INSTITUTO DE HIGIENE E
MEDICINA TROPICAL
DESDE 1902

Universidade Nova de Lisboa Instituto de Higiene e Medicina Tropical

Dynamics of HIV-1 transmission in Europe: a guidance for evidence-based prevention strategies

Mafalda Nunes Da Silva Miranda

Tese para a obtenção do grau de Doutor em Saúde Pública Global

Julho 2022

INSTITUTO DE HIGIENE E
MEDICINA TROPICAL
DESDE 1902

Universidade Nova de Lisboa Instituto de Higiene e Medicina Tropical

Dynamics of HIV-1 transmission in Europe: a guidance for evidence-based prevention strategies

Autor: Mafalda Nunes Da Silva Miranda

Orientador: Ana Barroso Abecasis

Coorientador: Marta Pingarilho

Tese apresentada para cumprimento dos requisitos necessários à obtenção do grau Doutor em Saúde Pública Global

Esta tese teve o apoio financeiro da Fundação para a Ciência e Tecnologia através das bolsas PD/BD/135714/2018 e COVID/BD/152613/2022

Publications that are part of this dissertation

1. Ana Cláudia Miranda*, **Mafalda Miranda***, Marta Pingarilho, Victor Pimentel, João Torres, Susana Peres, Teresa Baptista Alberto, Perpetua Gomes, Ana Abecasis, and Kamal Mansinho. Determinants of HIV-1 Late Presentation in a Cohort of Portuguese HIV-1 Patients. AIDS Research and Human Retroviruses.Nov 2021.846-851.http://doi.org/10.1089/aid.2020.0175 (shared first author)
2. **Miranda MNS**, Pingarilho M, Pimentel V, Martins MdRO, Vandamme A-M, Bobkova M, Böhm M, Seguin-Devaux C, Paredes R, Rubio R, Zazzi M, Incardona F, Abecasis A. Determinants of HIV-1 Late Presentation in Patients Followed in Europe. Pathogens. 2021; 10(7):835. https://doi.org/10.3390/pathogens10070835
3. **Miranda MNS**, Pingarilho M, Pimentel V, Martins MdRO, Kaiser R, Seguin-Devaux C, Paredes R, Zazzi M, Incardona F and Abecasis AB (2022) Trends of Transmitted and Acquired Drug Resistance in Europe From 1981 to 2019: A Comparison Between the Populations of Late Presenters and Non-late Presenters. Front. Microbiol. 13:846943. doi:10.3389/fmicb.2022.846943
4. **Miranda MNS**, Pingarilho M, Pimentel V, Martins MdRO, Seabra SG, Gomes P, Kaiser R, Böhn M, Seguin-Devaux C, Paredes R, Bobkova M, Zazzi M, Incardona F and Abecasis AB The role of Late Presenters on HIV-1 transmission clusters in Europe (in submission)

Oral Presentations

1. **Miranda, M**; Pingarilho, M; Pimentel, V; O. Martins, MR; Vandamme, A; Bobkova, M; Böhm, M; Devaux, C; Paredes, R; Rubio, R; Zazzi, M; Incardona, F; Abecasis, A "Determinants of HIV-1 late presentation in patients followed in Europe" (Oral presentation at 19th European Meeting on HIV & Hepatitis 2021, Virtual meeting)
2. **Miranda, MNS**; Pingarilho, M; Pimentel, V; O. Martins, MR; Kaiser, R; Seguin-Devaux, C; Paredes, R; Zazzi, M; Incardona, F; Abecasis, A "Prevalence of transmitted drug resistance among late presenters in Europe" (Oral presentation

at 32nd European Congress of Clinical Microbiology and Infectious Diseases (ECCMID) April 2022, Lisbon, Portugal)

3. **Miranda, MNS**; Pingarilho, M; Pimentel, V; O. Martins, MR; Seabra, SG; Kaiser, R; Seguin-Devaux, C; Paredes, R; Bobkova, M; Zazzi, M; Incardona, F; Abecasis, A "HIV transmission clusters in Europe: A perspective view of late presenters and non-late presenters" (Oral presentation at 20th European Meeting on HIV & Hepatitis June 2022, Paris, France)

Other manuscripts for which I have contributed during my PhD but that are not included as part of this thesis:

1. Pimentel V, Pingarilho M, Alves D, Diogo I, Fernandes S, **Miranda M**, Pineda-Peña A-C, Libin P, Martins MRO, Vandamme A-M, Camacho R, Gomes P, Abecasis A. Molecular Epidemiology of HIV-1 Infected Migrants Followed Up in Portugal: Trends between 2001– 2017. Viruses. 2020; 12(3):268. https://doi.org/10.3390/v12030268
2. Pingarilho M, Pimentel V, Diogo I, Fernandes S, **Miranda M**, Pineda- Pena A, Libin P, Theys K, O. Martins MR, Vandamme A-M, Camacho R, Gomes P, Abecasis A, on behalf of the Portuguese HIV-1 Resistance Study Group. Increasing Prevalence of HIV-1 Transmitted Drug Resistance in Portugal: Implications for First Line Treatment Recommendations. Viruses. 2020; 12(11):1238. https://doi.org/10.3390/v12111238
3. Pimentel, Victor Figueiredo; Pingarilho, Marta; Sole, Giordano; Alves, Daniela; **Miranda, Mafalda**; Diogo, Isabel; Fernandes, Sandra; Pineda-Pena, Andrea; Martins, M. Rosário O.; Camacho, Ricardo; Gomes, Perpétua; Abecasis, Ana B. on behalf of the Portuguese HIV-1 Resistance Study Group Differential patterns of post-migration HIV-1 infection acquisition among Portuguese immigrants of different geographical origin, AIDS: February 25, 2022 - Volume - Issue - doi: 10.1097/QAD.0000000000003203
4. Pingarilho M, Pimentel V, **Miranda MNS**, Silva AR, Diniz A, Ascenção BB, Piñeiro C, Koch C, Rodrigues C, Caldas C, Morais C, Faria D, da Silva EG, Teófilo E, Monteiro F, Roxo F, Maltez F, Rodrigues F, Gaião G, Ramos H, Costa I, Germano I, Simões J, Oliveira J, Ferreira J, Poças J, da Cunha JS, Soares J,

Henriques J, Mansinho K, Pedro L, Aleixo MJ, Gonçalves MJ, Manata MJ, Mouro M, Serrado M, Caixeiro M, Marques N, Costa O, Pacheco P, Proença P, Rodrigues P, Pinho R, Tavares R, de Abreu RC, Côrte-Real R, Serrão R, Castro RS, Nunes S, Faria T, Baptista T, Martins MRO, Gomes P, Mendão L, Simões D and Abecasis A (2022) HIV- 1-Transmitted Drug Resistance and Transmission Clusters in Newly Diagnosed Patients in Portugal Between 2014 and 2019. Front. Microbiol. 13:823208. doi: 10.3389/fmicb.2022.823208

5. **Miranda MNS**, Pingarilho M, Pimentel V, Torneri A, Seabra SG, Libin PJK and Abecasis AB (2022) A Tale of Three Recent Pandemics: Influenza, HIV and SARS-CoV-2. Front. Microbiol. 13:889643. doi: 10.3389/fmicb.2022.889643

This Doctoral programme is an association between:
Instituto de Higiene e Medicina Tropical, Universidade Nova de Lisboa - IHMT/UNL
Escola Nacional de Saúde Pública/ Universidade Nova de Lisboa - ENSP/UNL
NOVA Medical School | Faculdade de Ciências Médicas - NMS|FCM/UNL
Faculdade de Medicina, Universidade do Porto - FM/UP

“Happiness can be found even in the darkest of times, if one only remembers to turn on the light.”

— Albus Dumbledore, *Harry Potter and the Prisoner of Azkaban*

Acknowledgments

First, I would like to thank my supervisor, Professor Ana Abecasis, who took me in when no one else would give me an opportunity to pursue a career in investigation and for guiding me throughout these years, not only in the course of my PhD but also outside the academic world, where I consider her a dear friend of mine. I would really like to express my deepest gratitude for her support and ongoing faith in my work and more important to make me feel that I am a part of her team. Since I started my PhD, she was also one of the people who I see as an inspiration.

Second, I would like to thank my co-supervisor, Doctor Marta Pingarilho, who helped me immensely to make this PhD possible. I know sometimes I was too much to deal with, but I hope that even though she looks at me as a colleague and friend that she can rely on. I would also want to highlight what a great hardworking professional she is, and her example makes me want to always be better. Marta was a key element in the development and delivery of my PhD thesis, and for that I will be forever grateful. I am also lucky to consider her a dear friend.

To Doctor Victor Pimentel, one dear friend and colleague, who taught me a lot during the course of the PhD, who was always ready to lighten up the room and was present in every step I gave with all his honesty and advice. He is also a big part of my PhD path and it would have not been the same without his presence.

To Professor Maria do Rosário Oliveira Martins, for her acceptance of being part of my Tutorial Commission and for her support during my PhD and for her availability to help every time I needed.

To Professor Anne-Mieke Vandamme, for being part of my Tutorial Commission.

To Professor Henrique Barros, for making this PhD possible.

To the EuResist network, especially to the CEO Francesca Incardona, for allowing me to use their database and that made my research possible.

To my family, especially my mother and sister, who supported me throughout these years and never let me give up or give it to pressure. To my father, that helped me to get my masters degree and therefore to be able to do a PhD. To my grandma Olimpia, who is here in body but not in mind and I know she would be very proud to see me achieve another step in my education. To my grandparents, Carlos e Fabiana, both still very proud of what I am and what I am doing. And finally, to my grandpa, Alfredo, who is among us anymore, but has been looking over me until now.

To all my colleagues from the Global Public Health PhD Program, Luis, Miguel, Patrícia, Kelli, Claúdia and Sousan, who accompanied me in this journey and challenged me to be a better professional and person.

To all my colleagues from the IHMT, especially Daniela Alves and Francisco Merca, with whom I share a lot of memories and adventures, and I am glad to call friends.

To all my friends, for continuous support and understanding. Highlighting my best friends, Andreia Antunes, Maria Inês Soares, Fátima Evangelista, Tatiana Mendes, Filipa Romão, Beatriz Crespo and Marta Taleto, who have been with me in every step of the way and helped me achieve clairvoyance and strength to pursue my dreams even though they might be hard to achieve. I would also like to thank especially to Diogo Pereira and Miguel Simplício, my oldest friends, who have been an important part of my life and had the patience to keep me straight to the right path and always cheered my accomplishments, I don't know how I would have survived without them.

Resumo

Introdução: Para controlar a pandemia de VIH a UNAIDS desenvolveu os objetivos 95-95-95 para serem atingidos até 2030. A concretização destes objetivos pode ser dificultada pelo aparecimento de mutações de resistência devido ao uso intensivo da terapia antirretroviral ou também à existência de indivíduos com apresentação tardia (IAT) ao diagnóstico. Estes IAT não só impactam os resultados dos seus próprios tratamentos como são também uma ameaça para alcançar os objetivos da UNAIDS, uma vez que podem potenciar, de forma inconsciente, a transmissão do VIH.

Objetivos: Primeiro, identificar as caraterísticas sociodemográficas e clínicas dos indivíduos infetados com VIH-1, bem como identificar os determinantes da apresentação tardia em Portugal e na Europa. Segundo, descrever os padrões de resistência transmitida (TDR) e de resistência adquirida (ADR) em indivíduos infetados com VIH-1 seguidos na Europa, comparar os seus padrões de resistência IAT e indivíduos com apresentação não-tardia (IANT) e analisar as mutações de resistência aos antirretrovirais nos diferentes subtipos de VIH-1. Para finalizar, descrever e caraterizar os clusters de transmissão (CT) de VIH-1 na Europa e comparar o papel dos IAT com os IANT nos CT do VIH-1.

Metodologia: No primeiro estudo, a base de dados utilizada incluiu dados clínicos e sociodemográficos de indivíduos infetados com VIH-1 seguidos no Hospital Egas Moniz, Centro Hospitalar de Lisboa Ocidental (CHLO), Lisboa, Portugal, entre1984 e 2017. Nos outros estudos, utilizou-se a EuResist Integrated Database (EIDB) que inclui dados sociodemográficos, clínicos e genómicos de indivíduos infetados com VIH-1 seguidos na Europa (Portugal, Espanha, Alemanha, Luxemburgo, Rússia, Reino Unido e Itália) entre 1981 e 2019. Para a análise dos CTs, foram utilizadas informações de indivíduos infetados pelos subtipos mais prevalentes (B, A e G).

Resultados: No primeiro estudo, 68,7% dos indivíduos infetados com VIH-1 eram homens com uma mediana de idade de 37 anos (IQR 30–47). 50,6% destes indivíduos tinham apresentação tardia (AT) e desses 61,9% tinham apresentação tardia com doença avançada. Os determinantes associados à AT foram idade ao diagnóstico superior a 30 anos e origem em países da África subsaariana. No segundo estudo, entre os indivíduos incluídos na análise a mediana de idade foi igual a 33 anos (IQR: 27,0–41,0) e 74,4% eram homens. 50,4% destes indivíduos tinham AT e os determinantes associados foram idade acima de 56 anos, heterossexuais, indivíduos com origem em países africanos e com carga viral abaixo de 4,1cópias/mL. No terceiro estudo, a mediana de idades obtida foi igual a 37 anos (IQR: 27,0–45,0) e 72,2% eram homens. 71,9% dos indivíduos tinham sido infetados pelo subtipo B e 54,8% foram classificados com AT. Para AT e apresentação não-tardia (ANT) a prevalência de TDR foi 12,3% e 12,6% respetivamente, e a de ADR foi de 69,9% e de 68,2% respetivamente. As mutações mais prevalentes observadas em IAT e IANT foram K103N/S, T215rev, T215FY, M184I/V, M41I/L, M46I/L e L90M. No quarto estudo, o subtipo mais prevalente nos indivíduos infetados com VIH-1 foi o subtipo B (84,7%) seguido do subtipo G (9,4%) e subtipo A (5,9%). A idade mediana foi de 33 anos (IQR: 26,0-41,0) e 75,5% eram homens. 51,4% dos

indivíduos infetados com VIH-1 tinham AT e 21,6% estavam dentro de CTs. As análises filogenéticas demonstraram que apenas 17,6% dos IAT estavam dentro de CTs comparados com 20,2% dos IANT. Para os subtipos A e B, verificou-se que os IAT dentro de CTs foram caracterizados por uma menor percentagem de homens e por uma maior percentagem de indivíduos mais velhos comparativamente aos IANT. Para os subtipos B e G, os IAT dentro de CTs apresentaram maior percentagem de tratados comparativamente com os IANT. No subtipo G, os IAT dentro de CTs, eram maioritariamente utilizadores de drogas intravenosas comparativamente com os IANT. Analisando o tamanho dos CTs, verificou-se que os IANT pertenciam a grandes CTs (>8 indivíduos) comparativamente aos IAT.

Conclusão: A AT é considerada um dos grandes obstáculos para travar a epidemia do VIH e uma ameaça à transmissão do mesmo. Os nossos resultados apresentam as características sociodemográficas e clínicas dos IAT na Europa e indicam que estes não contribuem, de forma significativa, para a transmissão dos VIH-1. Os resultados encontrados podem contribuir para o desenvolvimento de medidas de prevenção e para uma melhor compreensão das mutações de resistência e falhas terapêuticas nesta população de indivíduos.

Palavras-Chave: VIH-1; Apresentação tardia; Europa; Resistências; Clusters de transmissão

Abstract

Background: To control the HIV pandemic, the UNAIDS set the 95-95-95 targets to be reached by 2030. These targets can be more difficult to achieve, whether due to the appearance of drug resistance mutations regarding the increasing use of antiretroviral therapy (ART) or due to individuals who present late at diagnosis (late presenters-LP). These individuals can not only impact treatment outcomes, but also threat UNAIDS goals, as well as potentiate the spread of HIV.

Aims: First, to identify clinical and sociodemographic characteristics of HIV-1 infected patients, as well as to identify determinants of late presentation in Portugal and in Europe. Second, to describe the patterns of transmitted drug resistance (TDR) and acquired drug resistance (ADR) in HIV-1 infected patients followed in Europe (Portugal, Spain, Germany, Luxembourg, United Kingdom, Russia and Italy), to compare its patterns in late presenters (LP) vs non-late presenters (NLP), and to analyze the most prevalent drug resistance mutations among HIV-1 subtypes. And finally, to describe and characterize HIV-1 transmission clusters in Europe and to compare the role of LP vs NLP populations on HIV-1 transmission clusters (TC).

Methods: For the first study, the database included clinical and sociodemographic information from HIV-1-infected patients followed in Hospital Egas Moniz, Centro Hospitalar de Lisboa Ocidental (CHLO), Lisbon, Portugal, between 1984 and 2017. For the other studies, the EuResist Integrated Database (EIDB) included socio-demographic, clinical, and genomic information from HIV-1 infected patients followed between 1981 and 2019. For the analysis of TC, information from patients infected with the most prevalent subtypes B, A and G was analyzed.

Results: In the first study, 68.7% of patients were males and the median age was 37 years (IQR 30–47). 50.6% patients were LP and, of those, 61.9% were late presenters with advanced disease (LPAD). The determinants associated with LP were age at diagnosis higher than 30 years and origin from sub-Saharan Africa. In the second study, among the HIV-1 infected patients included in the analysis, the median age was 33 (IQR: 27.0–41.0) years and 74.4% were males. 50.4% were late presenters and the determinants associated with late presentation were older patients (>56), heterosexuals, patients originated from Africa and patients presenting with log VL >4.1. In the third study, the median age of HIV-1 infected individuals was 37 (IQR: 27.0–45.0) years old and 72.6% were males. 71.9% of patients were infected by subtype B and 54.8% of patients were classified as LP. For LP and NLP, the TDR prevalence was 12.3% and 12.6%, respectively, while ADR, was 69.9% and 68.2%, respectively. The most prevalent TDR drug resistance mutations, in both LP and NLP, were K103N/S, T215rev, T215FY, M184I/V, M41I/L, M46I/L, and L90M. In the fourth study, the most prevalent subtype among those infected with HIV-1 was subtype B (84.7%), followed by subtype G (9.4%) and subtype A (5.9%). The median age was 33 (IQR: 26.0-41.0) years old and 75.5% of patients were males. 51.4% of patients were classified as LP and 21.6% of patients were inside TCs.

Phylogenetic analyses showed that only 17.6% of LPs were inside clusters compared to 20.2% of NLPs. For subtypes A and B, we found that LP inside clusters were less frequently males and were older than NLPs. For subtypes B and G, LP inside clusters were more frequently treated than NLP. In subtype G, LP inside clusters more frequently had IDU transmission route than NLP. Finally, when analyzing cluster size, we found that NLP more frequently belonged to large clusters (>8 patients) when compared to LP.

Conclusion: Late presentation is a major obstacle to halt the HIV epidemic and could be a threat to HIV-1 transmission. Our results characterize the socio-demographic and clinical characteristics of LPs in Europe and, all together, indicate that LPs are not important contributors to forward HIV-1 transmission. These results help to direct prevention measures for this population and to better understand drug resistance mutations and therapeutic failure in this population of patients.

Keywords: HIV-1; Late presentation; Europe; Drug Resistance; Transmission Clusters

Table of Contents

List of Figures

List of Tables

Abbreviations

3TC - lamivudine

ABC - abacavir

ADR - acquired drug resistance

AIDS - acquired immunodeficiency syndrome

ART - antiretroviral therapy

ARV - antiretroviral drugs

ATV - atazanavir

AZT - zidovudine

CA - capsid

CAB - cabotegravir

CCR5 - chemokine (CC motif) reeptor 5

CDC - centers for disease control e prevention

CRF - circulating recombinant form

CXCR4 - chemokine (C-X-C motif) receptor 4

d4T - stavudine

DDI - didanosine

DNA - deoxyribonucleic acid

DOR – doravirine

DRM - drug resistance mutations

DRV - darunavir

DTG - dolutegravir

EFV - efavirenz

ETR - etravirine

FDA - food and drug administration

FI - fusion inhibitors

FPV - fosamprenavir

FTC - emtricitabine

HAART - highly active antiretroviral therapy

HIV- human immunodeficiency virus

IDU- intravenous drug user

IDV- indinavir

IN - integrase

INSTI - integrase inhibitors

LAV - lymphadenopathy-associated virus

LP - late presenter

LPV - lopinavir

MA - matrix

MSM - men who have sex with men

MVC - maraviroc

MRCA - most recent common ancestor

NC - nucleocapsid

NFV- nelfinavir

NLP- non-late presenter

NNRTI - non-nucleoside reverse transcriptase inhibitors

NRTI - nucleoside reverse transcriptase inhibitors

NVP - nevirapine

PI - protease inhibitors

PR - protease

PrEP - pre-exposure prophylaxis

RAL - raltegravir

RNA - ribonucleic acid

RPV - rilpivirine

RT - reverse transcriptase

RTV - ritonavir

SDRM - surveillance drug resistance mutations

SIV - simian immunodeficiency virus

STDs - sexual transmitted diseases

SQV - saquinavir

T-20 - enfuvirtide

TC - transmission cluster

TDF - tenofovir

TDR - transmitted drug resistance

TPV – tipranavir

UNAIDS- joint united nations programme on HIV/AIDS

WHO - world health organization

1. Introduction

1.1. HIV-1 Epidemiology

In 1981, the first cases concerning young homosexual men with depleted T-lymphocytes, which died of opportunistic infections, were reported in the United States. This disease would later be known to the world as AIDS (1). The cause of AIDS, in 1983 was still unknown, however the number of people with AIDS in the US continued to grow. In February of the same year, the Centers for Disease Control e Prevention (CDC) reported 1000 cases of AIDS (2). AIDS was identified as a transmissible syndrome that could be transmitted between individuals, and it was observed that the infectious agent could be spread via distinct transmission routes, including sexual transmission, vertical transmission and blood-borne transmission (i.e., intravenous drug user (IDU) and blood products) (2). To unravel the pathogen responsible for AIDS, scientists studied the immune response of individuals that exhibited AIDS-related symptoms, and in late 1983, a new human retrovirus, dubbed lymphadenopathy-associated virus (LAV), was isolated from a patient in France at the Institut Pasteur (3). This virus was confirmed as the cause of AIDS and was later renamed as human immunodeficiency virus (HIV) (3).

