS-cohort. *Overall and Cause-Specific Mortality in HIV-positive Subjects Compared to the General Population.* HIV Therapy Glasgow 2014. (Poster)

B. Alejos, V. Hernando, J. López-Aldeguer, F. Segura, J. Antonio Oteo, R. Rubio, A. Sanvisens, P. Sobrino, J. del Amo, CoRIS-cohort. *Overall and Cause-Specific Mortality in HIV-positive Subjects Compared to the General Population*. VI Congreso Nacional de Gesida. y comunicación oral en Reunión Docente de la Red de Investigación en Sida. Malaga 2014. (Oral poster)

B. Alejos, S. Monge, J. Antonio Iribarren, M. Riera, C. Rodriguez, F. Pulido, Jesus Reparaz, J. Portilla, J. del Amo, P. Sobrino, CoRIS. *Rationale to Use Multiple Imputation by Chained Equations (MICE) to Deal with Incomplete Data in a Cohort of HIV-positive Patients in Spain*. 18th International Workshop on HIV Observational Databases. Sitges 2014. (Poster)

B. Alejos, J. del Amo, P. Sobrino, I. Jarrín on behalf of CoRIS. *Different approaches to account for missing Cause of Death in a cohort of HIV-positive patients*

to estimate mortality excess due to Non-AIDS defining Malignancies. 19th International Workshop on HIV Observational Databases. Catany 2015. (Poster)

10.1.2 Scientific publications

Berenguer J, **Alejos B**, Hernando V, Viciana P, Salavert M, Santos I, Gómez-Sirvent JL, Vidal F, Portilla J, Del Amo J; CoRIS (AIDS Research Network Cohort). *Trends in mortality according to hepatitis C virus serostatus in the era of combination antiretroviral therapy*. AIDS. 2012 Nov 13;26(17):2241-2246.

Hernando Sebastián V, **Alejos B**, et al. *All-cause mortality in the cohorts of the Spanish AIDS Research Network (RIS) compared with the general population: 1997-2010*. BMC Infect Dis. 2013 Aug 20;13(1):382.

10.2 Annex 2: Articles published

CONCISE COMMUNICATION

Trends in mortality according to hepatitis C virus serostatus in the era of combination antiretroviral therapy

Juan Berenguer[a,b], Belén Alejos[c], Victoria Hernando[c], Pompeyo Viciana[d], Miguel Salavert[e], Ignacio Santos[f], Juan L. Gómez-Sirvent[g], Francesc Vidal[h], Joaquín Portilla[i], Julia Del Amo[c], CoRIS (AIDS Research Network Cohort)

Objective: To study trends in overall deaths and cause-specific deaths stratified by hepatitis C virus (HCV) serostatus in a cohort of combination antiretroviral (cART)-naive HIV-infected patients in Spain.

Methods: We analyzed data from 1997 to 2008 in two calendar periods: 1997–2003 and 2004–2008. Deaths were ascertained through cohort reporting and a cross-match with the Spanish National Death Index. We used Poisson regression to model mortality rates and risk factors.

Results: We analyzed 5974 HIV-positive cART-naive patients: 2471 (1497 HCV positive) in the period 1997–2003, and 3503 (689 HCV positive) in the period 2004–2008. A total of 232 deaths (158 during the first period, and 74 during the second period) were detected during 19 416 person-years of follow-up; the death rate was 12.9 of 1000 person-years. Crude overall death rates [95% confidence interval (CI)] were 16.5 (14.2–19.1) in 1997–2003 and 8.5 (6.7–10.6) in 2004–2008. The incidence rate ratio (IRR) (95%CI) in 2004–2008 taking 1997–2003 as a reference was 0.51 (0.39–0.67). When we stratified by HCV serostatus, the overall death IRR (95% CI) taking 1997–2003 as reference was 0.52 (0.32–0.85) for HCV-negative patients and 1.27 (0.90–1.79) for HCV-positive patients. When we considered cause-specific deaths (liver-related, AIDS-related, and nonliver-related/non-AIDS-related), findings were similar to those for overall-deaths.

Conclusion: Taking the first years of the cART era as a reference, we observed a decrease in overall and cause-specific mortality. This decrease was only observed in HCV-negative patients.

AIDS 2012, **26**:2241–2246

Keywords: AIDS-related opportunistic infections, antiretroviral therapy, cause of death, cohort studies, hepatitis C, highly active, HIV, prospective studies

[a]Hospital General Universitario Gregorio Marañón, [b]Instituto de Investigación Sanitaria Gregorio Marañón, [c]National Center of Epidemiology, Instituto de Salud Carlos III, Madrid, [d]Hospital Universitario Virgen del Rocío, Sevilla, [e]Hospital Universitario La Fe, Valencia, [f]Hospital Universitario de la Princesa, Madrid, [g]Hospital Universitario de Canarias, Santa Cruz de Tenerife, [h]Hospital Universitari Joan XXIII, IISPV, Universitat Rovira i Virgili, Tarragona, and [i]Hospital General Universitario de Alicante, Alicante, Spain.

Correspondence to Juan Berenguer, MD, PhD, Infectious Diseases and HIV Unit, Hospital General Universitario Gregorio Marañón, Doctor Esquerdo, 46, Madrid 28007, Spain.

Tel: +34 91 586 8592; fax: +34 91 426 5177; e-mail: jbb4@me.com

Received: 22 March 2012; revised: 17 June 2012; accepted: 22 June 2012.

DOI:10.1097/QAD.0b013e3283574e94

2242 AIDS 2012, Vol 26 No 17

Introduction

The introduction of combination antiretroviral therapy (cART) for HIV infection was followed by a sharp decline in the frequency of AIDS-related complications and increased survival [1–4]. Coinfection by hepatitis C virus (HCV) is by far one of the most common comorbidities in HIV-infected patients [5]. HIV infection modifies the natural history of chronic hepatitis C, promoting more rapid progression to fibrosis, cirrhosis and end-stage liver disease [6], and liver-related mortality [7–10]. Data on the effect of HCV on the progression of HIV infection are inconsistent [11,12].

The factors affecting mortality in the HIV–HCV-coinfected population over the last 10 years include the use of more efficacious and safer cART regimens for all clinical scenarios and the availability of anti-HCV treatment based on interferon (IFN)–ribavirin (RBV) [13,14], the response to which is associated with a reduced frequency of liver-related complications and mortality [15]. Data on the impact of HCV coinfection on the mortality in HIV-infected patients are limited [16]. In this study, we aimed to assess trends in overall mortality and cause-specific mortality stratified by HCV serostatus in a cohort of cART-naive HIV-infected patients in Spain.

Methods

Study population and design

HIV-infected and cART-naive patients were included in one of two different cohorts of the AIDS Research Network in Spain (RIS): CoRIS-MD and CoRIS [17,18]. CoRIS-MD is a retrospectively assembled multicenter cohort, patients were incorporated at any time from 1 January 1997 to 30 June 2003, and the cohort was assembled in 2004. CoRIS is an ongoing prospective cohort, patients were incorporated at any time from 1 January 2004 to 30 June 2008, and recruitment is ongoing.

The variables recorded in both cohorts at entry were sex, age, HIV transmission category, Centers for Disease Control and Prevention category, $CD4^+$ T-cell count, plasma HIV RNA, and HCV serologic status (categorized as positive or negative). No information was collected on HCV RNA. We also had information about initiation of cART during follow-up, new AIDS-defining conditions, $CD4^+$ T-cell count, viral load, and vital status. cART was modeled as a time-dependent covariate in an intention-to-treat analysis (once patients started cART, they were assumed to remain on it).

To assess trends in overall and cause-specific mortality, follow-up was divided into two calendar periods, 1997–2003 and 2004–2008, which correspond to the follow-up period of each cohort. The date of censoring for these analyses was the date of the last visit or cohort-specific administrative censoring date (31 December 2003 and 31 December 2008 for CoRIS-MD and CoRIS, respectively).

Causes of death

Deaths were ascertained through cohort reporting and a cross-match with the Spanish National Death Index. Causes of deaths for deceased patients were obtained from the National Basic Death File coded by the underlying cause of death in accordance with the International Classification of Diseases-10. The underlying cause of death is coded as the initial or basic cause as noted on the death certificate. Data were cross-matched in the first quarter of 2010 for patients who had died between January 1997 and December 2008. Cause of death was categorized as follows: AIDS-related death (ARD) (codes A02-A029, A073, A15-A19, A30-A31, A812, B00-B009, B20-B24, B250-B259, B371, B383-B389, B393-B399, B451-B459, B582, C53, C83); liver-related death (LRD) (codes B15-B19, C22-C229, K70-K77, K922, K65, R18); and non-LRD–non-ARDs (NLRD–NARD).

Statistical analysis

We calculated overall and cause-specific crude death rates per 1000 person-years stratified by calendar period and HCV status. We used Poisson regression to calculate crude and adjusted incidence rate ratios of death in the second period, taking the first period as a reference and categorizing deaths by HCV serostatus. Multivariate models were adjusted for the potential confounders. Robust methods were used to estimate standard errors and calculate confidence intervals. All statistical analyses were performed using Stata version 11 (College Station, Texas, USA)

Results

Patients

We analyzed 5974 HIV-infected cART-naive patients: 2471 in the 1997–2003 period with a median follow-up of 4.3 years [interquartile range (IQR), 2.3–6.21] and 3503 in the 2004–2008 period with a median follow-up of 2.4 years (IQR, 1.34–3.6). The characteristics of the patients at inclusion are shown in Table 1.

Causes of death

During follow-up, 232 deaths were identified, 158 during the first period (85 ARD, 54 NLRD-NARD, and 19 LRD) and 74 during the second period (42 ARD, 26 NLRD-NARD, and 6 LRD). There were 58 NLRD-NARD among HCV-positive patients (cardiovascular disease $N=14$; accidental poisoning $N=9$; accidental/traumatic events $N=7$; suicide $N=6$; non-AIDS-defining malignancy $N=4$; non-AIDS-defining infections $N=4$; and other causes $N=14$). There were 22

Table 1. Baseline characteristics of 5974 HIV-positive combination antiretroviral therapy-naive patients at the inclusion in the study.

	1997–2003 ($N=2471$)	2004–2008 ($N=3503$)	*P*
Male sex *n* (%)	1791 (72.5)	2737 (78.1)	<0.001
Age at entry-years *n* (%)			
<30	563 (22.8)	876 (25.0)	<0.001
30–40	1432 (58.0)	1425 (40.7)	
>40	476 (19.3)	1202 (34.3)	
HIV-transmission category *n* (%)			
MSM	361 (14.6)	1583 (45.2)	<0.001
IDU	1329 (53.8)	530 (15.1)	
Heterosexual	575 (23.27)	1259 (35.9)	
Other/unknown	206 (8.34)	131 (3.7)	
CDC disease category C *n* (%)	282 (11.4)	490 (14.0)	0.003
$CD4^+$ cells/μl *n* (%)			
<200	608 (24.6)	1047 (29.9)	<0.001
201–350	372 (15.1)	713 (20.4)	
>350	881 (35.7)	1662 (47.4)	
Unknown	610 (24.7)	81 (2.3)	
HIV RNA copies/ml *n* (%)			
<100 000 copies/ml	1076 (43.6)	2242 (64.0)	<0.001
≥100 000 copies/ml	434 (17.6)	1178 (33.6)	
Unknown	961 (39.0)	83 (2.4)	
HCV serostatus *n* (%)			
Negative	974 (39.4)	2814 (80.3)	<0.001
Positive	1497 (60.6)	689 (19.7)	

CDC, Centers for Disease Control and Prevention; HCV, hepatitis C virus; IDU, injection drug use; MSM, men who have sex with men.

