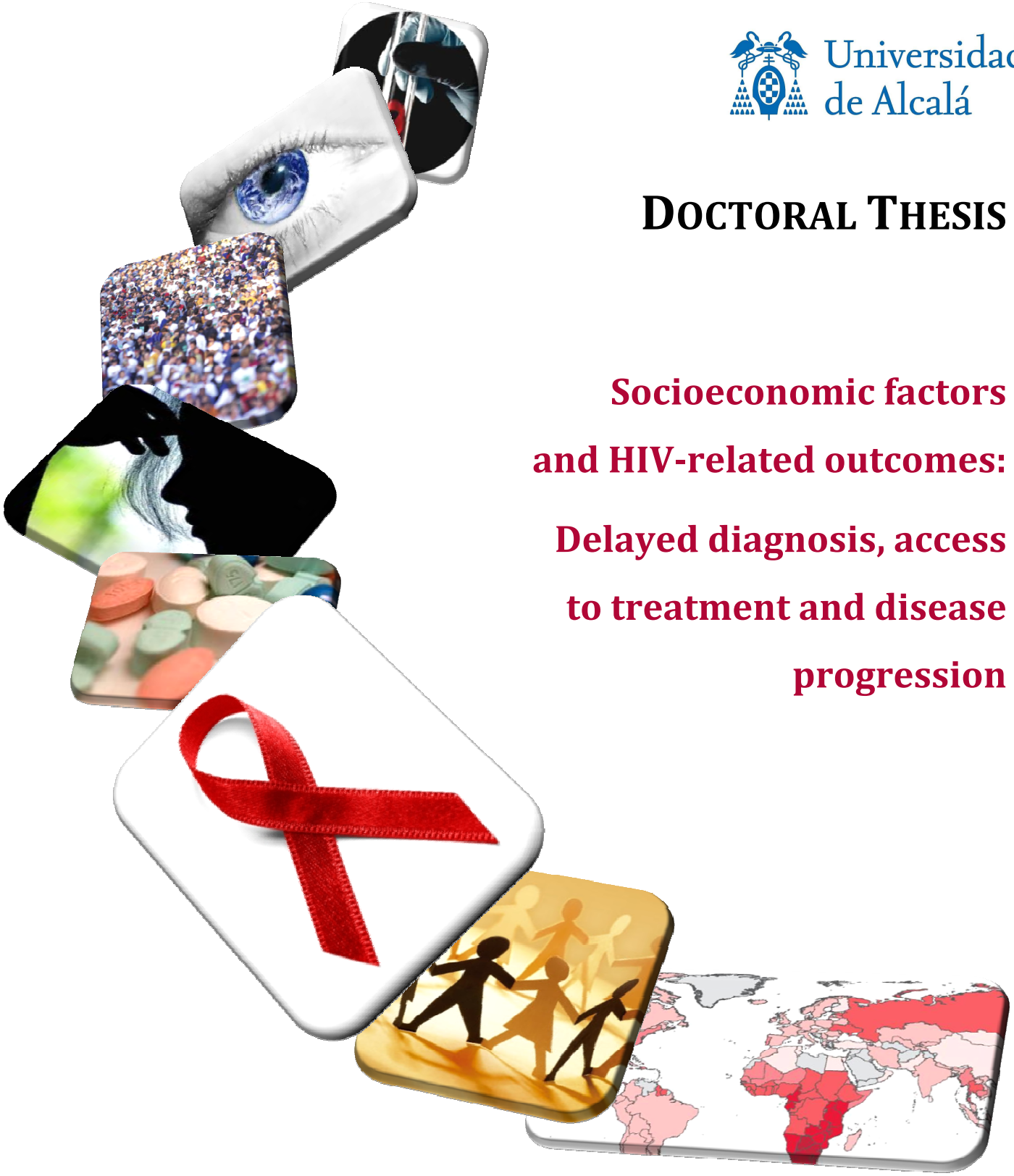


Universidad de Alcalá

DOCTORAL THESIS

Socioeconomic factors and HIV-related outcomes: Delayed diagnosis, access to treatment and disease progression

Susana Monge Corella

Directors: Francisco Bolúmar Montrull and Julia del Amo Valero

DEPARTAMENTO DE CIENCIAS SANITARIAS Y MÉDICOSOCIALES

UNIVERSIDAD DE ALCALÁ

2012

Universidad de Alcalá

DOCTORAL THESIS

SOCIOECONOMIC FACTORS AND HIV-RELATED OUTCOMES: DELAYED DIAGNOSIS, ACCESS TO TREATMENT AND DISEASE PROGRESSION

FACTORES SOCIOECONÓMICOS Y RESULTADOS EN SALUD RELACIONADOS CON EL VIH: RETRASO DIAGNÓSTICO, ACCESO A TRATAMIENTO Y PROGRESIÓN DE LA ENFERMEDAD

Susana Monge Corella

DEPARTAMENTO DE CIENCIAS SANITARIAS Y MÉDICOSOCIALES

UNIVERSIDAD DE ALCALÁ

MADRID, 2012

Directors:

Francisco Bolúmar Montrull

Julia del Amo Valero

To my family

TABLE OF CONTENTS

ACKNOWLEDGEMENTS

In the first place, I would like to thank all the patients who have participated in the cohorts in a voluntary and disinterested way, because it is them who make research possible and who most contribute to the better understanding of HIV disease and to the efforts to improve the lives of people living with HIV/AIDS.

I would also like to thank all the doctors, nurses and assistants from the various participating sites of the cohorts, who, with their daily work and dedication, collect all the information and send it to the respective coordinating centres, making it possible for these cohorts to exist and to have up-to-date, high quality data.

To all the clinicians and researchers who have contributed to the analysis that conform this doctoral thesis, and who are co-authors of the related papers and communications.

To my colleagues at the Centro Nacional de Epidemiología, the best team I will ever work with, who help me every day, from whom I continuously learn new things, and who have also become my friends: Paz, Vicky, Belén, Débora, Inma, Cristina, Sara; and those who have moved on and are no longer part of our team: Ana María and Adela.

To Paco, because he was the first person from whom I heard of Hazard Ratios and made me like epidemiology in the first place, because he believed in me and accepted to direct my thesis.

To Julia, for accepting me as a student and granting me the privilege of working under her supervision during these years, for always having time for me, for teaching

me, for giving me so many opportunities and new challenges and for her constant support and leadership.

To my partners in the discovery of epidemiology and public health whose support has been a cornerstone throughout the years: my colleagues during my medical residency, especially to Esther, Inés and Cristina; and others who shared important parts of this trip: Oriana and Nuria.

And finally, I would like to thank my family, specially my parents, who have been a constant example in life; and my husband Borja, who supports me and makes me tea, and whose unique understanding of life is something worth waking up to every morning.

LIST OF ABBREVIATIONS

ADI	AIDS Defining Illness
AIDS	Acquired Immunodeficiency Syndrome
ART	Antiretroviral Therapy
CD4	T4 Lymphocytes
CoRIS	Spanish AIDS Research Network adult cohort
DD	Delayed Diagnosis
ECDC	European Centre for Disease Control and Prevention
GEMES	Spanish Multicenter Study Group of Seroconverters
HAART	Highly Active Antiretroviral Therapy
HIV	Human Immunodeficiency Virus
IDU	Injecting Drug User/s
LAC	Latin America and the Spanish-speaking Caribbean
LTFU	Lost to Follow-Up
MSM	Men who have sex with men
NGO	Non-Governmental Organisation
NSP	Native Spanish
RIS	Spanish AIDS Research Network
SINIVIH	National information system on new HIV diagnosis
SSA	Sub-Saharan Africa
STD	Sexually Transmitted Diseases
TARGA	Tratamiento Antirretroviral de Gran Efectividad
PY	Person-Years
UNAIDS	Joint United Nations Programme on HIV/AIDS
VL	Viral Load

1. SUMMARY/*RESUMEN*

1.1. Summary (English version)

As in other industrialised countries, the HIV epidemic in Spain is not generalised but concentrates predominantly in vulnerable groups. Some of the factors that increase vulnerability are migrant and ethnic minority status, poor economic situation, and low socioeconomic level. Moreover, there is a risk that the most vulnerable populations may be the very ones who are also least likely to benefit from available treatment resources, contributing doubly to increased health inequalities.

Delayed diagnosis, worse access to treatment, limited effectiveness of available therapies and worse disease progression (i.e., higher incidence of AIDS and higher mortality) are ways in which inequalities may be expressed. The objective of this Doctoral Thesis is to study the effect of educational level and migration status over key HIV-related outcomes, in order to assess the impact of social inequalities over the course of HIV disease.

Data from three open, multicentre, prospective cohorts of patients over 13 years of age with confirmed HIV infection have been analysed. Two of them are seroconverter cohorts: the Madrid Seroconverter cohort recruited in the Centro Sanitario Sandoval, and the GEMES cohort. The third one, the CoRIS cohort, is a seroprevalent cohort of subjects naïve to antiretroviral treatment at entry.

A survival analysis was carried out. Kaplan-Meier function and Cox proportional hazards model were used to study overall survival and to evaluate its associated factors. For other outcomes, a competing risk methodology was used, including multiple decrements method and Fine & Gray regression. Censoring strategies varied between

cohorts and between different outcomes, and sensitivity analyses under different assumptions were performed.

Availability of HAART after 1996 showed great population effectiveness in reducing AIDS incidence and mortality in HIV infected patients in the Spanish context. However, people of low educational level benefit to a lesser extent, which reflects in a higher risk of progression to AIDS that becomes evident after HAART became available; and the inequality gap is further widened as treatments become more effective. Higher all-cause mortality is also evident, but is not affected by the availability of HAART. However, no difference is found on access to HAART by educational level.

Compared to native Spanish, migrants from Latin America and Sub-Saharan Africa experience a higher risk of delayed diagnosis of HIV infection, specially the younger subjects, probably reflecting the existence of barriers to HIV testing. However no meaningful delays in treatment initiation are identified, showing no further barriers for migrants once they have accessed the system. Migrants from Latin America and Sub-Saharan Africa aged between 35 to 50 years progress faster to an AIDS diagnosis, at the expense of a higher incidence rate of tuberculosis. In contrast, mortality in these groups tends to be lower, compatible with the healthy migrant effect. Immunological and virological response to antiretroviral treatment is poorer for Sub-Saharan Africans, but not for Latin Americans.

These results are important to inform appropriate preventive and health care services as well as health programmes and policies to better respond to challenges posed by social inequalities and to reduce their impact on health.

1.2. Resumen (versión en Español)

Al igual que en otros países industrializados, la epidemia del VIH en España no está generalizada, sino predominantemente concentrada en grupos vulnerables. Factores que aumentan la vulnerabilidad son el estatus de migrante y minorías étnicas, la mala situación económica y el bajo nivel socioeconómico. Además, existe el riesgo de que precisamente las poblaciones más vulnerables se beneficien menos de los recursos disponibles, contribuyendo doblemente a aumentar las desigualdades en salud.

Estas desigualdades pueden manifestarse como mayor retraso diagnóstico, barreras para el acceso al tratamiento, efectividad limitada del mismo y peor progresión de la infección, es decir, una mayor incidencia de SIDA y de mortalidad. El objetivo de esta Tesis Doctoral es estudiar el efecto del nivel educativo y el estatus migratorio sobre resultados clave relacionados con el VIH, con el fin de evaluar el impacto de las desigualdades sociales sobre el curso de la infección por VIH.

Se han analizado datos de tres cohortes abiertas, multicéntricas y prospectivas de sujetos mayores de 13 años con infección VIH confirmada. Dos de ellas son cohortes de seroconvertores: la cohorte de seroconvertores de Madrid, reclutada en el Centro Sanitario Sandoval, y la cohorte GEMES. La tercera, la cohorte CoRIS, es una cohorte de sujetos seroprevalentes naïve a tratamiento antirretroviral al reclutamiento.

Se realizó un análisis de supervivencia. Para la supervivencia global y los factores asociados se usaron la función de Kaplan-Meier y el modelo de riesgos proporcionales de Cox. Para otros eventos, se usaron métodos de riesgos competitivos, incluyendo el método de decrementos múltiples y la regresión de Fine&Gray. La estrategia

de censura varió en función de la cohorte y del evento y se realizaron análisis de sensibilidad con distintas asunciones.

La disponibilidad de TARGA después de 1996 ha mostrado gran efectividad poblacional en la reducción de la incidencia de SIDA y la mortalidad de las personas infectadas por el VIH en el contexto español. Sin embargo, las personas con bajo nivel educativo se benefician en menor medida, lo que se refleja en mayor riesgo de progresión a SIDA tras la disponibilidad de TARGA; y la brecha de desigualdad aumenta conforme los tratamientos son más eficaces. También se evidencia una mayor mortalidad global, que no se ve afectada por la disponibilidad de TARGA. Sin embargo, no se encuentran diferencias en el acceso al tratamiento en función del nivel educativo.

En comparación con los españoles, los migrantes de América Latina y África Subsahariana presentan mayor riesgo de retraso diagnóstico, especialmente los más jóvenes, reflejando probablemente la existencia de barreras a la prueba del VIH. Sin embargo, no se han detectado retrasos significativos en el inicio del tratamiento, mostrando la ausencia de barreras adicionales una vez han accedido al sistema. Los migrantes de ambos orígenes entre 35 y 50 años progresan más rápidamente a SIDA, debido a la tuberculosis. Al contrario, la mortalidad tiende a ser menor, compatible con el efecto del inmigrante sano. La respuesta inmunológica y virológica al tratamiento es peor en las personas de África Subsahariana, pero no en los Latinoamericanos.

Estos resultados son importantes para orientar los servicios preventivos y asistenciales, así como los programas y políticas de salud, para responder mejor a los retos planteados por las desigualdades sociales y para ser capaces de reducir su impacto en salud.

2. BACKGROUND

2.1. Epidemiology of HIV infection

2.1.1. A look at the global HIV epidemic

After 30 years of ongoing HIV/AIDS epidemic and according to 2010 UNAIDS data [1, 2], the overall growth of the global AIDS epidemic appears to have stabilized in an estimated prevalence of 0.8%. The annual number of new HIV infections has fallen by 19% since 1999, the year in which it is thought that the epidemic peaked, and there are fewer AIDS-related deaths due to the significant scaling up of antiretroviral therapy over the past few years. However, access to treatment still needs to be expanded: of the estimated 15 million people living with HIV in low- and middle-income countries who need treatment, only 5.2 million have access to it.

Despite the optimistic global figures, HIV/AIDS pandemic reflects the socio-economic and health inequalities between industrialised and non-industrialised countries very clearly. In 2009, 68% of the 33.3 million people infected with HIV worldwide lived in Sub-Saharan Africa, and over 97% of new HIV infections took place in low- and middle-income countries.

Sub-Saharan Africa (SSA) is the area most severely hit by the HIV/AIDS pandemic, where prevalence rates in adult population aged 15–49 were estimated to be 5% in 2009. Of the 1.8 million people who died of AIDS that year, 1.3 million lived in SSA where universal access to HAART is still far from being achieved. The epidemic in SSA is largely a heterosexual epidemic, with an increasing number of women being infected, and a significant number of vertically infected children.

The Caribbean, with an adult HIV prevalence of 1%, is the second most affected region. There is, nevertheless, substantial heterogeneity within the islands, the Dominican Republic being the most affected one. HIV in the Caribbean is predominantly transmitted through heterosexual intercourse and shows characteristics of a generalised epidemic, very much like Sub-Saharan Africa's. Moreover, the Caribbean is the only region outside SSA where HIV infected women outnumber men.

Eastern Europe and Central Asia rank third in worldwide HIV prevalence (of 0.8%), with an incidence rate in Eastern Europe over twice the rate of Western Europe. 90% of HIV-infected people in this region live in Russia or Ukraine. This region has experienced one of the most recent and explosive epidemics, largely driven by injecting drug use (IDU). The number of AIDS cases in this region also continues to increase[3].

In Latin America, HIV prevalence is around 0.5%, again with some differences between countries, being higher in countries surrounding the Caribbean. Transmission in the region is mainly through sex between men, although a meaningful transmission through injecting drug use was observed in the Southern Cone, mainly at the beginning of the AIDS epidemic. Consequently, the HIV/AIDS epidemic in this region is highly masculinised.

Among high-income regions, North America reported an adult prevalence of 0.5%, more than double than Western and Central Europe with a 0.2%. Epidemics within industrialised countries are transmitted mainly through sex between men, are highly masculinised and show a concentrated pattern. Large inequalities have been reported within countries: in 2008 the United States found HIV prevalence in Afro-Americans (1.8%) and Hispanics (0.6%) to be significantly higher than in Caucasians

(0.2%)[4]. In Europe, migrant populations, largely from Sub-Saharan Africa, represented a considerable and growing proportion of both the AIDS cases and HIV infections reported in the 27 EU countries plus Norway and Iceland during 1999–2006[5, 6].

2.1.2. The HIV epidemic in Spain: the 80's and the 90's

During the last two decades of the 20th Century, Spain suffered one of the largest HIV epidemics in Europe. Against a background of an injected heroin consumption boom that took place in the country in the late '70 and beginning of the '80, HIV expanded rapidly among drug users, new infections peaking between 1984 and 1987. The epidemic in men who have sex with men (MSM) was less pronounced, in contrast with the situation in other industrialized countries, where homosexual intercourse was the predominant route of transmission. Injecting drug users (IDU) were mainly young, sexually active individuals, so heterosexual transmission started to rise, together with cases of mother to child transmission. The implementation of harm reduction programmes for drug users, the abandonment of injection as the preferred route for drug uptake and other prevention measures managed to change the course of the epidemic in Spain, and transmission started to decrease in the late 80's[7].

AIDS cases rose accordingly until the mid 90's, when AIDS came to be the first cause of death in subjects from 25-44 years old[8], with incidence rates higher than anywhere else in Europe. The generalisation of HAART in 1997 completely changed the course of the HIV infection, drastically reducing the incidence of AIDS and the AIDS related mortality [9-11].

2.1.3. Current HIV epidemic figures and challenges

HIV epidemic profile in Spain has changed very much in the last decade, levelling with the epidemics of its neighbouring countries. New diagnoses have been continuously decreasing to reach 88.5 cases per million in 2010 and, from 1996 to 2010, an 83% reduction in AIDS cases has been observed. Due to the decrease in the number of new infections and the drop in mortality, the number of persons living with HIV in Spain has stabilised in 120,000 – 150,000 subjects, for an estimated prevalence of 3 per 1,000.

As for the current epidemic profile, the national information system on new HIV diagnosis (SINIVIH) has been widening its population coverage and it currently represents 71% of the Spanish population. According to SINIVIH data, a shift has been observed towards a predominantly sexually transmitted epidemic[12]. Sex between men is the main route of infection (46% of new diagnoses in 2010), and this is the only group where the rate of new diagnoses has been increasing in the last decade. This determines a highly masculinised epidemic, with an 80% of males. In contrast, nowadays IDU stand for as low as 6% of HIV diagnoses. The proportion of people originating from countries outside Spain has increased during the last decade to reach 38.4% of HIV diagnoses[12]. The fact that the majority of migrants originate from countries where the predominant routes of infection are sexual has also contributed to the sexualisation of the Spanish epidemic.

In recent years, however, the decline in AIDS cases has been less steep, probably because we are close to reaching the maximum benefits of HAART availability, but also because efficacy of treatment is dependent on starting treatment on time. Late

presentation of patients for care may limit effectiveness and makes early diagnosis of infection a cornerstone of HIV management. In 2010 in Spain, 45.4% of patients were diagnosed with less than 350 CD4 cells/μL[12], which is considered a late diagnosis[13] and involves a worse prognosis at the individual level.

But late diagnosis is also key from the public health perspective, as it has been estimated that undiagnosed HIV-positive subjects have a transmission rate 3.5 times higher than those who are aware of their HIV-infection[14]. In Spain, an approximate 30% of HIV infected persons are unaware of their sero-status[7]. All this makes diagnosis of HIV infection one of the main challenges for the HIV epidemic control nowadays.

Finally, as in other industrialised countries, HIV epidemic in Spain is not generalised, but concentrates predominantly in vulnerable groups. Some of the factors that increase vulnerability are migrant and ethnic minority status, poor economic situation, and low socioeconomic level [15-19]. Moreover, there is a risk that persons in the most vulnerable populations may be less likely to benefit from available treatment resources, contributing doubly to increased health inequalities. Reducing the impact of social inequalities is one of the challenges that industrialised countries need to respond to, for health in general, and for HIV in particular.

2.2. Socioeconomic vulnerability and HIV

2.2.1. Social class and HIV

2.2.1.1. Measuring social class in health research

Social class is a complex phenomenon and its characterisation and measurement has been a debated issue within health sciences. Different variables have been used as *proxies* or best representatives of social class.

Educational level is one of the most widely used variable because it has been shown to be a good approximation, especially with regard to psychosocial and behavioural factors associated with social class [20-22], and when the study populations is comprised of young adults [23, 24]. It also has the advantage of being easy to define and obtain, and of being relatively stable; whereas others, such as employment status, are more likely to change over time and are harder to collect in the context of observational cohorts. On the other hand, educational level faces some limitations that need to be accounted for, such as a narrower variation span compared to other variables like income or wealth, and its inability to capture further changes in economic well-being in adulthood [25-27]. Finally, education curriculum varies greatly from one country to another, making it difficult to establish comparisons between people of different origins and in different settings.

Socioeconomic characteristics, like social class, ethnic origin or marital status have been associated with a poorer health status and with a higher mortality in persons affected by chronic diseases [28-30]. Studies in Europe have shown a widening

inequality gap in mortality by educational level over the last decades, despite expansion of the welfare state [31, 32].

2.2.1.2. Social class and vulnerability to HIV

The HIV epidemic in the industrialised countries appears to have shifted towards the more socially vulnerable population [28]. Vulnerability to HIV can translate into higher incidence of infection, but it is also possible that persons in the most vulnerable populations may be less likely to benefit from available diagnostic and treatment resources. Delayed diagnosis, worse access to treatment, limited effectiveness of available therapies and worse disease progression (i.e., higher incidence of AIDS and higher mortality), are ways in which inequalities may be expressed.

Different studies have found association of low socioeconomic level with higher rates of mortality in HIV patients[33-38], although all were performed before 1999 and on seroprevalent subjects. Potential causes have been pointed out, including different disease stage at diagnosis and at treatment initiation, inequalities in access to HAART, the type of care and drugs prescribed, adherence to treatment or associated comorbidity, particularly in the case of latent tuberculosis and hepatitis C virus infections.

For all these reasons Dray-Spira et al.[28] argue that the benefits of the most recent advances in HIV infection management may differ among groups of patients independently of their HIV-related clinical characteristics, and this can lead to inequalities in the consequences of disease in settings where HAART is widely used.