Since the beginning of the pandemic around 36 million people have died from AIDS-related illness, while around 79 million people have been infected with HIV. In 2020, there were 37.7 million individuals living with HIV, 1.5 million were new infections and 680000 individuals died from AIDS-related illness and in June 2021, 28.2 million individuals were under antiretroviral therapy (ART) regimens (4). Graphs in Figure 1 show the evolution through time of the people living with HIV, the AIDS-related deaths and the new HIV-1 infections between 1990 and 2020.

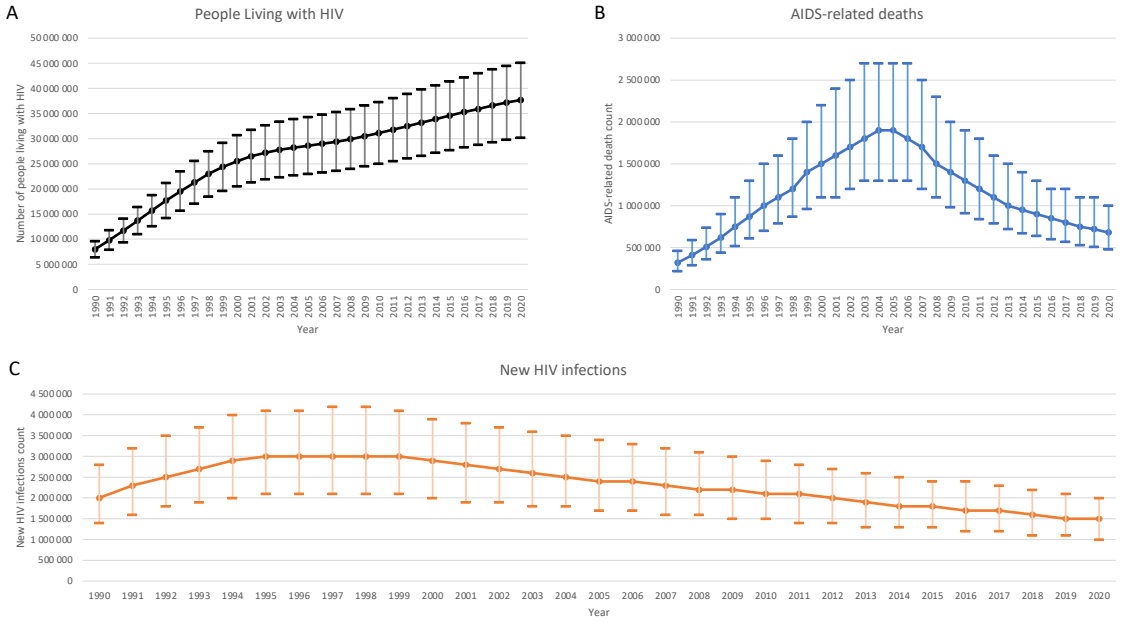

Figure 1. HIV-1 graphs for the number of people living with HIV-1 per year (A), AIDS-related death count per year (B) and new HIV-1 infections count per year (C) between 1990 and 2020.

Data source: UNAIDS (5).

In 2014 the Joint United Nation Programme on HIV/AIDS (UNAIDS) had set some targets to be attained until 2020 with the objective to control the HIV pandemic. The 90-90-90 targets state that 90% of people living with HIV know their status, of those 90% are receiving Antiretroviral Therapy (ART) and of those 90% achieve viral suppression. By the end of 2019, according to UNAIDS, globally, there were 81% of people living with HIV who knew their status. Of those, 67% were receiving antiretroviral therapy and of those 59% had reached HIV viral suppression. Moreover, between 2010 and 2019, the percentage of new infections dropped by 31% (6). After that, the 95-95-95 targets were defined based on the same definition of the previous ones and were set to end the pandemic by 2030 (7).

1.2. Etiological agent and origin of HIV-1 infection

Orthogonal to outbreak analyses, determining HIV-1 genomic sequences also enabled researchers to study the origin of HIV, which led to the discovery that the different HIV-1 groups originated from a series of distinct zoonotic transmission event from non-human primates of simian immunodeficiency virus (SIV) (8).

HIV-1 is a part of the *Retroviridae* family and *Lentivirus* genus (9). SIV was transmitted from chimpanzees subspecies *Pan troglodytes troglodytes* and gorillas to humans (9) as shown in Figure 2.

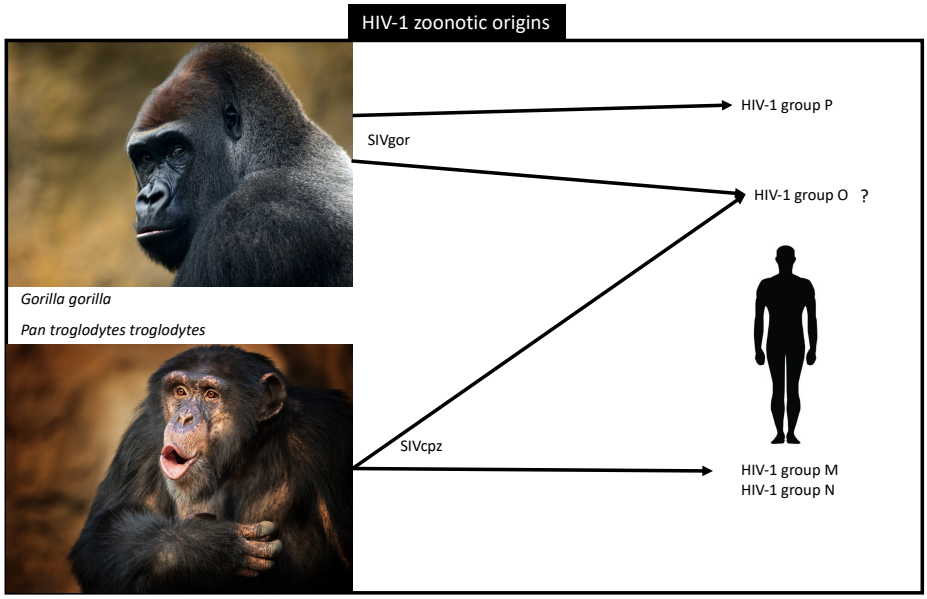

Figure 2. HIV-1 zoonotic origins.

In this figure we show the zoonotic transmission of simian immunodeficiency viruses (SIV) from non-human primates to humans, leading to the development of the human immunodeficiency virus (Adapted from Tebit DM and Arts EJ (2011) (10)).

Early in the 20th century, four independent zoonotic transmission events from these primates led to the origin of four HIV-1 groups. The most common and the oldest being group major (M), responsible for the HIV-1 global pandemic, directly originated from the chimpanzee *Pan troglodytes troglodytes* (11). HIV-1 is composed by four groups (M,N,O and P) and 10 subtypes (A,B,C,D,F,G,H,J,K and L) from group M, and

at least 132 circulating recombinant forms (CRFs), as well as some unique recombinant forms (URFs), have been documented (9,12–14). CRFs are viruses characterized as having a genome with identical mosaic patterns and different clustering regions in phylogenetic trees of genomic sequences of individuals who are epidemiological unlinked. They are a combination of two or more different pure subtypes of the virus and are a result of viral replication and high levels of mutation and recombination through the reverse transcriptase enzyme inside the host infected cell (12,15,16).

The consensus on the origin of HIV group-M is that the zoonotic transmission event took place in Central Africa, and the epidemic ignition occurred in Kinshasa (Democratic Republic of Congo) region. From there it could spread to other regions in Africa and subsequently the rest of the World (17), as shown in Figure 3.

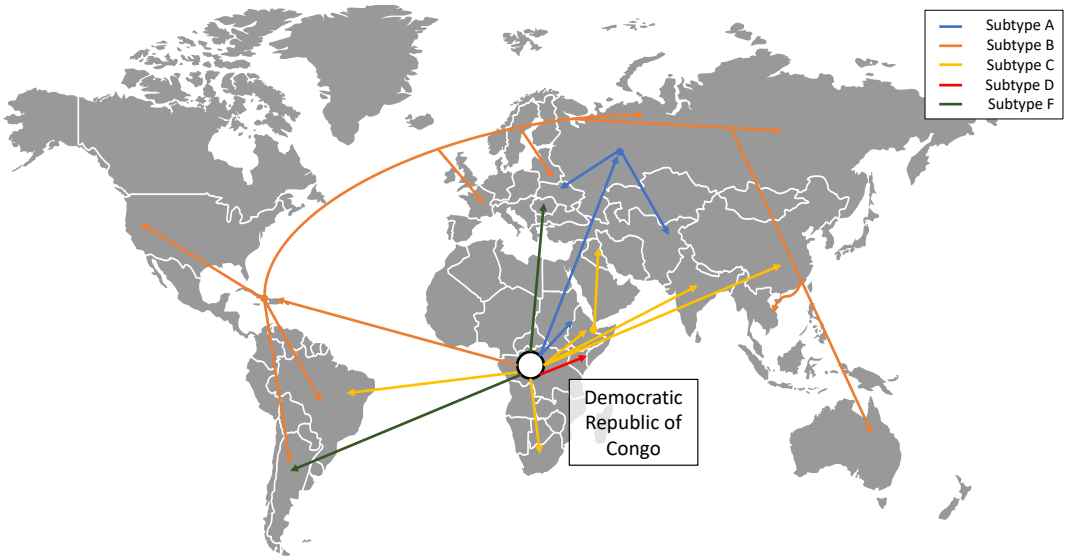

Figure 3. HIV-1 group M global dispersion patterns.

In this figure is shown how HIV- 1 group M disseminated from its original epidemic location (Kinshasa, Democratic Republic of Congo) to other regions of the globe. It is clear how subtype B is the most widely spread subtype globally. It is also demonstrated how subtype A disseminated mostly to the east regions of Africa, Europe and Asia and subtype C disseminated widely through Brazil, South Africa and Southeast Asia. (Adapted from Tebit DM and Arts EJ (2011) (10)).

The group outlier (O) was related to SIV from gorillas in Cameroon and caused infections in West-Central Africa, especially in regions like Cameroon, Gabon and

Equatorial Guinea, although it is though that the original hosts might have been chimpanzees (10,18). The group nonmajor and nonoutlier (N) was originated from natural reservoirs of SIV from chimpanzee *Pan troglodytes troglodytes* and caused infections in a small proportion of individuals in Cameroon (11) as we can see in Figure 4. The more recently identified group (P), only caused infections in a few individuals in Cameroon and is thought to be originated from SIV from gorillas (19).

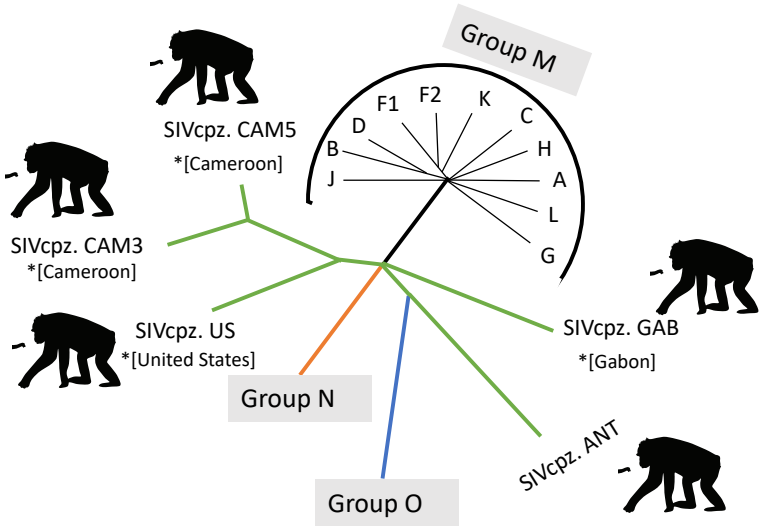

Figure 4. Phylogenetic tree for HIV-1.

The HIV-1 phylogenetic tree shows sequences that demonstrate the zoonotic jump of the distinct HIV-1 groups (Adapted from Thomson MM, Pérez-Álvarez L and Nájera R (2002) (20)).

1.3. Viral particle

Characteristic of the genus *lentivirus*, the HIV is enveloped by a lipid bilayer, derived from the membrane host cell, where are anchored the surface and transmembrane glycoproteins gp120 and gp41, respectively (21). Internally, the particle is composed by the matrix shell which includes the matrix protein (p17), a nucleocapsid (NC) and a conical capsid core particle, which includes the capsid protein (p24), which is located the

HIV genome (22). The HIV genome consists of two copies of a single-stranded, positive-sense segment of Ribonucleic Acid (RNA) as shown in Figure 5 (23).

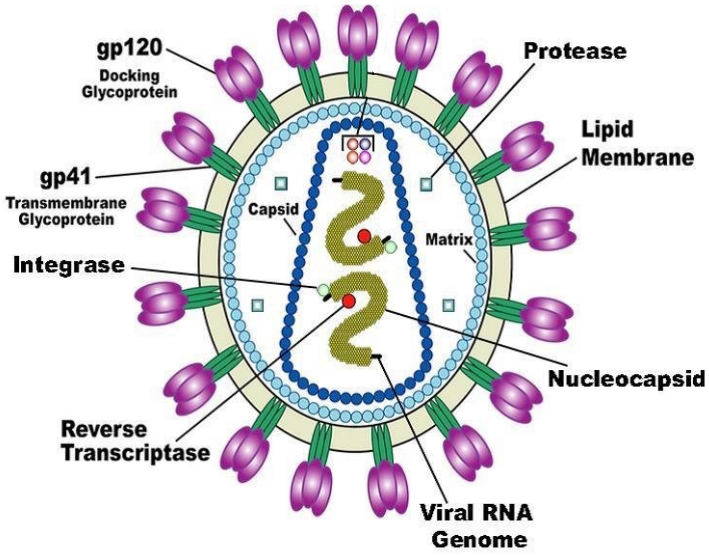

Figure 5. HIV viral particle.

RNA: Ribonucleic Acid; gp120: glycoprotein 120; gp41: glycoprotein 41. (Adapted from Drug Discovery and Development/ Discovery of Key Component of HIV Virus Yields Drug Target (24)).

1.4. HIV-1 genomic structure

This virus genome is composed by the structural genes *gag* (group-specific antigen), *env* (envelope glycoprotein) and *pol* (polymerase). Additionally, the HIV-1 genomic structure is composed by four accessory genes (*vif*, *vpr*, *vpu* and *nef*) and two regulatory genes (*tat* and *rev*), besides the main structural genes, as shown in Figure 6 (25).

For the expression of the virus genes and the formation of hybrid structures of DNA-RNA, the regions of Long Terminal Repeats (LTR) are essential. These regions are hundreds of nucleotides-long and are comprised at the extremities 5' and 3' of the proviral Deoxyribonucleic Acid (DNA) as shown in Figure 6. These regions are involved in the expression of the enhancer/promoter proximal, transcription and insertion processes (26,27).

The *gag* gene is located immediately downstream of LTR 5'. The *gag* gene encodes for capsid (CA), matrix (MA), nucleocapsid (NC) and p6 proteins. These proteins are responsible for the assembly of the virion (28). From 5' to 3', next to the *gag* gene there is the *pol* gene, which encodes for protease (PR), reverse transcriptase (RT), and integrase (IN) associated with the genetic material. The PR is responsible for the maturation of viral proteins, the RT is responsible for the transcription of the viral RNA into double-stranded DNA and the IN is responsible for the integration of the viral DNA into the host cell (29). On the 3' extremity there is the *env* gene, which encodes for glycoprotein gp120 and glycoprotein gp41 and are responsible for the binding of the virus to the CD4 cell receptor and the envelope fusion process, respectively (30).

The accessory protein *tat* contributes to the HIV-1 viral replication, increasing the production of viral RNA by enhancing the rate of transcription (31). The *rev* is a viral protein regulator and helps the transportation of viral mRNA from the nucleus into the cytoplasm (22). The *vif*, virion infectivity factor, encodes a cytoplasmatic protein with an essential role in replication of highly infectious mature virions. The *vpr* encodes the viral protein R that integrates the mature virion, and its role has been difficult to describe, although it is known to enhance the expression of HIV proteins and could induce apoptosis. The gene *vpu* encodes the viral protein U and is responsible for cell surface modulation and influence viral replication and dissemination. The *nef* gene encodes one of the first viral proteins expressed after infection (22,25,29).

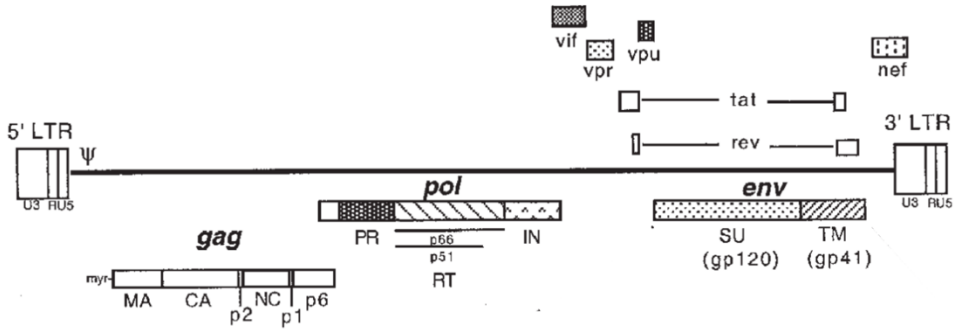

Figure 6. Schematic representation of HIV-1 proviral genome.

Different boxes indicate different genes. MA-Matrix; CA-Capsid; NC-Nucleocapsid; PR-Protease; RT-Reverse Transcriptase; IN-Integrase; SU-Surface; TM-Transmembrane (Source: Adapted from *Freed, E.O. HIV-1 Replication* (27)).

1.5. Replicative cycle of HIV

The replicative cycle starts when the HIV-1 envelope spikes are triggered, and it initiates a cascade of changes that culminate in the fusion of the viral and host cell membranes. Therefore, there is a release of the viral nucleus into the cytoplasm that goes to infect the TCD4+ cells and macrophages first (30).

Regarding the infectious mechanism of HIV, it is divided into two phases (21). The early phase and the late phase. The early phase initiates with the entry of the virus into the host cell mediated by an interaction between the HIV envelope and the CD4 and a co-receptor, usually chemokine (CC motif) receptor 5 (CCR5) or chemokine (C-X-C motif) receptor 4 (CXCR4), of the host cell (32). After that, the CD4 cell receptor binds to the gp120 of the virus which leads to the fusion peptide (in the terminus of gp41) to be inserted into the cell plasma membrane (33). After the release into the cytoplasm, the HIV-1 core undergoes a uncoating event which is poorly understood and then the process of reverse transcription is initiated and catalyzed by RT (27). The transcription process of the viral RNA begins from the 5' extremity, and through the action of the RT the DNA segment is synthetized in direction to that same extremity. In the extremity 3', the DNA

segment, going in the opposite direction of the previous one described, completes the synthetizes into a single strand of DNA. At the same time, the RNase H enzyme, part of RT, is necessary to digest the RNA from the hybrid DNA-RNA. Alongside the DNA-polymerase RNA dependent of RT transform the complementary DNA single strand, resulting in the double-stranded DNA (34). After, the double-stranded DNA is directed to the HIV nucleus where it is integrated, by IN, into the host genome (35).

The late phase starts after the HIV integration, the RNA polymerase II of the host cell uses the viral DNA as a base to the syntheses of viral proteins, mRNA transcripts. Consequently, the mRNAs are exported to the cytoplasm, where they are translated in the ribosomes. The final step of the late phase is the virus assembly and release, where the genomic RNAs move to the plasma membrane and acquires a portion of it containing gp120 and gp41. The next step is the maturation of HIV virus, that happens through the action of the PR enzyme, that cleaves the proteins into their final forms, therefore forming the infectious and mature viral particle (21,28,36). The Figure 7 summarizes the replication process.

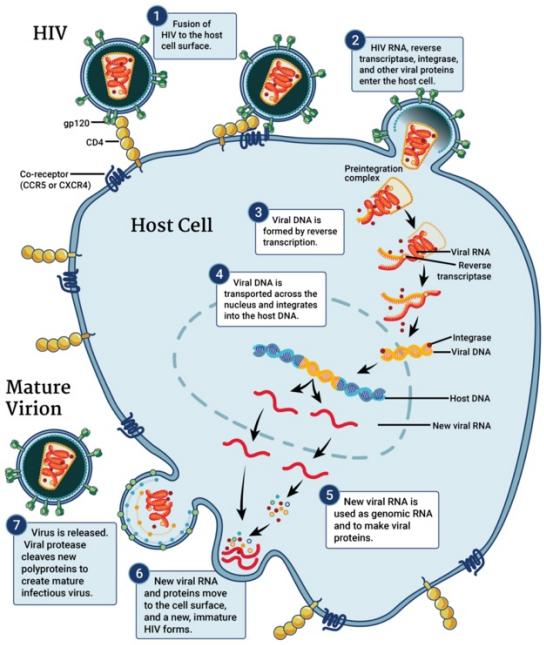

Figure 7. HIV-1 replicative cycle.

HIV: Human Immunodeficiency Virus; gp120: glycoprotein 120; RNA: Ribonucleic Acid; DNA: Deoxyribonucleic Acid. (Source: Adapted from National Institute of Allergy and Infectious Diseases, HIV Replication Cycle (37)).

1.6. Routes of transmission

HIV can be transmitted via different routes, the most common being sexual contacts (via heterosexual contact and between men who have sex with men (MSM)). There is also transmission via IDU, transmission via blood products (i.e., contact with blood products) and vertical transmission (mother-to-child) (38). The transmission of HIV-1 depends on its concentration in the body fluid (i.e., blood or genital secretions) and the susceptibility of the human cells to virus-specific determinants. Furthermore, HIV-1 transmission is linked to other Sexual Transmitted Diseases (STDs), as they can increase the efficiency of transmission, such as gonorrhea, chlamydia and trichomoniasis (39). The transmission route of HIV differs considering different geographical areas. In Sub-Saharan Africa, the main route of transmission is via heterosexual contact followed by vertical transmission (40). In Latin America, including South and Central America and Mexico, new cases are mainly generated between MSM and via IDU (40). In North America and Western Europe, the main route is sexual transmission, where MSM are

most at risk (41). In Eastern Europe and Central Asia, the most prevalent mode of transmission remains IDU (42).