NLRD–NARD among HCV-negative patients (non-AIDS-defining malignancy $N=7$; cardiovascular disease $N=5$; accidental/traumatic events $N=1$; non-AIDS-defining infections $N=1$; and other causes $N=8$).

Mortality rates

During 19 117 person-years of follow-up, 232 patients died (death rate: 12.1 per 1000 person-years).

The crude overall death rate declined significantly (44%) from 1997–2003 to 2004–2008. When we stratified by HCV serostatus, a significant 48% decrease in the overall death rate was observed among HCV-negative patients, whereas a nonsignificant 27% increase was observed among HCV-positive patients (Fig. 1).

The crude ARD rate declined significantly (41%) from 1997–2003 to 2004–2008. When we stratified by HCV serostatus, the ARD rate declined from 1997–2003 to 2004–2008 among HCV-negative patients, although the difference did not reach statistical significance. The ARD rate remained virtually unchanged among HCV-positive patients (Fig. 1).

The crude LRD death rate declined significantly (62%) from 1997–2003 to 2004–2008. When we stratified by HCV serostatus, the LRD death rate declined non-significantly from 1997–2003 to 2004–2008 among HCV-negative patients and increased nonsignificantly among HCV-positive patients (Fig. 1).

The crude NLRD–NARD rate declined significantly (43%) from 1997–2003 to 2004–2008. When we stratified by HCV serostatus, NLRD-NARD rates decreased significantly by 64% among HCV-negative patients but increased nonsignificantly among HCV-positive patients (Fig. 1).

The effects seen in the crude analyses are maintained in the multivariate models presented in Fig. 1. A statistically significant 48% reduction in all-cause mortality was observed for HCV-negative patients, whereas no reductions were seen for HCV-positive individuals.

Discussion

We analyzed mortality trends in the cART era in a cohort of cART-naive HIV-infected patients and found a decline in overall mortality, taking the first years of the cART era as reference. The trend was also identified for different death categories. Of note, the decreases in mortality in our cohort were largely at the expense of the HCV-negative patients, as declines were not observed among HCV-positive patients. Interestingly, among HCV-positive patients, we identified a trend not only toward an increasing incidence in LRD, but also toward an increasing incidence in ARD and NLRD–NARD.

In our study, the LRD rate among HCV-positive patients did not decline between the two periods. This finding is of interest because IFN–RBV has been available in Spain for coinfected patients since 2000. Although our data do not enable us to provide an explanation for this observation, we believe it is due to barriers that limit the initiation and completion of anti-HCV therapy in this population [19,20].

2244 AIDS 2012, Vol 26 No 17

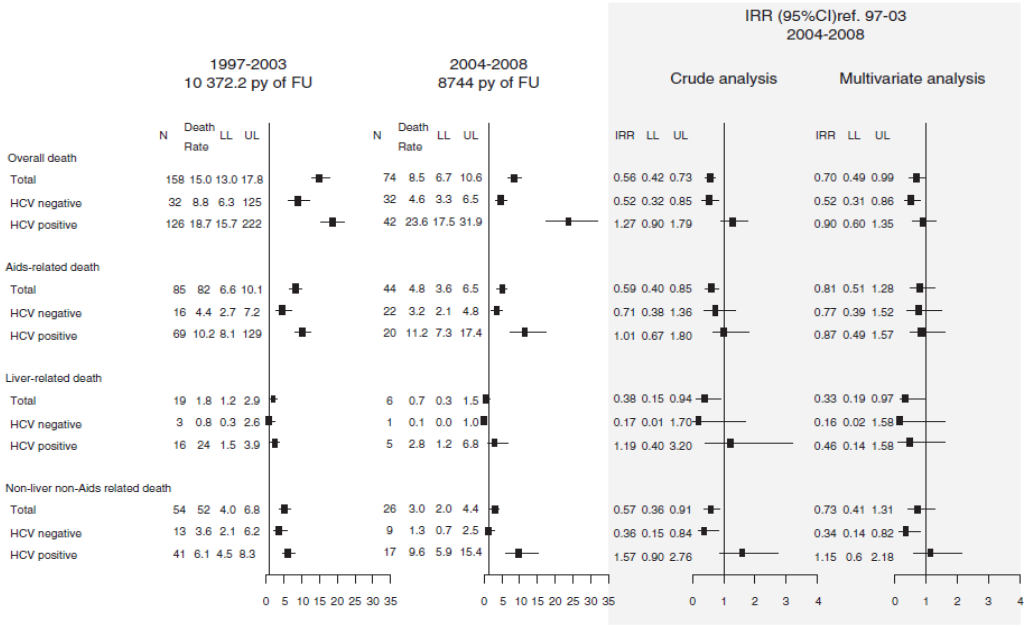


Fig. 1. Effect of calendar period (2004–2008 vs. 1997–2003) on cause-specific mortality and mortality rates (and 95% confidence intervals) for each HCV serostatus. For each death category and hepatitis C virus serostatus, the IRRs of the calendar period were calculated considering the 1997–2003 as the reference category. One separate multivariate Poisson model for each cause of death was performed, adjusted for baseline $CD4^+$ cell count, age at enrollment, HIV transmission category, and active cART as a time-updated variable. CI, confidence interval; FU, follow-up; HCV, hepatitis C virus; IRR, incidence rate ratio; LL, lower limit of 95% CI; py, person-years; UL, upper limit of 95% CI.

In the cART era, AIDS-defining conditions have been the first cause of death in some cohorts [21] and institutions [22], but not in others, wherein the first cause of death is non-AIDS-defining conditions [23]. ARD was the most frequent cause of death in both periods in our study. Interestingly, the ARD rate declined between the two periods among HCV-negative patients but remained virtually unchanged among HCV-positive patients. This finding again raises the question of whether HCV infection can negatively affect the natural history of HIV infection. This question is difficult to answer because injection drug use (IDU) is the primary risk factor among HCV-positive patients, and despite the fact that in some studies IDU has not been associated with decreased survival among HIV-infected patients initiating cART [24], in other studies IDU has been associated with poorer access to care [25] and increased mortality [26]. This may explain in part the differences found in studies that have studied the influence of HCV on HIV from a clinical [11,27] and an immunological perspective [12,28]. However, recent works suggest that HCV coinfection can contribute to immune activation [29] and increases the risk for progression to AIDS and mortality [30,31].

In our cohort, NLRD–NARD ranked second in the causes of death in both periods. We found that the rate of NLRD–NARD declined significantly from 1997–2003 to 2004–2008; however, when we stratified by HCV serostatus, the decrease was observed among HCV-negative patients only. Some of the causes of death in this category – substance abuse, accidents, and suicide – were particularly common among HCV-positive patients and virtually nonexistent among HCV-negative patients. These causes of deaths have frequently been reported among IDUs in several studies [32–35] and, consequently, seem more related to lifestyle factors than to HCV-associated biological factors. However, a large proportion of NLRD–NARD among HCV-positive patients was caused by cardiovascular disease and non-AIDS-defining malignancies, conditions that in addition to end-stage liver disease, now play a major role in most of the fatal conditions classified as non-AIDS-related events as a consequence of increased survival in this population group [36–40].

Our study has several limitations. The first 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 it is due to being a young and postcART cohort. We tried to minimize the limitations of the joint analysis of two different cohorts with slightly different inclusion criteria by restricting the analyses to those patients who fulfilled the inclusion criteria for both cohorts. The second limitation is that overall periods of follow-up are short, particularly in the 2004–2008 period. As HCV-related liver disease requires time to progress, this may result in underestimating its contribution to mortality. The third limitation is that no determination of HCV-RNA was performed; this could have better sorted out the contribution of active infection to mortality. Finally, as has been the case with other studies of similar characteristics, some possible confounding factors (e.g. adherence to cART, active mental illness, alcohol use, access/adherence to IFN–RBV) were not collected and consequently their impact in the results could not be analyzed.

The prevalence of HCV infection in newly HIV-infected patients has decreased sharply in parallel with a decrease in IDU as a mechanism of HIV transmission in Spain [41]. Nevertheless, our findings suggest that coinfection with HCV is still an important factor in overall and cause-specific mortality and should influence changes in the clinical management of HIV-infected patients in order to further diminish mortality in HIV–HCV-coinfected patients.

Acknowledgements

This study would not have been possible without the collaboration of the patients, medical and nursing staff, and data managers.

We are grateful to Thomas O'Boyle for proofreading the manuscript.

Study concept and design by J.B., B.A., J.D.A.

Acquisition of data by J.B., P.V., M.S., I.S., J.L.G.S., F.V., J.P.

Administrative, technical, or material support by B.A., V.H., J.D.A.

Analysis and interpretation of data by B.A., J.D.A., J.B.

Drafting of the manuscript by J.B.

Critical revision of the manuscript for important intellectual content by B.A., J.D.A.

Final revision and approval of the manuscript by all of the authors

Study supervision by J.B., J.D.A.

This work was partially funded by grant 04/0900 from FIS (Spanish Networks for Research on AIDS and Public Health) and RD06/0006 from the RIS (Spanish HIV Research Network for excellence). We are grateful for funds provided by CIBERESP (Ciber de Epidemiologia y Salud Pública). Dr J.B. is an investigator from the Intensification of Research Activity Program of the Spanish National Health Service (I3SNS). B.A. is an investigator supported by the Predoctoral Training in Health Research Program (National Plan for Scientific Research, Development and Technological Innovation).

Conflicts of interest

There are no conflicts of interest.