A recent study conducted in the CoRIS cohort [39] has found association of low educational level with a higher risk of delayed diagnosis, of delayed HAART initiation and a worse virological and immunological response to treatment which could partially

explain the higher mortality. However, it found no differences in delayed initiation of treatment (defined as CD4 count below 200 cells/µl at HAART initiation) by educational level in the group of patients timely diagnosed, pointing HIV diagnosis as the bottleneck to timely access to treatment. So the existence of specific barriers to access treatment, independently of barriers to HIV diagnosis, in the low socioeconomic strata is still to be clarified. The few studies that have included educational level as an explanatory variable for progression to AIDS have yielded contradictory results [38, 40, 41].

2.2.2. Migration background and HIV

2.2.2.1. Introduction to the migration phenomenon

The United Nations defines a migrant as "any person who lives temporarily or permanently in a country where he or she was not born, and has acquired significant social ties to this country". However the definition of "migration" or "migrant" has been highly discussed and no authoritative definition has been put forward. This reflects the complexity of the multidimensional migratory phenomena and the difficulties that arise from the first moment when trying to study it.

According to figures from the International Organization for Migration[42] the number of migrants at the global level has increased significantly in the last decade, to reach an estimated 214 million people by 2010, 3.1% of the world's population and 9.5% of the European Union's. Poverty, the search for economic improvement and better standards of living are significant drivers. The majority of migrants originate from developing countries and move to developed countries whose economies de-

mand labour that cannot be met by the local workforce. In the last 10 years, the education sector has also become a major driver for mobility.

In the European Union, non-EU immigrants are made up fairly equally of citizens of European non-EU countries, Asian, American and African countries, with each comprising between 13% and 16% of the total[43]. However, their distribution between countries is uneven, with subjects from particular origins concentrating in specific countries. Factors related to the historical and economic relations between countries and the colonial past play an important role in shaping the specific migration profile in different countries [5, 44].

Compared to other European countries, Spain only started to receive significant economic immigration in the 21st Century, and the number of migrants living in Spain rose 5-fold from the year 2000 to 2010. Excluding people from the European Union and North America, migrants represent 8.8% of the Spanish population. The migration profile in Spain also differs from its neighbouring countries, with the largest share of economic migrants coming from Latin America and the Spanish-speaking Caribbean (LAC), who have virtually no language barriers and a smaller cultural distance than migrants from other regions. Sub-Saharan Africans (SSA) represent a small proportion of migrants in Spain[45], unlike other European countries in which SSA are one of the largest shares of economic migrants.

2.2.2.2. Socio-economic vulnerability of migrants

Migration involves a process that can be traumatic for individuals, even under the best of circumstances. The migrating person leaves behind a social network and a family, a full set of learned values, social rules and behavioural patterns, and arrives

into an alien culture which works under different coordinates. Moreover, this arrival usually brings along a loss of social status in the host country, stigma and discrimination. Socioeconomic conditions after arrival can also be precarious, especially in the group of undocumented migrants. Stressors like unemployment or poor working conditions, poverty, or legal circumstances can negatively affect migrants. This situation has been further worsened by the economic crisis during the last years. In Spain, at the end of 2007, 12.4% of immigrants were unemployed, compared with 7.9% of native-born Spaniards, but, by mid-2010, those figures had gone up to 30.2% and 18.1%, respectively[42].

All these circumstances make it necessary to study migrants from the perspective of social inequality, and to focus on the impact that this may have on their physical and psychological health. Migration involves several stages, each of which presents strategic opportunities for prevention and disease control. There is a pre-entry phase, where a migrant's health reflects the disease profile of his or her country of origin. There is then a transitional phase, where the process of moving, sometimes through intermediate countries, can influence a migrant's health[46]. Finally, there is a post-entry phase, where the process of adapting to working and living conditions in the host country can also influence a migrant's health. Migrants are often confronted with poor social support and discrimination in host countries which, together with language, cultural or legal barriers to health care can have a negative impact on their health, even in universal health care contexts, such as the Spanish one.

The health of migrant populations is a fundamental aspect for social integration, public health policies and health services planning and delivery, as expressed in recent EU policy documents[44].

2.2.2.3. Migrants as a most-at-risk population for HIV/AIDS

Migration and social exclusion make migrants highly vulnerable to HIV/AIDS and their related complications[47-49]. As an ECDC Report on HIV epidemiology in migrant populations in Europe points out, migrants, largely from Sub-Saharan Africa, represent a considerable and growing proportion of both AIDS cases and HIV infections reported in the 27 EU countries plus Norway and Iceland during 1999– 2006[5]. Although the proportions of migrants from Sub-Saharan Africa among heterosexual and mother-to-child HIV transmission reports are very high, a significant percentage of diagnoses in men who have sex with men are also made up of migrants, largely from Western Europe, Latin America and the Caribbean. The contribution of migrant populations to the AIDS and HIV epidemics is notably higher among female reports, highlighting the feminisation of the HIV/AIDS migrant epidemic in Europe, in contrast to the largely male autochthonous HIV epidemics.

Regarding Spain, several information sources document the increasing representation of migrant population in the Spanish HIV/AIDS epidemic[50]. According to SINIVIH data, in Spain in the year 2010, 38.4% of the new HIV diagnoses were in people who originated from countries different from Spain. The most common origin was Latin America and the Caribbean, comprising 21.4% of all diagnosis followed by an 8.0% of Sub-Saharan Africans[12], a unique pattern within the EU where SSA account by far for the largest proportion of new HIV diagnoses among migrants[3, 5, 50].The

percentage of migrants was larger among women and in heterosexually transmitted cases, where more than 50% of diagnoses were from non-Spanish origin. In each of the migrant groups the HIV epidemic greatly reflects the prevalent transmission patterns in their countries of origin. Latin Americans were infected through sex between men in 59% of cases, compared to Sub-Saharan Africans in whom 84% of new diagnoses had been infected through heterosexual intercourse [12].

Information on specific HIV prevalence in migrants in Spain is scarce. A study in patients who attended a sexually transmitted infections clinic to voluntarily take an HIV test, found higher HIV prevalence in migrants from Sub-Saharan Africa and Latin America, compared to Spanish population[51], as other studies have also pointed out[52]. This same study group found that the excess prevalence depended on the region of origin and the specific risk for HIV infection. Compared to Spaniards, prevalence of HIV was 4 times higher in Latin American men who have sex with men (MSM), and 9.4 times and 19.3 times higher, respectively, in Latin American and Sub-Saharan African heterosexual males. Regarding heterosexual women, the higher prevalence was found in Sub-Saharan Africans (16.9 times more) and North Africans (15.3 times more) [53].

One relevant aspect is whether higher HIV prevalence in migrants is explained because they are already infected on arrival, or because they are at a higher risk of infection once in the host country. This question has been extremely controversial and has favoured racist reactions to the HIV epidemic in migrants. However, it is very relevant from the Public Health perspective, as it would mean a failure of prevention

strategies in this group, and would point out the need to reinforce access to information and prevention, together with improving access to HIV testing.

It is a fact that migrants coming from countries with generalised epidemics, mainly Sub-Saharan Africa, as a group, arrive with a prevalence of HIV more similar to their countries of origin than to the host country[54].

However, several studies have pointed out that once in Europe, they would be experiencing a higher risk of infection than native-born population, either when they travel to their countries for visiting friends and relatives[55], or in the host country[56-58]. In Spain, a study in people who visited a network of 19 STI clinics between 2003 and 2004 and who performed repeated HIV testing found incidence rates of new HIV diagnosis were 8 times higher in Sub-Saharan Africans compared to Spaniards, and 2.7 times higher in Easter Europeans, with no difference for other groups of migrants[59]. For 25% of patients the most probable country of infection was their country of origin, with a median stay in Spain of 7 months; 33% probably would have acquired their infection in Spain, where they had been living a median of 48 months; for 42% the likely country of infection could not be determined[59, 60].

All these figures show the fact that migrants are a population that is being disproportionately hit by the HIV epidemic and that they comprise a highly heterogeneous group who suffer very distinct epidemics, making it necessary to study migrant groups independently.

2.2.2.4. Other aspects of HIV vulnerability in migrants

Apart from the higher rates of HIV infection in migrants, other aspects of increased vulnerability can be described regarding HIV. As reported by several studies

and information sources, migrants tend to be diagnosed in later stages compared to the native population[61-63]. In Spain, delayed diagnosis defined as a CD4 count of less than 350 cells/µl, was experienced in 42% of new HIV diagnosis in native Spaniards as of 2010, but was higher for most migrant groups, reaching up to 49% in the case of Latin Americans and 61% in Sub-Saharan Africans[12]. Risk of delayed diagnosis in migrants compared to autochthonous population has been found to be higher in cohorts studies in Spain, but has not been estimated specifically for the different migrant groups[50, 63].

Delayed initiation of treatment has also not been studied in Spain in different migrant groups, although studies under different designs have found a similar need and access to treatment for migrants as a whole [52, 64, 65]. Until year 2012, the Spanish health care system has been *de facto* universal: almost every patient is entitled to receive medical care and antiretroviral treatment, and those who are not, namely the migrants of uncertain status without any official identification, effectively access through the NGO network. So it would be expected that no barriers exist for migrants to access treatment, but this hypothesis has not been properly contrasted across the different migrant groups.

Regarding prognosis of HIV infection, the majority of publications in Europe have shown no major differences in response to antiretroviral treatment (ART), risk of AIDS or survival in immigrants, although discordant results have been found regarding immunological response to treatment [66-77]. However, most studies in Europe have focused on Sub-Saharan African population, while risk of AIDS and survival in Latin Americans has not been studied other than in the United States. In Spain, no cohort

studies have focused on studying the risk of AIDS or death in migrants from different regions of origin compared to the native population[64].

2.3. Cohorts of HIV infected subjects and their role in research

2.3.1. Definition and classification

Cohort studies are observational studies where a group of patients, free from an event of interest, are recruited and followed up over time to observe the occurrence of the event. Cohort studies make it possible to analyse the occurrence of one or several events or outcomes in subjects who have been under one or several exposure variables, and whose assignment was not randomized [78, 79]. Cohort studies have been fundamental for describing the natural history of HIV and are the most suitable design to study the effect of different exposures over HIV disease progression.

Among the observational studies, cohort design is also the most sound for causal inference. In experimental studies exposure is assigned in a randomized way, and if the sample is large enough, randomization ensures that exposed and unexposed subjects are comparable in all characteristics except the exposition whose effect is under analysis. This is the basis of a randomized clinical trial, the gold standard design for causal inference in health sciences. However, experimental studies are limited to exposures that are of a nature that can be controlled and which is ethically correct to randomise. Other drawbacks are that they operate under a controlled environment, select subjects with restrictive inclusion criteria, and are under the volunteer bias (subjects who self-select to participate in trials are often healthier than the average). This limits the external validity of clinical trials. Cohorts operate under real conditions,

which means that expositions are assigned according to medical practice, and so causal inference is not straightforward, but the profile of patients who participate are more representative of the real target population.

Cohorts can be classified according to a number of characteristics, as defined by Jarrín et al.[78], but one of the more relevant classifications in HIV cohorts depends on the characteristics of the subjects included in the study and differentiate between seroconverter and seroprevalent cohorts.

Seroconverter cohorts include subjects whose date of HIV infection is known or can be estimated with acceptable accuracy. As described by Jarrín et al., seroconverter cohorts can be further classified into incident cohorts, if they identify seronegative individuals and follow them up over time to observe their seroconversion; or prevalent cohorts with known date of infection, if they identify HIV positive subjects and assess their seroconversion date retrospectively based on previous information, generally a previous negative test or an undetermined Western Blot.[80]. On the other hand seroprevalent cohorts include subjects who are HIV-positive at recruitment but there is no information available on when they got infected.

2.3.2. Analytical approaches for cohort data

2.3.2.1. Statistical Methods for cohort analysis

Normally, in a cohort study analysis the main interest will be in estimating the incidence rate of a given event, or the time it takes a certain percentage of patients to be affected by it and/or in measuring the effect of a given set of variables over the risk of the event.

Classical statistical methods used to respond to these questions are Kaplan-Meier method, which allows us to estimate the survival curve or its complementary, the cumulative incidence curve; log-rank test, used to compare survival curves over different levels of an explanatory variable; and Cox proportional hazard models, to estimate the effect of a set of variables over the hazard, i.e. over the instant risk of suffering the event. Due to correspondence between survival and hazard functions, results of Cox proportional hazard models can be interpreted in terms of effect over cumulative incidence of the event.

However, these methods are limited in several situations, two of them relevant for the analysis performed in this Doctoral Thesis: in the presence of competing events and in the presence of informative censoring, as will be discussed below. Other potential biases and methodological approaches are discussed under the following headings.

2.3.2.2. Common biases and analytical approaches in cohort studies

In an ideal setting, researchers would follow a cohort of patients for an infinite time and thus would be able to observe the occurrence of the event of interest in every subject in the study. In a real setting, the time under observation is finite, and events may not occur within the observed period, producing a so called right censoring. There are three main reasons for right censoring: administrative censoring (i.e. end of the study), loss to follow-up (LTFU) of the patient, or occurrence of a secondary event that prevents the event of interest from happening (i.e. death of the patient prevents him/her from being diagnosed AIDS or being prescribed treatment).

Most methods for statistical analysis, like the Kaplan-Meier estimate or the Cox proportional hazards model, work under the assumption that time to event and time to censoring are independent. In the first case, where subjects are censored at the date of the end of the study, we can easily assume that the end of the study is not associated to the event of interest. But in the other two situations, independence can be harder to assume.

Losses to follow-up are one of the most important sources of bias in cohort studies. If subjects are lost to follow-up randomly, for reasons not associated to the event of interest, those individuals who remain in the study can be assumed to represent missing subjects. But it is possible that subjects who leave the study are different from those who remain. For example, if people with a slower disease progression happen to miss follow-up visits while those with a faster disease progression are more adherent to clinical visits, those who experience the event would be over-represented and we would overestimate incidence. Several HIV cohorts have described that subjects lost to follow-up have better clinical characteristics[81, 82], confirming this situation. If in addition, some variable of interest, like the country of origin of the patient[81, 82] determine different rates of losses to follow up, we can introduce a bias by not taking into account this informative censoring. Several methods have been described to deal with this possible source of bias.

One method is to weight each observation by the probability of remaining uncensored given a set of explanatory variables. For that purpose and as described by Fewell et al.[83] each patient has to be split up into as many registers in the database as months is under observation. The cumulative number of months under follow-up at

every consecutive month ('month'), together with splines or another function of 'month' are introduced in the model to allow for a non-linear relation between outcome and follow-up time. A pooled logistic regression is then performed, thus simulating a Cox model where the estimated function of 'month' is equal to the baseline hazard function.

For each particular month, probability of not being LTFU is estimated using a logistic regression where baseline and current month variables are taken into account, so that we allow for past and present clinical conditions, markers, treatment, etc. to influence the probability of not leaving the study. Any other fixed variables such as educational level, sex or region of origin, among others, can also be introduced into the model. The inverse of this probability can be used to make each patient-month under observation represent those patient-months of similar characteristics which were not observed due to censoring. However, use of stabilised weights is recommended by Cole S.R. and Hernán M.A.[84]. Stabilised weights are calculated as the probability of being not LTFU estimated out from a model that takes only into account baseline variables, divided by the probability of being not LTFU derived from a model that also considers time-varying variables updated for each particular month. These stabilised weights are then introduced in the pooled logistic regression considering each patient as a cluster.

Another usual option in cohorts is to cross cohort patients with AIDS and death registries, if available. This allows retrieving information about events of interest in patients that were LTFU. Assumptions have to be made that registries are complete, that patients not appearing in the registry are free from the event, and that every pa-

tient has the same probability to appear in the registry, or otherwise a bias can be introduced. If this strategy is used, follow-up of every patient LTFU has to be artificially expanded up to the closing date of the registry. Otherwise, patients with the event would be represented, but not the free-of-event time of patients not experiencing the event. Generally, the last one or two years before the closing date of the registry are not considered to allow for the usual delays in notification.

As previously mentioned, the third reason for right censoring is the presence of a competing event, an event that, when occurring, prevents the event of interest from happening. Again, if there is no association between the time to the event of interest and the time to a competing event, those who experience the competing event would virtually remain represented by those who do not, and a standard analysis which censors patients who experience a competing event, would give us an unbiased estimate of the incidence rate in a counterfactual world where nobody experienced the competing event. But this assumption cannot be tested in the observed data and in this situation, Kaplan Meier estimates can only be interpreted as the instant probability of experiencing the event conditioned to the probability of being alive and event-free at that instant. Also, as argued by Putter at al.[85], a subject that is censored because of failure from a competing risk will with certainty not experience the event of interest. Since subjects that will never fail are treated as if they could fail (they are censored), the Kaplan–Meier function overestimates the probability of failure (and hence underestimates the corresponding survival probability).

Obtaining an unbiased estimation of the cumulative incidence of a certain event and its association with a range of independent variables, acknowledging the

existence of competing events, requires alternative methods. Putter et al. described a method to estimate cumulative incidence in the presence of competing events called the multiple decrements method[85], that estimates a cumulative incidence curve for the event of interest and each of the competing events; And Fine and Gray[86] described a method to model the effect of a given covariate over the sub-distribution of the risk of the event of interest. In practical terms, both methods are based in a simple idea: in a world of existing competing events, those individuals who experience the competing event will never experience the event of interest, so they are not censored, but are kept "alive" in the data set as event-free follow-up time. Results from these models can be interpreted directly as effect over the cumulative incidence, or risk of the event.

2.3.2.3. Seroconverter cohorts: strengths and potential biases

Studies of seroconverters are the best way to investigate the natural history of HIV[87], as they have information on the exact origin of risk, i.e. when the HIV infection took place. However, they may need a longer follow-up time to observe events that can take a long time to occur, as is the case of AIDS incubation period, which was estimated to be around 10 years in the absence of treatment[88, 89].

Also, the profile of patients recruited in a seroconverter cohort is very specific and does not represent all the HIV-infected population. Subjects who undergo HIV testing repeatedly, and thus are self-perceived to be at risk of infection, are at an increased probability of being recruited into a seroconverter cohort. Also, the requirements to be a seroconverter exclude subjects who experience diagnostic delay. This type of cohort design therefore ignores patients who may have either a low self-

perceived risk or experience barriers to access testing. This is known as the *seroconverter bias* and limits the external validity of these studies. Also because heterosexually infected individuals tend to have a lower risk perception, they are often underrepresented in these cohorts, as often are women, so studies that aim at exploring the influence of any of these variables would probably find small sample sizes under this design.

From the analysis perspective, it is important to define if the seroconverter cohort has a prevalent design –where HIV positive subjects are recruited and the infection date is assessed retrospectively- because they are affected by the so called left truncation or delayed entry. This means that the event that marks the risk origin (in our case, the date of HIV infection) is not observed, but the patient is only recruited and starts to be observed after some time after infection has passed. During the time after HIV infection and before cohort recruitment, some patients could experience adverse events or die. This can lead to a *survivor bias*, as only subjects who survive long enough have the opportunity of being recruited into the cohort; but those who died too soon to be identified in the cohort and recruited are not represented in the dataset. So if we analysed data directly taking into account only the time from HIV infection to the outcome, we would be underestimating its incidence, as subjects with faster occurrence of the event are missing.

To minimize this source of bias, it is necessary to incorporate methods that correct for the fact that some of the time since risk origin was in fact not observed. Under this method, patients only start to contribute to cumulative incidence calculation since the date they start their follow-up in the cohort. However, time since HIV infection

(risk origin) is also considered, in a way that subjects are always compared with subjects with similar total duration of infection [90].

Finally, other possible bias can occur under this design since we estimate the date of seroconversion based on the existence of a previous negative test. The exact date of seroconversion will be comprised between this last negative and the first positive test, generating a so called *interval censored* event. Usually, the middle point between both tests is estimated to be the date when the subject got infected, but this assumes every subject has a constant risk during the interval and that subjects perform HIV test with similar frequency, independently of their risk. As described by Law et al.[91] these facts don't affect liability as long as only intervals between both tests of 3 years or less are considered. Also, it has been pointed out that testing dates need to be documented, to avoid a *memory bias* in non-documented seroconverters[90].

2.3.2.4. Seroprevalent cohorts: strengths and potential biases

Seroprevalent cohorts yield less accurate estimates of the natural history of disease, as no direct information exists on the date when subjects got infected by HIV. However, they have several advantages. As included subjects have normally gone through part of the natural history of disease, follow-up time needed to observe events is generally shorter and recruiting subjects is easier, so they are more efficient studies. Also, they do not tend to include a special patient profile, and inclusion criteria are generally less restrictive, so included patients are more varied and representative of the whole HIV epidemic and external validity is improved.

The main source of bias in seroprevalent cohorts relates to the fact that the origin of risk (i.e. HIV infection) is unknown [92], leading to a so called left censoring. If

we analysed time from recruitment in the cohort to the outcome, we can introduce a bias if subjects of a certain characteristic –for example, being from a non-Spanish origin- is related to being recruited in more advanced stages of disease. In this case, risk of suffering the outcome would appear higher in this group even if there was no real difference. The most accepted solution to minimize this problem is to adjust the model by a progression marker measured at recruitment that would account for duration of infection, normally, CD4 count and viral load (VL).

Finally, for studies whose objective is to analyse response to treatment –and so the risk origin would be treatment initiation- a seroprevalent cohort composed of antiretroviral naïve individuals who start treatment over their follow-up is the most efficient and the ideal design, as it allows us to observe the exact risk origin. In this particular case, seroconverter cohorts do not offer any advantages, but in fact only the inconveniences associated with this type of design.

2.3.2.5. Studying migrants: methodological considerations

Some considerations have already been made about problems arising when working with migrants. For example, the fact that they show higher rates of loss to follow-up in cohort studies makes it necessary to account for this informative censoring, preferably through IPW. Crossing with registries may not be such a good option in this case, as migrants are more likely to return to their home country or further immigrate to another country in Europe, so they would be less likely to appear in Spanish AIDS and death registries. Crossing with national registries would then underestimate AIDS and death risk in this group.

There are also two well-described phenomena in studies involving migrants that tend to result in lower mortality rates in migrants compared to host country population: the healthy migrant effect and the salmon bias. The first one refers to the fact that, at the time of arrival into the country of emigration, most migrants tend to have better health than host country populations. This phenomenon is a particular form of selection bias attributed to the various processes that labour migrants undergo before coming into the country of destination. Since most people go to another country expecting to work, those who most frequently migrate are the fittest, best able to survive the journey and pass the medical examinations they may have to undergo. This effect could account for better health and lower mortality in migrants, but it is not a bias, as differences found are in fact present. With notable exceptions, immigrants' and nationals' health patterns tend to converge after some years after migration and, for some health conditions, immigrants fare worse[93-95].