1.7. Clinical phases of infection

Without treatment, HIV infection deteriorates the immune system, which lead to a variety of symptoms and therefore progressing through different stages, getting worse over time. There are three stages of HIV infection, the first stage, i.e., acute HIV infection stage; the second stage, i.e., clinical latency, and the third stage, i.e., AIDS (43). Figure 8 illustrates the stages of HIV-1 infection (44).

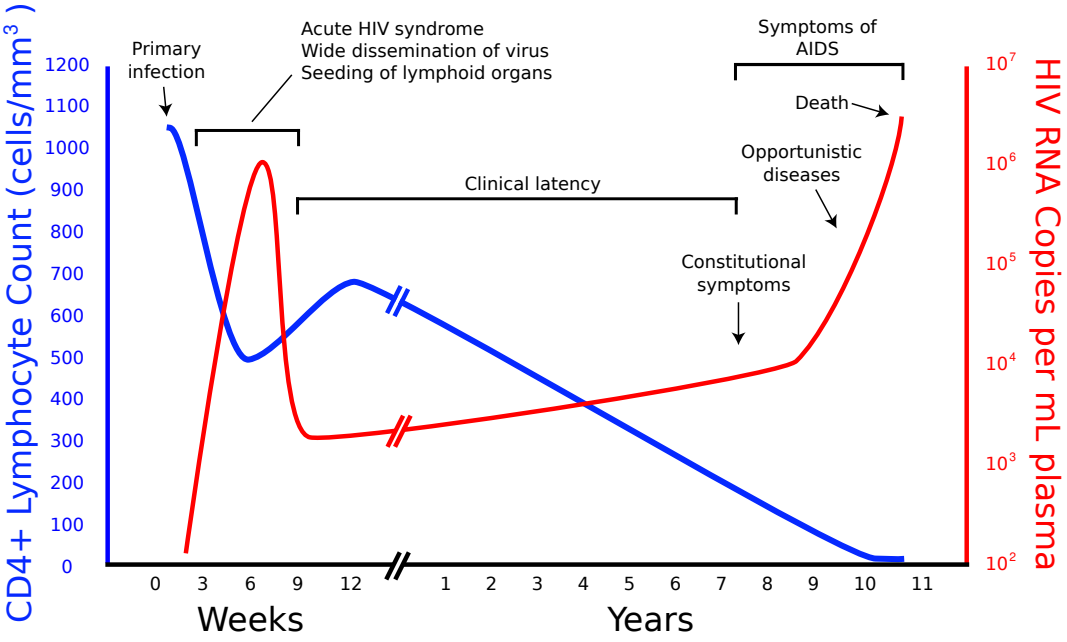

Figure 8. Clinical stages of HIV-1 infection.

Data source: Figure from https://commons.wikimedia.org/wiki/File:Hiv-timecourse.png (Accessed on 23 March 2022).

The first stage occurs between 2-4 weeks post HIV infection. Infected individuals typically start to develop flu-like symptoms, yet a proportion of individuals remains asymptomatic. This stage only lasts several days to few weeks due to high values of viral load and infectiousness (45,46).

During the second stage, the virus multiplies at a low rate and infected individuals will typically not experience any symptoms. This stage can last 10-15 years without treatment and has an epidemiological importance as most HIV infected individuals are not aware of their infection state, enabling them to generate new infections (45). This signifies the importance of reducing the delayed diagnosis in HIV infected individuals. Also, early diagnosis prevent the onward transmission of the virus and increases the chances of treatment success (47). At the end of this stage, the viral load goes up and the CD4 count drops, which constitutes the prelude to the third and final stage, i.e., the AIDS stage (43).

In this last stage, as the viral load increases and the drop in the number of CD4+ cells weaken the immune system of the patient. In this phase the immune system is highly affected, for that reason some opportunistic infections, like pneumonia, can be developed and without treatment the survival time is around three years (48). The main symptoms in this stage can be weight loss, recurring fever, extreme tiredness, prolonged swelling of the lymph nodes, prolonged diarrhea, sores of the mouth, anus or genitals, red, brown, pink or purplish blotches on or under the skin, memory loss, depression and other neurologic affections (45,48).

1.8. Late Presentation at diagnosis

Late presentation is an important clinical condition and can impact health and treatment of infected individuals, which leads to poorer outcomes and increased health care costs (49). Late presenter (LP) is defined, according to the European Late Presenter Consensus working group, as an individual presenting a TCD4+ count lower than 350 cells/mm3 at diagnosis or an AIDS-defining event at diagnosis, regardless of TCD4+ cell count (50).

Late presentation has an impact at individual and at population level, that way negatively impacting the control of the pandemic. LP could also increase the risk of

onward HIV transmission by those who are not aware of their HIV status, considered a public health problem for that reason (51).

Moreover, the cost of care for late presenters has remained higher for the past years achieving values at least twice higher than for non-late presenters (NLP). These costs not only include ART use but also out and in-patient care and other therapeutics related to different diseases. Furthermore, the hospital admission rate is higher for LP whether presenting HIV or non-HIV related conditions (52,53). Treatment failure is more common among LP since they are more difficult to treat, probably because of the toxic effects of ART when a patients present a low CD4 cell count (53). It has also been shown that late presenters above 50 years old are at higher risk for developing non-infectious co-morbidities and complex multimorbidity (54). Usually, LP are individuals that belong to vulnerable groups, as migrants, or without proper access to HIV testing or care and mainly infected by an heterosexual route of transmission (55,56). For prevention and treatment of HIV infection, timely diagnosis and linkage to health care are essential strategies (57).

It is estimated that LP account for 50.4% of HIV cases in Europe (47), in Asia the percentage of LP range from 72 to 83% (58), in Africa range from 35 to 89% (59) and in Brazil, it is estimated that the percentage is around 45-55% (60). As we can see LP represent more than half of HIV infected patients from different regions, which is motive of concern not only in Europe, but globally as well. This population remains understudied, and it can have a high impact on HIV-1 transmission and therapy.

1.9. Antiretroviral Therapy

Highly active antiretroviral therapy (HAART) is used to achieve and maintain viral suppression, making the viral load in the individual infected plasma undetected, which increase CD4 cell count helping to recover the immune system function, thereby improving the clinical status of the patient (61,62). At an individual level, HAART increases quality of life, reduces morbidity and mortality and prevents transmission when an infected individual presents undetectable viral load (61).

Currently, there are seven classes of antiretrovirals drugs (ARV) (63):

- Nucleoside reverse transcriptase inhibitors (NRTI)
- Non-nucleoside reverse transcriptase inhibitors (NNRTI)
- Protease inhibitors (PI)
- Fusion inhibitor (FI)
- Integrase inhibitors (INSTI)
- CCR5 antagonist
- Attachment inhibitor

The NRTIs class inhibits the activity of the reverse transcriptase, and its main function is to translate the virus' RNA to DNA (64). The first approved antiretroviral agent for HIV-1 treatment was Zidovudine (AZT), which is a thymidine analogue and belongs to the class of NRTIs, in 1987 by the Food Drug Administration (FDA) (65). The class of NNRTIs reduces virus replication through the inhibition of the reverse transcriptase, the first NNRTI nevirapine (NVP) was approved in 1996 (66). The class of PIs reduces virus replication by inhibiting the activity of viral enzyme protease, and consequently inhibiting the maturation of virions, and the first PI saquinavir (SQV) was approved in 1995 (65). In the FI class the first and only so far drug approved was Enfuvirtide (T-20) in 2003, which blocks the fusion of the HIV envelope to the cell membrane (61). In the INSTI class, the first drug approved was Raltegravir (RAL) in 2007, and it inhibits the integration of HIV viral genome into the DNA of the host cell, by inhibiting the IN (66). The class of CCR5 antagonist, having Maraviroc (MVC) as the first and only so far drug approved in 2007, inhibits the entry of HIV viral through the blocking of the CCR5 co-receptors (67).

After the first ARV approved, ART was always used in monotherapy even though there were new compounds being developed and approved. However, these compounds used in monotherapy regimen experienced therapeutic failure. (65). It was only in 1996 that ARV drugs could be combined, resulting in HAART, a breakthrough in management of HIV patients, therefore decreasing the mortality rate. The first HAART combinations included two NRTIs plus a protease inhibitor (PI) (62). Figure 9 shows the chronological FDA approval of antiretroviral drugs.

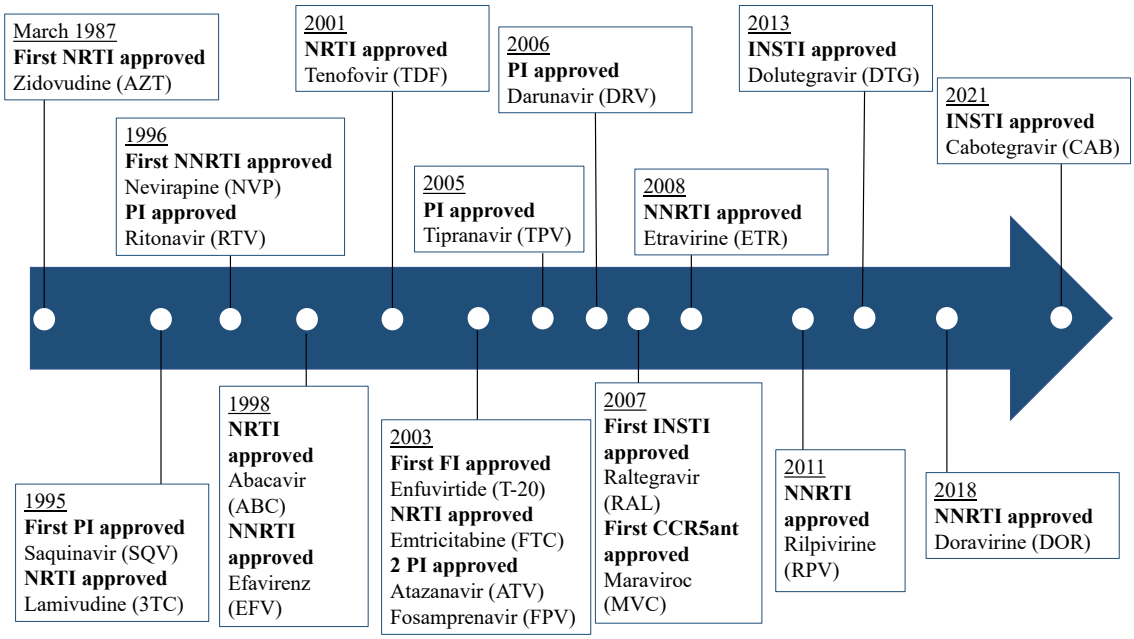

Figure 9. Antiretroviral Drugs by year of FDA approval.

NRTI: nucleoside reverse transcriptase inhibitors; PI: protease inhibitors; NNRTI: non-nucleoside reverse transcriptase inhibitors; FI: fusion inhibitors; INSTI: integrase inhibitors; CCR5ant: CCR5 antagonist. Data source: https://hivinfo.nih.gov/understanding-hiv/fact-sheets/fda-approved-hiv-medicines

In 2016, the guidelines from the World Health Organization (WHO) recommended as first-line ART regimens the combination of two NRTIs, such as tenofovir (TDF) and lamivudine (3TC) or emtricitabine (FTC), plus an INSTI, such as dolutegravir (DTG), or instead of DTG the combination with the NNRTI efavirenz (EFV) (68). There has been some improvements to minimize adherence problems that have been arising through the years, and one strategy developed was the use of injectable long-acting HIV medication, which showed acceptance and effectiveness (69). It is very important to mention that more recent antiretrovirals have a higher genetic barrier, which means that the virus is less likely to escape from selective pressure (70).

In 2015, WHO launched the "Treatment for all" recommendation, which recommended immediate treatment for all HIV infected individuals at any CD4 cell count (71) and in 2020, 73% of people living with HIV had access to ART (4). The observation of this improvement in quality of life and clinical status led to the use of antiviral therapy in prevention, and in the same year, 2015, WHO also issued a recommendation regarding

the use of pre-exposure prophylaxis (PrEP) as a choice for people at risk of contracting HIV (72). The most common PrEP is a combination of TDF and FTC and it has been demonstrated effective in reducing the risk of HIV-1 infection (73). ART is also used as a post-exposure prophylaxis (PEP), when there is a recent exposure to HIV (74).

While at the start of the HIV pandemic a limited number of ARVs were available, nowadays, the options for HIV treatment are wider, with new generation drugs and antiretroviral classes (75). However, resistance to ARV, that could be related to poor treatment adherence, is still a global reality and it is a major barrier to end HIV/AIDS pandemic (76).

1.9.1. Transmitted and Acquired Drug Resistance to ART

Since the introduction of HAART, the clinical outcomes of HIV-1 infected individuals have significantly improved, thereby decreasing both mortality and morbidity rates. Although, due to the fast evolutionary rate of HIV and the selective pressure that is induced by HAART, the emergence of HIV drug resistance can compromise the effectiveness of antiretroviral drugs. The development of antiretroviral resistance occurs when the virus develops mutations that escape the inhibitory properties of the antiviral drugs (77). However, resistance to ART can manifest by two different ways, as a result of selective pressure of antiretrovirals in individuals or as a result of an infection with a virus strain that carries drug resistance mutations (DRM). The first is called acquired drug resistance (ADR) and the second is called transmitted drug resistance (TDR) (78,79). The overall prevalence of ADR in Europe between 1991 and 2019 was 68.5% and TDR between 1995 and 2019 was 12.8% (80). Drug resistance testing is a necessary tool to detect DRMs in newly diagnosed patients in order to guide the selection of ART, to minimize the risks of virological failure (81). DRMs can be categorized as primary and accessory. A primary DRM is a single mutation that has sufficient capacity to reduce the virus' susceptibility to ARV, while an accessory DRM is a mutation that can enhance the capacity of a strain carrying also a primary DRM or contribute to reduce virus susceptibility. There are specific DRMs associated with the different antiretroviral drug classes. The most common DRMs that are associated with higher resistance rates to

NRTIs, NNRTIs and PIs are M41L and M184V; K103N and L90M and M46IL, respectively (82–87).

To acknowledge the importance of drug resistance, WHO released in 2009 a surveillance list with a standard list of mutations with the objective to compare the prevalence of transmitted resistance from different times and regions (88). In this list the surveillance drug resistance mutations (SDRMs) are from the four ARV classes, NRTIs, NNRTIs, PIs and INSTI (89). These DRMs can have different clinical impacts on the different ARV drugs used for HIV therapy. Clinical impact can present different levels of resistance, e.g. high, medium, low, depending on which mutation or a combination of mutations a viral strain of the infected individual is carrying and their association with the ARV drug used in the same individual therapy. DRMs can present low-level, intermediate or high level of clinical impact (90). Table 1 shows the most prevalent SDRMs for each drug class.

Table 1. SDRMs according to ARV class (91–93).

In this table it is described the most prevalent SDRM and which ARV drugs they affect.

ARV Class	**SDRMs**	**ARV Drug**
NRTIs	M41L	AZT, D4T
		ABC, DDI, TDF
	M184VI	FTC, 3TC
		ABC
NNRTIs	K103NS	NVP
		EFV
PIs	M46IL	NFV
		ATV, FPV, IDV, LPV, SQV
	L90M	NFV
		IDV, SQV
		ATV, FPV, LPV

Drug resistance testing is recommended for all patients who are ART-naïve, and usually involves testing for mutations in the RT, PR and IN coding sequences. For patients ART-experienced, genotypic and phenotypic resistance testing also involves testing for mutations in the RT and PR genes. Testing is recommended in individuals carrying mutations suspected of multi drug-resistance and virological failure or suboptimal viral load reduction. Drug resistance testing for mutations in the IN gene are performed if there is virological failure while receiving a regimen including an INSTI drug (94).

1.10. Genetic diversity and Molecular Epidemiology of HIV-1

The genetic diversity of HIV is a result of various factors, not only patterns of human migration and globalization, but also due to mutations, replication cycles and recombination, that are different according to the disparities in sub-epidemics, and heterogeneous in nature (18,95). The CRFs are the cause of the HIV genetic diversity that continues to increase globally, the virus can range between $5x10^{-6}$ and $9x10^{-5}$ mutations per nucleotide, per cycle of virus replication (15). Besides the virus' high rates of genetic diversity, the within-host and between-host dynamics can impact on the evolutionary process as the first dynamic implies competitive fitness and selective forces, and the second implies HIV-1 strains co-existing epidemiologically (95). This evolution of viral strains of HIV-1 is closely related to their subtype and this can be noticeable regarding their differences in route of transmission (96), pathogenicity (97), transmissibility (98) and susceptibility to ART (99).

Africa is still the most affected region, globally, where HIV had its origin (100). The United States, on the other hand, is one of the most affected country within high-income countries, where AIDS was first detected (101). Subtype C is the most prevalent subtype worldwide and account for 50% of infections concentrated in Southern and East Africa, as well as India (98). Even though subtype B accounts for 12% of infections, this subtype is the most spread globally and is highly dispersed in Europe, North America, Australia and South America (97). Subtype A is the third most prevalent accounting for 10% of

infections and is concentrated in Eastern Africa region. Subtype A is followed by the recombinant strains CRF02_AG (7.7%) mainly dispersed in the West and Central Africa regions and CRF01_AE (5.3%) mainly dispersed in South and East Asia (97,102). Following the CRFs strains, the most prevalent subtypes worldwide are subtype G (4.6%) and D (2.7%). The four subtypes B, A, G and D are associated with higher pathogenicity (102). The genetically different viruses have been spread throughout geographical areas (103).

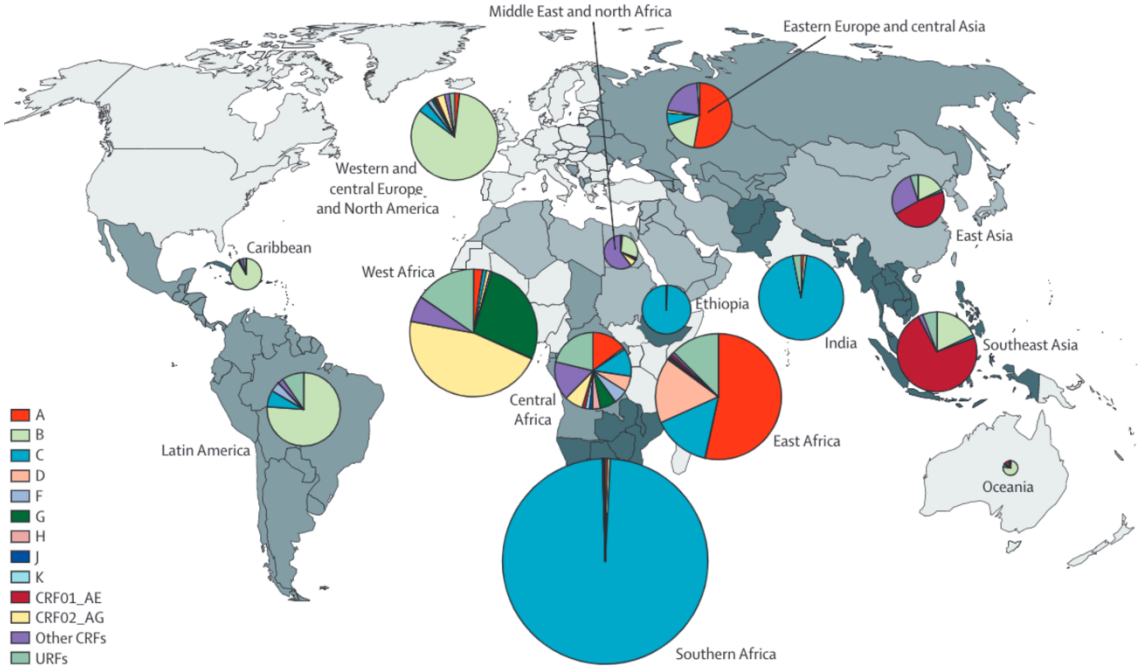

Figure 10. Global distribution of HIV-1 subtypes and CRFs.

(Adapted from Hemelaar *et al.* (2019) (102)).

Focusing on Europe region, the prevalence of HIV-1 subtypes differs from Western to Eastern regions, where in the first region subtype B is the most prevalent and in the second region subtype A is the most prevalent (104).

As it was mentioned above, HIV-1 has a fast evolutionary rate and for that reason the study of transmission patterns and clusters is important (105). One major strategy to understand transmission patterns and be able to characterize populations is phylogenetic analysis through transmission cluster (TC). HIV-1 TCs are defined as an nonrandomly

aggregation of HIV-1 sequences linked to their epidemiology (106). From TCs it is possible to understand the origins, evolution, interactions and geographical location of HIV-1 virus. It is important to combine epidemiologic and clinical data to the information given by TCs in order to not only have a better knowledge of the patterns and subgroups of the HIV-1 epidemic, but also to give information for better HIV prevention strategies (107). HIV-1 transmission is highly influenced by migration and globalization patterns and this can be shown in TCs, since individuals that share a most recent common ancestor (MRCA) might be linked epidemiologically (107,108).

As previously mentioned, the concern around HIV late presenters population is centred in their clinical presentation to healthcare and their ability to carry on HIV transmission by those who are not aware of their HIV status. Late presenters are a population understudied and might have a great impact on HIV transmission. Until now to our knowledge there are no studies regarding TCs in this population. This thesis proposes to explore the characteristic of this population referring to the determinants associated with their late presentation, to analyse drug resistance mutations that occur in this population, and finally understand their dynamics of HIV transmission. With this thesis we would like to expand the knowledge about this population and give inputs to public health experts for new preventive health strategies and guidance based in evidence.