References

1. Palella FJ Jr, Delaney KM, Moorman AC, Loveless MO, Fuhrer J, Satten GA, *et al.* **Declining morbidity and mortality among patients with advanced human immunodeficiency virus infection. HIV Outpatient Study Investigators.** *N Engl J Med* 1998; **338**:853–860.
2. Detels R, Munoz A, McFarlane G, Kingsley LA, Margolick JB, Giorgi J, *et al.* **Effectiveness of potent antiretroviral therapy on time to AIDS and death in men with known HIV infection duration. Multicenter AIDS Cohort Study Investigators.** *JAMA* 1998; **280**:1497–1503.
3. CASCADE Collaboration. **Survival after introduction of HAART in people with known duration of HIV-1 infection. The CASCADE Collaboration. Concerted Action on SeroConversion to AIDS and Death in Europe.** *Lancet* 2000; **355**:1158–1159.
4. Mocroft A, Brettle R, Kirk O, Blaxhult A, Parkin JM, Antunes F, *et al.* **Changes in the cause of death among HIV positive subjects across Europe: results from the EuroSIDA study.** *AIDS* 2002; **16**:1663–1671.
5. Alter MJ. **Epidemiology of viral hepatitis and HIV co-infection.** *J Hepatol* 2006; **44**:S6–S9.
6. Graham CS, Baden LR, Yu E, Mrus JM, Carnie J, Heeren T, *et al.* **Influence of human immunodeficiency virus infection on the course of hepatitis C virus infection: a meta-analysis.** *Clin Infect Dis* 2001; **33**:562–569.
7. Puoti M, Spinetti A, Ghezzi A, Donato F, Zaltron S, Putzolu V, *et al.* **Mortality for liver disease in patients with HIV infection: a cohort study.** *J Acquir Immune Defic Syndr* 2000; **24**:211–217.
8. Martin-Carbonero L, Soriano V, Valencia E, Garcia-Samaniego J, Lopez M, Gonzalez-Lahoz J. **Increasing impact of chronic viral hepatitis on hospital admissions and mortality among HIV-infected patients.** *AIDS Res Hum Retroviruses* 2001; **17**:1467–1471.
9. Soriano V, Garcia-Samaniego J, Valencia E, Rodriguez-Rosado R, Munoz F, Gonzalez-Lahoz J. **Impact of chronic liver disease due to hepatitis viruses as cause of hospital admission and death in HIV-infected drug users.** *Eur J Epidemiol* 1999; **15**:1–4.
10. Bica I, McGovern B, Dhar R, Stone D, McGowan K, Scheib R, *et al.* **Increasing mortality due to end-stage liver disease in patients with human immunodeficiency virus infection.** *Clin Infect Dis* 2001; **32**:492–497.
11. Sulkowski MS, Moore RD, Mehta SH, Chaisson RE, Thomas DL. **Hepatitis C and progression of HIV disease.** *JAMA* 2002; **288**:199–206.
12. Peters L, Mocroft A, Soriano V, Rockstroh JK, Losso M, Valerio L, *et al.* **Hepatitis C virus coinfection does not influence the CD4 cell recovery in HIV-1-infected patients with maximum virologic suppression.** *J Acquir Immune Defic Syndr* 2009; **50**:457–463.
13. Carrat F, Bani-Sadr F, Pol S, Rosenthal E, Lunel-Fabiani F, Benzekri A, *et al.* **Pegylated interferon alfa-2b vs standard interferon alfa-2b, plus ribavirin, for chronic hepatitis C in HIV-infected patients: a randomized controlled trial.** *JAMA* 2004; **292**:2839–2848.

14. Torriani FJ, Rodriguez-Torres M, Rockstroh JK, Lissen E, Gonzalez-Garcia J, Lazzarin A, *et al.* **Peginterferon Alfa-2a plus ribavirin for chronic hepatitis C virus infection in HIV-infected patients.** *N Engl J Med* 2004; **351**:438–450.
15. Berenguer J, Alvarez-Pellicer J, Martin PM, Lopez-Aldeguer J, Von-Wichmann MA, Quereda C, *et al.* **Sustained virological response to interferon plus ribavirin reduces liver-related complications and mortality in patients coinfected with human immunodeficiency virus and hepatitis C virus.** *Hepatology* 2009; **50**:407–413.
16. Chen TY, Ding EL, Seage Iii GR, Kim AY. **Meta-analysis: increased mortality associated with hepatitis C in HIV-infected persons is unrelated to HIV disease progression.** *Clin Infect Dis* 2009; **49**:1605–1615.
17. Gutierrez F, Padilla S, Masia M, Iribarren JA, Moreno S, Viciana P, *et al.* **Patients' characteristics and clinical implications of suboptimal CD4 T-cell gains after 1 year of successful antiretroviral therapy.** *Curr HIV Res* 2008; **6**:100–107.
18. Caro-Murillo AM, Castilla J, Perez-Hoyos S, Miro JM, Podzamczer D, Rubio R, *et al.* **[Spanish cohort of naive HIV-infected patients (CoRIS): rationale, organization and initial results].** *Enferm Infecc Microbiol Clin* 2007; **25**:23–31.
19. Fleming CA, Craven DE, Thornton D, Tumilty S, Nunes D. **Hepatitis C virus and human immunodeficiency virus coinfection in an urban population: low eligibility for interferon treatment.** *Clin Infect Dis* 2003; **36**:97–100.
20. Fleming CA, Tumilty S, Murray JE, Nunes D. **Challenges in the treatment of patients coinfected with HIV and hepatitis C virus: need for team care.** *Clin Infect Dis* 2005; **40 (Suppl 5)**:S349–S354.
21. Antiretroviral Therapy Cohort Collaboration. **Causes of death in HIV-1-infected patients treated with antiretroviral therapy, 1996–2006: collaborative analysis of 13 HIV cohort studies.** *Clin Infect Dis* 2010; **50**:1387–1396.
22. Martinez E, Milinkovic A, Buira E, de Lazzari E, Leon A, Larrousse M, *et al.* **Incidence and causes of death in HIV-infected persons receiving highly active antiretroviral therapy compared with estimates for the general population of similar age and from the same geographical area.** *HIV Med* 2007; **8**:251–258.
23. Bonnet F, Morlat P, Chene G, Mercie P, Neau D, Chossat I, *et al.* **Causes of death among HIV-infected patients in the era of highly active antiretroviral therapy, Bordeaux, France, 1998–1999.** *HIV Med* 2002; **3**:195–199.
24. Wood E, Hogg RS, Lima VD, Kerr T, Yip B, Marshall BD, *et al.* **Highly active antiretroviral therapy and survival in HIV-infected injection drug users.** *JAMA* 2008; **300**:550–554.
25. Celentano DD, Galai N, Sethi AK, Shah NG, Strathdee SA, Vlahov D, *et al.* **Time to initiating highly active antiretroviral therapy among HIV-infected injection drug users.** *AIDS* 2001; **15**:1707–1715.
26. van Asten LC, Boufassa F, Schiffer V, Brettle RP, Robertson JR, Hernandez Aguado I, *et al.* **Limited effect of highly active antiretroviral therapy among HIV-positive injecting drug users on the population level.** *Eur J Public Health* 2003; **13**:347–349.
27. Greub G, Ledergerber B, Battegay M, Grob P, Perrin L, Furrer H, *et al.* **Clinical progression, survival, and immune recovery during antiretroviral therapy in patients with HIV-1 and hepatitis C virus coinfection: the Swiss HIV Cohort Study.** *Lancet* 2000; **356**:1800–1805.
28. Potter M, Odueyungbo A, Yang H, Saeed S, Klein MB. **Impact of hepatitis C viral replication on CD4+ T-lymphocyte progression in HIV-HCV coinfection before and after antiretroviral therapy.** *AIDS* 2010; **24**:1857–1865.
29. Kovacs A, Al-Harthi L, Christensen S, Mack W, Cohen M, Landay A. **$CD8^{(+)}$ T cell activation in women coinfected with human immunodeficiency virus type 1 and hepatitis C virus.** *J Infect Dis* 2008; **197**:1402–1407.
30. Daar ES, Lynn H, Donfield S, Gomperts E, O'Brien SJ, Hilgartner MW, *et al.* **Hepatitis C virus load is associated with human immunodeficiency virus type 1 disease progression in hemophiliacs.** *J Infect Dis* 2001; **183**:589–595.
31. Kovacs A, Karim R, Mack WJ, Xu J, Chen Z, Operskalski E, *et al.* **Activation of CD8 T cells predicts progression of HIV infection in women coinfected with hepatitis C virus.** *J Infect Dis* 2010; **201**:823–834.
32. Prins M, Hernandez Aguado IH, Brettle RP, Robertson JR, Broers B, Carre N, *et al.* **Pre-AIDS mortality from natural causes associated with HIV disease progression: evidence from the European Seroconverter Study among injecting drug users.** *AIDS* 1997; **11**:1747–1756.
33. Tyndall MW, Craib KJ, Currie S, Li K, O'Shaughnessy MV, Schechter MT. **Impact of HIV infection on mortality in a cohort of injection drug users.** *J Acquir Immune Defic Syndr* 2001; **28**:351–357.
34. Copeland L, Budd J, Robertson JR, Elton RA. **Changing patterns in causes of death in a cohort of injecting drug users, 1980–2001.** *Arch Intern Med* 2004; **164**:1214–1220.
35. Kohli R, Lo Y, Howard AA, Buono D, Floris-Moore M, Klein RS, *et al.* **Mortality in an urban cohort of HIV-infected and at-risk drug users in the era of highly active antiretroviral therapy.** *Clin Infect Dis* 2005; **41**:864–872.
36. Lewden C, Salmon D, Morlat P, Bevilacqua S, Jougla E, Bonnet F, *et al.* **Causes of death among human immunodeficiency virus (HIV)-infected adults in the era of potent antiretroviral therapy: emerging role of hepatitis and cancers, persistent role of AIDS.** *Int J Epidemiol* 2005; **34**:121–130.
37. Cohen MH, French AL, Benning L, Kovacs A, Anastos K, Young M, *et al.* **Causes of death among women with human immunodeficiency virus infection in the era of combination antiretroviral therapy.** *Am J Med* 2002; **113**:91–98.
38. Louie JK, Hsu LC, Osmond DH, Katz MH, Schwarcz SK. **Trends in causes of death among persons with acquired immunodeficiency syndrome in the era of highly active antiretroviral therapy, San Francisco, 1994–1998.** *J Infect Dis* 2002; **186**:1023–1027.
39. Marin B, Thiebaut R, Bucher HC, Rondeau V, Costagliola D, Dorrucci M, *et al.* **Non-AIDS-defining deaths and immunodeficiency in the era of combination antiretroviral therapy.** *AIDS* 2009; **23**:1743–1753.
40. Palella FJ Jr, Baker RK, Moorman AC, Chmiel JS, Wood KC, Brooks JT, *et al.* **Mortality in the highly active antiretroviral therapy era: changing causes of death and disease in the HIV outpatient study.** *J Acquir Immune Defic Syndr* 2006; **43**:27–34.
41. Perez Cachafeiro S, Del Amo J, Iribarren JA, Salavert Lleti M, Gutierrez F, Moreno A, *et al.* **Decrease in serial prevalence of coinfection with hepatitis C virus among HIV-infected patients in Spain, 1997–2006.** *Clin Infect Dis* 2009; **48**:1467–1470.

Hernando *et al. BMC Infectious Diseases* 2013, **13**:382
http://www.biomedcentral.com/1471-2334/13/382

BMC Infectious Diseases

RESEARCH ARTICLE Open Access

All-cause mortality in the cohorts of the Spanish AIDS Research Network (RIS) compared with the general population: 1997–2010

Victoria Hernando[1,2*], Belen Alejos[1,2], Susana Monge[1,2], Juan Berenguer[3], Lourdes Anta[4], David Vinuesa[5], Rosario Palacios[6], Roberto Muga[7], Santiago Moreno[8], Inma Jarrin[1,2] and for CoRIS cohort

Abstract

Background: Combination antiretroviral therapy (cART) has produced significant changes in mortality of HIV-infected persons. Our objective was to estimate mortality rates, standardized mortality ratios and excess mortality rates of cohorts of the AIDS Research Network (RIS) (CoRIS-MD and CoRIS) compared to the general population.

Methods: We analysed data of CoRIS-MD and CoRIS cohorts from 1997 to 2010. We calculated: (i) all-cause mortality rates, (ii) standardized mortality ratio (SMR) and (iii) excess mortality rates for both cohort for 100 person-years (py) of follow-up, comparing all-cause mortality with that of the general population of similar age and gender.

Results: Between 1997 and 2010, 8,214 HIV positive subjects were included, 2,453 (29.9%) in CoRIS-MD and 5,761 (70.1%) in CoRIS and 294 deaths were registered. All-cause mortality rate was 1.02 (95% CI 0.91-1.15) per 100 py, SMR was 6.8 (95% CI 5.9-7.9) and excess mortality rate was 0.8 (95% CI 0.7-0.9) per 100 py. Mortality was higher in patients with AIDS, hepatitis C virus (HCV) co-infection, and those from CoRIS-MD cohort (1997–2003).