On the other hand, the so-called salmon bias refers to the fact that migrants may want to return to their home countries when feeling chronically and/or severely sick, and can result in artificial estimates of a lower mortality in migrants even in the case where no differences existed. It may be explained by a hampered ability to remain employed in adverse health conditions, the search for a context of better social and family support to cope with disease, or simply the desire to die in one's birthplace. This would imply that migrants who abandon the country (and the study) will be those in worse health conditions, and thus the most likely to suffer a health event or to decease, which produces an underestimation of risk [96, 97].

2.3.2.6. On exposure variables in cohort studies

In HIV cohort studies, we are able to measure a number of variables at recruitment and during follow-up. From the analysis perspective, it is important to notice the difference between fixed variables and time-dependent variables.

Fixed variables are either variables that don't change over time like sex, region of origin, likely route of HIV infection, or delayed diagnosis or variables that can vary with time, but that are however collected at recruitment and treated as fixed, such as educational level, CDC stage, CD4 count and Viral Load at recruitment, etc.

On the contrary, we would like to allow some variables to change with time, as CD4 count and Viral Load, antiretroviral treatment, AIDS defining illnesses suffered, etc. These variables that change overtime in each individual are known as *internal varying variables*[78]. But variables can also change over time in the same way for all the subjects in the study. For example, in HIV, the generalization of HAART in 1997 marked a prognostic turnover and completely changed the natural history of HIV infection, and so the risk of developing AIDS and death for patients before 1997 was much higher than after this date. Calendar period can be used as a *proxy* of HAART availability, and in this case it is called an *external varying covariate*.

2.3.2.7. Individual and population effectiveness

The diffusion of HAART has drastically slowed down the rate of progression from HIV infection to AIDS and death, transforming HIV into a chronic disease. Most conclusive results on treatment efficacy have been obtained trough clinical trials, which operate under ideal conditions. However, the measurement of the effect of a

given treatment under real life conditions and routine medical practice gives complementary information and is called effectiveness. Two different effectiveness measures have been defined by Muñoz et al. [98]

Individual effectiveness analyses try to replicate a randomized clinical trial out of observational data, and compares outcomes in treated vs. non treated subjects, adjusting for any characteristics that are associated with receiving vs. not receiving treatment to avoid an *indication bias*.

On the other hand, population effectiveness analyses show the benefits of existing treatments when they penetrate a given population, and are especially important to analyse its real impact. It compares a population where treatments are available and where subjects who need treatment will receive it, to a population where treatment is not available for anybody, even for those who may need it. Considering calendar period as an explanatory variable, as previously mentioned, we can compare AIDS and death incidence in the population before 1997 with that of the population after 1997, which will yield us results on population effectiveness of HAART.

In this case, an individual can be under follow-up during both periods and thus be present in both populations that are being compared. Using the calendar period as an explanatory variable means we need to split patients up and assign observed patient-time to the corresponding period, ignoring the time observed outside the period, but taking into account the total duration of infection, so that comparisons are always established between patients with same length of infection. The result would be similar to performing an administrative censoring at the end of each period and a left truncation in the starting point of the next one.

3. JUSTIFICATION

Monitoring equal access to resources and treatment effectiveness in different socioeconomic groups is fundamental to ensure that appropriate interventions are put in place and respond to specific challenges, especially in the context of a universal, free health care system like the one existing in Spain.

Despite several studies have been performed addressing the issue of social inequalities in HIV, some clues are missing to better characterise the influence of socioeconomic determinants over the complex dimensions of HIV disease. Due to its relevance and its ability to capture important aspects of socioeconomic level, in this doctoral thesis we decided to focus on studying the influence of educational level and country of origin.

However, a first approach was performed to the study subject, evaluating a wide range of socioeconomic and demographic variables in a predictive approach over risk of HIV progression. A cohort in Madrid was chosen for this study outcome, as a previous similar study had been carried out with 1999 data [40], and so we would be able to fulfil our objective while contrasting or confirming data from previous epidemic stages.

Regarding the effect of educational level over HIV-related outcomes, previously published articles point out the existence of inequalities in different educational groups, specifically a higher mortality in people of low educational level. However, all studies have been carried out in seroprevalent subjects, with their intrinsic limitations when studying natural history of disease. Also, the few studies that have included educational level as an explanatory variable for progression to AIDS, all of which were conducted in the pre-HAART era, have yielded contradictory results. On the other hand, recent studies have shown a higher risk of delayed diagnosis and delayed initia-

tion of treatment in subjects of low educational level. However, whether a higher risk of delayed initiation of treatment actually exists in those timely diagnosed is still to be clarified. To address these issues, we analyzed a national seroconverter cohort, which is the best design to study natural history of disease and which includes subjects that, by definition, are not subjected to delayed diagnosis.

Another source of socioeconomic inequalities, as argued in the introduction, is the migratory status. Migrants are a most-at-risk population regarding HIV/AIDS and studies in Europe have shown a higher prevalence and a higher risk of delayed diagnosis in this group. However, in Spain, no cohort studies have focused on estimating the risk of delayed diagnosis, of delayed access to treatment and risk of AIDS and death in migrants of different regions of origin. Also, little data exists in our context regarding response to treatment of HIV-infected migrants according to specific regions of origin. For all these reasons we decided to study migrants from different origins and their risk of a wide set of HIV-related end points. For this purpose, we decided to analyse a seroprevalent cohort, which is a more efficient design and provided us with an appropriate sample size of Sub-Saharan Africans and Latin Americans. Other regions of origin were represented in too small numbers and could not be studied.

4. OBJECTIVES

Objective 1: To identify socioeconomic and demographic variables that predict disease progression since HIV seroconversion in Madrid (Spain), before and after the generalisation of HAART.

Objective 1.1: To evaluate association of age, sex, educational level, transmission category, and region of origin with HIV disease progression to AIDS and all-cause mortality, before and after the generalisation of HAART.

Objective 2: To estimate the effect of educational level over HIV-disease progression since HIV seroconversion and access to treatment at different periods of the HIV epidemic in Spain.

Objective 2.1: To estimate the effect of educational level over risk of AIDS at different periods of the HIV epidemic in Spain.

Objective 2.2: To estimate the effect of educational level over survival at different periods of the HIV epidemic in Spain.

Objective 2.3: To estimate the effect of educational level over time to HAART requirement and time to HAART initiation as *proxies* for access to treatment in the period after 1997 in Spain.

Objective 3: To analyse key HIV-related outcomes for migrants originating from Latin America and Sub-Saharan Africa living in Spain compared to the native population.

Objective 3.1: To analyse if any differences exist by geographical origin in risk of delayed diagnosis of HIV infection.

Objective 3.2: To analyse if any differences exist by geographical origin in time to cART requirement and time to cART initiation as *proxies* for access to treatment.

Objective 3.3: To analyse if any differences exist by geographical origin in virological and immunological response to cART once initiated.

Objective 3.4: To analyse if any differences exist by geographical origin in risks of AIDS and overall survival.

5. METHODS

5.1. Methods for Objective 1

5.1.1. Study Population

We analysed data from the Seroconverter Madrid Cohort, whose patients have been identified at the "Centro Sanitario Sandoval" from 1985 onwards. Recruitment is still ongoing. The Centro Sanitario Sandoval is an ambulatory STD clinic and HIV screening centre whose access is open, free, and anonymous and has been a pioneering centre in HIV prevention in Madrid. HIV negative subjects are invited to come back after 6 months for follow-up HIV tests. If they become HIV positive, patients are followed up in the centre until they require antiretroviral treatment and/or hospital admission. Patients are then referred to various hospitals in Madrid for clinical follow-up and antiretroviral treatment.

In this cohort, the definition of a seroconverter was an individual aged 16 or over who had a negative HIV test within the 3 years before the first HIV positive test, or who had a positive ELISA test with an undetermined result in the Western Blot. Patients who could provide a documented HIV negative test or undetermined WB performed outside the Centro Sanitario Sandoval were also included in the cohort. Complete ascertainment of all seroconverters seen in the recruiting centre was carefully sought. The reconstruction of the cohort was done in 1997, but all people who seroconverted before that date were selected independently from their outcome.

Seroconversion date was estimated as the mid-point between the last HIV negative and the first HIV positive tests, or the date of the undetermined WB test, as appropriate. Administrative censoring for this analysis was February 2009.

5.1.2. Follow-up and variables

Socio-demographic and epidemiological patient characteristics were collected upon recruitment, including age, sex, mode of HIV transmission, educational level and country of origin.

Educational level was recorded in four categories: below primary education; primary education completed, being subjects who had finalised the compulsory education, usually remaining on formal education until 14 years of age; secondary education, which included subjects who had completed high school degree or equivalent; and tertiary education, if the subject had completed university studies. The variable was grouped into two levels for this analysis: persons with a low educational level (no education or only primary education completed) and persons with a high educational level (secondary or university studies completed).

Due to the low number of subjects with heterosexual transmission, the transmission category was grouped into two mutually exclusive categories: injecting drug users (IDU); and sexual transmission and other, which included heterosexual transmission, men who have sex with men and people with other transmission routes.

Due to small sample size, country of origin could only be categorised into three regions: Spain, Latin America and the Caribbean, and other regions.

During the recruitment visit, clinical and analytical baseline variables were also collected.

Patient's follow-up was done according to routine medical practice and, during visits, any relevant clinical and analytical information was collected, along with HIV treatment status. Follow-up information of seroconverters was updated in the database yearly, both in the recruiting centre and in the referring hospitals. For patients lost to follow up, cross-checks using name, surname, and date of birth were performed with the databases from the 12 participating hospitals within the Community of Madrid.

Patients were also cross-checked with the national AIDS registry in 2007 and with the national death registry from the National Institute of Statistics in 2006. Patients not appearing in the registries were assumed to be AIDS and/or death free by 31st December 2005 and 31st December 2004 respectively, acknowledging the usual reporting delays and selecting a period where completeness of the registries could be assumed[90]. After those dates, subjects were censored in the date of their last follow-up visit in either the recruiting centre or the referral hospitals.

Sensitivity analyses were performed with different censoring strategies to allow for different yield of the registries. The first strategy considered patients not appearing in the registries were event-free up to one year before the registry closing date. The second strategy considered registries could only be assumed as complete up to three years before their closing date, and censored patients as event-free on that date. The third strategy made no assumptions beyond the time patients were last seen on a

clinical follow-up visit, and thus censored every patient at their last visit. Results remained largely unchanged under the different models.

To estimate population effectiveness of HAART, calendar period was used as a *proxy* for HAART availability and introduced in the analysis as a time-changing variable with two categories: pre-HAART era (up to 1996) and HAART era (1997 onwards). Each patient contributed to the analyses with as many registries as periods at risk he/she contributed to the study, and patients with an equal duration of infection were compared for each period.

5.1.3. Statistical Analysis

Characteristics of the sample were described using proportion or median (and Inter-quartile Range), as appropriate. Chi2 test for categorical variables and Kruskal-Wallis test for continuous variables were used for bivariate analysis.

A survival analysis was performed, considering seroconversion date as the onset of risk, but with a delayed entry (or left truncation) to the date of first HIV positive test in the recruiting centre, to minimize the survivor bias. A predictive step-forward modelling was performed to identify variables associated to cumulative risk of AIDS and cumulative risk of death. Independent variables assessed were sex, age at seroconversion, educational level, region of origin, transmission category and calendar period under observation. No adjustment for CD4 counts or Viral Load was needed at baseline, as all patients were seroconverters and baseline values were very homogeneous. To allow the effect of variables to vary among the different calendar periods,

interaction between each variable in the final model and calendar period was systematically explored.

Cumulative risk of all-cause mortality was calculated by Kaplan-Meier estimates and Cox proportional hazards model was used to assess adjusted effects of independent variables over survival, estimated through Hazard Ratios (HR).

For cumulative risk of AIDS, death was considered as a competing event, as patients who died before being diagnosed of AIDS had no longer the opportunity of experiencing the event of interest. Cumulative incidence function was therefore calculated through the multiple decrements method, and a Fine and Gray model was used for multivariate analysis, and resulted in estimation of the so-called sub-Hazard Ratios (sHR). Sensitivity analyses were performed under different assumptions of what happened to those suffering the competing event: the first strategy censored patients on their date of death; and the second strategy censored them at the end of the calendar period where they were deceased. Results did not vary under any of these assumptions.

Robust methods were used for standard error estimation and statistical significance was evaluated using Wald's test. All analyses were carried out using Stata Software (version 11.1, Stata Corporation, College Station, Texas).

5.2. Methods for Objective 2

5.2.1. Study Population

We analysed data from the Spanish Multicenter Study Group of Seroconverters (Spanish acronym GEMES), an open, prospective, multicentre cohort of HIV positive subjects with a known date of seroconversion.

GEMES comprises nine individual cohorts from different recruiting centres over Spain: a cohort from Centres for Information and Prevention of AIDS (CIPS, by the Spanish acronym) in the Comunidad Valenciana located in the cities of Alicante, Castellón and Valencia; cohorts from the Centres for Prevention and Care of AIDS (CAPS) in the city of Barcelona; the Madrid cohort recruited in the Centro Sanitario Sandoval, previously described in section 5.1.1; the cohort from the Prisons Health Service in Barcelona; the seroconverter cohort from Navarra; the cohort of the HIV Unit in the Germans Trias i Pujol Hospital, in Badalona; and three Haemophilia Unit cohorts from Hospitals in Barcelona, Seville and Madrid. More information on individual cohorts characteristics can be found in individual publications [99-102]. For the purpose of our analysis, patients recruited in the Haemophilia Units were excluded from the study population as they had been infected in a good proportion during childhood and reverse causality between HIV progression and educational attainment could not be ruled out in that case.

A seroconverter was defined as mentioned for Objective 1 (section 5.1.1), and the seroconversion date was estimated in a similar way. Recruiting sites identified se-

roconverters from 1983 onwards, both prospectively and retrospectively, so both seroincident and seroprevalent subjects were included as long as the necessary clinical information was available.

Methods for identification of seroconverters in each individual cohort were different according to specific characteristics of each site, but in every case it was done independently of the subsequent progression of disease. Administrative censoring for this analysis was February 2009.

5.2.2. Follow-up and variables

A standardised protocol for data collection was followed by the different participating cohorts. Baseline information was collected at recruitment date. Educational level was measured at cohort entry in four categories and categorised in two levels for analysis, in the same way as exposed for Objective 1 (section 5.1.2). HIV transmission route was also grouped into two categories, as done in the previous analysis. Data on CD4 levels and viral load at diagnosis, sex and age at time of seroconversion, patient's region of origin, and the method used to estimate the date of seroconversion were also collected. Region of origin was not introduced in the analysis as there were too few subjects to create meaningful groups that could adequately account for the heterogeneity of the migrant population in Spain.

Patients were followed up according to clinical practice with further collection of clinical, laboratory and treatment variables. Many of the recruiting centres refer patients to Hospital Units when they experience an AIDS-defining illness or other conditions that need specialised care, or fulfil criteria to start antiretroviral treatment.

These patients were followed up in coordination with the referral centres or with the patient him/herself.

To minimize losses to follow-up, patients were cross-matched with the National Registry of AIDS cases in the year 2004. Patients not appearing in the registry were considered AIDS-free as of 31st December 2002, leaving two washout years to correct for reporting delays[90]. For dates after the cross-match, only AIDS events occurring during clinical follow-up were registered, and those lost to follow-up were censored as event-free on the last date they had been seen. Three of the six cohorts under analysis were cross-matched with the mortality registry of the National Statistics Institute: two by 2006 and one by 2008. Subjects not recorded as deceased in the registries were considered alive until 2 years before the closing date of the registry. After the cross-match date and for cohorts not cross-matched, losses to follow-up were censored and considered alive on the date of their last visit. Sensitivity analysis were performed with different censoring strategies to allow for different yield of the registries, under the same assumptions described in section 5.1.2, and results remained largely unchanged.

Follow-up was divided into three periods that reflected the different availability and changes in HAART recommendations: Years before 1996 formed the pre-HAART era and were used as reference for comparison; years from 1997 to 2003 were considered the first part of the HAART era, and comprised the second period for analysis; finally, years 2004 onwards, were considered as a third and independent period based the recommendation to start treatment at an earlier stage of disease and on a significant increase in HAART efficacy. Calendar period was introduced as a time-dependent covariable, as exposed for Objective 1.

Analysed outcomes were risk of progression to AIDS, death, HAART initiation and HAART requirement, defined as the time until a patient reached a CD4 count under 350cells/mm3 or developed an AIDS defining condition.

5.2.3. Statistical Analysis

The characteristics of the sample were described and nonparametric tests – Chi2 for categorical data and Kruskal-Wallis for continuous variables- were applied to evaluate clinical and socio-demographic differences by educational level. Only patients with information on educational level were included in subsequent analysis.

Survival analyses were conducted taking the date of seroconversion as the risk origin, but with delayed entry to the time of the first positive diagnosis in the recruitment centre to eliminate possible survival bias. The association between the exposure of interest, 'educational level' and the various outcomes was examined in crude analyses, and a multivariate analysis was carried out with an estimative strategy to adjust for all potential confounders, where variables that produced a change higher than 10% in the HR of interest were retained in the model. Proportionality of hazards assumption was tested for educational level and every outcome. Interaction between educational level and the different confounding variables was tested in every model.

To evaluate the risk of progression to AIDS, death was considered as a competing event; therefore, the cumulative incidence function was estimated by the multiple decrements method. To evaluate the adjusted effect of educational level on the risk of AIDS the Fine and Gray method was used. For overall survival, the cumulative incidence function was calculated using the Kaplan-Meier method. To evaluate the ad-

justed effect of educational level on the risk of death, Cox's proportional hazards model was used to estimate the HR.

The analysis of time to requirement of HAART and time to HAART initiation was restricted to persons seroconverting after 1996, time at which HAART became universally available in Spain, and each individual was censored on the date of the last visit. Death was considered as a competing event, and the multiple decrements method and the Fine and Gray model were used for respective estimations.

Sensitivity analysis were performed under different assumptions of what happened to those suffering the competing event, as previously described in section 5.1.3, and results did not vary with under any of the assumptions.

Robust methods were used for standard error estimation, and statistical significance was evaluated using the Wald test. All the analyses were conducted using Stata software (V.11.1, Stata Corporation, College Station, Texas, USA).

5.3. Methods for Objective 3

5.3.1. Study Population

We analysed data from the Cohort of the Spanish AIDS Research Network (CoRIS) an open, multicentre, prospective cohort of patients over 13 years-old with confirmed HIV infection and naïve to antiretroviral treatment at entry. CoRIS is a seroprevalent cohort, although seroconverter subjects can also be recruited.

CoRIS is a national cohort under the umbrella of the Spanish AIDS Research Network (Spanish acronym, RIS). 31 HIV units from 28 Centres in 13 of the 17 Autonomous Communities of Spain participate in CoRIS. After a given Centre is included as a CoRIS participating site, it is required to recruit every patient who consults for the first time in the centre and who fulfils inclusion criteria as mentioned above. A semi-anonymous code is given to every patient containing the first two letters of his/her two surnames, the birth date and the sex, so patients transferred between participating sites can be univocally identified, follow-up continued and duplicates are detected. Information is annually sent to the coordinating centre of the cohort to undergo several internal quality controls.

The CoRIS cohort began in 2004 and recruitment is ongoing. Ethics approval was obtained and every patient signed a written informed consent. English and French versions were available to encourage participation of migrants with poor knowledge of Spanish. Detailed descriptions of the cohort can be found in the literature[81, 103]. Administrative censoring for this study was October 2010.

5.3.2. Follow-up and variables

All CoRIS sites record variables in a standardised way following the cohort protocol and data quality control procedures are performed every year. Different socio-demographical and epidemiological variables are collected at recruitment including sex, age, most probable route of HIV transmission, educational level, self-referred region of origin, CDC Stage, date of HIV diagnosis and previous negative tests if available, along with other clinical and analytical baseline data.

Patients are followed up according to routine clinical practice, and relevant information on CD4 counts and Viral Load (VL) is recorded in each visit, along with clinical, analytical and treatment information. Follow-up of patients ends when the patient dies, or when the patient changes his follow-up to a centre that does not belong to CoRIS and thus is lost to follow-up. In this case, no cross-match with registries was performed, to avoid bias associated with the different probability of appearing for the different regions of origin, so patients were censored in their last follow-up visit in a CoRIS centre.

To characterise access to ART and prognosis of HIV infection across different regions of origin, we analysed seven different outcomes: HIV diagnostic delay, time to ART requirement, time to ART initiation, immunological and virological response to ART, time to AIDS diagnosis and time to death.

Diagnostic delay (DD) was defined as having a CD4 count below 350cells/mm3 or an AIDS-defining illness in the first year following HIV diagnosis[13, 63], so only patients with an available CD4 count within that period could contribute to the denominator.

ART requirement was defined according to minimum criteria for ART initiation in Spain throughout the study period[104]; that is, CD4 count ≤350cells/mm3 or an AIDS diagnosis. These criteria have been expanded in the last years to include new treatment indications. These were not included in the analysis, and thus there are a proportion of subjects initiating treatment before fulfilling initiation criteria. Date of HAART initiation was considered to be the date where the first antiretroviral drug was prescribed to the patient.

For analysis of immunological and virological response to treatment, included patients were those who had at least one CD4 count of any value, or a VL measurement over 50copies/ml, respectively, in the 6 months prior to ART initiation and at least 2 post-ART determinations, one within the first year after ART initiation. Time to virological response was analysed as time from ART initiation until the first of two consecutive VLs was below 50copies/ml; and immunological response was the first of two consecutive CD4 counts of at least 100cells/mm3 higher than pre-ART determination. Patients not experiencing the event were censored in their last CD4 or VL assessment respectively.

5.3.3. Statistical Analysis

Only patients originating from Spain, Latin America and the Caribbean (LAC) or Sub-Saharan Africa (SSA) were included in the analysis. Chi2 and Kruskal-Wallis tests were used to evaluate differences according to the region of origin. The association between the exposure of interest, 'region of origin' and the various outcomes was examined in crude analyses, and a multivariate analysis was carried out with an estima-

tive strategy to adjust for all potential confounders, where variables that produced a change higher than 10% in the HR of interest were retained in the model.

For analysis of Delayed Diagnosis (DD), a multivariate logistic regression was fitted to analyse the association between region of origin and DD after adjustment for any confounding variables.

To assess differences in ART requirement across regions, a survival analysis was performed, taking into account that death and initiating ART were competing events for ART requirement, as those initiating ART could not experience the event of fulfilling criteria to start ART for the first time. Thus patients for this analysis were censored at the date of the last follow-up visit or at date of ART initiation, whatever came first, and a multivariate Fine and Gray regression was fitted to estimate the effect of region of origin, accounting for competing risks and adjusting for any confounders.

Time to ART initiation and time to AIDS diagnosis were analysed using a similar method, but censoring each patient at the date of the last follow-up visit. In this case, only death was recorded as a competing event. Finally, time to all-cause mortality was analysed using a standard multivariate Cox regression model.