1.11. Introduction References

1. Greene WC. A history of AIDS: Looking back to see ahead. Eur J Immunol [Internet]. 2007 Nov;37(S1):S94–102. Available from: https://onlinelibrary.wiley.com/doi/10.1002/eji.200737441
2. De Cock KM, Jaffe HW, Curran JW. The evolving epidemiology of HIV/AIDS. AIDS [Internet]. 2012 Jun 19;26(10):1205–13. Available from: https://journals.lww.com/00002030-201206190-00009
3. Montagnier L, Chermann JC, Barré-Sinoussi F, Klatzmann D, Wain-Hobson S, Alizon M, et al. Lymphadenopathy associated virus and its etiological role in AIDS. Princess Takamatsu Symp [Internet]. 1984;15:319—331. Available from: http://europepmc.org/abstract/MED/6100650
4. Global HIV & AIDS statistics — Fact sheet | UNAIDS [Internet]. [cited 2022 Mar 14]. Available from: https://www.unaids.org/en/resources/fact-sheet
5. AIDSinfo | UNAIDS [Internet]. [cited 2022 Mar 23]. Available from: https://aidsinfo.unaids.org/
6. 90-90-90 treatment target | UNAIDS [Internet]. [cited 2022 Mar 17]. Available from: https://www.unaids.org/en/90-90-90
7. Heath K, Levi J, Hill A. The Joint United Nations Programme on HIV/AIDS 95–95–95 targets: worldwide clinical and cost benefits of generic manufacture. AIDS [Internet]. 2021 Dec 15;35(Supplement 2):S197–203. Available from: https://journals.lww.com/10.1097/QAD.0000000000002983
8. Barré-Sinoussi F, Ross AL, Delfraissy JF. Past, present and future: 30 years of HIV research. Nat Rev Microbiol [Internet]. 2013;11(12):877–83. Available from: http://dx.doi.org/10.1038/nrmicro3132
9. Requejo HIZ. Worldwide molecular epidemiology of HIV. Rev Saude Publica [Internet]. 2006 Apr;40(2):331–45. Available from: http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0034-89102006000200023&lng=en&tlng=en
10. Tebit DM, Arts EJ. Tracking a century of global expansion and evolution of HIV to drive understanding and to combat disease. Lancet Infect Dis [Internet]. 2011 Jan;11(1):45–56. Available from: http://dx.doi.org/10.1016/S1473-3099(10)70186-9
11. Korber B, Muldoon M, Theiler J, Gao F, Gupta R, Lapedes A, et al. Timing the Ancestor of the HIV-1 Pandemic Strains. Science (80-) [Internet]. 2000 Jun 9;288(5472):1789–96. Available from: https://www.science.org/doi/10.1126/science.288.5472.1789
12. Tang R, Yu Z, Ma Y, Wu Y, Phoebe Chen Y-P, Wong L, et al. Genetic source completeness of HIV-1 circulating recombinant forms (CRFs) predicted by multi-label learning. Alfonso V, editor. Bioinformatics [Internet]. 2021 May 5;37(6):750–8. Available from: https://academic.oup.com/bioinformatics/article/37/6/750/5924547
13. Mendes Da Silva RK, Monteiro de Pina Araujo II, Venegas Maciera K, Gonçalves Morgado M, Lindenmeyer Guimarães M. Genetic Characterization of a New HIV-1 Sub-Subtype A in Cabo Verde, Denominated A8. Viruses [Internet]. 2021 Jun 8;13(6):1093. Available from: https://www.mdpi.com/1999-4915/13/6/1093
14. HIV Circulating Recombinant Forms (CRFs) [Internet]. [cited 2022 Jun 23]. Available from: https://www.hiv.lanl.gov/content/sequence/HIV/CRFs/CRFs.html
15. Burke D. Recombination in HIV: An Important Viral Evolutionary Strategy. Emerg Infect Dis [Internet]. 1997 Sep;3(3):253–9. Available from:

http://www.cdc.gov/ncidod/eid/vol3no3/burke.htm
16. Cañada JE, Delgado E, Gil H, Sánchez M, Benito S, García-Bodas E, et al. Identification of a New HIV-1 BC Intersubtype Circulating Recombinant Form (CRF108_BC) in Spain. Viruses [Internet]. 2021 Jan 12;13(1):93. Available from: https://www.mdpi.com/1999-4915/13/1/93
17. Faria NR, Rambaut A, Suchard MA, Baele G, Bedford T, Ward MJ, et al. The early spread and epidemic ignition of HIV-1 in human populations. Science (80-) [Internet]. 2014 Oct 3;346(6205):56–61. Available from: https://www.science.org/doi/10.1126/science.1256739
18. Hemelaar J. The origin and diversity of the HIV-1 pandemic. Trends Mol Med [Internet]. 2012;18(3):182–92. Available from: http://dx.doi.org/10.1016/j.molmed.2011.12.001
19. Sharp PM, Hahn BH. Origins of HIV and the AIDS Pandemic. Cold Spring Harb Perspect Med [Internet]. 2011 Sep 1;1(1):a006841–a006841. Available from: http://perspectivesinmedicine.cshlp.org/lookup/doi/10.1101/cshperspect.a006841
20. Thomson MM, Pérez-Álvarez L, Nájera R. Molecular epidemiology of HIV-1 genetic forms and its significance for vaccine development and therapy. Lancet Infect Dis [Internet]. 2002 Aug;2(8):461–71. Available from: https://linkinghub.elsevier.com/retrieve/pii/S1473309902003432
21. Turner BG, Summers MF. Structural biology of HIV 1 1Edited by P. E. Wright. J Mol Biol [Internet]. 1999 Jan 1;285(1):1–32. Available from: https://portlandpress.com/biochemsoctrans/article/30/6/1001/63827/Structural-biology-of-Cl
22. Frankel AD, Young JAT. HIV-1: Fifteen Proteins and an RNA. Annu Rev Biochem [Internet]. 1998 Jun;67(1):1–25. Available from: https://www.annualreviews.org/doi/10.1146/annurev.biochem.67.1.1
23. Kirchhoff F. HIV Life Cycle: Overview. In: Hope TJ, Stevenson M, Richman D, editors. Encyclopedia of AIDS [Internet]. New York, NY: Springer New York; 2013. p. 1–9. Available from: http://link.springer.com/10.1007/978-1-4614-9610-6
24. Discovery of Key Component of HIV Virus Yields Drug Target - Drug Discovery and Development [Internet]. [cited 2022 Mar 23]. Available from: https://www.drugdiscoverytrends.com/discovery-of-key-component-of-hiv-virus-yields-drug-target/
25. Malim MH, Emerman M. HIV-1 Accessory Proteins—Ensuring Viral Survival in a Hostile Environment. Cell Host Microbe [Internet]. 2008 Jun;3(6):388–98. Available from: https://linkinghub.elsevier.com/retrieve/pii/S1931312808001261
26. Reed-Inderbitzin E, Maury W. Cellular specificity of HIV-1 replication can be controlled by LTR sequences. Virology [Internet]. 2003 Sep;314(2):680–95. Available from: https://linkinghub.elsevier.com/retrieve/pii/S0042682203005087
27. Freed EO. HIV-1 Replication. Somat Cell Mol Genet [Internet]. 2001;26(1):13–33. Available from: https://doi.org/10.1023/A:1021070512287
28. Ganser-Pornillos BK, Yeager M, Sundquist WI. The structural biology of HIV assembly. Curr Opin Struct Biol [Internet]. 2008 Apr;18(2):203–17. Available from: http://www.smltsa.org.za/journal/archive/vol21no13.pdf
29. Jacks T, Power MD, Masiarz FR, Luciw PA, Barr PJ, Varmus HE. Characterization of ribosomal frameshifting in HIV-1 gag-pol expression. Nature [Internet]. 1988 Jan;331(6153):280–3. Available from: https://www.infodesign.org.br/infodesign/article/view/355%0Ahttp://www.abergo.org.b

r/revista/index.php/ae/article/view/731%0Ahttp://www.abergo.org.br/revista/index.php/ae/article/view/269%0Ahttp://www.abergo.org.br/revista/index.php/ae/article/view/106
30. Engelman A, Cherepanov P. The structural biology of HIV-1: mechanistic and therapeutic insights. Nat Rev Microbiol [Internet]. 2012 Apr 16;10(4):279–90. Available from: http://www.nature.com/articles/nrmicro2747
31. Rice AP. The HIV-1 Tat Protein: Mechanism of Action and Target for HIV-1 Cure Strategies. Curr Pharm Des [Internet]. 2017 Nov 2;23(28):4098–102. Available from: http://www.eurekaselect.com/153740/article
32. Zaitseva M, Blauvelt A, Lee S, Lapham CK, Kiaus-Kovrun V, Mostowski H, et al. Expression and function of CCR5 and CXCR4 on human Langerhans cells and macrophages: Implications for HIV primary infection. Nat Med [Internet]. 1997;3(12):1369–75. Available from: https://doi.org/10.1038/nm1297-1369
33. Wyatt R, Sodroski J. The HIV-1 Envelope Glycoproteins: Fusogens, Antigens, and Immunogens. Science (80-) [Internet]. 1998 Jun 19;280(5371):1884–8. Available from: https://www.science.org/doi/10.1126/science.280.5371.1884
34. Isel C, Ehresmann C, Marquet R. Initiation of HIV Reverse Transcription. Viruses [Internet]. 2010 Jan 18;2(1):213–43. Available from: http://www.mdpi.com/1999-4915/2/1/213
35. Blumenthal R, Durell S, Viard M. HIV Entry and Envelope Glycoprotein-mediated Fusion. J Biol Chem [Internet]. 2012 Nov;287(49):40841–9. Available from: https://linkinghub.elsevier.com/retrieve/pii/S0021925820438876
36. Wang WK, Chen MY, Chuang CY, Jeang KT, Huang LM. Molecular biology of human immunodeficiency virus type 1. J Microbiol Immunol Infect [Internet]. 2000 Sep;33(3):131–40. Available from: http://www.ncbi.nlm.nih.gov/pubmed/11045374
37. HIV Replication Cycle | NIH: National Institute of Allergy and Infectious Diseases [Internet]. [cited 2022 Mar 23]. Available from: https://www.niaid.nih.gov/diseases-conditions/hiv-replication-cycle
38. Moir S, Chun T-W, Fauci AS. Pathogenic Mechanisms of HIV Disease. Annu Rev Pathol Mech Dis [Internet]. 2011 Feb 28;6(1):223–48. Available from: https://www.annualreviews.org/doi/10.1146/annurev-pathol-011110-130254
39. Galvin SR, Cohen MS. The role of sexually transmitted diseases in HIV transmission. Nat Rev Microbiol [Internet]. 2004 Jan;2(1):33–42. Available from: http://www.nature.com/articles/nrmicro794
40. Fettig J, Swaminathan M, Murrill CS, Kaplan JE. Global Epidemiology of HIV. Infect Dis Clin North Am [Internet]. 2014 Sep;28(3):323–37. Available from: https://linkinghub.elsevier.com/retrieve/pii/S0891552014000300
41. van de Vijver DAMC, Boucher CAB. Insights on transmission of HIV from phylogenetic analysis to locally optimize HIV prevention strategies. Curr Opin HIV AIDS [Internet]. 2018 Mar;13(2):95–101. Available from: https://journals.lww.com/01222929-201803000-00002
42. Paraschiv S, Banica L, Nicolae I, Niculescu I, Abagiu A, Jipa R, et al. Epidemic dispersion of HIV and HCV in a population of co-infected Romanian injecting drug users. Meng Z, editor. PLoS One [Internet]. 2017 Oct 9;12(10):e0185866. Available from: https://dx.plos.org/10.1371/journal.pone.0185866
43. Enger C, Graham N, Peng Y, Chmiel JS, Kingsley LA, Detels R, et al. Survival From Early, Intermediate, and Late Stages of HIV Infection. JAMA [Internet]. 1996 May 1;275(17):1329–34. Available from: https://doi.org/10.1001/jama.1996.03530410043031

44. Global HIV & AIDS statistics — 2020 fact sheet | UNAIDS [Internet]. [cited 2021 Jan 4]. Available from: https://www.unaids.org/en/resources/fact-sheet
45. Symptoms of HIV | HIV.gov [Internet]. [cited 2021 Jun 30]. Available from: https://www.hiv.gov/hiv-basics/overview/about-hiv-and-aids/symptoms-of-hiv
46. Kahn JO, Walker BD. Acute Human Immunodeficiency Virus Type 1 Infection. N Engl J Med [Internet]. 1998 Jul 2;339(1):33–9. Available from: http://www.nejm.org/doi/abs/10.1056/NEJM199807023390107
47. Miranda MNS, Pingarilho M, Pimentel V, Martins M do RO, Vandamme A-M, Bobkova M, et al. Determinants of HIV-1 Late Presentation in Patients Followed in Europe. Pathogens [Internet]. 2021;10(7). Available from: https://www.mdpi.com/2076-0817/10/7/835
48. About HIV/AIDS | HIV Basics | HIV/AIDS | CDC [Internet]. [cited 2021 Jun 30]. Available from: https://www.cdc.gov/hiv/basics/whatishiv.html
49. Guaraldi G, Zona S, Menozzi M, Brothers TD, Carli F, Stentarelli C, et al. Late presentation increases risk and costs of non-infectious comorbidities in people with HIV: an Italian cost impact study. AIDS Res Ther [Internet]. 2017 Dec 16;14(1):8. Available from: http://aidsrestherapy.biomedcentral.com/articles/10.1186/s12981-016-0129-4
50. Antinori A, Coenen T, Costagiola D, Dedes N, Ellefson M, Gatell J, et al. Late presentation of HIV infection: a consensus definition. HIV Med [Internet]. 2011 Jan;12(1):61–4. Available from: https://onlinelibrary.wiley.com/doi/10.1111/j.1468-1293.2010.00857.x
51. Gesesew HA, Ward P, Woldemichael K, Mwanri L. Late presentation for HIV care in Southwest Ethiopia in 2003–2015: prevalence, trend, outcomes and risk factors. BMC Infect Dis [Internet]. 2018 Dec 30;18(1):59. Available from: https://bmcinfectdis.biomedcentral.com/articles/10.1186/s12879-018-2971-6
52. Krentz HB, Gill MJ. The Direct Medical Costs ofLate Presentation (<350/mm3)of HIV Infection over a 15-Year Period. AIDS Res Treat [Internet]. 2012;2012:1–8. Available from: http://www.hindawi.com/journals/art/2012/757135/
53. Moreno S, Mocroft A, Monforte A d'Arminio. Medical and Societal Consequences of Late Presentation. Antivir Ther [Internet]. 2010 Jan 1;15(1_suppl):9–15. Available from: http://journals.sagepub.com/doi/10.3851/IMP1523
54. Conway AS, Esteve A, Fernández-Quevedo M, Casabona J. Determinants and Outcomes of Late Presentation of HIV Infection in Migrants in Catalonia, Spain: PISCIS Cohort 2004–2016. J Immigr Minor Heal [Internet]. 2019 Oct 30;21(5):920–30. Available from: http://dx.doi.org/10.1007/s10903-018-0834-2
55. Socio-economic Inequalities and HIV Writing Group for Collaboration of Observational HIV Epidemiological Research in Europe (COHERE) in EuroCoord, Lodi S, Dray-Spira R, Touloumi G, Braun D, Teira R, et al. Delayed HIV diagnosis and initiation of antiretroviral therapy: inequalities by educational level, COHERE in EuroCoord. AIDS [Internet]. 2014 Sep 24;28(15):2297–306. Available from: https://journals.lww.com/00002030-201409240-00013
56. Wójcik-cichy K, Jabłonowska O, Piekarska A. The high incidence of late presenters for HIV / AIDS infection in the Lodz province , Poland in the years 2009 – 2016 : we are still far from the UNAIDS 90 % target. AIDS Care [Internet]. 2018;0(0):1–4. Available from: https://doi.org/10.1080/09540121.2018.1470306
57. Wilton J, Light L, Gardner S, Rachlis B, Conway T, Cooper C, et al. Late diagnosis, delayed presentation and late presentation among persons enrolled in a clinical HIV cohort in Ontario, Canada (1999-2013). HIV Med [Internet]. 2019 Feb;20(2):110–

20. Available from: https://onlinelibrary.wiley.com/doi/10.1111/hiv.12686
58. Hu X, Liang B, Zhou C, Jiang J, Huang J, Ning C, et al. HIV late presentation and advanced HIV disease among patients with newly diagnosed HIV/AIDS in Southwestern China: a large-scale cross-sectional study. AIDS Res Ther [Internet]. 2019 Dec 16;16(1):6. Available from: https://aidsrestherapy.biomedcentral.com/articles/10.1186/s12981-019-0221-7
59. Luma HN, Jua P, Donfack O-T, Kamdem F, Ngouadjeu E, Mbatchou HB, et al. Late presentation to HIV/AIDS care at the Douala general hospital, Cameroon: its associated factors, and consequences. BMC Infect Dis [Internet]. 2018 Dec 3;18(1):298. Available from: https://bmcinfectdis.biomedcentral.com/articles/10.1186/s12879-018-3204-8
60. Diaz RS, Inocêncio LA, Sucupira MCA, Pereira AA, Hunter J, Ferreira JE, et al. The Virological and Immunological Characteristics of the HIV-1-Infected Population in Brazil: From Initial Diagnosis to Impact of Antiretroviral Use. López-Galíndez C, editor. PLoS One [Internet]. 2015 Oct 28;10(10):e0139677. Available from: https://dx.plos.org/10.1371/journal.pone.0139677
61. Yeni P. Update on HAART in HIV. J Hepatol [Internet]. 2006 Jan;44(SUPPL. 1):S100–3. Available from: https://linkinghub.elsevier.com/retrieve/pii/S0168827805007464
62. Pau AK, George JM. Antiretroviral Therapy: Current Drugs. Zeichner SL, Read JS, editors. Infect Dis Clin North Am [Internet]. 2014 Sep;28(3):371–402. Available from: https://www.cambridge.org/core/product/identifier/CBO9780511544781A023/type/book_part
63. Arts EJ, Hazuda DJ. HIV-1 Antiretroviral Drug Therapy. Cold Spring Harb Perspect Med [Internet]. 2012 Apr 1;2(4):a007161–a007161. Available from: http://perspectivesinmedicine.cshlp.org/lookup/doi/10.1101/cshperspect.a007161
64. Maeda K, Das D, Kobayakawa T, Tamamura H, Takeuchi H. Discovery and Development of Anti-HIV Therapeutic Agents: Progress Towards Improved HIV Medication. Curr Top Med Chem [Internet]. 2019 Oct 9;19(18):1621–49. Available from: http://www.eurekaselect.com/173455/article
65. Vella S, Schwartländer B, Sow SP, Eholie SP, Murphy RL. The history of antiretroviral therapy and of its implementation in resource-limited areas of the world. AIDS [Internet]. 2012 Jun 19;26(10):1231–41. Available from: https://journals.lww.com/00002030-201206190-00012
66. De Clercq E. The history of antiretrovirals: key discoveries over the past 25 years. Rev Med Virol [Internet]. 2009 Sep;19(5):287–99. Available from: http://doi.wiley.com/10.1002/rmv.624
67. Lu D-Y, Wu H-Y, Yarla NS, Xu B, Ding J, Lu T-R. HAART in HIV/AIDS Treatments: Future Trends. Infect Disord - Drug Targets [Internet]. 2018 Mar 20;18(1):15–22. Available from: http://www.eurekaselect.com/152220/article
68. Updated recommendations on first-line and second-line antiretroviral regimens and post-exposure prophylaxis and recommendations on early infant diagnosis of HIV - Interim guidance. Geneva World Heal Organ 2018 [Internet]. :Licence: CC BY-NC-SA 3.0 IGO. Available from: http://apps.who.int/bookorders.%0Ahttps://www.who.int/hiv/pub/guidelines/ARV2018update/en/
69. Simoni JM, Tapia K, Lee S-J, Graham SM, Beima-Sofie K, Mohamed ZH, et al.