Conclusion: Mortality among HIV-positive persons remains higher than that of the general population of similar age and sex, with significant differences depending on the history of AIDS or HCV coinfection.

Keywords: Mortality rate, HIV infection, Standardized mortality ratios, Excess mortality

Background

Mortality of HIV-infected persons in Western countries has decreased significantly due to improvements in combined antiretroviral therapy (cART) [1,2]. Nevertheless it continues to be higher than in the general population [3-5], even in HIV-infected patients with good initial response to cART [6]. Global reduction in mortality has been achieved thanks to a decrease of AIDS-related deaths which has led to a greater relevance of other causes of death in relation to co-morbidities, such as hepatitis C virus (HCV) and/or hepatitis B virus (HBV) co-infections, drug abuse and cardiovascular diseases [2,7].

In Barcelona and Navarre, HIV-positive subjects were found to have a higher mortality compared to the general population [8,9] but no estimates are available for the whole country. Unlike other cohorts, in this work we have analyzed data of a cohort of persons with HIV infection recruited during a period where highly effective antiretroviral treatment is available and all patients are naïve to treatment. But we believe that even though these patients may be in a better starting point than patients in other similar studies, the risk of mortality compared with the general population is still higher.

Therefore, the objectives of this study were to calculate the overall mortality rates, standardized mortality ratios (SMR), and excess mortality rates in the cohorts of the Spanish AIDS Research Network (RIS) – CoRIS-MD and CoRIS, comparing the overall mortality rates observed in HIV positive subjects in both cohorts with the mortality rates of the general population of similar age and sex.

* Correspondence: vhernando@isciii.es
[1]Red de Investigación en Sida, Centro Nacional de Epidemiología, Instituto de Salud Carlos III, Avda. Monforte de Lemos, Madrid 5 28029, Spain
[2]CIBER de Epidemiología y Salud Pública (CIBERESP), Madrid, Spain
Full list of author information is available at the end of the article

Hernando et al. BMC Infectious Diseases 2013, 13:382
http://www.biomedcentral.com/1471-2334/13/382

Methods

Patients

We analyzed data from the cohorts of HIV-infected adults of the Spanish AIDS Research Network (RIS). CoRIS-MD is a multicenter cohort including data from 1997 to 2003 from 9 hospitals of 7 Spanish Autonomous regions assembled in 2003. CoRIS is a multicenter cohort which recruits patients from 2004 onwards from 28 health-care centers and hospitals in 12 of the 17 Autonomous regions that compose Spain [10,11]. Both cohorts recruit patients newly attended in any of the participating sites. Ethics approval was obtained from all hospitals Ethics' Committees (see Appendix 1 all hospitals participants) and every patient provides written informed consent to participate in the cohorts. For this analysis, we selected subjects who were naïve to cART at cohort entry, older than 20 years, had a follow up of more than 6 months and had had at least one diagnostic test for hepatitis C virus.

Variables

We considered the following variables: age at cohort entry (20–29; 30–39; 40–49; > = 50); gender (male, female); year of cohort entry; HIV transmission category, classified as injecting drugs users (IDUs), men who have sex with men (MSM), heterosexual contact and others or unknown risk category; AIDS before entry and changes in AIDS status during follow-up; CD4 count at entry (<200, 200–349, ≥350); HIV viral load at entry (<20000, 20000–100000, ≥100000); combined antiretroviral treatment (cART) initiation during follow-up; HCV serological status classified as positive or negative antibodies and vital status.

To calculate mortality rates, AIDS variable was classified as "Yes" when the person had AIDS before entering the cohort, AIDS at cohort entry or AIDS during follow-up and "No" when the person didn't develop AIDS at any moment during the study.

Statistical analyses

Descriptive analysis of patients' characteristics was carried out using frequency distribution for categorical variables and median (interquartile range -IQR) for continuous variables.

Individuals were followed up from study entry to death, last study contact or the administrative censoring date (31/12/2003 in CoRIS-MD and 31/12/2010 in CoRIS) whichever arose first. We calculated mortality rates, overall and according to socio-demographic and clinical characteristics, as the number of deaths by 100 persons-year (py) of follow-up with 95% confidence intervals (95% CI) calculated using the exact Poisson method.

Standardized mortality ratios (SMR) were estimated for all-cause mortality in CoRIS-MD and CoRIS, comparing with the overall mortality rates of the general population standardized by sex and age. SMR were estimated as the ratio of observed deaths to expected deaths, had our patients had the same distribution of mortality as the general population. SMR were calculated through Poisson models offsetting expected mortality rates, and adjusted for gender, age, category of transmission and HCV test. Mortality rates for general population, between 1997 and 2010, were obtained from the National Statistics Institute (www.ine.es), stratified by sex and age at 5 year intervals. A constant mortality rate within each 5 year stratum was assumed.

A sensitivity analysis was performed to assess a possible *selection bias*. The SMR was calculated for the first 12 months after cohort entry separately for all patients together. This was to determinate whether it is necessary to include a lag time to avoid an overestimation of SMR.

Excess Mortality Rates were calculated as the difference between observed and expected deaths according to mortality in the general population, divided by the number of persons-year (py) of follow-up. Confidence intervals for Excess Mortality Rates were estimated using Poisson's exact method.

All statistical analyses were performed by using Stata software (Version 11.0, College Station, Texas).

Results

Baseline characteristics of the study population

A total of 8,214 subjects were included in the study, 2,453 (29.9%) in CoRIS-MD and 5,761 (70.1%) in CoRIS, adding up to 28,743 persons-year of follow up, and 294 deaths.

Men represented 78.0% (n = 6,412) of the sample, and median age at the cohort entry was 35.0 years (interquartile range IQR: 30.2 – 41.0), 35.5 years (IQR: 30.2-41.7) for men and 34.2 years (IQR: 29.1-40.1) for women. Regarding transmission categories, the sample was distributed between injecting drugs users (IDUs) or ex-users, 25.0% (n = 2,050), men who have sex with men (MSM), 39.6% (n = 3,255), and heterosexuals, 30.7% (n = 2,524). A 20.4% of the subjects had a history of an AIDS defining illness (ADI), although for 59.4% (n = 994) of them the ADI diagnosis was previous to cohort entry. Median CD4 count at cohort entry was 350 cell/mm^3 (IQR 170 – 552), and median viral load was 39,811 copies/ml (IQR 7,520 – 135,988) (Table 1).

Among the 294 deceased subjects, 80.6% (n = 237) were men, and median age was 37.7 years (IQR 33.3 – 44.5). Some 60.2% (n = 177) were IDU or ex-IDU, 51.0% (n = 150) had an AIDS diagnosis and 67.4% (n = 198) were co-infected by HCV. Median CD4 count at entry was 154 cell/mm^3 (IQR 66 – 390) and median HIV viral load was 78,200 copies/ml (IQR 17,335 – 230,000) (Table 1).

Hernando *et al. BMC Infectious Diseases* 2013, **13**:382
http://www.biomedcentral.com/1471-2334/13/382

Table 1 Socio demographics and clinical characteristics at cohort entry for total of analyzed subjects and deceased subjects

		Total		Deaths	
	py	n	%	n	%
Total	28,743	8,214	100	294	100
Gender					
Males	21,903	6,412	78.0	237	80.6
Females	6,840	1,802	22.0	57	19.4
Age at cohort entry (years)					
20–29	6,945	2,064	25.1	34	11.6
30–39	13,778	3,722	45.3	145	49.3
40–49	5,584	1,705	13.1	71	24.1
> = 50	2,436	723	8.8	44	15.0
Median age (IQR)		35.0 (30.2–41.0)		37.7(33.5–44.5)	
Category of transmission					
IDUs	8,515	2,050	25.0	177	60.2
MSM	9,994	3,255	39.6	41	14.0
Heterosexual	8,909	2,524	30.7	67	22.8
Others/Unknown	1,325	385	4.7	9	3.0
AIDS					
No	22,255	6,542	79.6	144	49.0
AIDS before entry	3,667	994	12.1	72	24.5
AIDS after entry	2,821	678	8.3	78	26.5
CD4 count at entry (cel/mm^3)					
<200	7,525	2,217	27.0	141	48.0
200–349	5,191	1,567	19.1	39	13.3
> = 350	12,366	3,744	45.6	64	21.8
Unknown	3,661	686	8.4	50	17.0
Median (IQR)		350 (170–552)		154 (66–390)	
HIV viral load (copies/ml)					
<20.000	8,748	2,769	33.7	61	20.7
20.000-100.000	7,141	2,202	26.8	65	22.1
>100.000	7,181	2,196	26.7	89	30.3
Unknown	5,673	1,047	12.8	79	26.9
Median (IQR)		39,810 (7,520–135,988)		78,200 (17,335–230,000)	
Cohorts					
CoRIS (2004–2008)	18,447	5,761	70.1	137	46.6
CoRIS-MD (1997–2003)	10,296	2,453	29.9	157	53.4
HCV test					
Negative	18,332	5,673	69.1	96	32.6
Positive	10,411	2,541	30.9	198	67.4
Antiretroviral treatment during follow-up					
No	9,992	1,948	23.7	63	21.4
Yes	18,751	6,266	76.3	231	78.6

IDUs Injecting Drugs Users, *MSM* Men have Sex with Men, *HCV* Hepatitis C virus.

Hernando et al. BMC Infectious Diseases 2013, 13:382
http://www.biomedcentral.com/1471-2334/13/382

Mortality rates, standardized mortality ratios and excess mortality rates

Figure 1 shows mortality rates for 100 persons-year (py) of follow up, standardized mortality ratios and excess mortality rates for 100 py in both RIS cohorts.

Overall mortality rate was 1.02 (95% CI: 0.91-1.15) deaths for 100 py of follow up, higher for men (1.08; 95% CI: 0.95-1.23), for subjects over 50 years-old (1.81; 95% CI: 1.34-2.42), for IDU (2.08; 95% CI: 1.79-2.41) compared to both MSM (0.41; 95% CI: 0.30-0.56) and heterosexuals (0.75; 95% CI: 0.59-0.96) and for patients included in CoRIS-MD (1.52; 95% CI: 1.30-1.78). For patients who had an AIDS diagnosis, mortality rate was 2.06 (95% CI: 2.21-3.05), compared to 0.63 (95% CI: 0.54-0.74) for those who were AIDS-free. For HCV co-infected patients mortality rate rose up to 1.90 (95% CI: 1.65-2.19) in contrast with 0.52 (95% CI: 0.42-0.64) for those not co-infected.

Global mortality in both CoRIS cohorts was 6.8 (95% CI: 5.9-7.9) times higher than mortality of the general population of same age and sex. As opposed to the crude mortality rates, standardized mortality ratios were higher in women (10.5; 95% CI: 7.6-13.3) compared to men (5.6; 95% CI: 4.8-6.4). Still, a higher SMR was found for IDUs (9.7; 95% CI: 7.4-12.0), persons with an AIDS diagnosis (14.9; 95% CI: 12.0-17.9), persons co-infected with HCV (9.2; 95% CI: 7.1-11.2) and those receiving antiretroviral treatment (8.1; 95% CI: 6.8-9.4).

In the sensitivity analysis, considering only the first 12 month of follow-up, SMR is lower than in the complete analysis (4.0; 95% CI 2.4 -5.6).

Finally, regarding excess mortality rate, as an absolute estimator, results are similar to those observed for crude mortality rates (Figure 1).