In all four survival models, patients who fulfilled the outcome definition at recruitment were excluded from the analysis. Taking into account that all subjects are ART naïve at entry, for analysis of time to ART requirement and time to ART initiation, patients who started ART on the day of recruitment were artificially given one day of follow-up. Analyses were repeated excluding these patients as a sensitivity analysis, but no changes were observed. All four models were adjusted by CD4 count and VL at recruitment, to account for duration of infection prior to recruitment[78]. Sensitivity

analyses excluding patients with missing information in either CD4 count or VL were carried out, with no impact on results.

A pooled logistic regression was performed alternatively for all four outcomes in order to adjust for informative censoring trough inverse probability weighting. The current cumulative month of follow-up and the cubic splines of the months with five knots at percentiles 5, 22.5, 50, 72.5 and 95, were used to account for time and introduced in the model. Patients who had not had any follow-up visit in the year prior to administrative censoring were considered to be lost to follow-up. The probability of not being LTFU at each month of patient's follow-up was calculated taking into account all socio-demographical variables, CD4 count, VL, and clinical and treatment status at baseline and at every particular month. As a sensitivity analysis, LTFU was defined as subjects who had not had any follow-up visit two and three years prior to administrative censoring, and results remained unchanged.

Finally, time to virological and immunological response to treatment was analysed through a Fine and Gray regression, considering deaths occurred before response as a competing event and censoring each patient at their last Viral Load or CD4 count assessment, respectively.

For every outcome, interaction between region of origin and sex was systematically explored, to assess if any gender differences were observed. Other interactions could not be explored due to insufficient sample size. Proportionality of hazards assumption was tested for region of origin and every outcome.

Statistical significance was evaluated using the Wald test. All the analyses were conducted using Stata software (V.11.1, Stata Corporation, College Station, Texas, USA).

6. RESULTS

6.1. Results for Objective 1

6.1.1. Characteristics of the sample

From 1986 to 2009, 479 seroconverters were identified with a median seroconversion in August 1999 (IQR: December 1992- June 2005). 60% of the sample were subjects infected after 1996, in the HAART era. Considering patients with no clinical follow-up visit in the last three years as lost to follow up, the percentage of patients lost was 35%, and median follow-up time for the total cohort was 3.7 years (IQR:1.2-9.2).

In 433 subjects (90.4%) seroconversion was estimated as the mid-point between a previous negative and the first positive test, and median interval between both tests was 0.96 years (IQR: 0.57-1.57). In the remaining 46 subjects (9.6%) seroconversion was established on the date of a Western Blot test with an undetermined result. In this latter percentage of women was 19.9 points higher (CI: 0.06-0.34; p<0.01) and subjects were a mean of 3.0 years younger (CI: 0.70-5.30; p=0.01).

Out of the sample subjects, 80% were MSM, followed by 14% of IDU. The majority were males with high educational level (68.9% of the total sample). Median age at seroconversion was 29.4 (IQR: 25.1-34.5). Most of the patients were of Spanish origin (78.9%), although there was a considerable proportion of subjects coming from Latin America and the Caribbean (15.4%). Median CD4 count at diagnosis was 663cells/mm3 (IQR: 501-826). Significant differences were found in the sample accord-

ing to educational level, so Table 1 shows baseline characteristics of the sample stratified by this variable.

Table 1. Baseline characteristics of the sample by educational level.

	Low education N=105 (21.9%)		**High education** N=330 (68.9%)		**Education unknown** N=44 (9.2%)		**Total** N= 479		**P***
	n	%	n	%	n	%	n	%	
Transmission category									
Homo/bisexual	56	53.3	291	88.2	36	81.8	383	80.0	0.000
Heterosexual	8	7.6	18	5.4	1	2.3	27	5.6	
Injecting drug user	40	38.1	20	6.1	7	15.9	67	14.0	
Other	1	1.0	1	0.3	0	0.0	2	0.4	
Sex									
Males	88	83.8	311	94.2	41	93.2	440	91.9	0.003
Females	17	16.2	19	5.8	3	6.8	39	8.1	
Age at seroconversion									
# in each group	105		329		44		478		0.000
Median (IQR)	26.3 (22.3-32.8)		30.1 (26.3-34.8)		30.1 (25.8-35.2)		29.4 (25.1-34.5)		
Region of origin									
Spain	90	85.7	253	76.7	35	79.5	378	78.9	0.337
Europe	4	3.8	11	3.3	3	6.8	18	3.8	
Latin America	10	9.6	59	17.9	5	11.4	74	15.4	
Other	1	0.9	7	2.1	1	2.3	9	1.9	
Seroconversion date									
1986 – 1996	67	63.8	104	31.5	20	45.5	191	39.9	0.000
1997 – 2009	38	36.2	226	68.5	24	54.5	288	60.1	
Calendar period**									
1986 – 1996	64	39.7	99	23.7	20	31.7	183	28.5	0.001
1997 – 2009	97	60.3	319	76.3	43	68.2	459	71.5	
CD4 at diagnosis***									
# in each group	70		264		29		363		0.318
Median (IQR)	699 (544-885)		649 (488-823.5)		665 (540-756)		663 (501-826)		
VL at diagnosis * **									
# in each group	33		204		21		258		0.602
Median (IQR)	39603 (9219-82832)		32585.5 (10045-87284.5)		26192 (5886-64771)		32333 (9885-86172)		

*Statistical significance (p-value) for the difference between the three groups of educational level. Chi2 used for categorical variables and Kruskal-Wallis test for continuous. Results remain largely unchanged if comparison is established only between both groups with known educational level; **Subjects are considered to belong to a calendar period if he/she contributes any follow-up time to that period, so a patient can simultaneously contribute to both periods. Therefore, total number of patients for this variable is larger than total number of patients in the sample; *** For these variables, CD4 counts and VL are considered when performed within a 3-month interval from the first HIV positive test; IQR: Interquartile range.

Patients with a lower educational level comprised a higher proportion of IDU, of women, of native Spanish patients, were younger and corresponded to infections acquired in more early stages of the HIV epidemic, with a higher proportion of seroconversions in the period before 1996. 39.9% of patients were prescribed antiretroviral treatment at some point of their follow up, being HAART in an 89% of cases.

6.1.2. Cumulative incidence of AIDS

During the total follow-up time of 2,953 person-years (py), 59 cases of AIDS were diagnosed, for an incidence rate of 20.0 cases per 1000 py. Incidence rate was three times higher before 1997 (38.4/1000py) than after that date (13.3/1000py).

Figure 1 shows cumulative incidence of AIDS function in each calendar period. As for variables associated with risk of AIDS, in the crude analysis only educational level and calendar period showed any effect. Results for all evaluated variables can be found in Table 2 and Figure 2 shows cumulative incidence of AIDS by each group of educational level, being worthy of note that subjects of high educational level showed a lower risk of AIDS at all points of patients' follow up.

Multivariate modelling selected three independent predictors for risk of AIDS: calendar period, educational level and age at seroconversion. The latter could be introduced as a continuous variable in the model.

Adjusted HR showed that after 1997 risk of AIDS was reduced by a 78.5% (95%CI: 48.1-89.0; $p<0.01$). Effects of both age and educational level were found to be different in each calendar period: p-value of the interaction between period and age was 0.006; and p-value for interaction with educational level was not significant

(p=0.396), but HR for the two calendar periods differed by 33.4%, so a possible effect modification was taken into account.

Figure 1. Cumulative incidence of AIDS by calendar period

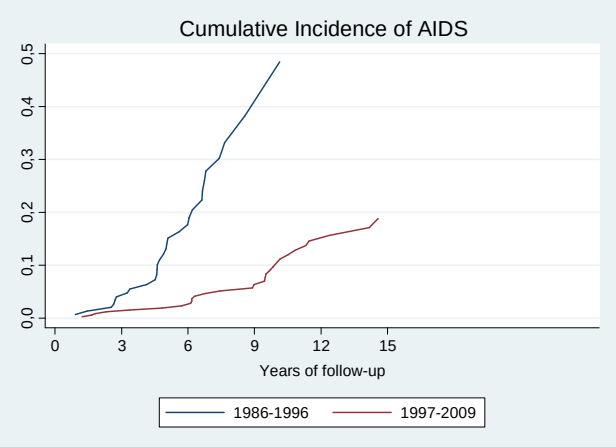


Figure 2. Cumulative incidence of AIDS by educational level

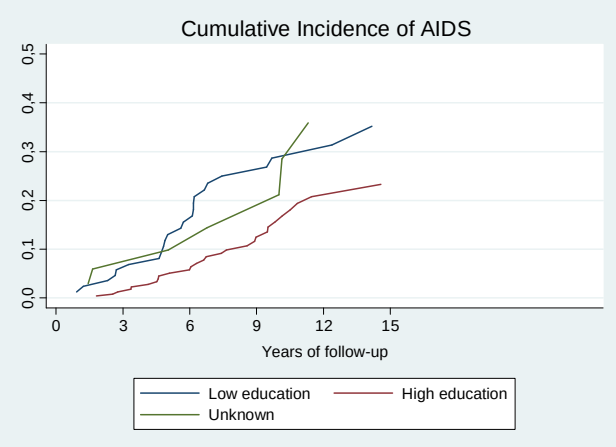


Table 2. Risk of AIDS. Results of univariate analysis.

	HR (CI)	p
Transmission category		
Sexual route and others	1	
Injecting drug user	1.349 (0.788-2.308)	0.275
Sex		
Males	1	
Females	1.269 (0.666- 2.419)	0.470
Region of origin		
Spain	1	
Latin America	1.440 (0.632- 3.283)	0.386
Other	0.524 (0.067- 4.103)	0.538
Calendar period		
1986 – 1996	1	
1997 – 2009	0.223 (0.122- 0.410)	0.000
Educational level		
Low	1	
High	0.511 (0.293-0 .891)	0.018
Unknown	0.874 (0.380- 2.013)	0.752
Age at seroconversion	1.000 (0.965-1.035)	0.980
CD4 at diagnosis (n=454)	0.999 (0.998-1.001)	0.541
CV at diagnosis (n=254)	1.000 (0.999-1.000)	0.905

HR: Hazard Ratio; CI: Confidence Interval

Table 3. Risk of AIDS stratified by period. Results of multivariate analysis.

		HR (CI)	p
Educational level			
1986 - 1996	Low	1	
	High	0.668 (0.314- 1.424)	0.297
	Unknown	0.547 (0.158- 1.897)	0.342
1997 - 2009	Low	1	
	High	0.445 (0.192- 1.031)	0.059
	Unknown	0.986 (0.319- 3.049)	0.980
Age at seroconversion			
1986 - 1996	Age (in years)	1.071 (1.038- 1.105)	0.000
1997 - 2009	Age (in years)	0.982 (0.936- 1.031)	0.465

HR: Hazard Ratio, CI: Confidence Interval

Table 3 shows the effect of both age and educational level stratified by calendar period. Before 1997, risk of AIDS increased by 7.0% (95%CI: 3.8-10.5) for each year of age at seroconversion, and in this period, no effect of educational level over risk of AIDS is evident. As opposed to these findings, in the HAART era (after 1997), individuals of high educational level came to have 55.5% lower risk of AIDS (95%CI: 80.8% lower-3.1% higher), and no effect of age is longer observed. Transmission category was not statistically significant in the adjusted analysis.

6.1.3. Cumulative incidence of all-cause mortality

During a total follow-up of 3,494 person-years (py) 27 deceases were observed, for an incidence rate of 7.7/1000py. Incidence rate was four times higher before 1997 (16.7/1000py) than after this year (4.9/1000py). Univariate analysis results are shown in Table 4. Variables with a crude effect over risk of death were transmission category, educational level and calendar period. Figure 3 shows risk of death by calendar period and points out a higher incidence of deaths in the period previous to 1997. Figure 4 shows cumulative risk of death by educational level.

Table 4. All-cause mortality risk. Results from univariate analysis.

	HR (CI)	p
Transmission category		
Sexual route and others	1	
Injecting drug user	2.282 (1.096-4.751)	0.027
Sex		
Males	1	
Females	1.440 (0.553-3.595)	0.472
Region of origin		
Spain	1	
Latin America	1.446 (0.438-4.773)	0.545
Other	1.410 (0.182-10.907)	0.742
Calendar period		
1986 – 1996	1	
1997 – 2009	0.148 (0.059-0.374)	0.000
Educational level		
Low	1	
High	0.357 (0.157-0.813)	0.014
Unknown	0.725 (0.203-2.581)	0.619
Age at seroconversion	1.018 (0.966- 1.073)	0.499
CD4 at diagnosis (n=454)	0.998 (.996-1.001)	0.135
CV at diagnosis (n=254)	Cannot be estimated*	

* In the group of patients with available information on VL at HIV diagnosis there was no decease observed, so HR could not be estimated. HR: Hazard Ratio, CI: Confidence Interval

Figure 3. Cumulative incidence of all-cause mortality by calendar period

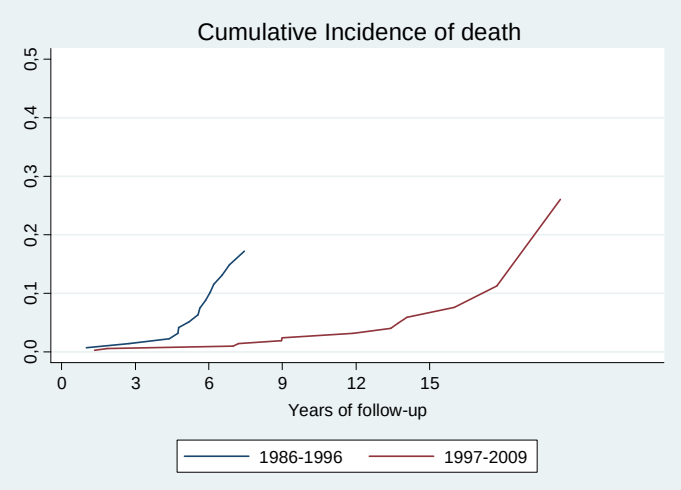
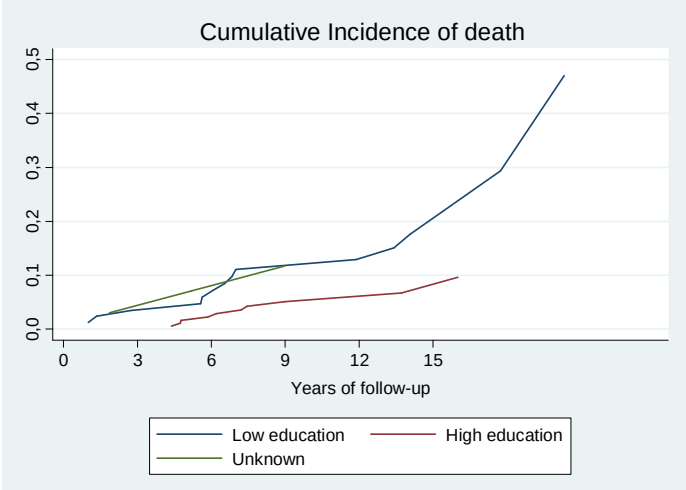


Figure 4. Cumulative incidence of all-cause mortality by educational level



Results of the adjusted analysis can be found in Table 5. Adjusted effect for calendar period shows survival after HAART availability was improved by 86.6% (95%CI: 65.4-94.8). People with high educational level had 61.7% (95%CI: 12.5-83.2) less risk of dying compared to those with low educational level. Effect of transmission category was no longer observed after adjusting for educational level. Risk of death increased by 4.8% (95%CI: 1.4-8.4) for each year of age at seroconversion. Both effects were homogeneous throughout the study period.

Table 5. All-cause mortality risk. Results from multivariate analysis.

	HR (CI)	p
Calendar Period		
1986 – 1996	1	
1997 – 2009	0.134 (0.052- 0.346)	0.000
Educational level		
Low	1	
High	0.383 (0.168- 0.875)	0.023
Unknown	0.693(0.184- 2.613)	0.588
Age at seroconversion	1.048 (1.014- 1.084)	0.006

HR: Hazard Ratio, CI: Confidence Interval

6.2. Results for Objective 2

6.2.1. Characteristics of the sample

For the period between April 1983 and February 2009, 1772 patients were included in the cohort, with a median seroconversion date on December 1994 (IQR): July 1991 - December 1999. The median seroconversion window (time between the last negative and first positive test) was 1.0 years (IQR: 0.6-1.7).

A total of 783 (44.2%) patients did not have information on educational level and were excluded. 67.4% of them corresponded to IDU from a prison cohort and a detoxification unit. Compared to those with available information, they were younger, with a higher proportion of IDU, women and people with missing region of origin, had been recruited in early calendar periods, had been imprisoned in their life-time and had lower CD4 counts at entry.

The final sample for analysis consisted of 989 persons. Of these, 9.7% had not completed primary education, 42.4% had only achieved primary education, 28.8% had secondary education and 19.1% had completed studies beyond secondary education. In all, 515 subjects (52.1%) were included in the category of low education and 474 (47.9%) in the category of high education. Some 52.4% were IDU, and the rest of the sample was composed of MSM (85.1%) and heterosexuals (11.7%), with a small proportion (3.2%) of unknowns.

Table 6 shows a description of the sample by educational level and calendar period. The low education group was composed mostly of IDU, had more women and

young people and fewer foreigners than the high education group. In more recent calendar periods, patients had higher education, were older, more likely to be foreigners and acquired the infection by sexual transmission. No differences were detected with regards to CD4 or viral load at the time of diagnosis by educational level.

Table 6. Socio-demographic and clinical characteristics by educational level

	Low education N=515 (100%)		High education N=474 (100%)		Total N=989 (100%)		p
	n	%	n	%	n	%	
Transmission category							
Sexual route and others	106	20.6	365	77.0	471	47.6	0.000
Injecting drug users	409	79.4	109	23.0	518	52.4	
Sex							
Males	394	76.5	418	88.2	812	82.1	0.000
Females	121	23.5	56	11.8	177	17.9	
Age at seroconversion							
Median (IQR)	25.6 (22.7-29.4)		29.8 (25.4-34.5)		27.4 (24.0-32.5)		0.000
Region of origin							
Spain	467	90.7	373	78.7	840	84.9	0.000
Europe	8	1.5	14	2.9	22	2.2	
Latin America and the Caribbean	11	2.1	64	13.5	75	7.6	
Other	3	0.6	7	1.5	10	1.0	
Unknown	26	5.1	16	3.4	42	4.3	
Method of estimating seroconversion							
Evidence of seroconversion	10	1.9	32	6.75	42	4.25	0.000
Mid-point between (-) and (+)	505	98.1	442	93.25	947	95.75	
Calendar period*							
<=1996	362	35.5	169	21.8	531	29.6	0.000
1997-2003	397	38.9	261	33.6	658	36,7	
>=2004	261	25.6	346	44.6	607	33,8	
CD4 at diagnosis** (cells/µl)							
N in each group	210		338		548		0.993
Median (IQR)	611 (464-857)		630 (459-817)		623 (462.5-831.5)		
Viral load at diagnosis ** (copies/µl)							
N in each group	77		250		327		0.326
Median (IQR)	46.6 (9.8-159.0)		35.8 (10.0- 98.2)		36.8 (10.0– 108.2)		

* Each subject is counted if he/she contributes periods of risk to the period, therefore the total for this variable is larger than the total number of subjects; ** Taken from tests made on the date of the first positive result plus/minus 3 months; IQR: Interquartile range; p: statistical significance, calculated by Chi square for categorical variables and Mann-Whitney test for continuous variable.

6.2.2. Effect of educational level on risk of AIDS

In the analysis of time to AIDS, 203 events were produced over 6783.1 person-years. The incidence rate was 37.7 per 1000 (95%CI: 32.1 to 44.2) in persons with low education versus 18.9 per 1000 (95%CI: 14.4 to 24.7) in those with high education.

In the adjusted analysis, transmission route and calendar period were confounding variables between educational level and the risk of AIDS, with a notable change in the crude HR observed after adjusting for these variables. A significant interaction was detected between educational level and period, so that no effect was observed before 1996, but between 1997 and 2003, persons with high education had a 42% lower risk of AIDS (95%CI: 1% to 46%); and this effect was larger for the period beginning in 2004 (74% lower; 95%CI: 32 to 90%; Table 6). No significant interaction was found with transmission category. Figure 5 shows the cumulative incidence function by educational level for each period.

Figure 5. Cumulative incidence of AIDS by educational level and period.

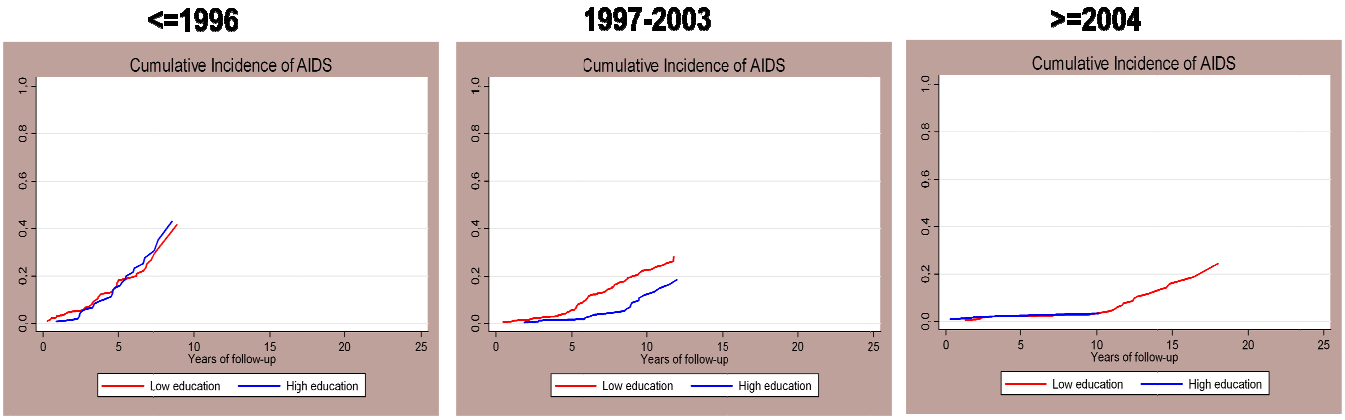


Table 7. Adjusted effect of high educational level on each of the events

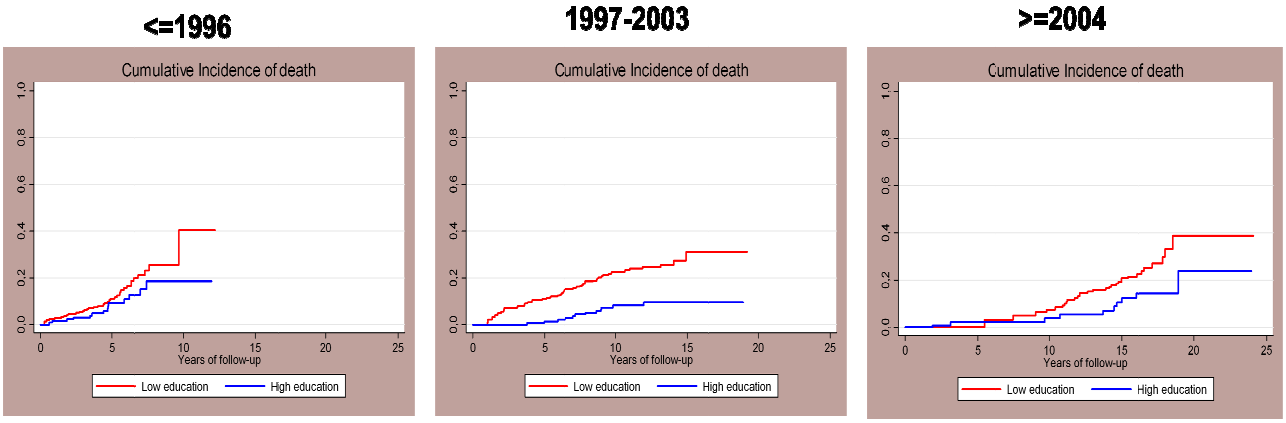
	AIDS*	DEATH**	HAART***
	sHR [CI]	HR [CI]	sHR [CI]
<=1996	1.47 [0.91-2.38] p=0.115	0.68 [0.45-1.03] p=0.069	0.84 [0.56-1.27] p= 0.41
1997-2003	0.58 [0.34-0.99] p=0.047		
>=2004	0.26 [0.10-0.68] p=0.006		

* Adjusted for confounding variables: transmission route and calendar period; ** Adjusted for confounding variables: transmission route and age at seroconversion; *** Adjusted for confounding variables: transmission route; HR: Hazard Ratio; sHR: sub hazard ratio.