A Conjoint Analysis of the Acceptability of Targeted Long-Acting Injectable Antiretroviral Therapy Among Persons Living with HIV in the U.S. AIDS Behav [Internet]. 2020 Apr 26;24(4):1226–36. Available from: https://doi.org/10.1007/s10461-019-02701-7
70. Theys K, Libin PJK, Van Laethem K, Abecasis AB. An Evolutionary Model-Based Approach To Quantify the Genetic Barrier to Drug Resistance in Fast-Evolving Viruses and Its Application to HIV-1 Subtypes and Integrase Inhibitors. Antimicrob Agents Chemother [Internet]. 2019 Aug;63(8). Available from: https://journals.asm.org/doi/10.1128/AAC.00539-19
71. Yotebieng M, Brazier E, Addison D, Kimmel AD, Cornell M, Keiser O, et al. Research priorities to inform "Treat All" policy implementation for people living with in sub-Saharan Africa: a consensus statement from the International epidemiology Databases to Evaluate AIDS (IeDEA). J Int AIDS Soc [Internet]. 2019 Jan 18;22(1):e25218. Available from: https://onlinelibrary.wiley.com/doi/10.1002/jia2.25218
72. Bavinton BR, Grulich AE. HIV pre-exposure prophylaxis: scaling up for impact now and in the future. Lancet Public Heal [Internet]. 2021 Jul;6(7):e528–33. Available from: http://dx.doi.org/10.1016/S2468-2667(21)00112-2
73. Seed CR, Styles CE, Hoad VC, Yang H, Thomas MJ, Gosbell IB. Effect of HIV pre-exposure prophylaxis (PrEP) on detection of early infection and its impact on the appropriate post-PrEP deferral period. Vox Sang [Internet]. 2021 Apr 23;116(4):379–87. Available from: https://onlinelibrary.wiley.com/doi/10.1111/vox.13011
74. World Health Organization. Guideline on When To Start Antiretroviral Therapy and on Pre-Exposure Prophylaxis for HIV. 2015;(September):1–76.
75. Schmidt D, Kollan C, Fätkenheuer G, Schülter E, Stellbrink H-J, Noah C, et al. Estimating Trends in the Proportion of Transmitted and Acquired HIV Drug Resistance in a Long Term Observational Cohort in Germany. Harrigan PR, editor. PLoS One [Internet]. 2014 Aug 22;9(8):e104474. Available from: https://dx.plos.org/10.1371/journal.pone.0104474
76. Kim J, Lee E, Park B-J, Bang JH, Lee JY. Adherence to antiretroviral therapy and factors affecting low medication adherence among incident HIV-infected individuals during 2009–2016: A nationwide study. Sci Rep [Internet]. 2018 Dec 16;8(1):3133. Available from: http://dx.doi.org/10.1038/s41598-018-21081-x
77. Pingarilho M, Pimentel V, Diogo I, Fernandes S, Miranda M, Pineda-Pena A, et al. Increasing Prevalence of HIV-1 Transmitted Drug Resistance in Portugal: Implications for First Line Treatment Recommendations. Viruses [Internet]. 2020 Oct 30;12(11):1238. Available from: https://www.mdpi.com/1999-4915/12/3/268
78. Baesi K, Ravanshad M, Ghanbarisafari M, Saberfar E, SeyedAlinaghi S, Volk JE. Antiretroviral drug resistance among antiretroviral-naïve and treatment experienced patients infected with HIV in Iran. J Med Virol [Internet]. 2014 Jul;86(7):1093–8. Available from: http://doi.wiley.com/10.1002/jmv.23898
79. Clutter DS, Jordan MR, Bertagnolio S, Shafer RW. HIV-1 drug resistance and resistance testing. Infect Genet Evol [Internet]. 2016 Dec;46(1):292–307. Available from: https://linkinghub.elsevier.com/retrieve/pii/S1567134816303690
80. Miranda MNS, Pingarilho M, Pimentel V, Martins M do RO, Kaiser R, Seguin-Devaux C, et al. Trends of Transmitted and Acquired Drug Resistance in Europe From 1981 to 2019: A Comparison Between the Populations of Late Presenters and Non-late Presenters. Front Microbiol [Internet]. 2022 Apr 13;13(April):1–12. Available from:

https://www.frontiersin.org/articles/10.3389/fmicb.2022.846943/full
81. Charles B. Hicks M. Antiretroviral Drug Resistance Testing — Updated Guidelines from the IAS–USA. NEJM J Watch [Internet]. 2008 Jul 14 [cited 2021 Mar 12];2008. Available from: https://www.jwatch.org/AC200807140000002/2008/07/14/antiretroviral-drug-resistance-testing-updated
82. Rossetti B, Di Giambenedetto S, Torti C, Postorino M, Punzi G, Saladini F, et al. Evolution of transmitted HIV-1 drug resistance and viral subtypes circulation in Italy from 2006 to 2016. HIV Med [Internet]. 2018 Oct;19(9):619–28. Available from: https://onlinelibrary.wiley.com/doi/10.1111/hiv.12640
83. Günthard HF, Calvez V, Paredes R, Pillay D, Shafer RW, Wensing AM, et al. Human Immunodeficiency Virus Drug Resistance: 2018 Recommendations of the International Antiviral Society–USA Panel. Clin Infect Dis [Internet]. 2019 Jan 7;68(2):177–87. Available from: https://academic.oup.com/cid/article/68/2/177/5055715
84. NRTI Resistance Notes - HIV Drug Resistance Database [Internet]. [cited 2022 Mar 24]. Available from: https://hivdb.stanford.edu/dr-summary/resistance-notes/NRTI/#thymidine.analog.mutations.tams.
85. NNRTI Resistance Notes - HIV Drug Resistance Database [Internet]. [cited 2022 Mar 24]. Available from: https://hivdb.stanford.edu/dr-summary/resistance-notes/NNRTI/
86. Zou X, He J, Zheng J, Malmgren R, Li W, Wei X, et al. Prevalence of acquired drug resistance mutations in antiretroviral- experiencing subjects from 2012 to 2017 in Hunan Province of central South China. Virol J [Internet]. 2020 Dec 17;17(1):38. Available from: https://virologyj.biomedcentral.com/articles/10.1186/s12985-020-01311-3
87. Clutter DS, Jordan MR, Bertagnolio S, Shafer RW. HIV-1 drug resistance and resistance testing. Infect Genet Evol [Internet]. 2016 Dec;46:292–307. Available from: http://dx.doi.org/10.1016/j.meegid.2016.08.031
88. Bennett DE, Camacho RJ, Otelea D, Kuritzkes DR, Fleury H, Kiuchi M, et al. Drug Resistance Mutations for Surveillance of Transmitted HIV-1 Drug-Resistance: 2009 Update. Nixon DF, editor. PLoS One [Internet]. 2009 Mar 6;4(3):e4724. Available from: https://dx.plos.org/10.1371/journal.pone.0004724
89. WHO SDRM List - HIV Drug Resistance Database [Internet]. [cited 2022 Mar 24]. Available from: https://hivdb.stanford.edu/page/who-sdrm-list/
90. Release Notes - HIV Drug Resistance Database [Internet]. [cited 2022 Mar 30]. Available from: https://hivdb.stanford.edu/page/release-notes/
91. NRTI Resistance Mutation Scores - HIV Drug Resistance Database [Internet]. [cited 2022 Mar 30]. Available from: https://hivdb.stanford.edu/dr-summary/mut-scores/NRTI/
92. NNRTI Resistance Mutation Scores - HIV Drug Resistance Database [Internet]. [cited 2022 Mar 30]. Available from: https://hivdb.stanford.edu/dr-summary/mut-scores/NNRTI/
93. PI Resistance Mutation Scores - HIV Drug Resistance Database [Internet]. [cited 2022 Mar 30]. Available from: https://hivdb.stanford.edu/dr-summary/mut-scores/PI/
94. Drug-Resistance Testing | NIH [Internet]. [cited 2021 Aug 3]. Available from: https://clinicalinfo.hiv.gov/en/guidelines/adult-and-adolescent-arv/drug-resistance-testing
95. Theys K, Libin P, Pineda-Peña A-C, Nowé A, Vandamme A-M, Abecasis AB.

The impact of HIV-1 within-host evolution on transmission dynamics. Curr Opin Virol [Internet]. 2018 Feb;28:92–101. Available from: https://linkinghub.elsevier.com/retrieve/pii/S1879625717301372
96. van Harmelen J, Wood R, Lambrick M, Rybicki EP, Williamson A-L, Williamson C. An association between HIV-1 subtypes and mode of transmission in Cape Town, South Africa. AIDS [Internet]. 1997 Jan;11(1):81–7. Available from: https://journals.lww.com/aidsonline/Fulltext/1997/01000/An_association_between_HIV_1_subtypes_and_mode_of.12.aspx
97. Pineda-Peña AC, Faria NR, Imbrechts S, Libin P, Abecasis AB, Deforche K, et al. Automated subtyping of HIV-1 genetic sequences for clinical and surveillance purposes: Performance evaluation of the new REGA version 3 and seven other tools. Infect Genet Evol [Internet]. 2013;19(100):337–48. Available from: http://dx.doi.org/10.1016/j.meegid.2013.04.032
98. Alcantara LCJ, Cassol S, Libin P, Deforche K, Pybus OG, Van Ranst M, et al. A standardized framework for accurate, high-throughput genotyping of recombinant and non-recombinant viral sequences. Nucleic Acids Res [Internet]. 2009 Jul 1;37(Web Server):W634–42. Available from: https://academic.oup.com/nar/article-lookup/doi/10.1093/nar/gkp455
99. Ngcapu S, Theys K, Libin P, Marconi V, Sunpath H, Ndung'u T, et al. Characterization of Nucleoside Reverse Transcriptase Inhibitor-Associated Mutations in the RNase H Region of HIV-1 Subtype C Infected Individuals. Viruses [Internet]. 2017 Nov 8;9(11):330. Available from: http://www.mdpi.com/1999-4915/9/11/330
100. Doat A, Negarandeh R, Hasanpour M. Disclosure of HIV Status to Children in Sub-Saharan Africa: A Systematic Review. Medicina (B Aires) [Internet]. 2019 Aug 2;55(8):433. Available from: https://www.mdpi.com/1648-9144/55/8/433
101. Cohen MS, Hellmann N, Levy JA, DeCock K, Lange J. The spread, treatment, and prevention of HIV-1: evolution of a global pandemic. J Clin Invest [Internet]. 2008 Apr 1;118(4):1244–54. Available from: http://www.jci.org/articles/view/34706
102. Hemelaar J, Elangovan R, Yun J, Dickson-Tetteh L, Fleminger I, Kirtley S, et al. Global and regional molecular epidemiology of HIV-1, 1990–2015: a systematic review, global survey, and trend analysis. Lancet Infect Dis [Internet]. 2019 Feb;19(2):143–55. Available from: https://linkinghub.elsevier.com/retrieve/pii/S1473309918306479
103. Vandamme A-M, Camacho RJ, Ceccherini-Silberstein F, de Luca A, Palmisano L, Paraskevis D, et al. European recommendations for the clinical use of HIV drug resistance testing: 2011 update. AIDS Rev [Internet]. 2011;13(2):77–108. Available from: http://www.ncbi.nlm.nih.gov/pubmed/21587341
104. Bbosa N, Kaleebu P, Ssemwanga D. HIV subtype diversity worldwide. Curr Opin HIV AIDS [Internet]. 2019 May;14(3):153–60. Available from: https://journals.lww.com/01222929-201905000-00003
105. Delgado E, Benito S, Montero V, Cuevas MT, Fernández-García A, Sánchez-Martínez M, et al. Diverse Large HIV-1 Non-subtype B Clusters Are Spreading Among Men Who Have Sex With Men in Spain. Front Microbiol [Internet]. 2019 Apr 3;10(APR):1–19. Available from: https://www.frontiersin.org/article/10.3389/fmicb.2019.00655/full
106. Abecasis AB, Pingarilho M, Vandamme A-M. Phylogenetic analysis as a forensic tool in HIV transmission investigations. AIDS [Internet]. 2018 Mar 13;32(5):543–54. Available from: https://journals.lww.com/00002030-201803130-00002
107. Hassan AS, Pybus OG, Sanders EJ, Albert J, Esbjörnsson J. Defining HIV-1

transmission clusters based on sequence data. AIDS [Internet]. 2017 Jun 1;31(9):1211–22. Available from: https://journals.lww.com/00002030-201706010-00003
108. Jovanović L, Šiljić M, Ćirković V, Salemović D, Pešić-Pavlović I, Todorović M, et al. Exploring Evolutionary and Transmission Dynamics of HIV Epidemic in Serbia: Bridging Socio-Demographic With Phylogenetic Approach. Front Microbiol [Internet]. 2019 Feb 25;10(February):1–15. Available from: https://www.frontiersin.org/article/10.3389/fmicb.2019.00287/full

2. Aims

The urge to study late presenters population came after a study made in Portugal about the same topic, where the proportion of late presenters was 50.6%, more than half the population (Manuscript I). This population could be accountable for spreading HIV-1, having a more difficult adherence to ART and higher health costs. For this reason, it is important to study the determinants associated with this population, to understand the patterns of resistance to ARV drugs and to construct transmission clusters in order to understand the transmission dynamics of HIV-1 among this specific population.

To be able to attain these characterization and analysis we set specific objectives:

1. To analyse the molecular epidemiology and identify clinical and sociodemographic characteristics of HIV-1 infected patients in Europe (Manuscript II).
2. To analyse the patterns of Transmitted and Acquired Drug Resistance and the most prevalent DRMs and subtypes among LP and NLP in Europe (Manuscript III).
3. To characterize HIV-1 transmission clusters and identify risk factors for HIV infection and vulnerable groups for late presentation (Manuscript IV)

3. Methods

3.1. Study Group (Manuscript I)

The database included clinical and sociodemographic information from HIV-1 infected patients followed in Hospital Egas Moniz, which is part of Centro Hospitalar de Lisboa Ocidental (CHLO), Lisbon, Portugal, and was collected during routine clinical care between 1984 and 2017.

3.2. Main Study Group (Manuscript II, III, IV)

The database we used to conduct our study was the EuResist Integrated Database (EIDB). This database has clinical, sociodemographic and genomic information from HIV-1 infected patients between 1981 and 2019. The EIDB is one of the largest existing datasets which integrate clinical, socio-demographic and viral genotypic information from HIV-1 patients. It integrates longitudinal, periodically updated data mainly from Italy (ARCA database), Germany (AREVIR database) Spain (CoRIS and IRSICAIXA), Sweden, Belgium, Portugal and Luxembourg. In our studies, information from the ARCA, AREVIR, Luxembourg, IRSICAIXA, Portugal, United Kingdom, Russia and CoRIS databases were used.

3.3. Study Variables (Manuscript II, III, IV)

For our study, we used information from the EuResist database to create new variables such as:

- **Migrant/Native**- Based on Country of Origin and Country of Follow-up (if country of origin and country of follow-up is the same, then patient is native; if country of origin and country of follow-up is not the same, then patient is migrant)
- **Age at Diagnosis**- Based on the difference between Year of Birth and Date of the first HIV Positive test;
- **Age at Drug Resistance test**- Based on the difference between Year of Birth and Date of the first drug resistance test;
- **Region of Origin**- Based on Country of Origin;

- **Treatment Status at date of first CD4 count**- Based on the difference between sample collection date of first CD4 count and first therapy date; for purposes of classification of Late Presentation, only patients naïve at date of first CD4 count were considered;
- **Treatment Status at date of first Drug Resistance Test** – based on the difference between sample collection date for first drug resistance test and date of start of first therapy:

  ART-naïve → patients who had a sample collection date for first drug resistance test before the date of start of first therapy

  ART-experienced → patients who had a sample collection date for first drug resistance test after the date of start of first therapy
- **Recentness of infection**- Based on ambiguity rate of genomic sequences. We defined Chronic infection as an ambiguity value higher than 0.45% and Recent infection as an ambiguity value equal or below 0.45%. Additionally, only genomic sequences larger than 500 nucleotides and with ambiguity rate lower than 2.5% were considered.
- **LP vs. NLP**- Based on CD4 count at diagnosis, LP were defined as patients with CD4 count lower than 350 cells/mm^3 and NLP were defined as patients with CD4 count higher than 350 cells/mm^3.

4. Results

4.1. Manuscript I

AIDS RESEARCH AND HUMAN RETROVIRUSES
Volume 37, Number 11, 2021

DOI: 10.1089/aid.2020.0175

Determinants of HIV-1 Late Presentation in a Cohort of Portuguese HIV-1 Patients

Ana Cláudia Miranda,[1,*] Mafalda Miranda,[2,*] Marta Pingarilho,[2] Victor Pimentel,[2] João Torres,[1] Susana Peres,[1] Teresa Baptista Alberto,[1] Perpetua Gomes,[3,4] Ana Abecasis,[2] and Kamal Mansinho[1]

Abstract

Undiagnosed HIV-1 patients still account for 25% of worldwide HIV patients. Studying late presenters (LPs) for HIV care may help to identify characteristics of such patients. The present study aims to identify factors associated with late presentation and late presentation with advanced disease based on a population of patients followed in a Portuguese hospital between 1984 and 2017. Sociodemographic and clinical data from infected patients with HIV-1 aged 18 years and older, followed in Egas Moniz Hospital, in Portugal were collected. Of the 907 patients included in this study, 68.7% were males and the median age was 37 years (interquartile range 30–47). Four hundred fifty-nine patients (50.6%) were LP and, of these, 284 patients (61.9%) were LPAD. The LP population mostly originated from Portugal and sub-Saharan Africa (64.4% and 28.8%; $p = .004$) and the HIV exposure category, mainly heterosexuals and men have sex with men (57.0% and 24.9%; $p < .001$). The stage of disease and viral load at diagnosis were significantly associated with both LP and LPAD ($p < .001$). Factors associated with LP in the logistic regression included age at diagnosis lower than 30 years (adjusted odds ratio [aOR] 0.34; 0.17–0.68; $p = .002$) and origin from sub-Saharan Africa (aOR 2.24; 1.44–3.50; $p < .001$). Late presentation is a major obstacle to halt the HIV epidemic. In this population, the majority of newly diagnosed HIV-infected individuals were LPs. Our results characterize vulnerable populations that should be frequently tested for HIV.

Keywords: HIV-1 infection, late presentation, late presentation with advanced disease

Introduction

HIV continues to be one of the main public health issues. In 2018, there were 1.7 million people newly infected worldwide and 973 new cases were reported in Portugal.[1,2] Early diagnosis is vital to achieve the objectives proposed by the WHO: the 95-95-95 target to end the pandemic by 2030: diagnosing 95% of people living with HIV, 95% of diagnosed on treatment, and 95% of people on treatment viral suppressed.[3] However, people living with HIV, who do not know their status, account for 25% of the total infected people worldwide (9.4 million people).[4] In the European Union, it is estimated that late presenters (LPs) represent around 49%–54% of cases and late presenters with advanced disease (LPAD) are around 33%–42% of HIV cases.[5] According to the last Portuguese report, in 2018, LP cases accounted for 55.8% and LPAD cases accounted for 34.3% of HIV infection.[2] Importantly, the proportion of LP cases among newly diagnosed is increasing, indicating that we are leaving some older cases of undiagnosed patients behind.

According to the European Late Presenter Consensus working group, LPs were defined as presenting a $TCD4^+$ count lower than 350 cells/mm^3 or an AIDS-defining event, regardless of $TCD4^+$ cell count.[6] A subgroup of LPs, called LPAD, is characterized by presenting a $TCD4^+$ count lower than 200 cells/mm^3 or an AIDS-defining event, regardless of $TCD4^+$ cell count. This latest subgroup particularly is at

[1]Department of Infectious Diseases, Hospital Center Lisboa Ocidental, Egas Moniz Hospital, Lisbon, Portugal.
[2]Global Health and Tropical Medicine (GHTM), Institute of Hygiene and Tropical Medicine/New University of Lisbon (IHMT/UNL), Lisbon, Portugal.
[3]Laboratory of Molecular Biology (LMCBM, SPC, CHLO-HEM), Lisbon, Portugal.
[4]Egas Moniz Interdisciplinary Research Center (CiiEM), Egas Moniz University Institute, Caparica, Portugal.
*These authors contributed equally for this work and are considered to be co-first authors.
A preliminary version of this article has been published as a preprint version in medRxiv (https://www.medrxiv.org/content/10.1101/2020.06.17.20133785v1) in June 20, 2020.

greater risk of severe disease and death.[7,8] The current guidelines for LP and LPAD patients suggest that they are unable to fully benefit from antiretroviral therapy, leading to poorer outcomes in treatment. This late entry to care can not only have an impact on an individual's morbidity and mortality but also can increase the risk of onward transmission due to unawareness of their HIV status, with an impact on the control of the pandemic.[9–11]

The present study has the objective of identifying determinants of late presentation and late presentation with advanced disease. To do this, we analyzed a population of patients followed-up in a Portuguese hospital, diagnosed between 1984 and 2017.

Methods

Study group

Clinical and sociodemographic information from 907 HIV-1-infected patients was collected during routine clinical care of patients followed in Hospital Egas Moniz, which is part of Centro Hospitalar de Lisboa Ocidental (CHLO), Lisbon, Portugal, between 1984 and 2017. Patients were aged 18 years and older.

Statistical analysis

A descriptive analysis was conducted. The proportion and median (interquartile range [IQR]) of LP, non-late presenters (NLP), and LPAD were calculated for every categorical and continuous variable, respectively. Our interest variables were compared with the categorical variables with chi-square test and continuous variables with t-test for independent samples.

To study the association between our dependent variable and the independent variables, logistic regression models were calculated. We first presented the logistic regression with the unadjusted odds ratios and confidence intervals at 95% (95% CI), variables with a p value $<.05$ were considered to enter the model. We calculated the logistic regression model with the stepwise mode, which lead to the construction of the final model. The final model for LP versus NLP was adjusted for gender, this variable was forced into the model regardless of its significance and the reference class was women, and the final model for LP versus LPAD was adjusted for gender and age at diagnosis. Results were considered statistically significant when $p<.05$. The odds ratio and 95% CIs were calculated for the variables of the final model. Data were analyzed using SPSS for Windows (Version 23.0).

Ethics

The protocol was in accordance with the Declaration of Helsinki and approved by the Ethics Committee of Centro Hospitalar de Lisboa Ocidental (108/CES-2014). This database contains anonymized patients' information, including demographic and clinical data from patients followed in Hospital Egas Moniz between 1984 and 2017.

Results

Among 907 HIV-1-infected patients included in the analysis, the median age was 37 (IQR: 30.0–47.0) years and 68.7% were males (Table 1). Four hundred fifty-nine patients (50.6%) were LP and, of these, 284 patients (61.9%) were LPAD. Heterosexuals and Portuguese originated patients presented a higher proportion in this study population, 52.3% and 67.7%, respectively. CD4 count at diagnosis and viral load at diagnosis ($\log_{10}$) presented a median of 342 cells/mm^3 (IQR 155–554) and 4.8 copies/mL (IQR 4.2–5.4), respectively. We also performed an analysis for CD4 count over time for LP and NLP (Supplementary Fig. S1) and we observed that there was no variation for either LP or NLP in CD4 counts at diagnosis over time. Patients at stage A of HIV infection accounted for 66.8% of the population of patients, when compared to stages B and C.

Characteristics of patients stratified according to time of presentation are presented in Table 1. In this sample, males accounted for the bigger proportion of LP (69.7%); 71.1% of those being LPAD. No statistical differences according to gender were found for LP and LPAD populations. The median age for LP was 39.5 (IQR 32–50; $p<.001$) and the age group between 31 and 55 years ($p<.001$) represented 64.2% of LP, significantly higher when compared with the other age groups (≤30 and ≥56; $p<.001$). The median age for LPAD, 41 (IQR 32.25–51; $p=.049$), was higher than LP. The median CD4 count for LP was 158 cell/mm^3 (IQR 59–250) and for LPAD was 83.5 cells/mm^3 (IQR 33–144). LP population was mainly from Portugal and sub-Saharan Africa (64.4% vs. 28.8%) and the HIV exposure category was mainly heterosexuals and men have sex with men (MSM) (57.0% vs. 24.9%). In the univariate analysis, both region of origin and HIV exposure category were associated with LP ($p<.001$), but not associated with LPAD. Clinical characteristics, as stage of disease and viral load at diagnosis were associated with both LP and LPAD ($p<.001$), for LP stage A had a higher proportion 48.6%, but for LPAD, the stage C had higher percentage, 46.4%.