Discussion and conclusion

Our results show that all-cause mortality in CoRIS-MD and CoRIS cohorts, between 1997 and 2010, is close to seven times higher than that of the general population of the same age and sex. Significant differences have been found depending on the history of AIDS and HCV co-infection.

A previously published study, carried out in similar cohorts in Europe and North America, found a lower global SMR, of 3.36 (95% CI: 3.16 – 3.56), but with a notable heterogeneity between cohorts depending on participant-specific characteristics, and being higher for cohorts with a greater representation of IDUs [12]. For example, Aldaz et al. found mortality of HIV-infected persons in Navarre (Spain) to be 14 times higher than mortality in general population; 63% of this cohort had been infected through the use of injected drugs [8].

These differences could also be related to the higher prevalence of HCV-co-infection as the standardized mortality in HCV co-infected subjects in our study was 9.2 times higher than the general population's. Similar results were found by Lewden et al., where SMR for HCV co-infected persons were 13.9 compared to 4.4 for the HCV negative subjects [4]. In a previous study of CoRIS-MD and CoRIS cohorts, an important increase of the risk of both all cause mortality and liver-related mortality was observed for HIV patients coinfected with HCV [13]. Berenguer et al. also found a decrease in overall mortality in HIV patients in cART era, but only in HCV negative subjects [14] and Chen et al. in a meta-analysis found that the risk of mortality was increased in HCV/HIV coinfected patients in HAART era [15].

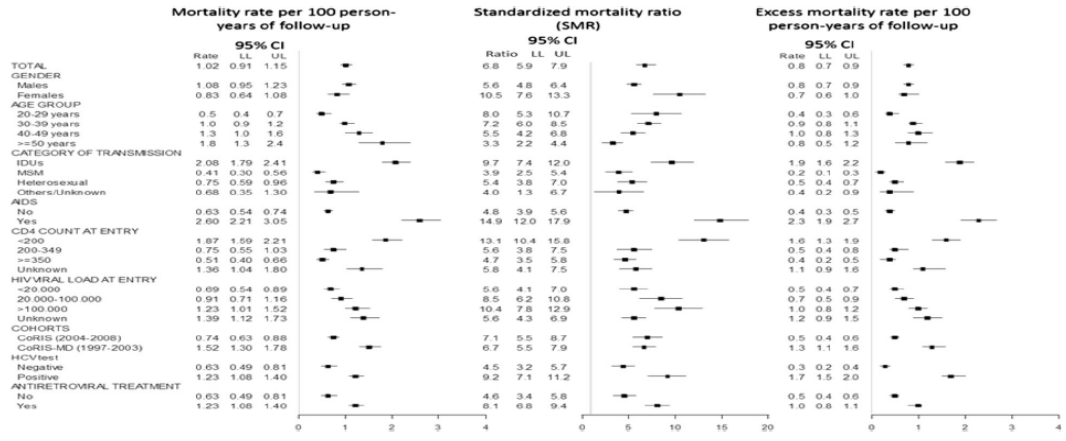


Figure 1 Mortality rates per 100 person-years of follow-up, Standardized mortality ratio (SMRs) and excess mortality rates per 100 person-years of follow-up according to sociodemographic, epidemiological and clinical characteristics.

In our study, we found a similar SMR for patients recruited in CoRIS, from 2004 onwards, and those recruited in CoRIS-MD, from 1997 to 2003, after adjustment for gender, age, transmission category and HCV infection. That is, the difference in the subject's characteristics along these years, the decrease in the representation of IDUs and the percentage of HCV co-infected subjects [11,16,17] were corrected after adjustment. Others studies observed a lower mortality in recent years with the improvement in antiretroviral therapies [18-21], although when specific groups were analyzed, for example: IDUs, found that mortality risk remain elevated [21].

We found non-statistically significant, lower mortality rates in women compared to men. Eventhough the women in our study showed a mortality ratio 10.5 times higher than women of the same age from the general population, and almost doubled the one from men in the cohorts. This higher relative mortality in women could be explained by the fact that women in the general population have a higher life expectancy than men, and specifically, mortality in the general population is very low in women between ages 30 to 40, where we find the majority of HIV-infected women [22]. The lower excess mortality rate in women is consistent with the higher proportion of HIV-infected men in the Spanish epidemic, and in our cohorts [23].

A possible limitation in the calculation of SMR could be using mortality rates in the general population to calculate the expected deaths, because this population contains HIV-related deaths. In our analysis, HIV-related mortality represents a small proportion of all-cause mortality in the general population of Spain, so therefore we consider correct to use the general population mortality rates to calculate the mortality rates in a non-HIV infected population.

The sensitivity analysis shows that when we establish as inclusion criteria to have at least 6 months of follow-up, we are introducing a time window to avoid the selection bias indirectly and overestimate SMRs.

To conclude, mortality in HIV-infected persons continues to be higher than that of the general population, although it has decreased in recent years. For future studies, we would highly recommend to consider, along with global mortality, excess mortality rate for specific causes of death, such as hepatic, non-aids related malignancies or drug-related, especially among IDUs.

Appendix 1: Centers and investigators involved in CoRIS

Executive committee: Juan Berenguer, Julia del Amo, Federico García, Félix Gutiérrez, Pablo Labarga, Santiago Moreno y María Ángeles Muñoz.

Fieldwork, data management and analysis: Paz Sobrino Vegas, Victoria Hernando Sebastián, Belén Alejos Ferreras, Débora Álvarez del Arco, Susana Monge Corella, Inma Jarrín Vera, Adela Castelló.

BioBank: M Ángeles Muñoz-Fernández, Isabel García-Merino, Coral Gómez Rico, Jorge Gallego de la Fuente and Almudena García Torre.

Participating centres:

Hospital General Universitario de Alicante (Alicante): Joaquín Portilla Sogorb, Esperanza Merino de Lucas, Sergio Reus Bañuls, Vicente Boix Martínez, Livia Giner Oncina, Carmen Gadea Pastor, Irene Portilla Tamarit, Patricia Arcaina Toledo.

Hospital Universitario de Canarias (Santa Cruz de Tenerife): Juan Luis Gómez Sirvent, Patricia Rodríguez Fortúnez, María Remedios Alemán Valls, María del Mar Alonso Socas, Ana María López Lirola, María Inmaculada Hernández Hernández, Felicitas Díaz-Flores.

Hospital Carlos III (Madrid): Vicente Soriano, Pablo Labarga, Pablo Barreiro, Pablo Rivas, Francisco Blanco, Luz Martín Carbonero, Eugenia Vispo, Carmen Solera.

Hospital Universitario Central de Asturias (Oviedo): Victor Asensi, Eulalia Valle, José Antonio Cartón

Hospital Clinic (Barcelona): José M. Miró, María López-Dieguez, Christian Manzardo, Laura Zamora, Iñaki Pérez, Mª Teresa García, Carmen Ligero, José Luis Blanco, Felipe García-Alcaide, Esteban Martínez, Josep Mallolas, José M. Gatell.

Hospital Doce de Octubre (Madrid): Rafael Rubio, Federico Pulido, Silvana Fiorante, Jara Llenas, Violeta Rodríguez, Mariano Matarranz.

Hospital Donostia (San Sebastián): José Antonio Iribarren, Julio Arrizabalaga, María José Aramburu, Xabier Camino, Francisco Rodríguez-Arrondo, Miguel Ángel von Wichmann, Lidia Pascual Tomé, Miguel Ángel Goenaga, Mª Jesús Bustinduy, Harkaitz Azkune Galparsoro.

Hospital General Universitario de Elche (Elche): Félix Gutiérrez, Mar Masiá, Cristina López Rodríguez, Sergio Padilla, Andrés Navarro, Fernando Montolio, Catalina Robledano García, Joan Gregori Colomé.

Hospital Germans Trías i Pujol (Badalona): Bonaventura Clotet, Cristina Tural, Lidia Ruiz, Cristina Miranda, Roberto Muga, Jordi Tor, Arantza Sanvisens.

Hospital General Universitario Gregorio Marañón (Madrid): Juan Berenguer, Juan Carlos López Bernaldo de Quirós, Pilar Miralles, Jaime Cosín Ochaíta, Isabel Gutiérrez Cuellar, Margarita Ramírez Schacke, Belén Padilla Ortega, Paloma Gijón Vidaurreta, Ana Carrero Gras, Teresa Aldamiz-Echevarría Lois y Francisco Tejerina Picado.

Hospital Universitari de Tarragona Joan XXIII, IISPV, Universitat Rovira i Virgili (Tarragona): Francesc Vidal, Joaquín Peraire, Consuelo Viladés, Sergio Veloso, Montserrat Vargas, Miguel López-Dupla, Montserrat

Olona, Alba Aguilar, Joan Josep Sirvent, Verónica Alba, Olga Calavia.

Hospital Universitario La Fe (Valencia): José López Aldeguer, Marino Blanes Juliá, José Lacruz Rodrigo, Miguel Salavert, Marta Montero, Eva Calabuig, Sandra Cuéllar.

Hospital Universitario La Paz (Madrid): Juan González García, Ignacio Bernardino de la Serna, José Ramón Arribas López, María Luisa Montes Ramírez, Jose Mª Peña, Blanca Arribas, Juan Miguel Castro, Fco Javier Zamora Vargas, Ignacio Pérez Valero, Miriam Estébanez, Silvia García Bujalance, Marta Díaz.

Hospital de la Princesa (Madrid): Ignacio de los Santos, Jesús Sanz Sanz, Ana Salas Aparicio, Cristina Sarriá Cepeda.

Hospital San Pedro-CIBIR (Logroño): José Antonio Oteo, José Ramón Blanco, Valvanera Ibarra, Luis Metola, Mercedes Sanz, Laura Pérez-Martínez.

Hospital San Pedro II (Logroño): Javier Pinilla Moraza.

Hospital Universitario Mutua de Terrassa (Terrassa): David Dalmau, Angels Jaén Manzanera, Mireia Cairó Llobell, Daniel Irigoyen Puig, Laura Ibáñez, Queralt Jordano Montañez, Mariona Xercavins Valls, Javier Martinez-Lacasa, Pablo Velli, Roser Font.

Hospital de Navarra (Pamplona): María Rivero, Marina Itziar Casado, Jorge Alberto Díaz González, Javier Uriz, Jesús Repáraz, Carmen Irigoyen, María Jesús Arraiza.

Hospital Parc Taulí (Sabadell): Ferrán Segura, María José Amengual, Eva Penelo, Gemma Navarro, Montserrat Sala, Manuel Cervantes, Valentín Pineda.

Hospital Ramón y Cajal (Madrid): Santiago Moreno, José Luis Casado, Fernando Dronda, Ana Moreno, María Jesús Pérez Elías, Dolores López, Carolina Gutiérrez, Beatriz Hernández, María Pumares, Paloma Martí.

Hospital Reina Sofía (Murcia): Alfredo Cano Sánchez, Enrique Bernal Morell, Ángeles Muñoz Pérez.

Hospital San Cecilio (Granada): Federico García García, José Hernández Quero, Alejandro Peña Monje, Leopoldo Muñoz Medina, Jorge Parra Ruiz.

Centro Sanitario Sandoval (Madrid): Jorge Del Romero Guerrero, Carmen Rodríguez Martín, Teresa Puerta López, Juan Carlos Carrió Montiel, Cristina González, Mar Vera.

Hospital Universitario Santiago de Compostela (Santiago de Compostela): Antonio Antela, Arturo Prieto, Elena Losada.