6.2.3. *Effect of educational level on overall survival*

There were 164 deaths over 8432.8 person-years. The incidence rate was 25.0 per 1000 (95%CI: 21.0 to 30.0) in persons with low education and 10.9 per 1000 (95%CI: 7.9 to 15.1) in those with high education. Figure 6 shows that the cumulative incidence of death in each period tends to be lower for persons with high education in all periods. After adjusting for confounding variables –transmission rate and age at seroconversion- persons with high education had 32% lower risk of death (95%CI: 55% lower to 3% higher), a result of borderline statistical significance (Table 7).

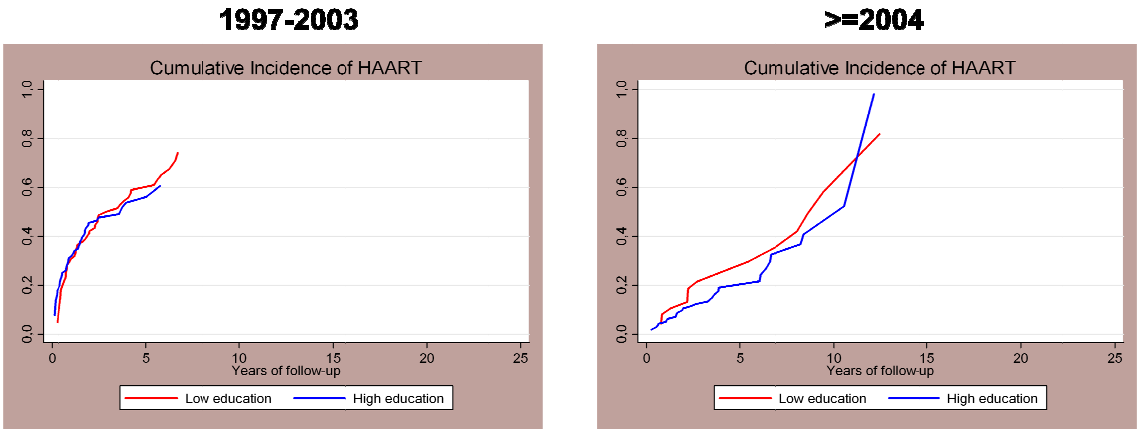
Figure 6. Cumulative incidence of death by educational level and period.



6.2.4. Effect of educational level on access to HAART

Access to HAART was measured by the combination of the information on two outcomes: time to HAART requirement and time to HAART initiation. This analysis was restricted to those seroconverting after 1996, a total of 435 persons. There were 186 patients who fulfilled HAART requirement criteria and 122 initiated HAART during 1149.8 and 914.8 person-years of follow-up, respectively. No crude effect in HAART requirement or initiation by period and educational level was observed. Figure 7 shows this similar cumulative incidence of HAART initiation both before and after 2003. After adjusting for confounders (transmission route and calendar period) no effect of educational level was seen in any of the two HAART outcomes and no interaction was found (Table 7).

Figure 7. Cumulative incidence of HAART initiation by educational level and period.



6.3. Results for Objective 3

6.3.1. Characteristics of the sample

Of the 6811 subjects recruited in CoRIS up to October 2010, 6278 (92.2%) were either native Spanish (NSP, n= 4657, 74.2%), Latin Americans and Caribbeans (LAC, n=1221, 19.4%) or Sub-Saharan Africans (SSA, n=400, 6.4%). NSP were older, more frequently infected through injecting drug use (IDU) and had higher viral loads, both at recruitment and at treatment initiation (Table 8). The majority of LAC were men who have sex with men (MSM), had higher educational level and higher percentage of CDC stage A at recruitment. SSA had the highest proportion of females, heterosexual transmission, lower education, delayed diagnosis, CDC stage C and lowest CD4 counts at recruitment. Both LAC and SSA had been enrolled in CoRIS more frequently in recent years. Importantly, no differences were found in CD4 counts at ART initiation.

Figure 8 shows a flow chart specifying which population contributed to the assessment of each of the outcomes and which criteria where applied for subjects exclusion. No differences were found by sex in any of the analysis.

6.3.2. HIV Diagnostic Delay by region of origin

Of the 4894 patients assessed, diagnostic delay (DD) of HIV infection was observed in 2434 (49.7%). Prevalence of DD was significantly lower in NSP (48%) compared to both LAC (51.6%) and SSA (62.8%).

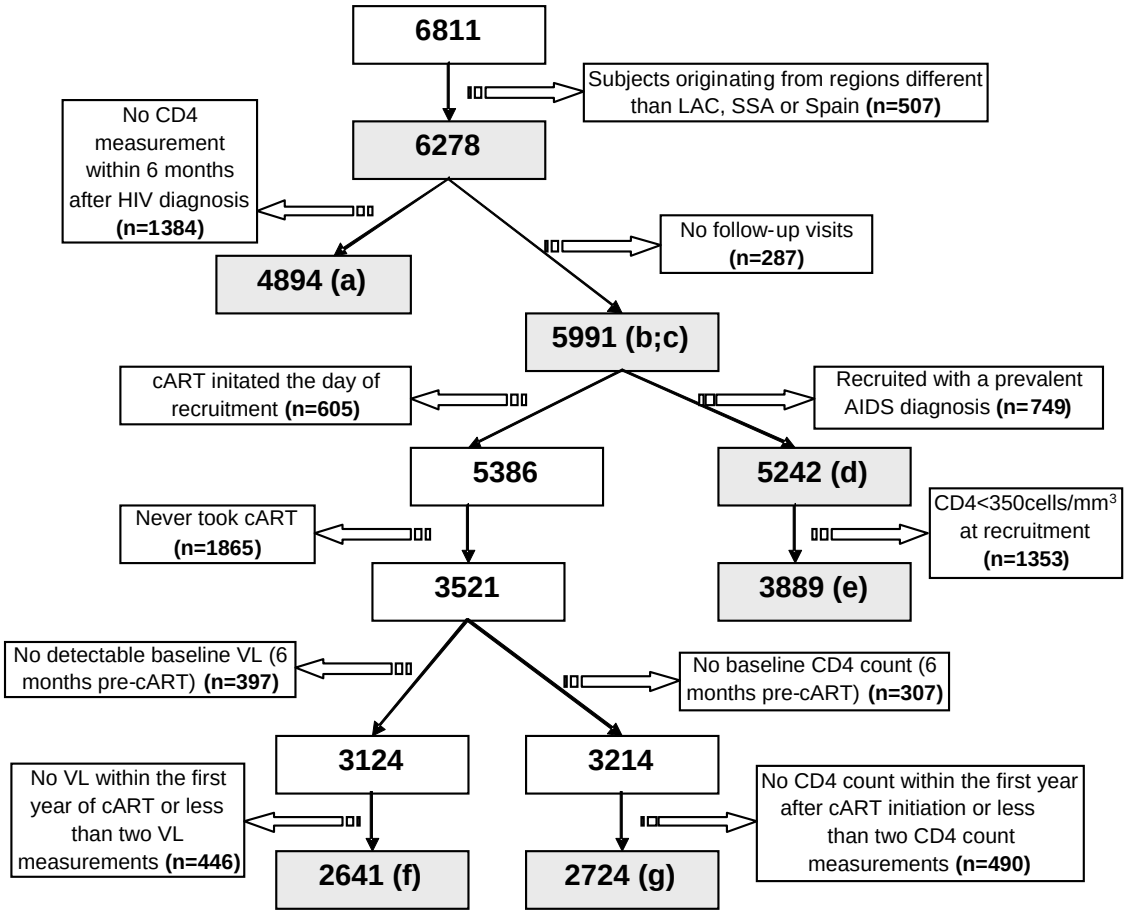
Table 8. Characteristics of the sample by geographical origin of the patient.

		NSP N=4657		LAC N=1221		SSA N=400		Total N=6278		P*
Sex	[n(%)]									
Male		3884	(83.4)	986	(80.7)	183	(45.8)	5053	(80.5)	<0.001
Female		773	(16.6)	235	(19.2)	217	(54.2)	1225	(19.5)	
Age										
	[Median (IQR)]	37 (30-43)		32 (27-38)		31 (27-38.5)		35 (29-42)		<0.001
Mode of transmission	[n(%)]									
MSM		2364	(50.8)	770	(63.1)	8	(2.0)	3142	(50.1)	<0.001
Heterosexual		1324	(28.4)	406	(33.2)	358	(89.5)	2088	(33.3)	
IDU		798	(17.1)	17	(1.4)	9	(2.3)	824	(13.1)	
Other/NA		171	(3.7)	28	(2.3)	25	(6.3)	224	(3.6)	
Educational level	[n(%)]									
Secondary or Higher		2281	(49.0)	664	(54.4)	87	(21.7)	3032	(48.3)	<0.001
Primary or none		1624	(34.9)	386	(31.6)	218	(54.5)	2228	(35.5)	
Unknown		752	(16.1)	171	(14.0)	95	(23.8)	1018	(16.2)	
CDC stage	[n(%)]									
A/P		3510	(75.4)	971	(79.5)	280	(70.0)	4761	(75.8)	<0.001
B		425	(9.1)	91	(7.5)	32	(8.0)	548	(8.7)	
C		599	(12.9)	133	(10.9)	81	(20.3)	813	(13.0)	
Unknown		123	(2.6)	26	(2.1)	7	(1.8)	156	(2.5)	
Year of recruitment	[n(%)]									
2004-2006		1872	(40.2)	408	(33.4)	151	(37.8)	2431	(38.7)	<0.001
2007-2010		2785	(59.8)	813	(66.6)	249	(62.3)	3847	(61.3)	
Delayed diagnosis	[n(%)]									
No		1865	(40.1)	472	(38.7)	123	(30.8)	2460	(39.2)	<0.001
Yes		1723	(37.0)	503	(41.2)	208	(52.0)	2434	(38.8)	
Unknown		1069	(23.0)	246	(20.2)	69	(17.3)	1384	(22.1)	
Losses to Follow-up	[n(IR)]	996	(8.1)	309	(11.1)	168	(19.7)	1473	(9.3)	<0.001
CD4 count at enrolment (cells/mm3)	n	4440		1160		379		5979		<0.001
	[Median (IQR)]	364 (180-572)		350.5 (171.5-527)		284 (130-474)		357 (175-555)		
Viral Load at enrolment (log copies/ml)	n	4420		1152		373		5945		<0.001
	[Median (IQR)]	4.66 (4.05-5.16)		4.56 (3.97-5.07)		4.58 (3.65-5.11)		4.64 (4.01-5.14)		
CD4 count at ART (cells/mm3)	n	2167		510		196		2873		0.945
	[Median (IQR)]	210 (85-306)		210 (97-302)		208.5 (92.5-287.5)		210 (89-302)		
Viral Load at ART (log copies/ml)	n	2129		501		193		2823		<0.001
	[Median (IQR)]	4.99 (4.44-5.40)		4.80 (4.27-5.28)		4.83 (4.06-5.18)		4.94 (4.37-5.37)		

NSP: Native Spanish; LAC: Latin America and the Caribbean; SSA: Sub-Saharan Africa; CDC: Centres for Disease Prevention and Control; IQR: Interquartile range; IR: Incidence Rate per 100 persons-year of follow up* p-values are for Chi-sq, Kruskal-Wallis or Log-Rank (as appropriate) for differences across regions.

Transmission route was a confounder of the association between region of origin and diagnostic delay, and an interaction was found between region of origin and age, so adjusted Odds Ratios (ORs) are shown stratified by age (Table 9). SSA below 35 years of age had a risk of DD which doubled that of young NSP, while older SSA showed no difference. For LAC, the excess risk compared to NSP was seen for subjects under 50, being more pronounced at ages under 35.

Figure 8. Flow chart of the sample. Population used to assess each outcome.



(a) Patients assessed for Delayed Diagnosis; (b) Patients assessed for risk of death; (c) Patients assessed for time until ART initiation; (d) Patients assessed for risk of AIDS; (e) Patients assessed for time until ART requirement; (f) Patients assessed for virological response to ART; (g) Patients assessed for immunological response to ART

Table 9. Results of the univariate and multivariate analysis for all the outcomes in the study

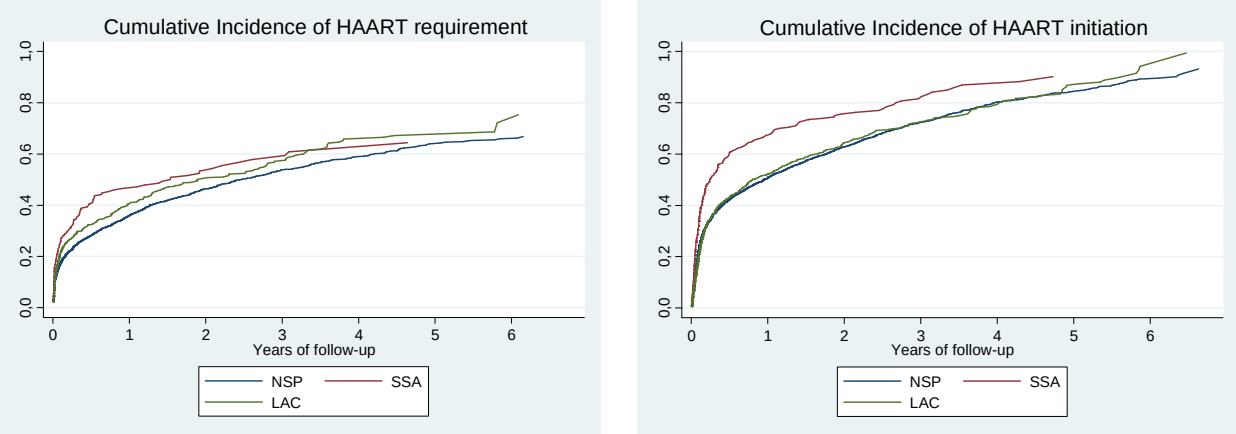
				SSA OR/HR [95%CI]		**LAC** OR/HR [95%CI]	
OUTCOME	# Events / N[a]		(%; IR[b])	Crude	Adjusted[c]	Crude	Adjusted[c]
Diagnostic Delay	Age			1.83 [1.45-2.31]**	Age	1.15 [1.00-1.33]*	Age
	<35	948 / 2382	(39.8)		<35 2.02 [1.47-2.78]**		<35 1.69 [1.39-2.06]**
	35-50	1116 / 1987	(56.2)		35-50 1.15 [0.78-1.72]		35-50 1.28 [1.00-1.64]*
	>50	370 / 525	(70.5)		>50 0.42 [0.14-1.24]		>50 0.84 [0.44-1.62]
ART requirement		1897 / 4306	(44.1)	1.22 [1.00-1.49]*	0.90 [0.67-1.21]	1.13 [1.01-1.26]*	1.02 [0.88-1.18]
ART initiation		3521 / 5765	(61.1)	1.43 [1.16-1.62]**	1.06 [0.91-1.24]	1.01 [0.93-1.09]	0.91 [0.84-1.00]*
Immuno. response		2226 / 2578	(86.3)	0.81 [0.69-0.96]*	0.80 [0.67-0.96]*	1.02 [0.92-1.13]	1.04 [0.94-1.16]
Viro. response		2064 / 2726	(75.7)	0.79 [0.65-0.94]*	0.74 [0.61-0.90]**	1.01 [0.90-1.13]	1.00 [0.89-1.13]
AIDS	Age			1.17 [0.77-1.78]	Age	0.80 [0.60-1.06]	Age
	<35	109 / 6422	(1.7)		<35 0.94 [0.46-1.91]		<35 0.71 [0.44-1.14]
	35-50	201 / 5334	(3.8)		35-50 2.05 [1.12-3.74]*		35-50 1.58 [1.05-2.40]*
	>50	64 / 1140	(5.6)		>50 1.47 [0.27-7.87]		>50 0.25 [0.04-1.79]
AIDS (excl. TB)		250 / 12897	(1.9)	1.04 [0.60-1.80]	1.26 [0.70-2.26]	0.85 [0.61-1.19]	1.19 [0.81-1.74]
Death		231 / 15868	(1.5)	0.62 [0.32-1.21]	0.66 [0.31-1.39]	0.54 [0.35-0.81]**	0.85 [0.55-1.30]

SSA: Sub-Saharan Africa; LAC: Latin America and the Caribbean; OR: Odds Ratio; HR: Hazard Ratio; IR: Incidence Rate per 100 person-years; [a] Denominator differs between outcomes: Diagnostic Delay (DD): number of subjects assessed for DD in each age group; The rest of outcomes: person-years at risk; [b] Percentage is shown for DD outcome. IR is shown for the other 7 outcomes; [c] For DD outcome, OR is shown adjusted by transmission route and age at recruitment; For time to ART requirement HR was adjusted by CD4 count and viral load (VL) at recruitment; For time to ART initiation, HR was adjusted by CD4 count and VL at recruitment and transmission route; For time to immunologic response and time to virological response, HR was adjusted by CD4 count and VL at treatment initiation; For time to AIDS, time to AIDS excluding TB and time to death, HR was adjusted for CD4 count and VL at recruitment, transmission route and age at recruitment; * $p<0.05$; ** $p<0.01$.

6.3.3. Access to treatment by region of origin

Of the 3889 subjects assessed for ART requirement (4306 person-years of follow-up), 1897 subjects required treatment (48.8%; Incidence rate 44.1/100py), higher for SSA (68.1/100py) and LAC (49.5/100py) compared to NSP (41.6/100py). Figure 9 shows the cumulative incidence of ART requirement by geographical origin of the patient. Crude Hazard Ratios (HRs) showed this faster progression to ART requirement for both groups, which disappeared after adjusting for CD4 counts and VL at recruitment (Table 9), suggesting the effect of recruitment in later stages of disease and therefore, shorter time to ART requirement at the crude level, but not after adjusting for duration of infection.

Figure 9. Cumulative incidence of ART requirement and ART initiation by region of origin



Overall, 3521 subjects initiated ART out of the 5991 patients (5765 py of follow-up) assessed for this outcome (58.8%, IR 61.1/100py). Incidence was highest for SSA (102.7/100py) compared to LAC (62.5/100py) and NSP (58.6/100py). Crude HR shows

faster ART initiation for SSA, which disappears after adjusting for CD4 count and VL at recruitment and transmission route. LAC do not show any difference in the crude analysis, but a higher risk of a delayed initiation of ART of small magnitude and clinical relevance is observed after adjustment (Table 9). Initial antiretroviral regimes for NSP, LAC and SSA are shown in Table 10, and no differences are observed.

Table 10. Initial Antiretroviral Therapy regime by region of origin

ART initial regime [n(%)]	NSP	SSA	LAC	Total
2 NRTI + 1 NNRTI	1448 (56.6)	134 (55.1)	376 (58.0)	1998 (56.8)
2 NRTI + 1 PI/r	877 (33.4)	85 (35.0)	202 (31.2)	1164 (33.1)
2 NRTI + 1PI	55 (2.1)	7 (2.9)	16 (2.5)	78 (2.2)
2 NRTI+ 1 Integrase Inhibitor	37 (1.4)	1 (0.4)	16 (2.5)	54 (1.5)
3 NRTI	42 (1.6)	3 (1.2)	7 (1.1)	52 (1.5)
Other	131 (4.9)	13 (5.3)	31 (4.8)	175 (5)
TOTAL	**2630 (100)**	**243 (100)**	**648 (100)**	**3521 (100)**

NSP: Natural Spanish; SSA: Sub-Saharan Africa; LAC: Latin America and The Caribbean; NRTI: nucleotide reverse transcriptase inhibitors; NNRTI: non-nucleotide reverse transcriptase inhibitors; PI: protease inhibitors. Fisher's exact p-value for the table=0.36.

6.3.4. Response to treatment by region of origin

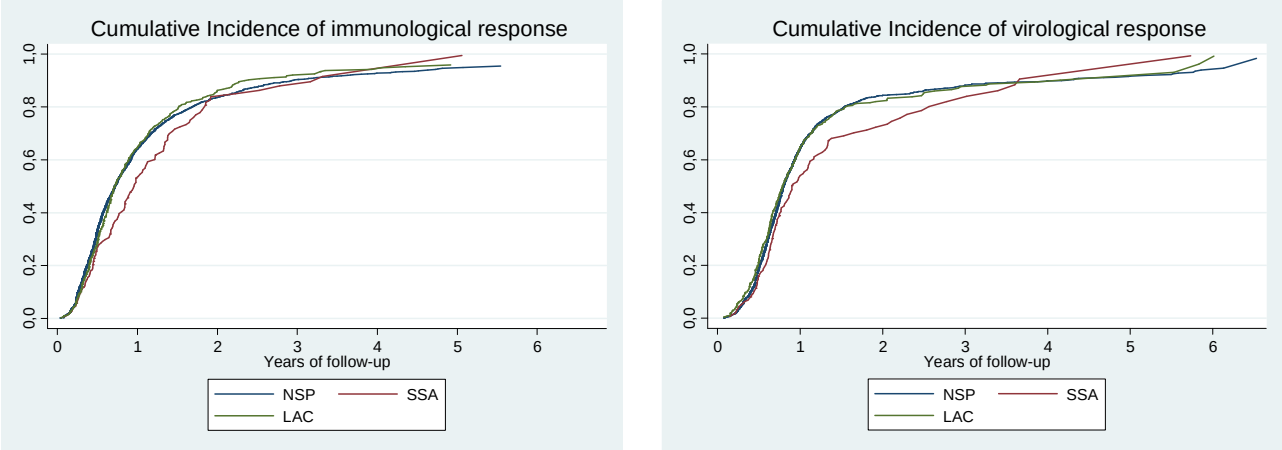
Of the 2724 subjects assessed for immunological response, 2226 (81.7%) responded at some point in time, 1896 (69.6%) within the first year of ART. Global response rate was 86.3/100 py, and was lower for SSA (71.8/100py) compared to NSP (86.9/100py) and LAC (89.4/100py).