In the unadjusted model (Table 2), for LP versus non-LP, no significant differences were found between gender. In the HIV exposure category, significant differences were found for MSM compared with heterosexuals, with a higher proportion of heterosexuals among LP. Furthermore, significantly more immigrants from sub-Saharan Africa were LP when compared to native Portuguese. In LP versus LPAD unadjusted model, significant differences were found in stage of infection, youngest age group compared to the older one (LPAD are older), and viral loads higher than 5.1 compared to viral loads lower than 4.0, as consistent with the evolution of infection. However, no significant differences were found between gender, HIV exposure category, and region of origin.

In the adjusted model, age at diagnosis is one of the factors associated with LP (Table 2), patients with <30 years old had lower probability of being LP than patients with >56 years old (adjusted odds ratio [aOR] 0.34; 0.17–0.68; $p=.002$). Patients from sub-Saharan Africa had 2.24 more probability of presenting late than those from Portugal [aOR 2.24; 1.44–3.50; $p<.001$), and patients presenting stage B or C had higher probability of being LP than those in stage A (aOR 2.95; 1.79–4.86 and aOR 10.16; 5.38–19.19); $p<.001$ and $p<.001$, respectively]. The last variable associated with LP was viral load at diagnosis, patients with a viral load between 4.1–5.0 and >5.1 had higher probability than those with a viral load of <4.0 (aOR 3.40; 2.00–5.79 and aOR 7.01; 4.03–12.20; $p<.001$ and $p<.001$, respectively).

Table 1. Demographic and Clinical Patient Characteristics

Patient characteristics	*Total*	*LPs*	*Non-LPs*	p	*LPAD*	p
Gender, *n* (%)	907 (100)	459 (50.6)	448 (49.4)	.499	284 (61.9)	.403[1]
Female	284 (31.3)	139 (30.3)	145 (32.4)		82 (28.9)	
Male	623 (68.7)	320 (69.7)	303 (67.6)		202 (71.1)	
Median age at diagnosis in years IQR, *n* (%)	806 (100) 37.0 (30.0–47.0)	430 (53.3) 39.5 (32.0–50.0)	376 (46.7) 33.5 (28.0–43.0)	**<.001**	264 (61.4) 41.0 (32.25–51.0)	**.049[2]**
<30	230 (28.5)	89 (20.7)	141 (37.5)	**<.001**	46 (17.4)	**.049[1]**
31–55	488 (60.5)	276 (64.2)	212 (56.4)		172 (65.2)	
>56	88 (10.9)	65 (15.1)	23 (6.1)		46 (17.4)	
Type of transmission, *n* (%)	894 (100)	453 (50.7)	441 (49.3)	**.001**	278 (61.4)	.496[1]
Heterosexual	468 (52.3)	258 (57.0)	210 (47.6)		165 (59.4)	
MSM	275 (30.8)	113 (24.9)	162 (36.7)		66 (23.7)	
IDU	141 (15.8)	75 (16.6)	66 (15.0)		44 (15.8)	
Other	10 (1.1)	7 (1.5)	3 (0.7)		3 (1.1)	
Region of origin, *n* (%)	899 (100)	455 (50.6)	444 (49.4)	**.004**	282 (62.0)	.893[1]
Portugal	609 (67.7)	293 (64.4)	316 (71.2)		181 (64.2)	
Sub-Saharan Africa	215 (23.9)	131 (28.8)	84 (18.9)		80 (28.4)	
Brazil	54 (6.0)	23 (5.1)	31 (7.0)		16 (5.7)	
Other	21 (2.3)	8 (1.8)	13 (2.9)		5 (1.8)	
Stage of infection at diagnosis, *n* (%)	895 (100)	453 (50.6)	442 (49.4)	**<.001**	276 (61.3)	**<.001[1]**
A	598 (66.8)	220 (48.6)	378 (85.5)		94 (34.1)	
B	131 (14.6)	84 (18.5)	47 (10.6)		54 (19.6)	
C	166 (18.5)	149 (32.9)	17 (3.8)		128 (46.4)	
Median CD4 count at diagnosis (cells/mL) IQR, *n* (%)	907 (100) 342.0 (155.0–554.0)	459 (50.6) 158.0 (59.0–250.0)	448 (49.4) 555.5 (444.0–712.0)		284 (61.9) 83.5 (33.0–144.0)	
Viral load at diagnosis (log_{10} copies/mL) IQR, *n* (%)	785 (100) 4.8 (4.2–5.4)	396 (50.4) 5.1 (4.7–5.6)	389 (49.6) 4.4 (3.8–4.9)	**<.001**	241 (60.9) 5.3 (4.9–5.7)	**<.001[2]**
>4.0	146 (18.6)	27 (6.8)	119 (30.6)	**<.001**	11 (4.6)	**<.001[1]**
4.1–5.0	322 (41.0)	139 (35.1)	183 (47.0)		61 (25.3)	
<5.1	317 (40.4)	230 (58.1)	87 (22.4)		169 (70.1)	

Bold text represents statistically significant *p* values.
[1]Chi-square test for variable comparison.
[2]*t*-test of independent samples for variable comparison.
Other in mode of transmission include transfusions; Other in region of origin include European and Latin and North American Countries; *p* values retrieved with *t*-test and chi-square test.
IDU, injection drug users; IQR, interquartile range; LP, late presenters; LPAD, late presenters with advanced disease; MSM, men have sex with men.

In the LPAD model (Table 2), the factors associated with LPAD included stage of infection at diagnosis—patients presenting stage B or C had higher probability than those in stage A (aOR 2.42; 1.35–4.35 and aOR 7.01; 3.79–12.98; $p = .003$ and $p < .001$, respectively) and viral load at diagnosis—patients with a viral load of >5.1 had higher probability of being LPAD than those with a viral load of <4.0 (aOR 3.10; 1.24–7.77; $p = .016$).

We also performed an analysis divided into three periods of time, which are consistent with important periods in the evolution of HIV treatment: (1) the pre-HAART, the time corresponding to the era before availability of Highly Active Antiretroviral Therapy (1984–1996); (2) the HAART period, before availability of single-tablet regimens (1997–2007), and (3) the introduction of a single-tablet regimens for treatment (2008–2017). The differences in the analysis of these time periods were not significant when compared to the continuous time analysis (Supplementary Tables S1–S3 and Table 2). In the first period (1984–1996), none of the variables was significant ($p < .05$), maybe due to the lower number of patients in this more distant time period. The last period (2008–2017) was the one presenting results more similar to our continuous analysis. The tables for this analysis were included in the Supplementary Tables S1–S3.

Discussion

This study had the goal of understanding the determinants of LP for HIV-1 infection.

In our population of patients, LPs represented 50.6% of the patients. Of these, 61.9% were LPAD, which is consistent with the last national report (2017).[2] Our results are in accordance with overall European data (2017), in which the LPs account for 49% of the HIV cases, and 28% were LPAD.[12]

The proportion of LPs were higher in male gender, patients with heterosexual transmission, immigrants originated from sub-Saharan Africa and patients aged between 31 and 55

Table 2. Univariate and Multivariate Logistic Regression Analysis of Factors Associated with Late Presentation and Late Presentation with Advanced Disease

	LP versus NLP				*LP versus LPAD*			
	Unadjusted		*Final model*		*Unadjusted*		*Final model*	
	OR (95% CI)	p	*aOR (95% CI)*	p	*OR (95% CI)*	p	*aOR (95% CI)*	p
Sex								
Female	Ref		Ref		Ref		Ref	
Male	1.10 (0.83–1.46)	0.499	1.07 (0.73–1.58)	0.732	1.19 (0.79–1.79)	0.403	0.91 (0.54–1.51)	0.706
Age at diagnosis	1.04 (1.03–1.05)	**<0.001**			1.02 (1.00–1.03)	**0.048**		
Age groups								
<30	0.22 (0.13–0.39)	**<0.001**	0.34 (0.17–0.68)	**0.002**	0.44 (0.22–0.87)	**0.018**	0.60 (0.27–1.35)	0.216
31–55	0.46 (0.28–0.77)	**0.003**	0.54 (0.19–1.03)	0.060	0.68 (0.38–1.23)	0.203	0.72 (0.36–1.43)	0.347
>56	Ref		Ref		Ref		Ref	
Type of transmission								
Heterosexual	Ref		Ref		Ref		Ref	
MSM	0.57 (0.42–0.77)	**<0.001**			0.79 (0.50–1.24)	0.311		
IDU	0.93 (0.63–1.35)	0.686			0.80 (0.47–1.35)	0.405		
Other	1.90 (0.49–7.44)	0.357			0.42 (0.09–1.93)	0.266		
Region of origin								
Portugal	Ref		Ref		Ref			
Sub-Saharan Africa	1.68 (1.23–2.31)	**0.001**	2.24 (1.44–3.50)	**<0.001**	0.97 (0.64–1.48)	0.890		
Brazil	0.80 (0.46–1.40)	0.437	0.68 (0.31–1.51)	0.349	1.41 (0.56–3.55)	0.460		
Other	0.66 (0.27–1.62)	0.369	0.47 (0.15–1.43)	0.181	1.03 (0.24–4.40)	0.967		
Stage of infection at diagnosis								
A	Ref		Ref		Ref		Ref	
B	3.07 (2.07–4.55)	**<0.001**	2.95 (1.79–4.86)	**<0.001**	2.41 (1.43–4.06)	**0.001**	2.42 (1.35–4.35)	**0.003**
C	15.06 (8.88–25.55)	**<0.001**	10.16 (5.38–19.19)	**<0.001**	9.53 (5.44–16.71)	**<0.001**	7.01 (3.79–12.98)	**<0.001**
Viral load at diagnosis	2.71 (2.23–3.30)	**<0.001**			2.08 (1.57–2.75)	**<0.001**		
Viral load groups								
<4.0	Ref		Ref		Ref		Ref	
4.1–5.0	3.35 (2.09–5.37)	**<0.001**	3.40 (2.00–5.79)	**<0.001**	1.14 (0.49–2.63)	0.763	1.14 (0.45–2.90)	0.788
>5.1	11.65 (7.17–18.93)	**<0.001**	7.01 (4.03–12.20)	**<0.001**	4.03 (1.77–9.16)	**0.001**	3.10 (1.24–7.77)	**0.016**

Bold text represents statistically significant *p* values.
Other in mode of transmission include transfusions; Other in region of origin include European and Latin and North American Countries.
aOR, adjusted odds ratio; NLP, non-late presenters; Ref, Reference category.

years. These results are consistent with other studies.[8,11,13] Patients originated from sub-Saharan Africa represent 23.9% of the study population; of those, 28.8% were LP, which is substantially lower when compared to a similar Belgian study, where the proportion of patients from sub-Saharan African represented 54.3% of the total.[14]

According to our results, the stage A of infection had higher proportions for the LP population, while in the LPAD population the higher proportion was for the stage C. This is expected and can be considered as a partial validation for our assignment of LP and LPAD to patients groups.[15]

The results from previous studies showed a statistically significant correlation between LP and HIV exposure category.[6,16,17] While we did not find this significant correlation in our logistic regression analysis, we did find it in the univariate analysis. However, recent changes in the proportion of HIV exposure categories among new diagnoses could confound such analysis.[18,19]

The main goal of this study was to identify factors associated with late presentation and late presentation with advanced disease. Those factors included age at diagnosis, region of origin, stage of infection at diagnosis, and $\log_{10}$ of the viral load at diagnosis and were in concordance with published studies.[10,17,20]

Conclusion

Even though Portugal has achieved the 90-90-90 objectives, the proportion of LPs is still very high, indicating that vulnerable populations are being left behind in screening protocols. Late presentation is a high impact issue at individual, economic, and social level. Our study highlighted the main factors associated with that condition. Targeted prevention and screening programs should be directed to this population.

Authors' Contributions

Conceptualization: M.M., M.P., and A.A.; data curation: A.C.M., J.T., S.P., T.B.A., P.G., and K.M.; data collection: A.C.M., J.T., S.P., T.B.A., P.G., and K.M.; data analyses: M.M., V.P., M.P., and A.A.; funding acquisition: A.A. and M.P.; study design: M.M., M.P., V.P., and A.A.; writing of original draft: M.M., M.P., and A.A.; and article review and editing: M.M., A.C.M., M.P., V.P., and A.A.

Author Disclosure Statement

No competing financial interests exist.

Funding Information

This study was financed by FCT through the following projects: MigrantsHIV PTDC/DTP-EPI/7066/2014, BESTHOPE HIVERA: Harmonizing Integrating Vitalizing European Research on HIV/Aids, grant 249697 and GHTM-UID/Multi/04413/2013 and GHTM-UID/04413/2020 and Gilead Génese HIVLatePresenters.

Supplementary Material

Supplementary Figure S1
Supplementary Table S1
Supplementary Table S2
Supplementary Table S3

References

1. HIV/AIDS. [Online]. Available at https://www.who.int/news-room/fact-sheets/detail/hiv-aids, accessed February 18, 2020.
2. Portugal: Ministério da Saúde. Direção-Geral da Saúde/Instituto Nacional de Saúde Doutor Ricardo Jorge. Infeção VIH e SIDA em Portugal 2019 http://repositorio.insa.pt/bitstream/10400.18/7093/1/DGS-INSA-RelatVIHSIDA2019.pdf Lisboa: DGS/INSA; 2019.
3. Understanding fast-track. Available at https://www.unaids.org/sites/default/files/media_asset/201506_JC2743_Understanding_FastTrack_en.pdf, accessed February 18, 2020.
4. UNAIDS: Knowledge is power—Know your status, know your viral load. Available at http://www.unaids.org/sites/default/files/media_asset/jc2940_knowledge-is-power-report_en.pdf, accessed February 18, 2020.
5. Buetikofer S, Wandeler G, Kouyos R, Weber R, Ledergerber B: Prevalence and risk factors of late presentation for HIV diagnosis and care in a tertiary referral centre in Switzerland. Swiss Med Wkly 2014;144:w13961
6. Antinori A, Coenen T, Costagiola D, *et al.*: Late presentation of HIV infection: A consensus definition. HIV Med 2011;12:61–64.
7. Luma HN, Jua P, Donfack O-T, *et al.*: Late presentation to HIV/AIDS care at the Douala general hospital, Cameroon: Its associated factors, and consequences. BMC Infect Dis 2018;18:1–9.
8. Iwuji CC, Churchill D, Gilleece Y, Weiss HA, Fisher M: Older HIV-infected individuals present late and have a higher mortality: Brighton, UK cohort study. BMC Public Health 2013;13:1–9.
9. Guelar A, Manzardo C, Sambeat MA, *et al.*: Epidemiological characteristics and predictors of late presentation of HIV infection in Barcelona (Spain) during the period 2001–2009. AIDS Res Ther 2011;8:22.
10. Darling KE, Hachfeld A, Cavassini M, Kirk O, Furrer H, Wandeler G: Late presentation to HIV care despite good access to health services: Current epidemiological trends and how to do better. Swiss Med Wkly 2016;146:w14348.
11. Op De Coul ELM, van Sighem A, Brinkman K, *et al.*: Factors associated with presenting late or with advanced HIV disease in the Netherlands, 1996 2014: Results from a national observational cohort. BMJ Open 2016;6:e009688.
12. HIV infection and AIDS methods. [Internet]. Available at https://ecdc.europa.eu/sites/portal/files/documents/AER_for_2017-hiv-infection-aids_1.pdf, accessed February 18, 2020.
13. Smith RD, Delpech VC, Brown AE, Rice BD: HIV transmission and high rates of late diagnoses among adults aged 50 years and over. Aids 2010;24:2109–2115.
14. Darcis G, Lambert I, Sauvage A-S, *et al.*: Factors associated with late presentation for HIV care in a single Belgian reference center: 2006–2017. Sci Rep 2018;8:8594.
15. Kigozi IM, Dobkin LM, Martin JN, *et al.*: Late-disease stage at presentation to an HIV clinic in the era of free antiretroviral therapy in Sub-Saharan Africa. J Acquir Immune Defic Syndr 2009;52:2.
16. Mocroft A, Lundgren JD, Sabin ML, *et al.*: Risk factors and outcomes for late presentation for HIV-positive persons in Europe: Results from the Collaboration of Observational

HIV Epidemiological Research Europe Study (COHERE). PLoS Med 2013;10:e1001510.
17. Wójcik-Cichy K, Jabłonowska O, Piekarska A: The high incidence of late presenters for HIV/AIDS infection in the Lodz province, Poland in the years 2009–2016: We are still far from the UNAIDS 90% target. AIDS Care 2018;0: 1–4.
18. Jenness SM, Murrill CS, Liu KL, Wendel T, Begier E, Hagan H: Missed opportunities for HIV testing among high-risk heterosexuals. Sex Transm Dis 2009;36:704–710.
19. Kellerman SE, Drake A, Lansky A, Klevens RM: Use of and exposure to HIV Prevention Programs and Services by Persons at High Risk for HIV. AIDS Patient Care STDS 2006;20:391–398.
20. Raffetti E, Postorino MC, Castelli F, *et al.*: The risk of late or advanced presentation of HIV infected patients is still high, associated factors evolve but impact on overall mortality is vanishing over calendar years: Results from the Italian MASTER Cohort. BMC Public Health 2016;16:878.

Address correspondence to:
Mafalda Miranda
Global Health and tropical Medicine (GHTM)
Institute of Hygiene and Tropical medicine (IHMT)
New University of Lisbon
Lisbon 1300-028
Portugal

E-mail: mafalda_nsm@hotmail.com

4.2. Manuscript II

pathogens

MDPI

Article

Determinants of HIV-1 Late Presentation in Patients Followed in Europe

Mafalda N. S. Miranda [1,*], Marta Pingarilho [1], Victor Pimentel [1], Maria do Rosário O. Martins [1], Anne-Mieke Vandamme [1,2], Marina Bobkova [3], Michael Böhm [4], Carole Seguin-Devaux [5], Roger Paredes [6], Rafael Rubio [7], Maurizio Zazzi [8], Francesca Incardona [9,10] and Ana Abecasis [1]

1 Global Health and Tropical Medicine (GHTM), Institute of Hygiene and Tropical Medicine, New University of Lisbon (IHMT/UNL), 1349-008 Lisbon, Portugal; martapingarilho@ihmt.unl.pt (M.P.); victor.pimentel@ihmt.unl.pt (V.P.); mrfom@ihmt.unl.pt (M.d.R.O.M.); annemie.vandamme@uzleuven.be (A.-M.V.); ana.abecasis@ihmt.unl.pt (A.A.)
2 Laboratory Clinical and Epidemiological Virology, Department of Microbiology and Immunology, KU Leuven, Rega Institute for Medical Research, 3000 Leuven, Belgium
3 Gamaleya Research Center of Epidemiology and Microbiology, Department of General Virology, Gamaleya Scientific Research Institute, 123098 Moscow, Russia; mrbobkova@mail.ru
4 Department of Medicine, Saarland University Hospital, 66421 Homburg, Germany; michael.boehm@uk-koeln.de
5 Laboratory of Retrovirology, Department of Infection and Immunity, Luxembourg Institute of Health, L-4354 Esch-sur-Alzette, Luxembourg; carole.devaux@lih.lu
6 Infectious Diseases Department and IrsiCaixa AIDS Research Institute, Hospital Universitari Germans Trias i Pujol, 08916 Badalona, Spain; rparedes@irsicaixa.es
7 Hospital Universitario 12 de Octubre, Universidad Complutense de Madrid, 28026 Madrid, Spain; rafaelrubiogarcia@ucm.es
8 Department of Medical Biotechnologies, University of Siena, 53100 Siena, Italy; maurizio.zazzi@unisi.it
9 IPRO—InformaPRO S.r.l., 98, 00152 Rome, Italy; f.incardona@informa.pro
10 EuResist Network, 98/100, 00152 Rome, Italy
* Correspondence: a21000919@ihmt.unl.pt; Tel.: +351-213-652-600

Abstract: To control the Human Immunodeficiency Virus (HIV) pandemic, the World Health Organization (WHO) set the 90-90-90 target to be reached by 2020. One major threat to those goals is late presentation, which is defined as an individual presenting a TCD4+ count lower than 350 cells/mm^3 or an AIDS-defining event. The present study aims to identify determinants of late presentation in Europe based on the EuResist database with HIV-1 infected patients followed-up between 1981 and 2019. Our study includes clinical and socio-demographic information from 89,851 HIV-1 infected patients. Statistical analysis was performed using RStudio and SPSS and a Bayesian network was constructed with the WEKA software to analyze the association between all variables. Among 89851 HIV-1 infected patients included in the analysis, the median age was 33 (IQR: 27.0–41.0) years and 74.4% were males. Of those, 28,889 patients (50.4%) were late presenters. Older patients (>56), heterosexuals, patients originated from Africa and patients presenting with log VL >4.1 had a higher probability of being late presenters ($p < 0.001$). Bayesian networks indicated VL, mode of transmission, age and recentness of infection as variables that were directly associated with LP. This study highlights the major determinants associated with late presentation in Europe. This study helps to direct prevention measures for this population.

Keywords: HIV-1 infection; late presentation; Europe

check for updates

Citation: Miranda, M.N.S.; Pingarilho, M.; Pimentel, V.; Martins, M.d.R.O.; Vandamme, A.-M.; Bobkova, M.; Böhm, M.; Seguin-Devaux, C.; Paredes, R.; Rubio, R.; et al. Determinants of HIV-1 Late Presentation in Patients Followed in Europe. *Pathogens* **2021**, *10*, 835. https://doi.org/10.3390/pathogens10070835

Academic Editor: Alessandra Borsetti

Received: 4 May 2021
Accepted: 30 June 2021
Published: 2 July 2021

1. Introduction

At the end of 2019, there were 38.0 million people living with the Human Immunodeficiency Virus (HIV) and 1.7 million people were newly infected worldwide. However, 7.1 million people were still unaware of their HIV status [1].