Hospital Son Espases (Palma de Mallorca): Melchor Riera, Javier Murillas, Maria Peñaranda, Maria Leyes, Mª Angels Ribas, Antoni Campins, Concepcion Villalonga, Carmen Vidal.

Hospital Universitario de Valme (Sevilla): Juan Antonio Pineda, Eva Recio Sánchez, Fernando Lozano de León, Juan Macías, José del Valle, Jesús Gómez-Mateos.

Hospital Virgen de la Victoria (Málaga): Jesús Santos González, Manuel Márquez Solero, Isabel Viciana Ramos, Rosario Palacios Muñoz.

Hospital Universitario Virgen del Rocío (Sevilla): Pompeyo Viciana, Manuel Leal, Luis Fernando López-Cortés, Mónica Trastoy.

Competing interests
The authors declare that they have no competing interests.

Authors' contributions
VH, BA, SM and IJ were involved in designing the study, participated in the collection and analysis of the data. VH, BA and IJ wrote the first draft if the manuscript. All authors contributed to data collection, reviewed draft of the manuscript and approved the final manuscript.

Acknowledgments
This study would not have been possible without the collaboration of all the patients, medical and nursing staff, and data mangers 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).

Financial support
This work has been partially funded by a grant from the FIS (Spanish Networks for Research on AIDS and Public Health), 04/0900 and RIS (Spanish HIV Research Network for excellence) RD06/0006. Susana Monge has been granted by Rio Ortega program (Nº exp CM10/00154).

Author details
[1]Red de Investigación en Sida, Centro Nacional de Epidemiología, Instituto de Salud Carlos III, Avda. Monforte de Lemos, Madrid 5 28029, Spain. [2]CIBER de Epidemiología y Salud Pública (CIBERESP), Madrid, Spain. [3]Hospital Universitario Gregorio Marañón, Madrid, Spain. [4]Hospital Universitario Carlos III, Madrid, Spain. [5]Hospital Universitario San Cecilio, Granada, Spain. [6]Hospital Universitario Virgen de la Victoria, Malaga, Spain. [7]Hospital Universitario Germans Trias I Pujol, Badalona, Spain. [8]Hospital Universitario Ramon y Cajal, Madrid, Spain.

Received: 19 February 2013 Accepted: 6 August 2013
Published: 20 August 2013

References

1. Krentz HB, Kliewer G, Gill MJ: **Changing mortality rates and causes of death for HIV-infected individuals living in Southern Alberta, Canada from 1984 to 2003.** *HIV Med* 2005, **6**:99–106.
2. Palella FJ Jr, Baker RK, Moorman AC, Chmiel JS, Wood KC, Brooks JT, Holmberg SD: **Mortality in the highly active antiretroviral therapy era: changing causes of death and disease in the HIV outpatient study.** *J Acquir Immune Defic Syndr* 2006, **43**:27–34.
3. Jaggy C, Von Overbeck J, Ledergerber B, Schwarz C, Egger M, Rickenbach M, Furrer HJ, Telenti A, Battegay M, Flepp M, *et al*: **Mortality in the Swiss HIV Cohort Study (SHCS) and the Swiss general population.** *Lancet* 2003, **362**:877–878.
4. Lewden C, Chene G, Morlat P, Raffi F, Dupon M, Dellamonica P, Pellegrin JL, Katlama C, Dabis F, Leport C: **HIV-infected adults with a CD4 cell count greater than 500 cells/mm3 on long-term combination antiretroviral therapy reach same mortality rates as the general population.** *J Acquir Immune Defic Syndr* 2007, **46**:72–77.
5. Lohse N, Hansen AB, Pedersen G, Kronborg G, Gerstoft J, Sorensen HT, Vaeth M, Obel N: **Survival of persons with and without HIV infection in Denmark, 1995–2005.** *Ann Intern Med* 2007, **146**:87–95.
6. Van Sighem A, Danner S, Ghani AC, Gras L, Anderson RM, De Wolf F: **Mortality in patients with successful initial response to highly active antiretroviral therapy is still higher than in non-HIV-infected individuals.** *J Acquir Immune Defic Syndr* 2005, **40**:212–218.
7. **Causes of death in HIV-1-infected patients treated with antiretroviral therapy, 1996–2006: collaborative analysis of 13 HIV cohort studies.** *Clin Infect Dis* 2010, **50**:1387–1396.
8. Aldaz P, Moreno-Iribas C, Egues N, Irisarri F, Floristan Y, Sola-Boneta J, Martinez-Artola V, Sagredo M, Castilla J: **Mortality by causes in HIV-infected adults: comparison with the general population.** *BMC Public Health* 2011, **11**:300.

9. Martinez E, Milinkovic A, Buira E, De Lazzari E, Leon A, Larrousse M, Lonca M, Laguno M, Blanco JL, Mallolas J, *et al*: **Incidence and causes of death in HIV-infected persons receiving highly active antiretroviral therapy compared with estimates for the general population of similar age and from the same geographical area.** *HIV Med* 2007, **8**:251–258.
10. Caro-Murillo AM, Castilla J, Perez-Hoyos S, Miro JM, Podzamczer D, Rubio R, Riera M, Viciana P, Lopez AJ, Iribarren JA, *et al*: **Spanish cohort of naive HIV-infected patients (CoRIS): rationale, organization and initial results.** *Enferm Infecc Microbiol Clin* 2007, **25**:23–31.
11. Sobrino-Vegas P, Gutierrez F, Berenguer J, Labarga P, Garcia F, Alejos-Ferreras B, Munoz MA, Moreno S, Del Amo J: **[The cohort of the spanish hiv research network (coris) and its associated biobank; organizational issues, main findings and losses to follow-up].** *Enferm Infecc Microbiol Clin* 2011, **29**:645–653.
12. Zwahlen M, Harris R, May M, Hogg R, Costagliola D, De Wolf F, Gill J, Fatkenheuer G, Lewden C, Saag M, *et al*: **Mortality of HIV-infected patients starting potent antiretroviral therapy: comparison with the general population in nine industrialized countries.** *Int J Epidemiol* 2009, **38**:1624–1633.
13. Hernando V, Perez-Cachafeiro S, Lewden C, Gonzalez J, Segura F, Oteo JA, Rubio R, Dalmau D, Moreno S, Amo JD: **All-cause and liver-related mortality in HIV positive subjects compared to the general population: differences by HCV co-infection.** *J Hepatol* 2012, **57**:743–751.
14. Berenguer J, Alejos B, Hernando V, Viciana P, Salavert M, Santos I, Gomez-Sirvent JL, Vidal F, Portilla J, Del AJ: **Trends in mortality according to hepatitis C virus serostatus in the era of combination antiretroviral therapy.** *AIDS* 2012, **26**(17):2241–2246.
15. Chen TY, Ding EL, Seage Iii GR, Kim AY: **Meta-analysis: increased mortality associated with hepatitis C in HIV-infected persons is unrelated to HIV disease progression.** *Clin Infect Dis* 2009, **49**:1605–1615.
16. Gutierrez F, Padilla S, Masia M, Iribarren JA, Moreno S, Viciana P, Munoz L, Gomez Sirvent JL, Vidal F, Lopez-Aldeguer J, *et al*: **Clinical outcome of HIV-infected patients with sustained virologic response to antiretroviral therapy: long-term follow-up of a multicenter cohort.** *PLoS One* 2006, **1**:e89.
17. Perez CS, Del Amo J, Iribarren JA, Salavert LM, Gutierrez F, Moreno A, Labarga P, Pineda JA, Vidal F, Berenguer J, *et al*: **Decrease in serial prevalence of coinfection with hepatitis C virus among HIV-infected patients in Spain, 1997–2006.** *Clin Infect Dis* 2009, **48**:1467–1470.
18. Ray M, Logan R, Sterne JA, Hernandez-Diaz S, Robins JM, Sabin C, Bansi L, Van Sighem A, De Wolf F, Costagliola D, *et al*: **The effect of combined antiretroviral therapy on the overall mortality of HIV-infected individuals.** *AIDS* 2010, **24**:123–137.
19. Smit C, Geskus R, Walker S, Sabin C, Coutinho R, Porter K, Prins M: **Effective therapy has altered the spectrum of cause-specific mortality following HIV seroconversion.** *AIDS* 2006, **20**:741–749.
20. Sobrino-Vegas P, Garcia-San Miguel L, Caro-Murillo AM, Miro JM, Viciana P, Tural C, Saumoy M, Santos I, Sola J, Del Amo J, *et al*: **Delayed diagnosis of HIV infection in a multicenter cohort: prevalence, risk factors, response to HAART and impact on mortality.** *Curr HIV Res* 2009, **7**:224–230.
21. Lewden C, Bouteloup V, De WS, Sabin C, Mocroft A, Wasmuth JC, Van SA, Kirk O, Obel N, Panos G, *et al*: **All-cause mortality in treated HIV-infected adults with CD4 >/=500/mm3 compared with the general population: evidence from a large European observational cohort collaboration.** *Int J Epidemiol* 2012, **41**:433–445.
22. Ministry of Health SPaE: *Patterns of mortality in Spain, 2008*; 2011. http://www.mspsi.gob.es/estadEstudios/estadisticas/estadisticas/estMinisterio/mortalidad/mortalidad.htm.
23. Centro Nacional de Epidemiología: *Vigilancia Epidemiológica del VIH/SIDA en España*; 2011. http://www.isciii.es/ISCIII/es/contenidos/fd-servicios-cientifico-tecnicos/fd-vigilancias-alertas/fd-enfermedades/Informe_VIH-sida_Junio_2011.pdf.

doi:10.1186/1471-2334-13-382
Cite this article as: Hernando *et al.*: All-cause mortality in the cohorts of the Spanish AIDS Research Network (RIS) compared with the general population: 1997–2010. *BMC Infectious Diseases* 2013 **13**:382.

10.3 Annex 3: Centres and researchers involved in analysed cohorts

10.3.1 Centres and researchers involved in period 1997-2008

Executive committee: Juan Berenguer, Julia del Amo, Federico García, Félix Gutiérrez, Pablo Labarga, Santiago Moreno y María Ángeles Muñoz.

Fieldwork, data management and analysis: Ana María Caro-Murillo, Paz Sobrino Vegas, Santiago Pérez-Cachafeiro, Victoria Hernando Sebastián, Belén Alejos Ferreras, Débora Álvarez, Susana Monge, Inma Jarrín, Mónica Trastoy.

BioBank: M Ángeles Muñoz-Fernández, Isabel García-Merino, Coral Gómez Rico, Jorge Gallego de la Fuente and Almudena García Torre.

Participating centres:

Hospital General Universitario de Alicante (Alicante): Joaquíçn Portilla Sogorb, Esperanza Merino de Lucas, Sergio Reus Bañuls, Vicente Boix Martínez, Livia Giner Oncina, Carmen Gadea Pastor, Irene Portilla Tamarit, Patricia Arcaina Toledo.

Hospital Universitario de Canarias (Santa Cruz de Tenerife): Juan Luis Gómez Sirvent, Patricia Rodríguez Fortúnez, María Remedios Alemán Valls, María del Mar Alonso Socas, Ana María López Lirola, María Inmaculada Hernández Hernández, Felicitas Díaz-Flores.

Hospital Carlos III (Madrid): Vicente Soriano, Pablo Labarga, Pablo Barreiro, Carol Castañares, Pablo Rivas, Andrés Ruiz, Francisco Blanco, Pilar García, Mercedes de Diego.