2641 patients were assessed for virological response, of which 2064 responded at some time (78.2%) and 1852 (70.1%) within the first year. Virological response rate

was 75.7/100py globally and again was significantly lower for SSA (61.5/100py) compared to NSP (76.5/100py) and LAC (77.6/100py).

Figure 10 shows cumulative incidence of immunological and virological response by region of origin. Higher risk of delayed immunological and virological response in SSA was observed in the crude analyses and after adjustment for CD4 count and VL at ART initiation, while no differences were found for LAC (Table 9).

Figure 10. Cumulative Incidence of immunological and virological response by region of origin

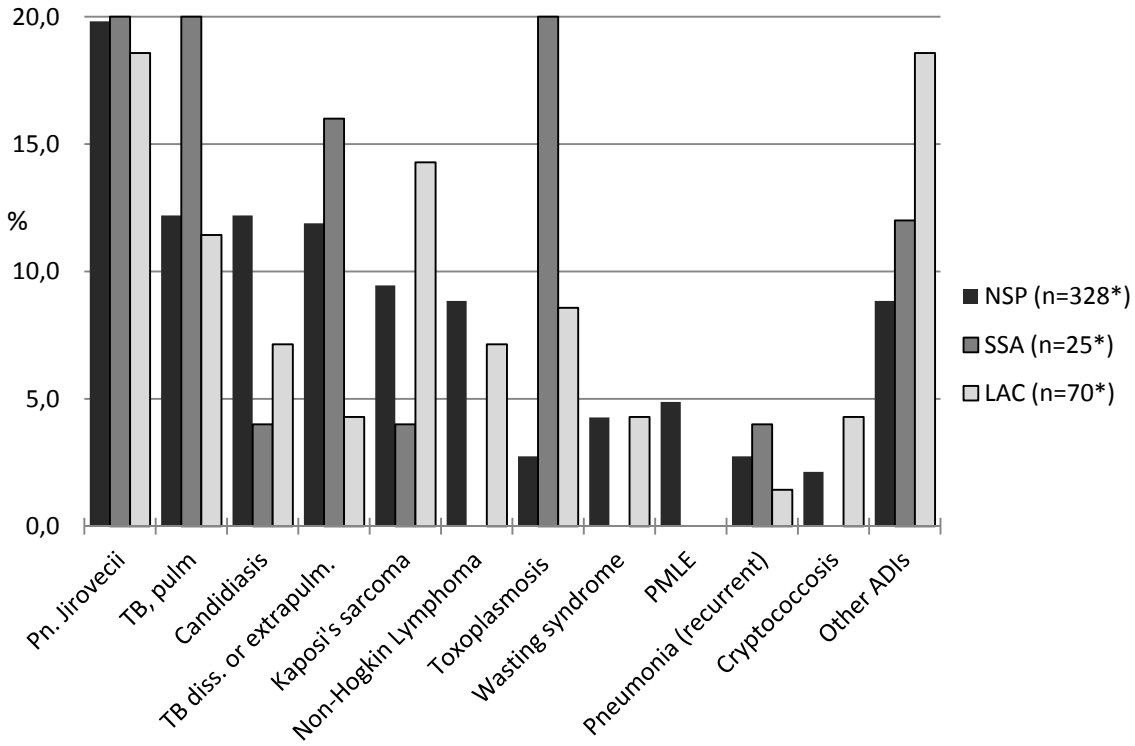


Total number of CD4 count determinations after treatment initiation was 8 (IQR:4-12), slightly higher for NSP (8; IQR:5-13) than for LAC (7; IQR:4-12) and for SSA (6; IQR:3-10) , with a p-value for differences across three groups <0.01. Number of VL determinations showed similar significant differences across regions (p<0.01), being for 8 determinations for NSP (IQR:5-13), 7 for LAC (IQR:4-11) and 6 for SSA (IQR:4-10). However, none of these variables showed any confounding effect between region of origin and immunological or virological response.

6.3.5. Risk of AIDS and death by region of origin

Of the 5242 patients evaluated -12898 py of follow up-, 374 were diagnosed of AIDS during their follow-up (7.1%, IR 2.9/100py), 280 of them within the first year. AIDS defining illnesses (ADIs) by region of origin are described in Figure 11. AIDS incidence was highest for SSA (3.7/100py), lowest for LAC (2.4/100py) and intermediate for NSP (3.0/100py). Figure 12 shows the cumulative incidence of AIDS by region of origin, and cumulative incidence when tuberculosis as an AIDS defining illness is excluded from the analysis.

Figure 11. Initial AIDS defining illnesses of incident AIDS diagnosis (n=374) by region of origin



Pn. Jirovecii: Pneumocystis jirovecii pneumonia; TB: Mycobacterium tuberculosis; pulm.: pulmonary; diss: disseminated; extrapulm.: extrapulmonary; PMLE: Progressive multifocal leukoencephalopathy; ADIs: AIDS defining illnesses. *One patient can simultaneously experience more than one ADI at AIDS diagnosis, so total "n" is larger than total number of patients.

Crude HR showed no significant differences across regions. Age was found to modify the effect of the region of origin, so HRs are shown stratified by age and adjusted for CD4 count and VL at recruitment and transmission route (Table 9). Adjustment revealed a higher risk of AIDS for both SSA and LAC in the medium ages (35 to 50 years old). A sub-analysis was performed excluding Tuberculosis (TB) and no differences in risk of AIDS were longer observed. Among the 5991 subjects -15868 py-, that were included for the overall survival analysis, 231 patients died (3.9%, IR 1.5/100py). Death rate was higher for NSP (1.6/100py), followed by LAC (1.1/100py) and lowest for SSA (0.9/100py).

Figure 12. Cumulative incidence of AIDS by region of origin, with and without including tuberculosis as an AIDS defining illness

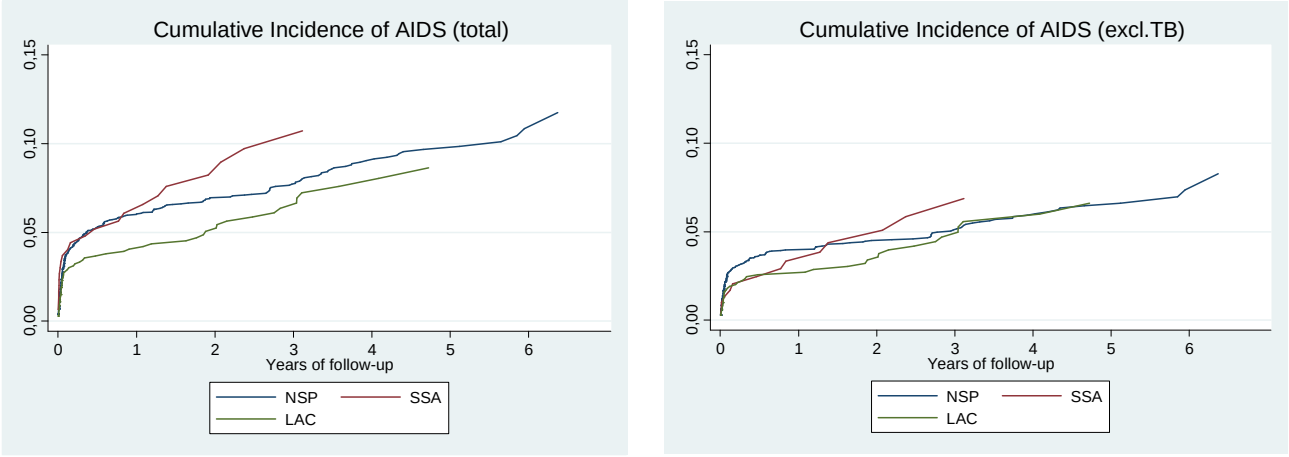
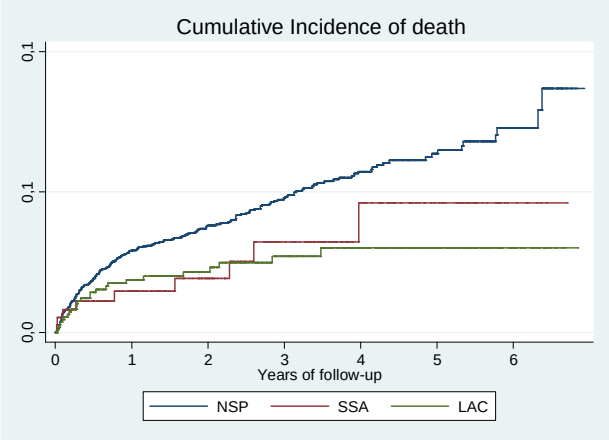


Figure 13 shows cumulative incidence by geographical origin. Crude analysis showed longer survival for LAC and SSA, reaching statistical significance for LAC. After adjusting for CD4 count, VL, transmission route and age, a non-significant lower risk of death was still observed, especially for SSA.

Figure 13. Cumulative incidence of death by region of origin



7. DISCUSSION

7.1. Discussion of results

Our study provides important evidence on the role of socioeconomic variables, specifically educational level and migration background on the progression of HIV infection. These results are important to inform appropriate preventive and health care services to respond to specific challenges posed by these groups, both at the clinical care level and from the perspective of health programmes and policies.

Regarding the **First study Objective, "To identify socioeconomic and demographic variables influencing prognosis since HIV seroconversion in Madrid (Spain), before and after the generalisation of HAART"**, risk of HIV disease progression to the endpoints of AIDS and death has shown to have decreased greatly after 1997, by 78.5% and 86.6% respectively. This data confirm HAART effectiveness in the population recruited and followed up in the Centro Sanitario Sandoval that receives treatment in different hospitals of the Autonomous Community of Madrid. Educational level and age at seroconversion have demonstrated to have an independent effect over disease progression, but their effect has been found to be different before and after the generalisation of HAART. This means that age and educational level determine groups where benefits of available resources are being unevenly experienced. These results are consistent with previous findings in this same cohort in the year 1999[40], although the longer follow-up has allowed us to better characterise the different effects.

Age at seroconversion determines a higher risk of both AIDS and death in years prior to generalisation of HAART. However, the effect over progression to AIDS disappears after 1997, pointing out a higher effectiveness of HAART at higher ages. This re-

sults are consistent with findings from other authors[88, 105], but are contrary to what one would expect based on studies that reveal that younger subjects have a better immune reconstitution after treatment, as this effect should have widened the age gap[106-108]. Possible explanations are a better acceptance and adherence to treatment in older subjects, a better virological suppression[107, 109], or simply a higher margin for AIDS-free survival improvement in older subjects.

The effect of age on overall survival was constant over the different periods, which is also compatible with previous studies[109].Possible explanations for this constantly higher mortality risk are more severe opportunistic diseases in older ages, more frequent co-morbidities or higher non-AIDS mortality. This last hypothesis could not be contrasted in our study, as the cross-match with the National mortality registry from the National Institute of Statistics could only retrieve death episodes, but causes of death were not provided. Importance of non-AIDS events is increasing due to the higher life expectancy of HIV infected patients, and to the higher proportion of subjects infected in older ages. Non-AIDS-defining events should constitute a priority area for research to better manage an increasingly aged HIV epidemic.

HIV transmission route was associated with cumulative incidence of death at a crude level, and being infected through IDU was a risk factor for higher mortality. However, this effect is no longer observed after adjustment by educational level. In fact, in our sample, the majority of IDUs present low educational level, whereas the majority of sexual infections occur within people of high education. Several previous works have explored the effect of exposure category over risk of AIDS, but results are very heterogeneous[88, 106, 110, 111]. In cases where MSMs outstand with a higher

risk of progression to AIDS, it has been attributed to a higher incidence of Kaposi's sarcoma, which usually appears at higher CD4 counts than other opportunistic infections. On the other hand, those studies that found IDU were the ones with an excess of risk of AIDS, explanations have included a worse access and adherence to treatment and a higher prevalence of co-infections. More studies are needed to characterise the effect of the transmission route of HIV in our context before and after the generalisation of HAART, and to clarify its relationship with educational level.

People of high educational level had a 56% lower risk of AIDS in the HAART era, although not before that time. A 62% less risk of death was also found, with no differences over periods. These results were further corroborated by the analysis performed to respond to the **Second study Objective "To estimate the effect of educational level over HIV-disease progression since HIV seroconversion and access to treatment at different periods of the HIV epidemic in Spain"**, and will be discussed together. In the second analysis, we found educational level was a determinant of progression to AIDS and death, but not to HAART requirement or initiation.

In the early stages of the epidemic, in the absence of effective therapeutic resources, no differences by educational level were seen in progression to AIDS, so lower education does not confer any disadvantage in the prognosis of HIV infection. However, after 1996 as effective treatments became available, those with lower education benefited to a smaller extent and higher educational level became a protective factor for AIDS. As we split the HAART era in two sub-periods for this analysis, we were also able to detect a gradient: the more efficacious the available treatments, the higher the

magnitude of the AIDS risk gap. Between 1997 and 2003, people of high educational level experienced 42% less risk of AIDS, while it reached 74% after 2004.

These results corroborate what has been previously found in Spain[40], but differ from what has been reported in other countries, although no study has used data after 1999 or has been performed in seroconverters. Schechter et al.[38] observed different progression to AIDS by socioeconomic level before HAART times, whereas Junghans et al.[41] did not find these differences in the years after HAART. The effect of educational level may vary from one country to another, and thus explain the different results.

In the case of progression to death, our study supports the hypothesis that HIV-infected persons with higher education have a reduced risk of death. In this analysis we found a 32% lower risk of dying for persons of high educational level, a smaller effect than the one found in the first analysis. Several studies in Spain have shown a greater risk of death in the general population associated with a low educational level, and have attributed the differences in the younger groups to infectious diseases and, specifically, HIV/AIDS[32, 112]. Jarrín et al.[34] found a protective effect of educational level on mortality from all causes in IDU after HAART availability, and several studies in Spain [33, 35, 39] and a study in British Columbia (Canada)[113], where access to HAART is also free and universal, also found that HIV-infected persons with a lower socioeconomic level experienced a higher mortality.

The effect of educational level on mortality was the same in all three calendar periods in our study, independent of the availability of HAART, which suggests the existence of other determinants of inequality besides HIV infection. Many causes of death,

like overdose or external causes are not affected by HAART; in fact, 23% of the deaths in our study were due to these causes. Attributing a death to HIV is complex, and a high degree of misclassification must be assumed, especially with regard to causes categorised as infectious diseases or tumours. Partly because of this limitation, and partly due to the small sample size, we did not specifically analyse the risk of death attributable to HIV.

Time to HAART initiation showed no difference by educational level, as no difference was found in time to HAART requirement, suggesting the lack of inequalities in access to treatment. This is to be expected in a health system with free, universal coverage, as is the case in Spain. Other studies however, some of them in similar contexts, have found differences in access according to socioeconomic level[41]. It is possible that in our cohort we did not capture the most marginalised individuals who may not have had access to the health system. Thus, the lack of differences in our study population does not rule out the possibility that such differences may exist in population subgroups that were not represented in our seroconverter sample.

As no marked differences in access to HAART were observed, the different progression to AIDS and death in our study could not, in principle, be attributed to that factor. But it may be that the effectiveness of HAART depends on educational level, because of either poor adherence or worse response to treatment.

Adherence to treatment is one of the key factors conditioning efficacy and has been a challenge since the beginning of the antiretroviral treatments, implying a heavy pill burden, coping with frequent adverse events and regular clinical follow-up visits. Adherence highly depends on psychosocial factors including emotional stress, quality

of the social support network, ability to understand and accept the disease and the medical recommendations, different attitudes and use of healthcare resources, and alcohol and drug abuse, as well as on clinical aspects like the frequency and severity of adverse events and drug toxicity[114-122]. The ability to cope with treatment side effects can be further influenced by the need to keep them hidden from the social environment and thus, the degree of disclosure of the HIV status and with the degree of understanding of the treatment and the symptoms.

However, association between educational level and adherence to ART has not been authoritatively established. A recent systematic review found that around one-third of studies on this subject had found a statistically significant association between educational level and adherence, while the rest had not[116]. On the other hand, some studies that have directly evaluated the virological, clinical and immunological responses to HAART have found lower effectiveness in persons with less education[39, 123], even after adjusting for adherence to treatment[124]; therefore, other factors may be implicated.

Some hypotheses of why low educational level could increase risk of AIDS and death through pathways not related to HAART access and adherence could be different lifestyles and nutritional habits[38], different attitudes towards prophylaxis, or other social and psychological resources needed to confront stressful life situations successfully. A higher co-morbidity in the low education group could also play a role, although adjusting for hepatitis C virus and hepatitis B virus co-infection in the subset of patients with this information did not modify the results.

Regarding results from the **Third study Objective "To analyse key HIV-related outcomes for migrants originating from Latin America and Sub-Saharan Africa living in Spain compared to the native population"**, we have found that HIV-positive migrants from Latin America and the Caribbean and from Sub-Saharan Africa show differences in the clinical presentation and the prognosis of HIV infection compared to native Spanish patients.

To our knowledge, this is the largest study carried out in HIV-positive migrants from Latin America in a European country. Cultural distance, legal status and language barriers affect LAC and SSA communities living in Spain differently, and can determine different vulnerability, even in the context of a universal health care system. Together with their different epidemiological and behavioural backgrounds, this shapes very different HIV profiles, as shown in our study and previously[12, 125] which calls for studying these groups independently to better identify their needs.

A higher risk of delayed diagnosis (DD) of HIV infection has been found in migrants in our study, as was previously reported by other authors [61, 62]. What our study adds to previous work is the identification of this delayed diagnosis in migrants from LAC under 50 years of age, compared to NSP. This can respond to initial legal, administrative or cultural barriers to access the system. The similar prevalence of DD over 50 years of age between migrants and native Spaniards could reflect, rather than an improved situation of the former, a much more frequent DD in the latter[63]. SSA also experience a higher frequency of DD below 35 years old, but not after that age. It is possible that their ethnic visibility leads physicians to comply with HIV testing rec-

ommendations when subjects contact the health system, thus protecting them from DD.

As a consequence of the higher frequency of delayed diagnosis in both migrant groups, time to requiring ART was shorter and a more prompt ART initiation was found in the crude analysis. However, these differences could not be attributed to the region of origin but to a more advanced stage of HIV disease at diagnosis. A delayed ART initiation was still found for LAC after adjustment, although its magnitude was small as well as probably its clinical relevance. These results are consistent with existing literature from European settings[70, 126].

Immunological and virological responses were poorer for SSA, but not for LAC. This difference could not be attributable to different initial ART regimes, which were independent of the geographical origin of the patients, as it was probably not attributable to a differential frequency of CD4 counts and VL measurements within the three groups compared. Some biological factors have been associated in the literature with a worse response to treatment, and thus could partially explain our results, such as specific viral subtypes and/or resistance mutation profiles[127-131]. There could also be a role for differences in laboratory reference parameters[132, 133] or in drug metabolism[134]. However, important factors from the socioeconomic arena are probably present [69, 75, 135], especially those related to ART adherence[136].

An increased rate of treatment adverse events has been described in SSA and could affect adherence[137]. Additionally, high geographical mobility, communication, language and administrative barriers, stigma and discrimination, different health beliefs and habits, low literacy, lack of social support and depression are circumstances

that affect SSA communities living in Spain and are negatively associated to adherence[138].

Consistently with our results, some studies have found more frequent virological and immunological failure in the long term in migrants, mainly from SSA[69, 76]. Further, previous studies found no difference in immunological and virological response in SSA when using a threshold of 400-500 copies/ml, but poorer response when considering a target VL of 50 copies/ml[67, 69, 71, 73, 139]. Achieving VL as low as 20 copies/ml has a demonstrated an impact on sustained virological response, so studies should consider target VL as low as possible to be able to capture clinically meaningful inequalities[140].

A higher risk of AIDS in ages between 35 and 50 was found for SSA and LAC, which was at the expense of a higher rate of tuberculosis. TB is a known socioeconomically sensitive disease, as it is concentrated in groups with poor living conditions[141, 142]. Further, migrants from SSA and LAC come from countries with higher TB endemicity than Spain, which poses them at a higher risk of TB latent infection and subsequent reactivation[143].

As opposed to our findings, published studies have found similar or slightly better HIV progression rates to AIDS in SSA[66, 68, 72]. A recently published study by Jarrín et al.[70] did not find any difference in the risk of progression to AIDS according to geographical origin in a large seroconverter cohort. Although this design is the most adequate to study the natural history of disease, it is not fully representative of the HIV epidemic, especially regarding migrants, as they include preferably migrants who were infected in Europe, who have a good access to the health care system and who,

by definition, have not had diagnostic delay. A previous work from Staehelin et al.[74] in a seroprevalent cohort also concluded that no differences had been found in HIV progression in SSA compared to North-western Europeans, but however showed a non-statistically significant 55% increased risk of AIDS, which was attenuated when excluding TB as first ADI, which could be explained by a lack of statistical power and which would be consistent with our results.

Finally, a lower mortality was observed in migrants from LAC and SSA compared to NSP but it did not reach statistical significance, probably due to a lack of statistical power. Other studies have also reported a similar or lower mortality in HIV-infected migrants compared to native population[66, 68, 70, 74]. Lower mortality in migrants suggests a healthy migrant pattern, responding to a self- selection of healthier subjects into migration as reported by other authors[94, 95].

7.2. Methodological discussion

The results of this Doctoral Thesis need to be interpreted in light of its methodological limitations that could affect its internal and external validity.

7.2.1. Internal Validity

The main biases that can affect internal validity of results in the study of the natural history and disease progression of HIV infection, are those derived from the following circumstances: lack of observation of the event that determines the risk origin, lack of observation of the event of interest, miss-classification in the independent variables, lack of information on variables that could be potential confounders or that are important for the discussion of results, possible ecological fallacy in the evaluation of the population effectiveness of therapies, and/or low precision of the estimations.