For the control of the HIV pandemic, the World Health Organization (WHO) had set a 90-90-90 target until 2020. 90% of people living with HIV know their status, of those 90% are receiving antiretroviral therapy (ART) and of those 90% achieve viral suppression. These targets had been successful in some countries. Globally, by the end of 2019, there were 81% of people living with HIV who knew their status. Of those, 67% were receiving antiretroviral therapy and of those 59% had reached HIV viral suppression. The success of these goals is dependent on the region of origin, the vulnerability of populations and on the national HIV programs that are implemented. Yet, between 2010 and 2019, the percentage of new infections dropped by 31% [2].

New goals were set to end the pandemic by 2030, the 95-95-95 targets, based on the same definition of the previous targets. In order to attain the WHO goals by 2030, early diagnosis is essential [3].

One major concern threatening those goals is late presentation. Late presentation can have consequences in the health and treatment of infected individuals, leading to poorer outcomes and increased health care costs, since it has been shown that late presenters, especially those aged above 50 years old, are at higher risk for developing non-infectious co-morbidities and complex multimorbidity [4]. In addition, late presentation can have a negative impact on the control of the pandemic, increasing the risk of onward HIV transmission in individuals that are not aware of their HIV status. Besides, late presentation to HIV care was shown to be the main reason for virological failure [5,6].

Late presentation is defined as an individual presenting a TCD4+ count lower than 350 cells/mm^3 or an AIDS-defining event, regardless of TCD4+ cell count. This is the definition according to the European Late Presenter Consensus working group [7]. It is estimated that Late Presenters (LP) account for 40–60% of HIV cases in Europe, in Asia the percentage of LP range from 72 to 83%, in Africa range from 35 to 89% and in Brazil, it is estimated that the percentage is near 40% [8–10]. For prevention and treatment of HIV, timely diagnosis and linkage to health care are essential tools [11].

The present study has the objective of identifying determinants of late presentation in Europe. To achieve this goal, we analyzed a population of patients from the EuResist database, a European database.

2. Results

2.1. Characteristics of European Population

Among 89851 HIV-1 infected patients included in the analysis, the median age was 33 (IQR: 27.0–41.0) years and 74.4% were males. From those 28889 patients (50.4%) were LP and 28388 (49.6%) were non-late presenters (NLP). The majority of patients with information about treatment status were naïve, 11487 (58.6%). 41.9% of patients were men who have sex with men (MSM) and 78.5% originated from Western Europe. The most prevalent subtype in this population was subtype B (64.4%), followed by Subtype G (20.4%), CRF 02_AG (15.9%) and Subtype A (13.5%). Most of the patients included in this study were classified as Native (75.4%) and as having Chronic Infection (59.8%) based on the ambiguity rate of the first genomic sequence. CD4 count at diagnosis and viral load at diagnosis (log10) presented a median of 348 cells/mm^3 (IQR 170-548) and 4.4 copies/mL (IQR 3.4–5.1), respectively.

50.4% of patients were classified as LP (CD4 < 350 cells/mm^3). Males accounted for the higher proportion of LP (74.9%). The median age of LPs was 34 years (IQR 28.0–43.0; $p < 0.001$). LPs were mainly from Western Europe and the HIV exposure category was mainly heterosexuals (77.4 and 37.1%; $p < 0.001$, respectively) (Table 1).

Table 1. Demographic and patient characteristics. Other in mode of transmission includes blood transfusions and vertical transmission. Other in region of origin and infection includes North and Central America, Asian and Oceania continents. *p*-values retrieved with chi-square test and Mann–Whitney U test.

Patient Characteristics	Total	Late Presenters	Non-Late Presenters	*p*-Value
Total, *n* (%)	89,851 (100%)	28,889 (50.4%)	28,388 (49.6%)	
Sex, *n* (%)	81,777 (91.0%)	27,972 (50.6%)	27,315 (49.4%)	
Male	60,852 (74.4%)	20,955 (74.9%)	20,969 (76.8%)	<0.001
Female	20,925 (25.6%)	7017 (25.1%)	6346 (23.2%)	
Treatment status, *n* (%)	19,605 (21.8%)	10,905 (55.6%)	8700 (44.4%)	
Naïve	11,487(58.6%)	6040 (55.4%)	5447 (62.6%)	<0.001
Treated	8118 (41.4%)	4865 (44.6%)	3253 (37.4%)	
Median age at diagnosis in years	25,530 (28.4%)	11,929 (52.3%)	10,897 (47.7%)	<0.001
IQR, *n* (%)	33.0 (27.0–41.0)	34.0 (28.0–43.0)	31.0 (26.0–39.0)	
≤18	700 (2.7%)	241 (2.0%)	340 (3.1%)	
19–30	9767 (38.3%)	4002 (33.5%)	4823 (44.3%)	<0.001
31–55	13,815 (54.1%)	6920 (58.0%)	5384 (49.4%)	
≥56	1248 (4.9%)	766 (6.4%)	350 (3.2%)	
Mode of transmission, *n* (%)	47,007 (52.3%)	21,283 (49.5%)	21,677 (50.5%)	
Heterosexual	15,165 (32.3%)	7894 (37.1%)	6071 (28.0%)	
MSM	19,696 (41.9%)	7657 (36.0%)	10,693 (49.3%)	<0.001
IDU	9532 (20.3%)	4453 (20.9%)	3896 (18.0%)	
Other	2614 (5.6%)	1279 (6.0%)	1017 (4.7%)	
Region of origin, *n* (%)	54,529 (60.7%)	21,584 (50.1%)	21,495 (49.9%)	
Western Europe	42,790 (78.5%)	16,693 (77.4%)	17,398 (81.0%)	
Eastern Europe	1862 (3.4%)	693 (3.2%)	672 (3.1%)	<0.001
Africa	5349 (9.8%)	2250 (10.4%)	1422 (6.6%)	
South America	3233 (5.9%)	1341 (6.2%)	1460 (6.8%)	
Other	1286 (2.4%)	607 (2.8%)	543 (2.5%)	
Subtype, *n* (%)	54,176 (60.3%)	17,449 (52.7%)	15,638 (47.3%)	
HIV-1 Subtype B	35,454 (64.4%)	11,966 (68.6%)	11,745 (75.1%)	<0.001
HIV-1 Subtype non-B	18,722 (34.6%)	5483 (31.4%)	3893 (24.9%)	
Distribution of non-B Subtypes				
HIV-1 CRF 01_AE	447 (2.4%)	183 (3.3%)	108 (2.8%)	
HIV-1 CRF 02_AG	2973 (15.9%)	871 (15.9%)	556 (14.3%)	
HIV-1 CRF 06_cpx	248 (1.3%)	81 (1.5%)	58 (1.5%)	
HIV-1 CRF 14_BG	1106 (5.9%)	337 (6.1%)	203 (5.2%)	
HIV-1 Subtype A	2521 (13.5%)	626 (11.4%)	527 (13.5%)	
HIV-1 Subtype C	1943 (10.4%)	550 (10.0%)	400 (10.3%)	
HIV-1 Subtype D	307 (1.6%)	102 (1.9%)	74 (1.9%)	
HIV-1 Subtype F	1619 (8.6%)	444 (8.1%)	362 (9.3%)	
HIV-1 Subtype G	3815 (20.4%)	1156 (21.1%)	701 (28.0%)	
Others	3743 (20.0%)	1133 (20.7%)	3893 (23.2%)	
Migrant status, *n* (%)	54,520 (60.7%)	21,584 (50.1%)	21,495 (49.9%)	
Migrant	13,408 (24.6%)	5588 (25.9%)	4895 (22.8%)	<0.001
Native	41,112 (75.4%)	15,996 (74.1%)	16,600 (77.2%)	
Recentness of infection, *n* (%)	50,132 (55.8%)	15,897 (52.6%)	14,304 (47.4%)	
Chronic	29,972 (59.8%)	11,069 (69.6%)	7803 (54.6%)	<0.001
Recent	20,160 (40.2%)	4828 (30.4%)	6501 (45.4%)	
Median CD4 count at diagnosis	57,277 (63.7%)	28,889 (50.4%)	28,388 (49.6%)	<0.001
(cells/mL) IQR, *n* (%)	348.0 (170.0–548.0)	172.0 (69.0–264.0)	550.0 (442.0–720.0)	
Viral Load at diagnosis (log10	34,046 (37.9%)	15,106 (50.8%)	14,605 (49.2%)	<0.001
copies/mL) IQR, n (%)	4.4 (3.4–5.1)	4.7 (3.8–5.3)	4.1 (3.1–4.8)	
≤4.0	12,994 (38.2%)	4485 (29.7%)	6819 (46.7%)	
4.1–5.0	11,715 (34.4%)	5034 (33.3%)	5295 (36.3%)	<0.001
≥5.1	9337 (27.4%)	5587 (37.0%)	2491 (17.1%)	

2.2. Determinants Associated with Late Presentation

In the unadjusted model (Table 2), sex was associated with LP. In the HIV exposure category, significant differences were found for MSM and Intravenous Drug Users (IDU) compared with heterosexuals. Significantly more LP were from Africa and other regions compared to Western Europe. In addition, the variables age at diagnosis, viral load, subtype, recentness of infection and migrant status were significantly associated with LP.

Table 2. Logistic Regression for determinants associated with late presentation. Ref—Reference category; aOR-adjusted Odds Ratio; Other in mode of transmission include transfusions; Other in region of origin include Latin and North American Countries; MSM- Men have sex with men; IDU- Injection drug users.

Late Presenters/ Non-Late Presenters	Unadjusted		Final Model	
Variable	**OR (95%CI)**	***p*-Value**	**aOR (95%CI)**	***p*-Value**
Sex				
Female	Ref	Ref	Ref	Ref
Male	0.90 (0.87–0.94)	<0.001	1.05 (0.91–1.21)	0.522
Age at diagnosis	1.03 (1.02–1.03)	<0.001		
Age groups				
<18	0.55 (0.47–0.65)	<0.001	0.48 (0.33–0.69)	<0.001
19–30	0.65 (0.61–0.68)	<0.001	0.70 (0.63–0.79)	<0.001
31–55	Ref	Ref	Ref	Ref
>56	1.70 (1.49–1.94)	<0.001	1.54 (1.15–2.06)	0.004
Mode of transmission				
Heterosexual	Ref	Ref	Ref	Ref
MSM	0.55 (0.53–0.58)	<0.001	0.74 (0.64–0.86)	<0.001
IDU	0.88 (0.83–0.93)	<0.001	1.12 (0.96–1.31)	0.137
Other	0.97 (0.88–1.06)	0.462	1.29 (0.99–1.70)	0.062
Region of Origin				
Western Europe	Ref	Ref	Ref	Ref
Eastern Europe	1.08 (0.97–1.20)	0.191	1.07 (0.78–1.48)	0.683
Africa	1.65 (1.54–1.77)	<0.001	1.76 (1.37–2.26)	<0.001
South America	0.96 (0.89–1.03)	0.267	1.41 (1.07–1.87)	0.015
Other	1.17 (1.04–1.31)	0.011	1.39 (0.92–2.09)	0.118
Subtype				
HIV-1 Subtype B	Ref	Ref		
HIV-1 Subtype non-B	1.38 (1.32–1.45)	<0.001		
Migrant Status				
Migrant	Ref	Ref		
Native	0.84 (0.81–0.88)	<0.001		
Recentness of infection				
Chronic	Ref	Ref	Ref	Ref
Recent	0.52 (0.50–0.55)	<0.001	0.61 (0.55–0.68)	<0.001
Log Viral load at diagnosis	1.45 (1.42–1.48)	<0.001		
Log Viral load groups				
<4.0	Ref	Ref	Ref	Ref
4.1–5.0	1.45 (1.37–1.53)	<0.001	1.37 (1.22–1.55)	<0.001
>5.1	3.41 (3.21–3.62)	<0.001	3.41 (2.96–3.91)	<0.001

Determinants associated with late presentation were age at diagnosis (Table 2): patients with less than 30yo had lower probability of being late presenters and patients aged above 56yo had higher probability of being late presenters when compared with patients aged between 31 and 55yo (>18yo: aOR 0.31 (0.20–0.49), $p < 0.001$; 19–30yo: aOR 0.46 (0.34–0.62), $p < 0.001$; >56: aOR 1.70 (1.49–1.94), $p = 0.004$), transmission via MSM had lower probability when compared with heterosexuals (aOR 0.74 (0.64–0.86); $p < 0.001$). Patients originating from Africa and South America had 1.76 and 1.41 more probability, respectively, of presenting late than those from Western Europe (aOR 1.76 (1.37–2.26), $p < 0.001$; aOR 1.41 (1.07–1.87), $p = 0.015$, respectively) and patients presenting with a viral load between

4.1 and 5.0 and higher than 5.1 had a higher probability of being LP than those with a viral load lower than 4.0 (aOR 1.45 (1.37–1.53) and aOR 3.41 (3.21–3.62); $p < 0.001$ and $p < 0.001$, respectively). As expected, but confirming the reliability of our classification of recentness of infection based on the ambiguity rate, patients with a recent infection—as classified based on the ambiguity rate of the genomic sequence from the first drug resistance test—had a lower probability of being LP than those classified as being chronically infected (aOR 0.61 (0.55–0.68); $p < 0.001$).

2.3. Bayesian Network

For the bayesian network, we used the HillClimber algorithm with nine as the maximum number of parents a node in BN can have. This algorithm is based on a "hill climbing adding and deleting arcs with no fixed ordering of variables" [12]. The BN had a LogScore Bayes of −35615.94 and an accuracy of 61%. In the BN (Figure 1), LPs are directly associated with the viral load, recentness of infection, mode of transmission and age, as we can see in the figure below, those were direct links between the nodes. The indirectly associated links were between LP and region of origin. As we can also see in the figure, there was no direct link between those two nodes. We can see that the mode of transmission is the variable with more direct associations and the variable sex is the only one that is not associated with LP. This BN is in accordance with our logistic regression model.

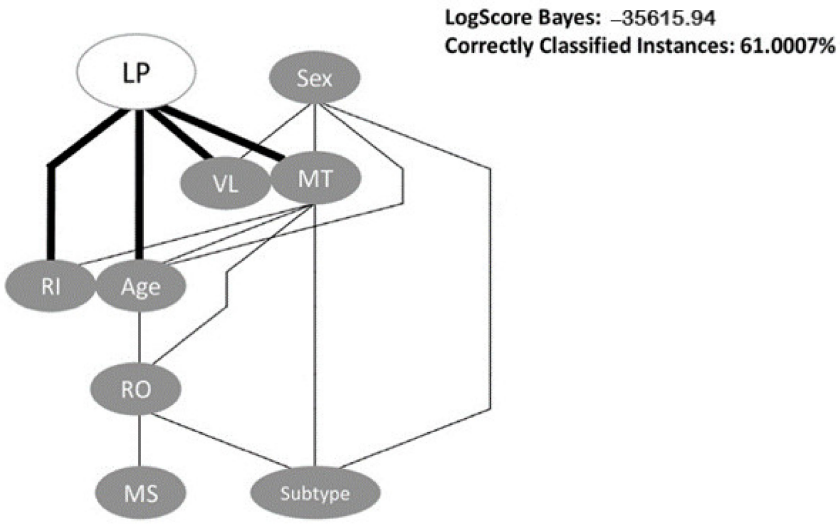

Figure 1. Bayesian Network analysis of association between variables investigated in the study. The BayesNet classifier and the HillClimber algorithm were used as implemented in the WEKA software. LP—Late Presenters; VL—Viral Load; MT—Mode of Transmission; RI—Recentness of infection; RO—Region of Origin; MS—Migrant Status.

The variables Subtype and Migrant status had been removed from the logistic regression model due to the conflict with the variable region of origin. As we can see in Figure 1, the region of origin is directly associated with those two variables and that the migrant status is only associated with region of origin.

2.4. Ambiguity Rate and CD4 Analysis

We performed an analysis to understand the association between CD4 count and the ambiguity rate overall and on subtype B, non-B and G. This association was inversely proportional in all correlations, this means that for higher values of CD4 count the ambiguity rate is lower. In this study, the LP population had higher ambiguity rates in their sequences, since their CD4 count is lower. We also performed a linear regression in order to explain how much of the CD4 count could the ambiguity rate explain. We divided that analysis in the same categories as mentioned above and the higher result was from only individuals with non-B subtype, in which the ambiguity rate explained 5% of the variation from CD4 count (Tables A1–A4).

2.5. Analysis of Late Presenters Rate over Time

We also constructed a graph to evaluate the evolution in time of the rate of LP (Figure 2). The confidence intervals were also calculated for each point. We did not include in the analysis the first three years (1981–1983) since the total number of patients in those years was low and the confidence intervals had high values. In 2019 the sample size was also small, but we included this year in the analysis to see the trend that LPs in Europe will have. As we can see in the graph, LPs have had constant values through the years. In 1984 we had 57.5% LPs, in 1991 we had the lowest value of LPs (45.1%). The evolution through the years maintained between 45 and 60% the rate of LP. Since 2017 the rate of LPs was growing until 2019 that peaked, beyond 60%.

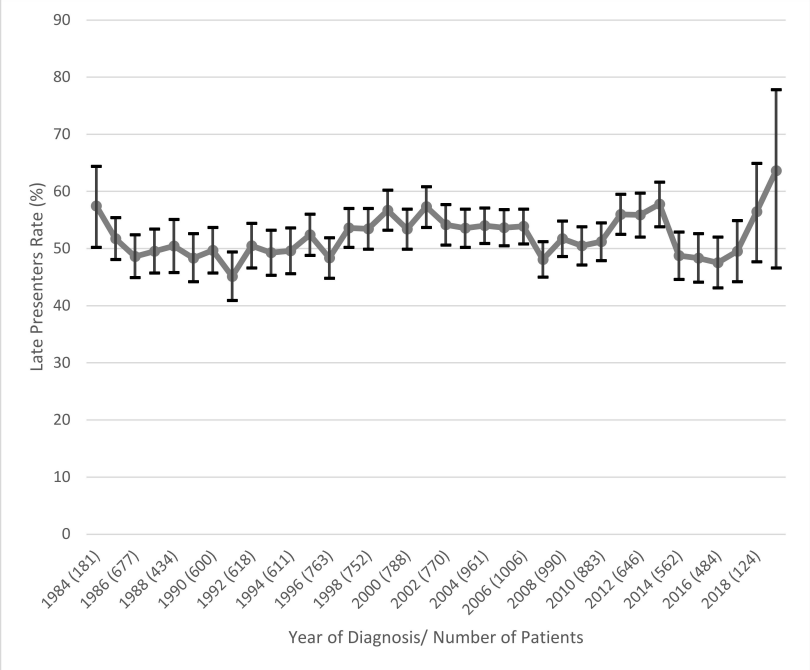

Figure 2. Evolution of Late Presenters rate per Year. Vertical black lines represent confidence intervals.

3. Discussion

This study had the goal of explaining the determinants of late presentation for HIV-1 infection in Europe.

In our population, late presenters represented 50.4% of the patients. A study in Georgia, using the same definition of late presentation as we used, reported 63.4% of late presenters. Another study analyzing late presentation in different settings indicated a rate of late presentation ranging between 40 and 67%, depending on the region of study. This study corresponded to the Swiss data incorporated in the COHERE study, a Collaboration of Observational HIV Epidemiological Research Europe Study. Our results are concordant with the results reported in these studies [13–15].

In our study, late presenters were more frequently males, with heterosexual transmission, from Western Europe and aged between 31 and 55 years old. In a study in East of England, the percentage of late presenters was higher in older patients and patients with heterosexual contact, when compared with homosexual and bisexual contact. Furthermore, according to other studies in Poland and the Netherlands, males were also more prevalent in the late presenters' population. These results are consistent with our study [16–18].

Patients originated from Africa had a higher probability of being LPs when compared to patients originated from Western Europe. This percentage of African migrants in the LP population can be explained by the lower access to health care. Furthermore, African migrants have a higher probability of being in conditions of unemployment, poverty and poorer household, which further increase their barriers to access to health care. A positive status for HIV also stigmatizes individuals, and they fear the reactions of their communities, since HIV is mostly associated among these communities with inappropriate and promiscuous sexual behavior [19]. The migrants of our study from South America were mainly from Brazil and the LP rate was lower than the NLPs. This can be explained in two ways: Brazil has a concentrated HIV epidemic among MSM population and that population is frequently tested [20]. These results are in accordance with HIV studies about the migrant population [21–23].

The results from a previous study showed a statistically significant correlation between late presentation and IDUs [24]. In our study, we found this significant association between LP and IDU in our univariate analysis, but in the logistic regression analysis, we only found significant the association between MSM when compared to heterosexuals. The prevalence of HIV-positive IDU population is mainly from Eastern Europe. In our study, the IDU group maybe underrepresented since the larger proportion of cases are from Western Europe, in which the major mode of transmission is through heterosexual and MSM contacts [25,26].

We also studied the association between CD4 count and the ambiguity rate of the sequences included in this study. Our results show a negative correlation between CD4 count and the ambiguity rate, for lower values of CD4 we had higher values of the ambiguity rate. There is still little information regarding this topic, but our results were in accordance with a study about sequence ambiguity and HIV incidence trends [27]. In fact, the ambiguity rate could be an alternative variable to be used for the definition of Late Presentation. As we know, the initial drop of CD4 count in the acute phase of HIV infection can be a cause of bias when we define Late Presentation based on a CD4 count lower than 350 cells/mm^3.

The results from the graph showed stable and high values for LPs rate. This indicates that LPs were and remain a big part of the HIV epidemic and represent a major threat to treatment and prevention strategies.

The main goal of this study was to identify determinants associated with late presentation. Those determinants included age at diagnosis, mode of transmission, region of origin, recentness of infection and viral load at diagnosis (Figure 3). Our results were in concordance with other previously published studies [13,28,29].