Hospital Universitario Central de Asturias (Oviedo): Victor Asensi, Eulalia Valle, José Antonio Cartón

Hospital Clinic (Barcelona): José M. Miró, María López-Dieguez, Christian Manzardo, Laura Zamora, Iñaki Pérez, Mª Teresa García, Carmen Ligero, José Luis Blanco, Felipe García-Alcaide, Esteban Martínez, Josep Mallolas, José M. Gatell

Hospital Doce de Octubre (Madrid): Rafael Rubio, Federico Pulido, Silvana Fiorante, Jara Llenas, Violeta Rodríguez, Mariano Matarranz.

Hospital Donostia (San Sebastián): José Antonio Iribarren, çJulio Arrizabalaga, María José Aramburu, Xabier Camino, Francisco Rodríguez-Arrondo, Miguel Ángel von Wichmann, Lidia Pascual Tomé, Miguel Ángel Goenaga, Mª Jesús Bustinduy, Harkaitz Azkune Galparsoro

Hospital General Universitario de Elche (Elche): Félix Gutiérrez, Mar Masiá, José Manuel Ramos, Sergio Padilla, Andrés Navarro, Fernando Montolio, Yolanda Peral, Catalina Robledano García

Hospital Germans Trías i Pujol (Badalona): Bonaventura Clotet, Cristina Tural, Lidia Ruiz, Cristina Miranda, Roberto Muga, Jordi Tor, Arantza Sanvisens

Hospital Gregorio Marañón (Madrid): Juan Berenguer, Juan Carlos López Bernaldo de Quirós, Pilar Miralles, Jaime Cosín Ochaíta, Matilde Sánchez Conde, Isabel Gutiérrez Cuellar, Margarita Ramírez Schacke, Belén Padilla Ortega, Paloma Gijón Vidaurreta.

Hospital Universitari de Tarragona Joan XXIII, IISPV, Universitat Rovira i Virgili (Tarragona): Francesc Vidal, Joaquín Peraire, Consuelo Viladés, Sergio Veloso, Montserrat Vargas, Miguel López-Dupla, Montserrat Olona, Joan-Josep Sirvent, Alba Aguilar, Antoni Soriano.

Hospital Universitario La Fe (Valencia): José López Aldeguer, Marino Blanes Juliá, José Lacruz Rodrigo, Miguel Salavert, Marta Montero, Eva Calabuig, Sandra Cuéllar.

Hospital Universitário La Paz (Madrid): Juan González García, Ignacio Bernardino de la Serna, José María Peña Sánchez de Rivera, Marta Mora Rillo, José Ramón Arribas López, María Luisa Montes Ramírez, José Francisco Pascual Pareja, Blanca Arribas, Juan Miguel Castro, Fco Javier Zamora Vargas, Ignacio Pérez Valero.

Hospital de la Princesa (Madrid): Ignacio de los Santos, Jesús Sanz Sanz, Johana Rodríguez, Ana Salas Aparicio, Cristina Sarriá Cepeda.

Hospital San Pedro-CIBIR (Logroño): José Antonio Oteo, José Ramón Blanco, Valvanera Ibarra, Luis Metola, Mercedes Sanz, Laura Pérez-Martínez

Hospital San Pedro II (Logroño): Javier Pinilla Moraza

Hospital Universitario Mutua de Terrassa (Terrassa): David Dalmau, Angels Jaén Manzanera, Mireia Cairó Llobell, Daniel Irigoyen Puig, Laura Ibáñez, Queralt Jordano Montañez, Mariona Xercavins Valls, Javier Martinez-Lacasa, Pablo Velli, Roser Font.

Hospital de Navarra (Pamplona): Julio Sola Boneta, Javier Uriz, Jesús Castiello, Jesús Reparaz, María Jesús Arraiza, Carmen Irigoyen, David Mozas,

Hospital Parc Taulí (Sabadell): Ferrán Segura, María José Amengual, Eva Penelo, Gemma Navarro, Montserrat Sala, Manuel Cervantes, Valentín Pineda.

Hospital Ramón y Cajal (Madrid): Santiago Moreno, José Luis Casado, Fernando Dronda, Ana Moreno, María Jesús Pérez Elías, Dolores López, Carolina Gutiérrez, Beatriz Hernández, María Pumares, Paloma Martí.

Hospital Reina Sofía (Murcia): Alfredo Cano Sánchez, EnRICue Bernal Morell, Ángeles Muñoz Pérez

Hospital San Cecilio (Granada): Federico García García, José Hernández Quero, Alejandro Peña Monje, Leopoldo Muñoz Medina, Jorge Parra Ruiz.

Centro Sanitario Sandoval (Madrid): Jorge Del Romero Guerrero, Carmen Rodríguez Martín, Teresa Puerta López, Juan Carlos Carrió Montiel, Cristina González.

Hospital Universitario Santiago de Compostela (Santiago de Compostela): Antonio Antela, Arturo Prieto, Elena Losada

Hospital Son Dureta (Palma de Mallorca): Melchor Riera, Javier Murillas, Maria Peñaranda, Maria Leyes, Mª Angels Ribas, Antoni Campins, Concepcion Villalonga.

Hospital Universitario de Valme (Sevilla): Juan Antonio Pineda, Eva Recio Sánchez, Fernando Lozano de León, Juan Macías, José del Valle, Jesús Gómez-Mateos, Rosario Mata.

Hospital Virgen de la Victoria (Málaga): Jesús Santos González, Manuel Márquez Solero, Isabel Viciana Ramos, Rosario Palacios Muñoz

Hospital Universitario Virgen del Rocío (Sevilla): Pompeyo Viciana, Manuel Leal, Luis Fernando López-Cortés, Mónica Trastoy

10.3.2 Centres and researchers involved in period 1997-2010

Executive committee: Juan Berenguer, Julia del Amo, Federico García, Félix Gutiérrez, Pablo Labarga, Santiago Moreno y María Ángeles Muñoz.

Fieldwork, data management and analysis: Paz Sobrino Vegas, Victoria Hernando Sebastián, Belén Alejos Ferreras, Débora Álvarez del Arco, Susana Monge Corella, Inma Jarrín Vera, Adela Castelló.

BioBank: M Ángeles Muñoz-Fernández, Isabel García-Merino, Coral Gómez Rico, Jorge Gallego de la Fuente and Almudena García Torre.

Participating centres:

Hospital General Universitario de Alicante (Alicante): Joaquín Portilla Sogorb, Esperanza Merino de Lucas, Sergio Reus Bañuls, Vicente Boix Martínez, Livia Giner Oncina, Carmen Gadea Pastor, Irene Portilla Tamarit, Patricia Arcaina Toledo.

Hospital Universitario de Canarias (Santa Cruz de Tenerife): Juan Luis Gómez Sirvent, Patricia Rodríguez Fortúnez, María Remedios Alemán Valls, María del Mar Alonso Socas, Ana María López Lirola, María Inmaculada Hernández Hernández, Felicitas Díaz-Flores.

Hospital Carlos III (Madrid): Vicente Soriano, Pablo Labarga, Pablo Barreiro, Pablo Rivas, Francisco Blanco, Luz Martín Carbonero, Eugenia Vispo, Carmen Solera.

Hospital Universitario Central de Asturias (Oviedo): Victor Asensi, Eulalia Valle, José Antonio Cartón

Hospital Clinic (Barcelona): José M. Miró, María López-Dieguez, Christian Manzardo, Laura Zamora, Iñaki Pérez, Mª Teresa García, Carmen Ligero, José Luis Blanco, Felipe García-Alcaide, Esteban Martínez, Josep Mallolas, José M. Gatell.

Hospital Doce de Octubre (Madrid): Rafael Rubio, Federico Pulido, Silvana Fiorante, Jara Llenas, Violeta Rodríguez, Mariano Matarranz.

Hospital Donostia (San Sebastián): José Antonio Iribarren, Julio Arrizabalaga, María José Aramburu, Xabier Camino, Francisco Rodríguez-Arrondo, Miguel Ángel von Wichmann, Lidia Pascual Tomé, Miguel Ángel Goenaga, Mª Jesús Bustinduy, Harkaitz Azkune Galparsoro.

Hospital General Universitario de Elche (Elche): Félix Gutiérrez, Mar Masiá, Cristina López Rodríguez, Sergio Padilla, Andrés Navarro, Fernando Montolio, Catalina Robledano García, Joan Gregori Colomé.

Hospital Germans Trías i Pujol (Badalona): Bonaventura Clotet, Cristina Tural, Lidia Ruiz, Cristina Miranda, Roberto Muga, Jordi Tor, Arantza Sanvisens.

Hospital General Universitario Gregorio Marañón (Madrid): Juan Berenguer, Juan Carlos López Bernaldo de Quirós, Pilar Miralles, Jaime Cosín Ochaíta, Isabel Gutiérrez Cuellar, Margarita Ramírez Schacke, Belén Padilla Ortega, Paloma Gijón Vidaurreta, Ana Carrero Gras, Teresa Aldamiz-Echevarría Lois y Francisco Tejerina Picado.

Hospital Universitari de Tarragona Joan XXIII, IISPV, Universitat Rovira i Virgili (Tarragona): Francesc Vidal, Joaquín Peraire, Consuelo Viladés, Sergio Veloso, Montserrat Vargas, Miguel López-Dupla, Montserrat Olona, Alba Aguilar, Joan Josep Sirvent, Verónica Alba, Olga Calavia.

Hospital Universitario La Fe (Valencia): José López Aldeguer, Marino Blanes Juliá, José Lacruz Rodrigo, Miguel Salavert, Marta Montero, Eva Calabuig, Sandra Cuéllar.

Hospital Universitario La Paz (Madrid): Juan González García, Ignacio Bernardino de la Serna, José Ramón Arribas López, María Luisa Montes Ramírez, Jose Mª Peña, Blanca Arribas, Juan Miguel Castro, Fco Javier Zamora Vargas, Ignacio Pérez Valero, Miriam Estébanez, Silvia García Bujalance, Marta Díaz.

Hospital de la Princesa (Madrid): Ignacio de los Santos, Jesús Sanz Sanz, Ana Salas Aparicio, Cristina Sarriá Cepeda.

Hospital San Pedro-CIBIR (Logroño): José Antonio Oteo, José Ramón Blanco, Valvanera Ibarra, Luis Metola, Mercedes Sanz, Laura Pérez-Martínez.

Hospital San Pedro II (Logroño): Javier Pinilla Moraza.

Hospital Universitario Mutua de Terrassa (Terrassa): David Dalmau, Angels Jaén Manzanera, Mireia Cairó Llobell, Daniel Irigoyen Puig, Laura Ibáñez, Queralt Jordano Montañez, Mariona Xercavins Valls, Javier Martinez-Lacasa, Pablo Velli, Roser Font.

Hospital de Navarra (Pamplona): María Rivero, Marina Itziar Casado, Jorge Alberto Díaz González, Javier Uriz, Jesús Repáraz, Carmen Irigoyen, María Jesús Arraiza.

Hospital Parc Taulí (Sabadell): Ferrán Segura, María José Amengual, Eva Penelo, Gemma Navarro, Montserrat Sala, Manuel Cervantes, Valentín Pineda.

Hospital Ramón y Cajal (Madrid): Santiago Moreno, José Luis Casado, Fernando Dronda, Ana Moreno, María Jesús Pérez Elías, Dolores López, Carolina Gutiérrez, Beatriz Hernández, María Pumares, Paloma Martí.