The possible bias caused by the non-observation of the risk origin was managed by left truncation techniques in the case of seroconverter cohorts analysis (objectives 1 and 2), and by adjustment by CD4 count and VL at recruitment when analysing data from a seroprevalent cohort (objective 3).

Potential biases caused by the absence of observation of the event of interest may also have been present. In Objectives 1 and 2, we have analysed data that mostly come from ambulatory health centres where HIV infected subjects are followed up until they develop any type of complication that requires specialised care or they need to start antiretroviral therapy, when they are transferred to hospital HIV units. This can prevent the cohort from detecting AIDS-defining events and deaths, as they usually

occur in the hospital setting, and thus are more difficult to detect in the ambulatory health centres that participate in the cohort. Follow-up of patients in the referral hospitals minimises this problem, and cross-matches with registries were also performed to further avoid this information bias.

However, the assumption that mortality and AIDS registries are complete and that persons who do not appear in them are event-free, may lead to an underestimation of the incidence rate. However, there is no reason to think this would differ by educational level, so we don't think that this may have biased our results. There is also no reason to think that differences in dates of cross-matches with registries and other differences in outcome ascertainment between different cohorts within GEMES will depend on individual educational level, so no information bias is likely to have been induced for this reason.

In the case of the analysis of the CoRIS cohort, this problem was not present, as it is a hospital-based cohort. However, an informative censoring was found, as characteristics of patients lost to follow-up were different from those who continued in the cohort. This means that the non-observation of the event of interest could have a different probability according to patients' characteristics and affect internal validity of results. Specifically, subjects of non-Spanish origin were more frequently LTFU compared to native Spanish. It is, however, unlikely that a *salmon bias* is present in our study, as migrants LTFU were those with better clinical characteristics. Inverse probability weighting for censoring was performed to minimize this source of bias, although results were largely unchanged under this approach, showing that the bias, if any, would be of small magnitude.

In the third place, a misclassification of independent variables could also have affected out study results. As for educational level, it was recorded at time of recruitment, which could lead to a classification bias, since some subjects may have completed higher-level studies during follow-up. However, the change in category would occur in subjects who completed secondary education after recruitment, which is uncommon in an adult cohort. Moreover, this bias would work in favour of the null hypothesis; therefore, it is unlikely to explain our results.

Selection bias could have also been introduced by the exclusion of those subjects with no information on educational level. Most were young patients infected in the early epidemic through IDU, probably of Spanish origin, who had been imprisoned in their life-time and had lower CD4 count at diagnosis; a high proportion of them probably correspond to people with a low educational level. Because of the poorer prognostic factors, we can assume the direction of this selection bias, if it existed at all, would also be towards the null. However, no differences were found in time to AIDS or death between people with or without information on educational level.

On the other hand, the way of recording migrant status can also be discussed, as it collects a self-referred region of origin, but does not capture other important dimensions such as differences within countries of the regions, different ethnical backgrounds, time since arrival in Spain, country of birth, nationality, legal status, etc. that determine different migration profiles with different socioeconomic vulnerability.

Another limitation stems from having combined MSM and cases of heterosexual transmission into a single category for the first and second analysis, as they have different demographic and socioeconomic characteristics. In Spain, some of the het-

erosexual cases are partners of IDU, with whom they share other determinants, among them, a lower educational level than MSM.

Another aspect that needs to be discussed is that important variables can be missing in our study. We have considered educational level as a proxy for social class, although it is very possible that this variable alone is not capturing important conditions associated to material and social deprivation with important consequences over HIV vulnerability. Educational level is best in measuring psychological and behavioural aspects of social class, and is a solid link between life conditions in early youth and adulthood. However, variables such as employment status or income, to mention some examples, can complement information provided by educational level and refine measurement of socioeconomic level in observational studies.

This situation is probably more evident in migrants, who are often faced with loss of social status in the host country, higher unemployment, legal precariousness that can result in their living conditions not matching with natives from the same educational level. So, in this particular group, educational level may be not the best *proxy* to adjust for socioeconomic conditions. Moreover, migrants from different origins may have undergone formal education in their countries of origin, which probably have different education curricula than Spain, and educational level categories may not be fully comparable. However, in every cohort educational level was registered by the physician or researcher, choosing the best equivalent in the Spanish educational curriculum for each case, so we expect this misclassification bias is not of significant magnitude.

Other unmeasured variables for the third objective were those related to biological factors, such as viral subtype or prevalence of resistance mutations, as were not explored adherence and other specific behavioural aspects with an impact on prognosis.

Another important consideration regarding internal validity of results lies in the fact that measuring effectiveness at the population level can be affected by an ecological fallacy. The reduced risk in AIDS and mortality observed after 1997 could, therefore, not be attributable, or at least not only, to the availability of HAART, but to other circumstances like an improvement in general life conditions or other health interventions. In addition, the socio-demographic profile of HIV patients has changed throughout the calendar periods, so we cannot rule out the possibility that changes in the profile of patients over time partially explain our results.

Finally, our study may be suffering from lack of statistical power to detect existing effects. The small proportion of heterosexually infected patients and women among the seroconverter subjects is especially notable. In the first case, this forced us into grouping sexual transmission categories and prevented us from studying differences within them, as previously mentioned. In the second case, it is possible that differences according to sex were not found even if present, due to insufficient power.

For the third objective, small sample sizes from regions of origin different to Latin America and Sub-Saharan Africa made including these subjects in the analysis impossible. The number of SSA could also have been insufficient to detect existing associations and, in fact, could explain the lack of statistical significance of some effects of considerable magnitude. We also did not observe any interaction between sex and

region of origin, but we cannot rule out the possibility that this is due to insufficient power.

7.2.2. External Validity

One of the main challenges of studies addressing the natural history and progression of HIV infection is recruiting samples who are representative of the general HIV-infected population. This one is not a delimited population where the sample frame is available, so probabilistic sample selection cannot be used. On the contrary, HIV remains to be a stigmatised disease and consequently affected populations are relatively hard to reach and recruiting subjects for research studies is a challenge. And this may be specially so in the case of migrants, who are a further discriminated population, highly mobile and heterogeneous and with a proportion of them facing legal barriers to stay in the country. For all the mentioned reasons, recruitment of HIV cohorts is usually done on a convenience sampling basis and in a health care setting, where these populations may be accessible. This can produce a certain degree of selection bias that can limit ability to generalise results to the general HIV-infected population.

In the first place, the Centro Sanitario Sandoval receives a large proportion of MSM, who are also the most likely to perform repeated HIV tests. Therefore, MSM are over-represented in this sample, as it is in GEMES cohort, at the expense of IDU and especially of heterosexually transmitted cases. This, together with the already mentioned *seroconverter bias*, can make the composition of the sample for the first two objectives non-representative of the real epidemic.

Also, the fact that recruitment is performed in the health care setting, selects with a higher probability patients with low barriers to access health care. And this is especially true for seroconverters, who are patients that access the system repeatedly. We cannot rule out the possibility that our results cannot be generalised to groups of patients who may be experiencing higher barriers to access health care.

Seroprevalent cohorts are more balanced in their composition and have a better external validity. However, a certain selection bias can be present, as inclusion in cohorts is never randomly assigned. Characteristics of CoRIS patients have been shown to be fairly similar to those reported by national surveillance systems, which suggests that external validity of data generated in CoRIS will probably be high.

8. CONCLUSIONS

Conclusion 1: Availability of HAART after 1996 has shown great population effectiveness in reducing AIDS incidence and mortality in HIV infected patients in the Spanish context. However, progression of infection is further affected by demographic and socioeconomic factors, specifically age, educational level and migration status, pointing out the existence of social inequalities with an impact on HIV prognosis.

Conclusion 2: In the absence of effective therapeutic resources, education does not determine prognosis of HIV-infected persons; but as effective treatments become available, those with lower education benefit to a smaller extent, which reflects in a higher risk of AIDS after 1996; and this inequality gap is further widened as treatments become more effective.

Conclusion 3: Higher all-cause mortality in subjects of low educational level is evident in the different calendar periods and is not affected by the availability of HAART. This points out an overall health vulnerability associated with lower social class which depends on factors beyond HIV infection.

Conclusion 4: No difference was found on access to HAART by educational level, so this factor probably does not explain the higher risk of AIDS and death in subjects of low socioeconomic condition. Other factors limiting treatment effectiveness in this group may be present probably including, but not limited to, a worse adherence to treatment.

Conclusion 5: Migrants from Latin America and Sub-Saharan Africa experience a higher risk of delayed diagnosis of HIV infection, specially the younger subjects, probably reflecting the existence of barriers to HIV testing for these communities. However no meaningful delays in treatment initiation are identified, showing an equitable access to

therapeutic resources and no further barriers for migrants once they have accessed the system.

Conclusion 6: Immunological and virological response to antiretroviral treatment is poorer for Sub-Saharan Africans, but not for Latin Americans. Possible explanatory variables were not measured in this study although they probably include a combination of biological and socioeconomic conditions, including adherence-related factors.

Conclusion 7: Migrants from Latin America and Sub-Saharan Africa between 35-50 years old progress faster to an AIDS diagnosis, and it is at the expense of a higher incidence rate of tuberculosis, which is also a socioeconomically sensitive disease. To the contrary, mortality in these groups tends to be lower, compatible with the healthy migrant effect.

Conclusion 8: These results show the impact of social inequalities on HIV-related outcomes and are important to inform appropriate preventive and health care services as well as health programmes and policies to better respond to challenges posed by social inequalities and to reduce their impact on health.

9. REFERENCES

Reference List

1. UNAIDS. Report on the global AIDS epidemic. In: 2010.
2. UNAIDS. AIDS epidemic update : November 2010. In: 2010.
3. European Centre for the Epidemiological Monitoring of AIDS. HIV/AIDS Surveillance in Europe. In: 2010.
4. Centres for Disease Control and Prevention (CDC). HIV Surveillance. United States, 1981-2008. In: 2011. pp. 689-693.
5. del Amo J, Pérez-Cachafeiro S, Hernando V, González C, Jarrin I, Bolúmar F. ECDC Technical Report. Migrant Health: Epidemiology of HIV and AIDS in migrant communities and ethnic minorities in EU/EEA countries. In: European Centre for Disease Prevention and Control; 2009.
6. del Amo J, Likatavicius G, Perez-Cachafeiro S, Hernando V, Gonzalez C, Jarrin I*, et al.* The epidemiology of HIV and AIDS reports in migrants in the 27 European Union countries, Norway and Iceland: 1999-2006. *Eur J Public Health* 2011; **21(5)**:620-626.
7. Ministerio de Sanidad y Política Social. Plan Multisectorial frente a la infección por VIH y el SIDA. España 2008-2010. In: 2011.
8. 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; **115(3)**:85-89.
9. Perez-Hoyos S, Ferreros I, del Amo J, Muga R, Del Romero J, de Olalla PG*, et al.* Survival and progression to AIDS in a seroconverter cohort in the post-highly active antiretroviral therapy era: effectiveness goes on. *AIDS* 2006; **20(2)**:289-291.
10. 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(17)**:1497-1503.
11. CASCADE. 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(9210)**:1158-1159.
12. Ministerio de Sanidad y Servicios Sociales e Igualdad. Vigilancia Epidemiológica del VIH/SIDA en España. Actualización 30 de junio de 2011. In: 2011.

13. 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; **12(1)**:61-64.

14. 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; **20(10)**:1447-1450.

15. UNAIDS. AIDS epidemic update : November 2009. In: 2009.

16. del Amo J, Pérez-Cachafeiro S, Hernando V, González C, Jarrin I, Bolúmar F. ECDC Technical Report. Migrant Health: Epidemiology of HIV and AIDS in migrant communities and ethnic minorities in EU/EEA countries. In: European Centre for Disease Prevention and Control; 2009.

17. Fournier AM, Carmichael C. Socioeconomic influences on the transmission of human immunodeficiency virus infection: the hidden risk. *Arch Fam Med* 1998; **7(3)**:214-217.

18. Moatti JP, Souteyrand Y. HIV/AIDS social and behavioural research: past advances and thoughts about the future. *Soc Sci Med* 2000; **50(11)**:1519-1532.

19. Dray-Spira R, Lert F. Social health inequalities during the course of chronic HIV disease in the era of highly active antiretroviral therapy. *AIDS* 2003; **17(3)**:283-290.

20. Pensola T, Martikainen P. Life-course experiences and mortality by adult social class among young men. *Soc Sci Med* 2004; **58(11)**:2149-2170.

21. Pensola TH, Martikainen P. Effect of living conditions in the parental home and youth paths on the social class differences in mortality among women. *Scand J Public Health* 2003; **31(6)**:428-438.

22. Skalicka V, van LF, Bambra C, Krokstad S, Mackenbach J. Material, psychosocial, behavioural and biomedical factors in the explanation of relative socioeconomic inequalities in mortality: evidence from the HUNT study. *Int J Epidemiol* 2009; **38(5)**:1272-1284.

23. Galobardes B, Shaw M, Lawlor DA, Lynch JW, Davey SG. Indicators of socioeconomic position (part 2). *J Epidemiol Community Health* 2006; **60(2)**:95-101.

24. Galobardes B, Shaw M, Lawlor DA, Lynch JW, Davey SG. Indicators of socioeconomic position (part 1). *J Epidemiol Community Health* 2006; **60(1)**:7-12.

25. Braveman PA, Cubbin C, Egerter S, Chideya S, Marchi KS, Metzler M*, et al.* Socioeconomic status in health research: one size does not fit all. *JAMA* 2005; **294(22)**:2879-2888.

26. Davey SG, Hart C, Hole D, MacKinnon P, Gillis C, Watt G*, et al.* Education and occupational social class: which is the more important indicator of mortality risk? *J Epidemiol Community Health* 1998; **52(3)**:153-160.

27. Krieger N, Williams DR, Moss NE. Measuring social class in US public health research: concepts, methodologies, and guidelines. *Annu Rev Public Health* 1997; **18**:341-378.

28. Dray-Spira R, Lert F. Social health inequalities during the course of chronic HIV disease in the era of highly active antiretroviral therapy. *AIDS* 2003; **17(3)**:283-290.

29. Koskenvuo M, Kaprio J, Romo M, Langinvainio H. Incidence and prognosis of ischaemic heart disease with respect to marital status and social class. A national record linkage study. *J Epidemiol Community Health* 1981; **35(3)**:192-196.

30. Vagero D, Persson G. Cancer survival and social class in Sweden. *J Epidemiol Community Health* 1987; **41(3)**:204-209.

31. Strand BH, Groholt EK, Steingrimsdottir OA, Blakely T, Graff-Iversen S, Naess O. Educational inequalities in mortality over four decades in Norway: prospective study of middle aged men and women followed for cause specific mortality, 1960-2000. *BMJ* 2010; **340**:c654.

32. Borrell C, Regidor E, Arias LC, Navarro P, Puigpinos R, Dominguez V*, et al.* Inequalities in mortality according to educational level in two large Southern European cities. *Int J Epidemiol* 1999; **28(1)**:58-63.

33. Borrell C, Rodriguez-Sanz M, Pasarin MI, Brugal MT, Garcia-de-Olalla P, Mari-Dell'olmo M*, et al.* AIDS mortality before and after the introduction of highly active antiretroviral therapy: does it vary with socioeconomic group in a country with a National Health System? *Eur J Public Health* 2006; **16(6)**:601-608.

34. Jarrin I, Lumbreras B, Ferreros I, Perez-Hoyos S, Hurtado I, Hernandez-Aguado I. Effect of education on overall and cause-specific mortality in injecting drug users, according to HIV and introduction of HAART. *Int J Epidemiol* 2007; **36(1)**:187-194.

35. Mari-Dell'olmo M, Rodriguez-Sanz M, Garcia-Olalla P, Pasarin MI, Brugal MT, Cayla JA*, et al.* Individual and community-level effects in the socioeconomic inequalities of AIDS-related mortality in an urban area of southern Europe. *J Epidemiol Community Health* 2007; **61(3)**:232-240.

36. McFarland W, Chen S, Hsu L, Schwarcz S, Katz M. Low socioeconomic status is associated with a higher rate of death in the era of highly active antiretroviral therapy, San Francisco. *J Acquir Immune Defic Syndr* 2003; **33(1)**:96-103.

37. Rapiti E, Porta D, Forastiere F, Fusco D, Perucci CA. Socioeconomic status and survival of persons with AIDS before and after the introduction of highly active

antiretroviral therapy. Lazio AIDS Surveillance Collaborative Group. *Epidemiology* 2000; **11(5)**:496-501.

38. Schechter MT, Hogg RS, Aylward B, Craib KJ, Le TN, Montaner JS. Higher socioeconomic status is associated with slower progression of HIV infection independent of access to health care. *J Clin Epidemiol* 1994; **47(1)**:59-67.

39. Sobrino-Vegas P, Rodriguez-Urrego J, Berenguer J, Caro-Murillo AM, Blanco JR, Viciana P*, et al.* Educational gradient in HIV diagnosis delay, mortality, antiretroviral treatment initiation and response in a country with universal health care. *Antivir Ther* 2012; **17(1)**:1-8.

40. del Amo J, del Romero J, Barrasa A, Perez-Hoyos S, Rodriguez C, Diez M*, et al.* Factors influencing HIV progression in a seroconverter cohort in Madrid from 1985 to 1999. *Sex Transm Infect* 2002; **78(4)**:255-260.

41. Junghans C, Low N, Chan P, Witschi A, Vernazza P, Egger M. Uniform risk of clinical progression despite differences in utilization of highly active antiretroviral therapy: Swiss HIV Cohort Study. *AIDS* 1999; **13(18)**:2547-2554.

42. International Organization for Migration (IOM). World Migration Report 2011 - Communicating Effectively about Migration. In: 2011.

43. Eurostat. Recent migration trends: citizens of EU-27 Member States become ever more mobile while EU remains attractive to non-EU citizens. In: 2008.

44. International Organization for Migration (IOM). Migration Health: Better Health for All in Europe. In. Brussels (Belgium); 2009.

45. Instituto Nacional de Estadística. INEbase: padrón municipal. In: 2011.

46. Ho MJ. Migratory journeys and tuberculosis risk. *Med Anthropol Q* 2003; **17(4)**:442-458.

47. del Amo J, Broring G, Hamers FF, Infuso A, Fenton K. Monitoring HIV/AIDS in Europe's migrant communities and ethnic minorities. *AIDS* 2004; **18(14)**:1867-1873.

48. del Amo J, Hamers F, Infuso A, Broring G, Fenton K. HIV/AIDS in migrants and ethnic minorities in Europe. In: *Ethnicity and HIV prevention and care in Europe and the USA.* International Medical Press Ltd. (editor). London; 2003. pp. 46-60.

49. Clarke K, Broring G. HIV/AIDS care and support for migrant and ethnic minority communities in Europe. In: AIDS & Mobility Europe; 2000.

50. Caro-Murillo AM, Castilla CJ, del A, V. [Epidemiology of HIV infection in immigrants in Spain: information sources, characteristics, magnitude and tendencies]. *Gac Sanit* 2010; **24(1)**:81-88.

51. Grupo EPI-VIH. Estudio prospectivo de prevalencia de VIH en pacientes de una red de
centros de diagnóstico de VIH, 2000-2009. In. Centro Nacional de Epidemiología (editor); 2011.

52. Hernando A, García-Cerradab C, Pulido F, Rubio R, Hervás R, Cepeda C*, et al.* Infección por el VIH en inmigrantes: diferencias clínico-epidemiológicas con la población autóctona en un área de salud de la Comunidad de Madrid (2001-2004). In: 2007. pp. 441-445.

53. Grupo EPI-VIH. HIV infection among people of foreign origin voluntarily tested in Spain. A comparison with national subjects. *Sex Transm Infect* 2002; **78(4)**:250-254.

54. Sinka K, Mortimer J, Evans B, Morgan D. Impact of the HIV epidemic in sub-Saharan Africa on the pattern of HIV in the UK. *AIDS* 2003; **17(11)**:1683-1690.

55. Fenton KA, Chinouya M, Davidson O, Copas A. HIV transmission risk among sub-Saharan Africans in London travelling to their countries of origin. *AIDS* 2001; **15(11)**:1442-1445.

56. Burns FM, Arthur G, Johnson AM, Nazroo J, Fenton KA. United Kingdom acquisition of HIV infection in African residents in London: more than previously thought. *AIDS* 2009; **23(2)**:262-266.

57. Dougan S, Gilbart VL, Sinka K, Evans BG. HIV infections acquired through heterosexual intercourse in the United Kingdom: findings from national surveillance. *BMJ* 2005; **330(7503)**:1303-1304.

58. Gras MJ, van Benthem BH, Coutinho RA, van den HA. Determinants of high-risk sexual behavior among immigrant groups in Amsterdam: implications for interventions. *J Acquir Immune Defic Syndr* 2001; **28(2)**:166-172.

59. Castilla J, Vall M, Belda J, Cámara M, Junquera M, Lorenzo J*, et al.* Incidencia de seroconversiones al VIH en inmigrantes en España [Abstract]. *Comunicación en el IX Congreso Nacional sobre el SIDA, Barcelona* 2006.

60. Noguer I, Sobrino-Vegas P, Barrasa A. La epidemia de VIH y otras infecciones de transmisión sexual en personas con país de origen distinto a España. In: *La prevención de la infección del VIH/sida en la población inmigrante*. Ministerio de Sanidad y Consumo (editor); 2006. pp. 49-63.

61. Burns FM, Fakoya AO, Copas AJ, French PD. Africans in London continue to present with advanced HIV disease in the era of highly active antiretroviral therapy. *AIDS* 2001; **15(18)**:2453-2455.

62. Zoufaly A, An der HM, Marcus U, Hoffmann C, Stellbrink H, Voss L*, et al.* Late presentation for HIV diagnosis and care in Germany. *HIV Med* 2012; **13(3)**:172-181.

63. 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; **7(2)**:224-230.

64. Ramos JM, Gutierrez F, Padilla S, Masia M, Escolano C. [Clinical and epidemiological characteristics of human immunodeficiency virus infection in foreigners residing in Elche, Spain (1998-2003)]. *Enferm Infecc Microbiol Clin* 2005; **23(8)**:469-473.

65. Perez-Molina JA, Lopez-Velez R, Navarro M, Perez-Elias MJ, Moreno S. Clinicoepidemiological characteristics of HIV-infected immigrants attended at a tropical medicine referral unit. *J Travel Med* 2009; **16(4)**:248-252.

66. Blaxhult A, Mocroft A, Phillips A, van LJ, Bentwich Z, Stergiou G*, et al.* Does European or non-European origin influence health care and prognosis for HIV patients in Europe? The EuroSIDA Study Group. *HIV Med* 1999; **1(1)**:2-9.

67. Breton G, Lewden C, Spire B, Salmon D, Brun-Vezinet F, Duong M*, et al.* Characteristics and response to antiretroviral therapy of HIV-1-infected patients born in Africa and living in France. *HIV Med* 2007; **8(3)**:164-170.