Hospital Reina Sofía (Murcia): Alfredo Cano Sánchez, Enrique Bernal Morell, Ángeles Muñoz Pérez.

Hospital San Cecilio (Granada): Federico García García, José Hernández Quero, Alejandro Peña Monje, Leopoldo Muñoz Medina, Jorge Parra Ruiz.

Centro Sanitario Sandoval (Madrid): Jorge Del Romero Guerrero, Carmen Rodríguez Martín, Teresa Puerta López, Juan Carlos Carrió Montiel, Cristina González, Mar Vera.

Hospital Universitario Santiago de Compostela (Santiago de Compostela): Antonio Antela, Arturo Prieto, Elena Losada.

Hospital Son Espases (Palma de Mallorca): Melchor Riera, Javier Murillas, Maria Peñaranda, Maria Leyes, Mª Angels Ribas, Antoni Campins, Concepcion Villalonga, Carmen Vidal.

Hospital Universitario de Valme (Sevilla): Juan Antonio Pineda, Eva Recio Sánchez, Fernando Lozano de León, Juan Macías, José del Valle, Jesús Gómez-Mateos.

Hospital Virgen de la Victoria (Málaga): Jesús Santos González, Manuel Márquez Solero, Isabel Viciana Ramos, Rosario Palacios Muñoz.

Hospital Universitario Virgen del Rocío (Sevilla): Pompeyo Viciana, Manuel Leal, Luis Fernando López-Cortés, Mónica Trastoy.

10.3.3 Centres and researchers involved in period 2004-2014

Executive committee: Santiago Moreno, Julia del Amo, David Dalmau, Maria Luisa Navarro, Maria Isabel González, Jose Luis Blanco, Federico Garcia, Rafael Rubio, Jose Antonio Iribarren, Félix Gutiérrez, Francesc Vidal, Juan Berenguer, Juan González.

Fieldwork, data management and analysis: Paz Sobrino, Victoria Hernando, Belén Alejos, Débora Álvarez, Inma Jarrín, Yaiza Rivero, Cristina González.

BioBanK HIV: M Ángeles Muñoz-Fernández, Isabel García-Merino, Coral Gómez Rico, Jorge Gallego de la Fuente y Almudena García Torre.

Participating centres:

1. Hospital General Universitario de Alicante (Alicante): Joaquín Portilla, Esperanza Merino, Sergio Reus, Vicente Boix, Livia Giner, Carmen Gadea, Irene Portilla, Maria Pampliega, Marcos Díez, Juan Carlos Rodríguez, Jose Sánchez-Payá.
2. Hospital Universitari de Bellvitge (Hospitalet de Llobregat): Daniel Podzamczer, Elena Ferrerm Arkaitz Imaz, Evan Van Den Eyncle, Silvana Di Yacovo, Maria Sumoy.
3. Hospital Universitario de Canarias (Santa Cruz de Tenerife): Juan Luis Gómez, Jehovana Hernández, María Remedios Alemán, María del Mar Alonso, María Inmaculada Hernández, Felicitas Díaz-Flores, Dácil García, Ricardo Pelazas.
4. Hospital Carlos III (Madrid): Vicente Soriano, Pablo Labarga, Pablo Barreiro, Pablo Rivas, Francisco Blanco, Luz Martín Carbonero, Eugenia Vispo, Carmen Solera.
5. Hospital Universitario Central de Asturias (Oviedo): Victor Asensi, Eulalia Valle, José Antonio Cartón
6. Hospital Clínico San Carlos (Madrid): Vicente Estrada Pérez, Maria Jesus Téllez Molina, Jorge Vergas García, Elisa Pérez-Cecila Carrera.
7. Hospital Doce de Octubre (Madrid): Rafael Rubio, Federico Pulido, Otilia Bisbal, Mariano Matarranz, Maria Lagarde, Rafael Rubio-Martín, Asunción Hernando, Angel Portillo.
8. Hospital Universitario Donostia (San Sebastián): José Antonio Iribarren, Julio Arrizabalaga, María José Aramburu, Xabier Camino, Francisco Rodríguez-Arrondo, Miguel Ángel von Wichmann, Lidia Pascual Tomé, Miguel Ángel Goenaga, Mª Jesús Bustinduy, Harkaitz Azkune Galparsoro. Maialen Ibarguren, Mirian Aguado.
9. Hospital General Universitario de Elche (Elche): Félix Gutiérrez, Mar Masiá, Cristina López, Sergio Padilla, Andrés Navarro, Fernando Montolio, Catalina Robledano, Joan Gregori Colomé, Araceli Adsuar, Rafael Pascual, Federico Carlos, Maravillas Martinez.
10. Hospital Germans Trías i Pujol (Badalona): Roberto Muga, Jordi Tor, Arantza Sanvisens
11. Hospital General Universitario Gregorio Marañón (Madrid): Juan Berenguer, Juan Carlos López Bernaldo de Quirós, Pilar Miralles, Isabel Gutiérrez, Margarita Ramírez, Belén Padilla, Paloma Gijón, Ana Carrero, Teresa Aldamiz-Echevarría, Francisco Tejerina, Francisco Jose Parras, Pascual Balsalobre, Cristina Diez.
12. Hospital Universitari de Tarragona Joan XXIII, IISPV, Universitat Rovira i Virgili (Tarragona): Francesc Vidal, Joaquín Peraire, Consuelo Viladés, Sergio Veloso, Montserrat Vargas, Miguel López-Dupla, Montserrat Olona, Alba Aguilar, Joan Josep Sirvent, Verónica Alba, Olga Calavia .
13. Hospital Universitario La Fe (Valencia): Marta Montero, José Lacruz, Marino Blanes, Eva Calabuig, Sandra Cuellar, José López, Miguel Salavert.
14. Hospital Universitario La Paz/IdiPaz (Madrid): Juan González, Ignacio Bernardino de la Serna, José Ramón Arribas, María Luisa Montes, Jose Mª Peña, Blanca Arribas, Juan Miguel Castro, Fco Javier Zamora, Ignacio Pérez, Miriam Estébanez, Silvia García, Marta Díaz,

Natalia Stella Alcáriz, Jesús Mingorance, Dolores Montero, Alicia González, Maria Isabel de José.

15. Hospital de la Princesa (Madrid): Ignacio de los Santos, Jesús Sanz, Ana Salas, Cristina Sarriá, Ana Gómez Berrocal.
16. Hospital San Pedro-CIBIR (Logroño): José Antonio Oteo, José Ramón Blanco, Valvanera Ibarra, Luis Metola, Mercedes Sanz, Laura Pérez-Martínez
17. Hospital Universitario Miguel Servet (Zaragoza): Ascensión Pascual, Carlos Ramos, Piedad Arazo, Desiré Gil.
18. Hospital Universitari Mutua de Terrassa (Terrassa): David Dalmau, Angels Jaén, Mireia Cairó, Daniel Irigoyen, Queralt Jordano, Mariona Xercavins, Javier Martinez-Lacasa, Pablo Velli, Roser Font, Montse Sanmartí, Laura Ibáñez.
19. Complejo Hospitalario de Navarra (Pamplona): María Rivero, Marina Itziar Casado, Jorge Alberto Díaz, Javier Uriz, Jesús Repáraz, Carmen Irigoyen, María Jesús Arraiza.
20. Hospital Parc Taulí (Sabadell): Ferrán Segura, María José Amengual, Gemma Navarro, Montserrat Sala, Manuel Cervantes, Valentín Pineda, Victor Segura, Marta Navarro, Esperanza Antón, Mª Merce Nogueras.
21. Hospital Ramón y Cajal (Madrid): Santiago Moreno, José Luis Casado, Fernando Dronda, Ana Moreno, María Jesús Pérez Elías, Dolores López, Carolina Gutiérrez, Nadia Madrid, Angel Lamas, Paloma Martí, Alberto de Diaz, Sergio Serrrano, Lucas Donat.
22. Hospital Reina Sofía (Murcia): Alfredo Cano, Enrique Bernal, Ángeles Muñoz.
23. Hospital San Cecilio (Granada): Federico García, José Hernández, Alejandro Peña, Leopoldo Muñoz, Jorge Parra, Marta Alvarez, Natalia Chueca, Vicente Guillot, David Vinuesa, Jose Angel Fernández.
24. Centro Sanitario Sandoval (Madrid): Jorge Del Romero, Carmen Rodríguez, Teresa Puerta, Juan Carlos Carrió, Cristina González, Mar Vera, Juan Ballesteros..
25. Hospital de la Santa Creu i Sant Pau (Barcelona): Pere Domingo, Mª Antonia Sambeat, Karuna Lamarca, Gracia Mateo, Mar Gutiérrez, Irene Fernández.
26. Hospital Universitario Santiago de Compostela (Santiago de Compostela): Antonio Antela, Elena Losada.
27. Hospital Son Espases (Palma de Mallorca): Melchor Riera, Maria Peñaranda, Maria Leyes, Mª Angels Ribas, Antoni A Campins, Carmen Vidal, Leire Gil, Francisco Fanjul, Carmen Marinescu.
28. Hospital Universitari Vall d´Hebron (Barcelona): Esteban Ribera
29. Hospital Virgen de la Victoria (Málaga): Jesús Santos, Manuel Márquez, Isabel Viciana, Rosario Palacios, Isabel Pérez, Carmen Maria González.
30. Hospital Universitario Virgen del Rocío (Sevilla): Pompeyo Viciana, Manuel Leal, Luis Fernando López-Cortés, Mónica Trastoy, Nuria Espinosa.
31. Hospital Universitario de Basurto (Bilbao): Josefa Muñoz, Miren Zuriñe Zubero, Josu Mirena Baraia-Etxaburu, Sofía Ibarra, Oscar Ferrero, Josefina López de Munain, Mª Mar Cámara. Iñigo López, Mireia de la Peña.
32. Hospital Universitario Infanta Sofía (San Sebastián de los Reyes): Inés Suárez-García, Eduardo Malmierca.
33. Hospital Universitario Costa del Sol (Marbella): Julián Olalla, Alfonso del Arco, Javier de la torre, José Luis Prada, Zaira Caracuel.
34. Hospital del Poniente (El Ejido): Ana Maria Lopez-Lirola, Ana Belén Lozano, Elisa Fernández, Inés Pérez, Juan Manuel Fernández

35. Hospital Universitario Santa Lucia (Cartagena): Onofre Juan Martínez, Francisco Jesús Vera, Lorena Martínez, Josefina García, Begoña Alcaraz, Amaya Jimeno.
36. INIBIC-Complejo Hospitalario Universitario de A Coruña (A Coruña): Eva Poveda, Berta Pernas, Álvaro Mena, Marta Grandal, Ángeles Castro, José D. Pedreira
37. Hospital Clínico Universitario Virgen de la Arrixaca (Murcia): Carlos Galera, Helena Albendin, Asunción Iborra, Antonio Moreno, Maria Angeles Campillo, Asunción vidal.
38. Hospital Marina Baixa (Villajoyosa): Concha Amador, Francisco Pasquau, Javier Ena, Concha Benito, Vicenta Fenoll.
39. Complejo Hospitalario de Jaén (Jaén): Mohamed Omar Mohamed-Balghata, Maria Amparo Gómez.
40. Hospital San Agustín de Aviles (Avilés): Miguel Alberto de Zarraga, Maria Eugenia Rivas.
41. Fundación Jiménez Diaz (Madrid): Miguel Cervero