68. del Amo J, Petruckevitch A, Phillips A, Johnson AM, Stephenson J, Desmond N*, et al.* Disease progression and survival in HIV-1-infected Africans in London. *AIDS* 1998; **12(10)**:1203-1209.

69. Frater AJ, Dunn DT, Beardall AJ, Ariyoshi K, Clarke JR, McClure MO*, et al.* Comparative response of African HIV-1-infected individuals to highly active antiretroviral therapy. *AIDS* 2002; **16(8)**:1139-1146.

70. Jarrin I, Pantazis N, Gill MJ, Geskus R, Perez-Hoyos S, Meyer L*, et al.* Uptake of combination antiretroviral therapy and HIV disease progression according to geographical origin in seroconverters in Europe, Canada, and Australia. *Clin Infect Dis* 2012; **54(1)**:111-118.

71. Jensen-Fangel S, Pedersen L, Pedersen C, Larsen CS, Tauris P, Moller A*, et al.* The effect of race/ethnicity on the outcome of highly active antiretroviral therapy for human immunodeficiency virus type 1-infected patients. *Clin Infect Dis* 2002; **35(12)**:1541-1548.

72. Muller V, von W, V, Yerly S, Boni J, Klimkait T, Burgisser P*, et al.* African descent is associated with slower CD4 cell count decline in treatment-naive patients of the Swiss HIV Cohort Study. *AIDS* 2009; **23(10)**:1269-1276.

73. Nellen JF, Wit FW, De WF, Jurriaans S, Lange JM, Prins JM. Virologic and immunologic response to highly active antiretroviral therapy in indigenous and non-indigenous HIV-1-infected patients in the Netherlands. *J Acquir Immune Defic Syndr* 2004; **36(4)**:943-950.

74. Staehelin C, Rickenbach M, Low N, Egger M, Ledergerber B, Hirschel B*, et al.* Migrants from Sub-Saharan Africa in the Swiss HIV Cohort Study: access to antiretroviral therapy, disease progression and survival. *AIDS* 2003; **17(15)**:2237-2244.

75. de Arellano ER, Benito JM, Soriano V, Lopez M, Holguin A. Impact of ethnicity and HIV type 1 subtype on response to first-line antiretroviral therapy. *AIDS Res Hum Retroviruses* 2007; **23(7)**:891-894.

76. Dray-Spira R, Spire B, Heard I, Lert F. Heterogeneous response to HAART across a diverse population of people living with HIV: results from the ANRS-EN12-VESPA Study. *AIDS* 2007; **21 Suppl 1**:S5-12.

77. Perez Molina JA, Rillo MM, Suarez-Lozano I, Casado Osorio JL, Cobo RT, Gonzalez PR*, et al.* Do HIV-Infected Immigrants Initiating HAART have Poorer Treatment-Related Outcomes than Autochthonous Patients in Spain? Results of the GESIDA 5808 Study. *Curr HIV Res* 2010; **8(7)**:521-530.

78. 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; **28(5)**:304-309.

79. MacMahon B, Trichopoulos D. *Epidemiology: Principles and methods.* Boston, MA: Little,Brown; 1996.

80. The UK register of HIV seroconverters: methods and analytical issues. UK register of HIV seroconverters (UKRHS) Steering Committee. *Epidemiol Infect* 1996; **117(2)**:305-312.

81. Sobrino-Vegas P, Gutierrez F, Berenguer J, Labarga P, Garcia F, ejos-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; **29(9)**:645-653.

82. Thierfelder C, Weber R, Elzi L, Furrer H, Cavassini M, Calmy A*, et al.* Participation, characteristics and retention rates of HIV-positive immigrants in the Swiss HIV Cohort Study. *HIV Med* 2012; **13(2)**:118-126.

83. Fewell z, Wolfe F, Choi H, Hernán M, Tilling K, Sterne J. Controlling for time-dependent confounding using marginal structural models. In: 2004. pp. 402-420.

84. Cole SR, Hernán MA. Constructing inverse probability weights for marginal structural models. In: 2008. pp. 656-664.

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

86. Fine JP, Gray RJ. A proportional hazards model for the subdistribution of a competing risk. In: 1999. pp. 496-509.

87. CASCADE. Effect of ignoring the time of HIV seroconversion in estimating changes in survival over calendar time in observational studies: results from CASCADE. *AIDS* 2000; **14(13)**:1899-1906.

88. CASCADE. Time from HIV-1 seroconversion to AIDS and death before widespread use of highly-active antiretroviral therapy: a collaborative re-analysis. Collaborative Group on AIDS Incubation and HIV Survival including the CASCADE EU Concerted Action. Concerted Action on SeroConversion to AIDS and Death in Europe. *Lancet* 2000; **355(9210)**:1131-1137.

89. GEMES (Grupo Español Multicéntrico para el Estudio de Seroconversores). El período de incubacion del sida en España antes de la terapia antirretroviral de alta eficacia. In: 2000. pp. 681-686.

90. Porter K, Johnson AM, Phillips AN, Darbyshire JH. The practical significance of potential biases in estimates of the AIDS incubation period distribution in the UK register of HIV seroconverters. *AIDS* 1999; **13(14)**:1943-1951.

91. Law CG, Brookmeyer R. Effects of mid-point imputation on the analysis of doubly censored data. *Stat Med* 1992; **11(12)**:1569-1578.

92. Brookmeyer R, Gail MH. Biases in prevalent cohorts. In: 1987. pp. 739-749.

93. Llacer A, Zunzunegui MV, del Amo J, Mazarrasa L, Bolumar F. The contribution of a gender perspective to the understanding of migrants' health. *J Epidemiol Community Health* 2007; **61 Suppl 2**:ii4-10.

94. Malmusi D, Borrell C, Benach J. Migration-related health inequalities: showing the complex interactions between gender, social class and place of origin. *Soc Sci Med* 2010; **71(9)**:1610-1619.

95. Razum O, Zeeb H, Rohrmann S. The 'healthy migrant effect'--not merely a fallacy of inaccurate denominator figures. *Int J Epidemiol* 2000; **29(1)**:191-192.

96. Turra CM, Elo IT. The Impact of Salmon Bias on the Hispanic Mortality Advantage: New Evidence from Social Security Data. *Popul Res Policy Rev* 2008; **27(5)**:515-530.

97. braido-Lanza AF, Dohrenwend BP, Ng-Mak DS, Turner JB. The Latino mortality paradox: a test of the "salmon bias" and healthy migrant hypotheses. *Am J Public Health* 1999; **89(10)**:1543-1548.

98. Munoz A, Gange SJ, Jacobson LP. Distinguishing efficacy, individual effectiveness and population effectiveness of therapies. *AIDS* 2000; **14(6)**:754-756.

99. Montoro JB, Oliveras J, Lorenzo JI, Tusell JM, Altisent C, Molina R*, et al.* An association between clotting factor concentrates use and mortality in human

immunodeficiency virus-infected hemophilic patients. *Blood* 1995; **86(6)**:2213-2219.

100. Perez-Hoyos S, Avino M, Hernandez I, Gonzalez J, Ruiz I. [AIDS-free time and survival of an injecting drug users HIV seroconvertors cohort]. *Gac Sanit* 1999; **13(5)**:337-345.

101. Brugal M, Mestres M, Díaz de Quijano E, García de Olalla P, Caylà J. Evolución temporal y factores asociados a la seroconversión por HIV en una cohorte de usuarios de drogas. In: 1998. pp. 142-143.

102. Muga R, Egea JM, Navio M, Sirera G, Vall M, Tor J. [Mortality in a cohort of intravenous drug users before the introduction of potent HIV therapy]. *Med Clin (Barc)* 1999; **112(19)**:721-725.

103. 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(1)**:23-31.

104. GESIDA. Recommendations of GESIDA/Spanish AIDS Plan on antiretroviral therapy in adults infected by the human immunodeficiency virus. Reports from 2005 to 2012. In: 2012.

105. Belanger F, Meyer L, Carre N, Coutellier A, Deveau C. Influence of age at infection on human immunodeficiency virus disease progression to different clinical endpoints: the SEROCO cohort (1988-1994). The Seroco Study Group. *Int J Epidemiol* 1997; **26(6)**:1340-1345.

106. Porter K, Babiker A, Bhaskaran K, Darbyshire J, Pezzotti P, Porter K*, et al.* Determinants of survival following HIV-1 seroconversion after the introduction of HAART. *Lancet* 2003; **362(9392)**:1267-1274.

107. Sabin CA, Smith CJ, d'Arminio Monforte A, Battegay M, Gabiano C, Galli L*, et al.* Response to combination antiretroviral therapy: variation by age. *AIDS* 2008; **22(12)**:1463-1473.

108. Yamashita TE, Phair JP, Munoz A, Margolick JB, Detels R, O'Brien SJ*, et al.* Immunologic and virologic response to highly active antiretroviral therapy in the Multicenter AIDS Cohort Study. *AIDS* 2001; **15(6)**:735-746.

109. Greenbaum AH, Wilson LE, Keruly JC, Moore RD, Gebo KA. Effect of age and HAART regimen on clinical response in an urban cohort of HIV-infected individuals. *AIDS* 2008; **22(17)**:2331-2339.

110. Alioum A, Leroy V, Commenges D, Dabis F, Salamon R. Effect of gender, age, transmission category, and antiretroviral therapy on the progression of human immunodeficiency virus infection using multistate Markov models. Groupe d'Epidemiologie Clinique du SIDA en Aquitaine. *Epidemiology* 1998; **9(6)**:605-612.

111. Pezzotti P, Phillips AN, Dorrucci M, Lepri AC, Galai N, Vlahov D*, et al.* Category of exposure to HIV and age in the progression to AIDS: longitudinal study of 1199 people with known dates of seroconversion. HIV Italian Seroconversion Study Group. *BMJ* 1996; **313(7057)**:583-586.

112. Regidor E, de Mateo S, Calle ME, Dominguez V. Educational level and mortality from infectious diseases. *J Epidemiol Community Health* 2002; **56(9)**:682-683.

113. Wood E, Montaner JS, Chan K, Tyndall MW, Schechter MT, Bangsberg D*, et al.* Socioeconomic status, access to triple therapy, and survival from HIV-disease since 1996. *AIDS* 2002; **16(15)**:2065-2072.

114. Catz SL, Kelly JA, Bogart LM, Benotsch EG, McAuliffe TL. Patterns, correlates, and barriers to medication adherence among persons prescribed new treatments for HIV disease. *Health Psychol* 2000; **19(2)**:124-133.

115. Chesney MA. Factors affecting adherence to antiretroviral therapy. *Clin Infect Dis* 2000; **30 Suppl 2**:S171-S176.

116. Falagas ME, Zarkadoulia EA, Pliatsika PA, Panos G. Socioeconomic status (SES) as a determinant of adherence to treatment in HIV infected patients: a systematic review of the literature. *Retrovirology* 2008; **5**:13.

117. Gordillo V, del Amo J, Soriano V, Gonzalez-Lahoz J. Sociodemographic and psychological variables influencing adherence to antiretroviral therapy. *AIDS* 1999; **13(13)**:1763-1769.

118. Cinti SK. Adherence to antiretrovirals in HIV disease. *AIDS Read* 2000; **10(12)**:709-717.

119. Ammassari A, Murri R, Pezzotti P, Trotta MP, Ravasio L, De LP*, et al.* Self-reported symptoms and medication side effects influence adherence to highly active antiretroviral therapy in persons with HIV infection. *J Acquir Immune Defic Syndr* 2001; **28(5)**:445-449.

120. Murphy DA, Roberts KJ, Martin DJ, Marelich W, Hoffman D. Barriers to antiretroviral adherence among HIV-infected adults. *AIDS Patient Care STDS* 2000; **14(1)**:47-58.

121. Roberts KJ, Mann T. Barriers to antiretroviral medication adherence in HIV-infected women. *AIDS Care* 2000; **12(4)**:377-386.

122. Lucas GM, Gebo KA, Chaisson RE, Moore RD. Longitudinal assessment of the effects of drug and alcohol abuse on HIV-1 treatment outcomes in an urban clinic. *AIDS* 2002; **16(5)**:767-774.

123. Collazos J, Asensi V, Carton JA, Ibarra S. The influence of the patients' educational levels on socioeconomic, clinical, immunological and virological endpoints. *AIDS Care* 2009; **21(4)**:511-519.

124. Zaragoza-Macias E, Cosco D, Nguyen ML, Del Rio C, Lennox J. Predictors of success with highly active antiretroviral therapy in an antiretroviral-naive urban population. *AIDS Res Hum Retroviruses* 2010; **26(2)**:133-138.

125. Caro-Murillo AM, Gutierrez F, Manuel RJ, Sobrino P, Miro JM, Lopez-Cortes LF*, et al.* [HIV infection in immigrants in Spain: Epidemiological characteristics and clinical presentation in the CoRIS Cohort (2004-2006)]. *Enferm Infecc Microbiol Clin* 2009; **27(7)**:380-388.

126. Staehelin C, Egloff N, Rickenbach M, Kopp C, Furrer H. Migrants from sub-Saharan Africa in the Swiss HIV Cohort Study: a single center study of epidemiologic migration-specific and clinical features. *AIDS Patient Care STDS* 2004; **18(11)**:665-675.

127. Garcia F, Perez-Cachafeiro S, Guillot V, Alvarez M, Perez-Romero P, Perez-Elias MJ*, et al.* Transmission of HIV drug resistance and non-B subtype distribution in the Spanish cohort of antiretroviral treatment naive HIV-infected individuals (CoRIS). *Antiviral Res* 2011; **91(2)**:150-153.

128. Martinez-Cajas JL, Pai NP, Klein MB, Wainberg MA. Differences in resistance mutations among HIV-1 non-subtype B infections: a systematic review of evidence (1996-2008). *J Int AIDS Soc* 2009; **12**:11.

129. Martinez-Cajas JL, Pant-Pai N, Klein MB, Wainberg MA. Role of genetic diversity amongst HIV-1 non-B subtypes in drug resistance: a systematic review of virologic and biochemical evidence. *AIDS Rev* 2008; **10(4)**:212-223.

130. Velazquez-Campoy A, Todd MJ, Vega S, Freire E. Catalytic efficiency and vitality of HIV-1 proteases from African viral subtypes. *Proc Natl Acad Sci U S A* 2001; **98(11)**:6062-6067.

131. Yebra G, de MM, Perez-Elias MJ, Perez-Molina JA, Galan JC, Llenas-Garcia J*, et al.* Increase of transmitted drug resistance among HIV-infected sub-Saharan Africans residing in Spain in contrast to the native population. *PLoS One* 2011; **6(10)**:e26757.

132. Eller LA, Eller MA, Ouma B, Kataaha P, Kyabaggu D, Tumusiime R*, et al.* Reference intervals in healthy adult Ugandan blood donors and their impact on conducting international vaccine trials. *PLoS One* 2008; **3(12)**:e3919.

133. Ngowi BJ, Mfinanga SG, Bruun JN, Morkve O. Immunohaematological reference values in human immunodeficiency virus-negative adolescent and adults in rural northern Tanzania. *BMC Infect Dis* 2009; **9**:1.

134. Stohr W, Back D, Dunn D, Sabin C, Winston A, Gilson R*, et al.* Factors influencing efavirenz and nevirapine plasma concentration: effect of ethnicity, weight and co-medication. *Antivir Ther* 2008; **13(5)**:675-685.

135. Anastos K, Schneider MF, Gange SJ, Minkoff H, Greenblatt RM, Feldman J*, et al.* The association of race, sociodemographic, and behavioral characteristics with response to highly active antiretroviral therapy in women. *J Acquir Immune Defic Syndr* 2005; **39(5)**:537-544.

136. Lima VD, Harrigan R, Murray M, Moore DM, Wood E, Hogg RS*, et al.* Differential impact of adherence on long-term treatment response among naive HIV-infected individuals. *AIDS* 2008; **22(17)**:2371-2380.

137. Tedaldi EM, Absalon J, Thomas AJ, Shlay JC, van dB-W. Ethnicity, race, and gender. Differences in serious adverse events among participants in an antiretroviral initiation trial: results of CPCRA 058 (FIRST Study). *J Acquir Immune Defic Syndr* 2008; **47(4)**:441-448.

138. Sumari-de Boer I, Sprangers MA, Prins JM, Nieuwkerk PT. HIV Stigma and Depressive Symptoms are Related to Adherence and Virological Response to Antiretroviral Treatment Among Immigrant and Indigenous HIV Infected Patients. *AIDS Behav* 2011.

139. Bocket L, Cheret A, uffic-Burban S, Choisy P, Gerard Y, de lT, X*, et al.* Impact of human immunodeficiency virus type 1 subtype on first-line antiretroviral therapy effectiveness. *Antivir Ther* 2005; **10(2)**:247-254.

140. Raboud JM, Montaner JS, Conway B, Rae S, Reiss P, Vella S*, et al.* Suppression of plasma viral load below 20 copies/ml is required to achieve a long-term response to therapy. *AIDS* 1998; **12(13)**:1619-1624.

141. Diaz deQuijano E, Brugal MT, Pasarin MI, Galdos-Tanguis H, Cayla J, Borrell C. [Influence of social inequality, social unrest and extreme poverty on tuberculosis morbidity in the City of Barcelona]. *Rev Esp Salud Publica* 2001; **75(6)**:517-527.

142. Diez M, Diaz A, Bleda MJ, Aldamiz M, Camafort M, Camino X*, et al.* Prevalence of M. tuberculosis infection and tuberculosis disease among HIV-infected people in Spain. *Int J Tuberc Lung Dis* 2007; **11(11)**:1196-1202.

143. Martin V, Garcia de OP, Orcau A, Cayla JA. Factors associated with tuberculosis as an AIDS-defining disease in an immigration setting. *J Epidemiol* 2011; **21(2)**:108-113.

10. ANNEXES

10.1. Scientific communications related to this doctoral thesis

10.1.1. Presentations at scientific meetings

POSTER COMMUNICATION

S Monge, Del Romero J; Rodríguez C; de Mendoza C; de Górgolas M; Cosín J; Dronda F; Pérez-Cecilia E; Peña JM; Santos I; Rubio R; del Amo J. **[Socio-demographic factors associated with the progression of HIV infection and the impact of HAART in a seroconverter cohort in Madrid (1986-2009)]** Spanish. XXIX Scientific Meeting of the Spanish Epidemiology Society (SEE). Madrid (Spain), 6-8 of October 2011.

ELECTRONIC POSTER

S Monge, J Del Romero, P García De Olalla, R Muga, I Ferreros, I Jarrín, S Pérez-Hoyos, J Del Amo, GEMES. **Educational level and HIV disease progression before and after the introduction of HAART. A study in 898 hiv seroconverters in Spain**. 6th IAS Conference on HIV Pathogenesis, Treatment and Prevention. Rome, Italy 17-20 July 2011.

ORAL PRESENTATION

S Monge, J Del Romero, P García De Olalla, R Muga, I Ferreros, I Jarrín, S Pérez-Hoyos, J Del Amo, GEMES. **[Educational level and HIV disease progression before and after the introduction of HAART. A study in 898 HIV Seroconverters in Spain]** Spanish. XVIII Scientific Meeting of the Spanish Society of Epidemiology. Valencia (Spain), 27-29 of October 2010.

10.1.2. Scientific publications

PUBLISHED PAPER

Monge S; Del Romero J; Rodríguez C; de Mendoza C; de Górgolas M; Cosín J; Dronda F; Pérez-Cecilia E; Peña JM; Santos I; Rubio R; del Amo J. **[Socio-demographic factors associated with the progression of HIV infection and the impact of HAART in a seroconverter cohort in Madrid (1986-2009).]** Spanish. Enferm Infecc Microbiol Clin. 2012; 30(3):117–123.

PUBLISHED PAPER

Monge S, Jarrín I, Pérez-Hoyos S, Ferreros I, García-Olalla P, Muga R, Del Romero J, Belda J, Castilla J, Bolumar F, del Amo J, GEMES. **Educational level and HIV disease progression before and after the introduction of HAART. A cohort study in 989 HIV seroconverters in Spain.** Journal: Sexually Transmitted Infections. Dec2011. 87(7): 571-6.

PAPER CURRENTLY UNDER REVIEW

Monge S, Alejos B, Dronda F, Del Romero J, Iribarren JA, Pulido F, Rubio R, Miró JM, Gutierrez F, Del Amo J, and CoRIS. **Access to antiretroviral treatment and prognosis of HIV infection in migrants from Latin America and Sub-Saharan Africa in Spain.**

10.2. Centres and researchers involved in analysed cohorts

Centres and researchers involved in the Grupo de Seroconvertores de la Comunidad de Madrid

Centro Sanitario Sandoval: Jorge del Romero, Juan Ballesteros, Carmen Rodríguez, Soledad García, Montserrat Raposo

Centro Nacional de Epidemiología: Julia del Amo, María Victoria Henando, Paz Sobrino, Ana María Caro, Susana Monge.

Hospital Carlos III: Vicente Soriano, Carmen De Mendoza

Hospital Ramón y Cajal: José María Pe˜na, Hospital de La Paz; Fernando Dronda, Santiago Moreno

Hospital Clínico San Carlos: Víctor Roca, Elisa Pérez-Cecilia

Hospital de La Princesa: Ignacio Santos, Jesús Sanz

Hospital Gregorio Marañón: Jaime Cosín

Fundación Jiménez Díaz: Miguel de Górgolas

Hospital 12 de Octubre: Rafael Rubio, Federico Pulido

Hospital de Móstoles: Carlos Barros

Hospital Severo Ochoa: Javier Gómez

Hospital Príncipe de Asturias: José Sanz

Hospital Puerta de Hierro: Teodoro Martín

Instituto de Salud Pública de Navarra: Jesús Castilla

Researchers involved in GEMES

Santiago Pérez-Hoyos, Inmaculada Ferreros, Julia del Amo, Cristina González, Ana María Caro, Inmaculada Jarrín, Isabel Hurtado, Josefina Belda, Elisa Fernández, Isabel Gómez, Ignacio Alastrue, Concha Santos, Teresa Tasa Zapater, Amparo Juan, Jose Trullen, Robert Muga, Arantza Sanvisens, Jorge del Romero, Paloma Raposo, Carmen Rodríguez, Soledad García, Patricia García de Olalla, Joan Cayla, Eva Masdeu, Rafael Guerrero, Manolo Quintana, Isabel Ruiz, Ramiro Núñez, Jesús Castilla, Marcela Guevara, Carmen de Mendoza and Natalia Zahonero.

Centres and researchers involved in CoRIS

Executive Board: Juan Berenguer, Julia del Amo, Federico García, Félix Gutiérrez, Pablo Labarga, Santiago Moreno y María Ángeles Muñoz.

Field Work, 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 y Almudena García Torre.

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 So-

cas, 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, 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, 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, Paloma Raposo, 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.