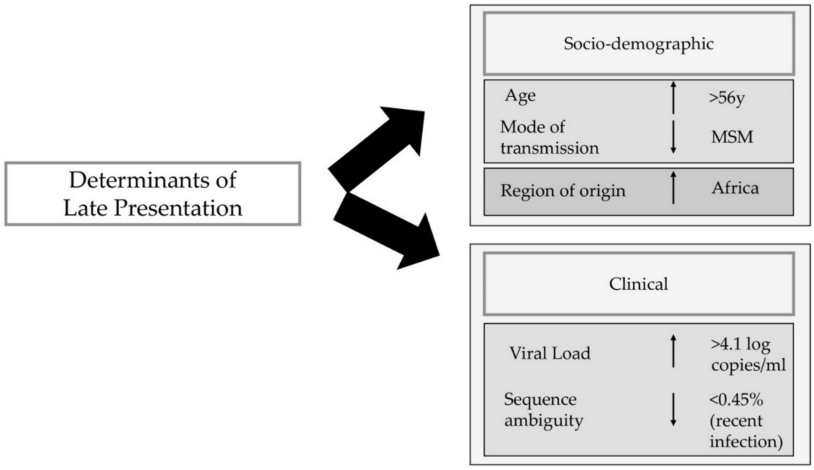

Figure 3. Schematic representation of determinants of late presentation based on logistic regression and Bayesian networks analysis. The light grey box indicates the direct determinants, and the dark grey box indicates the indirect determinants of late presentation. Individuals older than 56yo and originated from Africa had higher probability of being late presenters, Men who have sex with men were less likely to be late presenters. In the clinical determinants there were only direct associated determinants. Individuals with a viral load higher than 4.1 copies/mL had higher probability of being late presenters and individuals with a lower rate of sequence ambiguity had lower probability of being late presenters. MSM-Men who have sex with Men.

The last study about late presentation in Europe was published in 2015 and the timeline of the study was between 2010 and 2013. This was an update from the first study published in 2013, with a timeline of analysis between 2000 and 2011 [29,30]. Our study analyzes a European database with a timeline between 1981 and 2019. The main strength of our study was the database used, which is one of the largest datasets and integrates clinical, socio-demographic and viral genotypic information from HIV-1 patients from all over Europe. This large dataset allows for a robust analysis of the data, and up to date information regarding late presentation. In addition, we can analyze trends in the evolution of late presentation in Europe.

The major limitation of our study was the lack of information about the stage of HIV infection and AIDS-defining events. While we used the ambiguity rate to minimize this problem, we only used the definition of a CD4 count below 350 cells/mm^3 to define an individual as LP or NLP.

Yet, this study is the most recent update on the HIV epidemic of late presentation in Europe, since the last one was published in 2015.

Since late presentation is a major obstacle to the 95-95-95 targets, it is necessary to reinforce the follow-up of this population. Increased HIV testing is key to reduce late presentation since it results in earlier HIV diagnosis. Prevention measures like targeting the vulnerable populations and increasing screening programs for those populations are the most urgent strategies to halt and decrease the percentage of late presenters. In low- and middle-income countries, point-of-care testing would be a major advance to stop the spread of the virus by those who do not know their serological status and therefore decreasing late presentation at diagnosis.

4. Materials and Methods

4.1. Study Group

Our study includes clinical and socio-demographic information from 89851 HIV-1 infected patients from the EuResist Integrated Database (EIDB) between 1981 and 2019. The EuResist integrated database (EIDB) is one of the largest existing datasets which integrate clinical, socio-demographic and viral genotypic information from HIV-1 patients. It integrates longitudinal, periodically updated data mainly from Italy (ARCA database), Germany (AREVIR database) Spain (CoRIS and IRISCAIXA), Sweden, Belgium, Portugal and Luxembourg [31–33].

In this study, information from the ARCA, AREVIR, Luxembourg, IRISCAIXA, Portugal, Russia, United Kingdom and CoRIS databases were used.

4.2. Subtyping

The genomic data included HIV-1 protease and reverse transcriptase sequences, generated through routine drug resistance testing and as stored in the EuResist database. Only the first HIV genomic sequence per patient was considered.

HIV-1 subtyping was performed using the consensus of the result obtained through three different tools: Rega HIV Subtyping Tool (https://www.genomedetective.com/app/typingtool/hiv, accessed on 1 July 2021) [34], COMET: adaptive context-based modeling for HIV-1 (https://comet.lih.lu, accessed on 1 July 2021) [35] and SCUEAL (http://classic.datamonkey.org/dataupload_scueal.php, accessed on 1 July 2021).

4.3. Study Variables

We used the information from the EuResist database regarding the following variables: Country of follow-up, Year of Birth, Gender, Country of Origin, Mode of transmission, Date of the first HIV Positive test, Date and value of the first CD4 count, Date and value of the first Viral Load count, first genomic sequence and sample collection date, Date of therapy initiation.

With this information we created new variables such as:

- Migrant/Native-Based on Country of Origin and Country of Follow-up (if country of origin and country of follow-up is the same, then patient is native; if country of origin and country of follow-up is not the same, then patient is migrant)
- Age at Diagnosis-Based on the difference between Year of Birth and Date of the first HIV Positive test;
- Region of Origin- Based on Country of Origin;
- Treatment Status at date of first CD4 count-Based on the difference between sample collection date of first CD4 count and first therapy date; for purposes of classification of Late Presentation, only patients naïve at date of first CD4 count were considered;
- Treatment Status at date of first Drug Resistance Test-based on the difference between sample collection date for first drug resistance test and date of start of first therapy;

After creating these two variables, for quality control purposes, we only included in the analysis patients for which treatment status at date of first CD4 count and Treatment Status at date of first Drug Resistance test were consistent.

- Recentness of infection-Based on ambiguity rate of genomic sequences. We defined Chronic infection as an ambiguity value higher than 0.45% and Recent infection as an ambiguity value equal or below 0.45% [36]. Additionally, only genomic sequences larger than 500 nucleotides and with ambiguity rate lower than 2.5% were considered.
- LP vs. NLP- Based on CD4 count, LP were defined as patients with CD4 count lower than 350 cells/mm^3 and NLP were defined as patients with CD4 count higher than 350 cells/mm^3.

4.4. Statistical Analysis

The proportion and median (interquartile range, IQR) of LP and non- Late presenters (NLP) were calculated for every categorical and continuous variable, respectively. Our interest variables were compared with the categorical variables with Chi-square test, and continuous variables with Mann–Whitney U test.

To study the relationship between our dependent variable (LP or NLP) and the independent variables, logistic regression models were calculated. We first presented the logistic regression with the unadjusted odds ratios (uOR) and confidence intervals at 95% (95% CI), in order to see the probability of our event, the dependent variable, (late presentation vs non-late presentation) on the occurrence of each independent variable, individually, e.g., the probability of a woman being late presenter. Variables with a p-value < 0.05 were considered to enter the model since it is the most used threshold. The final model for LP vs. NLP was adjusted for sex, this variable was forced into the model regardless of its significance and the reference class was women. The final model included only the variables that were considered statistically significant ($p < 0.05$) and the variables that suited the best regression model according to the backward stepwise regression analysis through SPSS. The odds ratio and 95% confidence intervals were calculated for those variables as well. Data were analyzed using RStudio (Version 1.2.5033) and SPSS (Version 26.0.0.0).

4.5. Bayesian Networks

A Bayesian network (BN) is a tool that consists of a directed acyclic graph (DAG), made of nodes and directed links between the nodes, which allows us to understand the representation of a probabilistic distribution. Each node is a representation of a variable, and the links indicate that one node is directly influencing another. The lack of a direct link does not mean that one variable is not associated with another. These networks are able to intuitively create causal links between variables since they are built from probability distributions and for prediction [37].

We constructed a BN to analyze the association between all variables, specifically, we wanted to see how the variables were associated with one another. With the different levels and connections between variables, it is possible to see if they are directly or indirectly associated. We used the WEKA software version 3.8.5. WEKA stands for Waikato Environment for Knowledge Analysis. After the upload of the dataset, the first step is to choose a classifier to start the analysis. We used a statistical-based learning scheme, the Bayes classifier, specifically the BayesNet [38]. After choosing the classifier, we used different search algorithms as a local score structure learning. Our final choice of algorithm was based on the LogScore Bayes value and the percentage of correctly classified instances.

5. Conclusions

In summary, late presentation still accounts for 50% of the new diagnosis in Europe. Its most important determinants are age at diagnosis, mode of transmission, region of origin and viral load at diagnosis (Figure 3). In addition, the evolution of the rate of late presentation through the years was stable, except for the last two years analyzed (2018 and 2019) when that rate showed an increase. This study highlights the major determinants associated with late presenters in Europe, and this will help to strengthen some prevention measures.

Author Contributions: Conceptualization, M.N.S.M., M.P. and A.A.; Methodology, M.N.S.M., M.P., V.P., A.-M.V., M.d.R.O.M. and A.A.; Software, M.N.S.M., V.P.; Validation, M.N.S.M., M.P., F.I. and A.A.; Formal Analysis, M.N.S.M., V.P., M.P. and A.A.; Investigation, M.N.S.M., M.P., V.P., A.-M.V. and M.d.R.O.M.; Resources, M.B. (Marina Bobkova), M.B. (Michael Böhm), C.S.-D., R.P., R.R., M.Z. and F.I.; Data Curation, M.B. (Marina Bobkova), M.B. (Michael Böhm), C.S.-D., R.P., R.R., M.Z. and F.I.; Writing—Original Draft Preparation, M.N.S.M., M.P. and A.A.; Writing—Review and Editing, M.N.S.M., M.P., M.Z., F.I. and A.A.; Visualization, M.N.S.M., M.P., V.P., A.-M.V., M.d.R.O.M. and A.A.; Supervision, A.A.; Project Administration, A.A.; Funding Acquisition, A.A. All authors have read and agreed to the published version of the manuscript.

Funding: This study was financed by FCT through the following projects: GHTM- UID/04413/2020, INTEGRIV (PTDC/SAU-INF/31990/2017) and the scholarship PD/BD/135714/2018 and Gilead Geénese HIVLatePresenters.

Institutional Review Board Statement: The protocol was in accordance with the declaration of Helsinki. This database contains anonymized patients' information, including demographic and clinical data from patients from the EuResist Integrated Database (Date of approval: 15 January 2021).

Data Availability Statement: Restrictions apply to the availability of these data. Data was obtained from the EuResist Network and is available for request through a study application form at https://www.euresist.org/become-a-partner with the permission of the EuResist Network.

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A

Table A1. Linear Regression for CD4 count vs. Ambiguity rate.

CD4 Count	R^2	Unstandardized B	Standardized Coefficients Beta	*p*-Value	95% Confidence Interval for B		Spearman's Correlation	
					Lower Bound	Upper Bond	*p*-Value	Correlation Coefficient
Ambiguity Rate	0.023	−69.52	−0.152	<0.001	−74.61	−64.44	<0.001	−0.190

Table A2. Linear Regression for CD4 count vs. Ambiguity rate for Subtype B.

CD4 Count	R^2	Unstandardized B	Standardized Coefficients Beta	*p*-Value	95% Confidence Interval for B		Spearman's Correlation	
					Lower Bound	Upper Bond	*p*-Value	Correlation Coefficient
Ambiguity Rate	0.015	−57.11	−0.123	<0.001	−63.27	−50.95	<0.001	−0.159

Table A3. Linear Regression for CD4 count vs. Ambiguity rate for Subtype non-B.

CD4 Count	R^2	Unstandardized B	Standardized Coefficients Beta	*p*-Value	95% Confidence Interval for B		Spearman's Correlation	
					Lower Bound	Upper Bond	*p*-Value	Correlation Coefficient
Ambiguity Rate	0.051	−96.72	−0.226	<0.001	−105.54	−87.91	<0.001	−0.265

Table A4. Linear Regression for CD4 count vs. Ambiguity rate for Subtype G.

CD4 Count	R^2	Unstandardized B	Standardized Coefficients Beta	*p*-Value	95% Confidence Interval for B		Spearman's Correlation	
					Lower Bound	Upper Bond	*p*-Value	Correlation Coefficient
Ambiguity Rate	0.014	−52.78	−0.120	<0.001	−73.38	−32.17	<0.001	−0.144

References

1. WHO. HIV/AIDS. Available online: https://www.who.int/news-room/fact-sheets/detail/hiv-aids. (accessed on 4 January 2021).
2. Global HIV & AIDS Statistics — 2020 Fact Sheet | UNAIDS. Available online: https://www.unaids.org/en/resources/fact-sheet. (accessed on 4 January 2021).
3. Understanding Fast-Track. Available online: https://www.unaids.org/sites/default/files/media_asset/201506_JC2743_Understanding_FastTrack_en.pdf. (accessed on 18 May 2020).
4. Guaraldi, G.; Zona, S.; Menozzi, M.; Brothers, T.D.; Carli, F.; Stentarelli, C.; Dolci, G.; Santoro, A.; Da Silva, A.R.D.; Rossi, E.; et al. Late presentation increases risk and costs of non-infectious comorbidities in people with HIV: An Italian cost impact study. *AIDS Res. Ther.* **2017**, *14*, 1–7. [CrossRef]
5. Conway, A.S.; Esteve, A.; Fernández-Quevedo, M.; Casabona, J. Determinants and Outcomes of Late Presentation of HIV Infection in Migrants in Catalonia, Spain: PISCIS Cohort 2004–2016. *J. Immigr. Minor. Health* **2018**, *21*, 920–930. [CrossRef] [PubMed]
6. Gesesew, H.A.; Ward, P.; Woldemichael, K.; Mwanri, L. Late presentation for HIV care in Southwest Ethiopia in 2003–2015: Prevalence, trend, outcomes and risk factors. *BMC Infect. Dis.* **2018**, *18*, 59. [CrossRef]
7. Antinori, A.; Coenen, T.; Costagiola, D.; Dedes, N.; Ellefson, M.; Gatell, J.; Girardi, E.; Johnson, M.; Kirk, O.; Lundgren, J.; et al. Late presentation of HIV infection: A consensus definition. *HIV Med.* **2010**, *12*, 61–64. [CrossRef] [PubMed]
8. The Late Presentation Working Groups in EuroSIDA and COHERE. Estimating the burden of HIV late presentation and its attributable morbidity and mortality across Europe 2010–2016. *BMC Infect. Dis.* **2020**, *20*, 1–11. [CrossRef]
9. Hu, X.; Liang, B.; Zhou, C.; Jiang, J.; Huang, J.; Ning, C.; Liu, J.; Zhou, B.; Zang, N.; Lai, J.; et al. HIV late presentation and advanced HIV disease among patients with newly diagnosed HIV/AIDS in Southwestern China: A large-scale cross-sectional study. *AIDS Res. Ther.* **2019**, *16*, 6. [CrossRef]
10. Luma, H.N.; Jua, P.; Donfack, O.-T.; Kamdem, F.; Ngouadjeu, E.; Mbatchou, H.B.; Doualla, M.-S.; Mapoure, Y.N. Late presentation to HIV/AIDS care at the Douala general hospital, Cameroon: Its associated factors, and consequences. *BMC Infect. Dis.* **2018**, *18*, 298. [CrossRef]
11. Wilton, J.; Light, L.; Gardner, S.; Rachlis, B.; Conway, T.; Cooper, C.; Cupido, P.; Kendall, C.E.; Loutfy, M.; McGee, F.; et al. Late diagnosis, delayed presentation and late presentation among persons enrolled in a clinical HIV cohort in Ontario, Canada (1999–2013). *HIV Med.* **2019**, *20*, 110–120. [CrossRef] [PubMed]
12. Hall, M.; Frank, E.; Holmes, G.; Pfahringer, B.; Reutemann, P.; Witten, I.H. The WEKA Data Mining Software: An Update. *SIGKDD Explor.* **2009**, *11*, 10–18. [CrossRef]
13. Chkhartishvili, N.; Chokoshvili, O.; Bolokadze, N.; Tsintsadze, M.; Sharvadze, L.; Gabunia, P.; Dvali, N.; Abutidze, A.; Tsertsvadze, T. Late presentation of HIV infection in the country of Georgia: 2012–2015. *PLOS ONE* **2017**, *12*, e0186835. [CrossRef] [PubMed]
14. Infographic: HIV Infection - Late Diagnosis. Available online: https://www.ecdc.europa.eu/en/publications-data/infographic-hiv-infection-late-diagnosis (accessed on 4 January 2021).
15. Darling, K.E.; Hachfeld, A.; Cavassini, M.; Kirk, O.; Furrer, H.; Wandeler, G. Late presentation to HIV care despite good access to health services: Current epidemiological trends and how to do better. *Swiss Med. Wkly.* **2016**, *146*, 14348. [CrossRef]
16. Wójcik-Cichy, K.; Jabłonowska, O.; Piekarska, A.; Jabłonowska, E. The high incidence of late presenters for HIV/AIDS infection in the Lodz province, Poland in the years 2009–2016: We are still far from the UNAIDS 90% target. *AIDS Care* **2018**, *30*, 1538–1541. [CrossRef]
17. Van Opstal, S.E.M.; Van Der Zwan, J.S.; Wagener, M.N.; Been, S.K.; Miedema, H.S.; Roelofs, P.; Van Gorp, E.C.M. Late Presentation of HIV Infection in the Netherlands: Reasons for Late Diagnoses and Impact on Vocational Functioning. *AIDS Behav.* **2018**, *22*, 2593–2603. [CrossRef] [PubMed]
18. Bath, R.E.; Emmett, L.; Verlander, N.Q.; Reacher, M. Risk factors for late HIV diagnosis in the East of England: Evidence from national surveillance data and policy implications. *Int. J. STD AIDS* **2018**, *30*, 37–44. [CrossRef]
19. Fakoya, I.; Arco, D.; Álvarez, D.; Copas, A.J.; Teixeira, B.; Block, K.; Gennotte, A.-F.; Volny-Anne, A.; Bil, J.P.; Touloumi, G.; et al. Factors Associated With Access to HIV Testing and Primary Care Among Migrants Living in Europe: Cross-Sectional Survey. *JMIR Public Health Surveill.* **2017**, *3*, 84. [CrossRef] [PubMed]
20. MacCarthy, S.; Brignol, S.; Reddy, M.; Nunn, A.; Dourado, I. Late presentation to HIV/AIDS care in Brazil among men who self-identify as heterosexual. *Revista de Saúde Pública* **2016**, *50*, 54. [CrossRef]
21. Ross, J.; Cunningham, C.O.; Hanna, D.B. HIV outcomes among migrants from low-income and middle-income countries living in high-income countries. *Curr. Opin. Infect. Dis.* **2018**, *31*, 25–32. [CrossRef]
22. Pimentel, V.; Pingarilho, M.; Alves, D.; Diogo, I.; Fernandes, S.; Miranda, M.; Pineda-Pena, A.-C.; Libin, P.; Martins, M.R.O.; Vandamme, A.-M.; et al. Molecular Epidemiology of HIV-1 Infected Migrants Followed Up in Portugal: Trends between 2001–2017. *Viruses* **2020**, *12*, 268. [CrossRef]
23. Hachfeld, A.; Darling, K.; Calmy, A.; Ledergerber, B.; Weber, R.; Battegay, M.; Wissel, K.; Di Benedetto, C.; Fux, C.; E Tarr, P.; et al. Why do sub-Saharan Africans present late for HIV care in Switzerland? *HIV Med.* **2019**, *20*, 418–423. [CrossRef] [PubMed]
24. Suárez-García, I.; Sobrino-Vegas, P.; Dalmau, D.; Rubio, R.; Iribarren, J.A.; Blanco, J.R.; Gutierrez, F.; Alonso, M.M.; Bernal, E.; García, D.V.; et al. Clinical outcomes of patients infected with HIV through use of injected drugs compared to patients infected through sexual transmission: Late presentation, delayed anti-retroviral treatment and higher mortality. *Addiction* **2016**, *111*, 1235–1245. [CrossRef]

25. Degenhardt, L.; Peacock, A.; Colledge, S.; Leung, J.; Grebely, J.; Vickerman, P.; Stone, J.; Cunningham, E.B.; Trickey, A.; Dumchev, K.; et al. Global prevalence of injecting drug use and sociodemographic characteristics and prevalence of HIV, HBV, and HCV in people who inject drugs: A multistage systematic review. *Lancet Glob. Health* **2017**, *5*, 1192–1207. [CrossRef]
26. Balayan, T.; Oprea, C.; Yurin, O.; Jevtović, Đ.; Begovac, J.; Lakatos, B.; Sedlacek, D.; Karpov, I.; Horban, A.; Kowalska, J.D.; et al. People who inject drugs remain hard-to-reach population across all HIV continuum stages in Central, Eastern and South Eastern Europe – data from Euro-guidelines in Central and Eastern Europe Network. *Infect. Dis.* **2019**, *51*, 277–286. [CrossRef]
27. Lunar, M.M.; Lepej Židovec, S.; Poljak, M. Sequence ambiguity determined from routine pol sequencing is a reliable tool for real-time surveillance of HIV incidence trends. *Infect. Genet. Evol.* **2019**, *69*, 146–152. [CrossRef]
28. Darcis, G.; Lambert, I.; Sauvage, A.-S.; Frippiat, F.; Meuris, C.; Uurlings, F.; LeComte, M.; Léonard, P.; Giot, J.-B.; Fombellida, K.; et al. Factors associated with late presentation for HIV care in a single Belgian reference center: 2006–2017. *Sci. Rep.* **2018**, *8*, 8594. [CrossRef] [PubMed]
29. Mocroft, A.; Lundgren, J.D.; Sabin, M.; Monforte, A.D.; Brockmeyer, N.N.; Casabona, J.J.; Castagna, A.; Costagliola, D.; Dabis, F.; De Wit, S.; et al. Risk Factors and Outcomes for Late Presentation for HIV-Positive Persons in Europe: Results from the Collaboration of Observational HIV Epidemiological Research Europe Study (COHERE). *PLoS Med.* **2013**, *10*, e1001510. [CrossRef]
30. Mocroft, A.; Lundgren, J.D.; Antinori, A.; Ad, M. The late presenters working group in COHERE in EuroCoord Late presentation for HIV care across Europe: Update from the Collaboration of Observational HIV Epidemiological Research Europe (COHERE) study, 2010 to 2013. *Eurosurveillance* **2015**, *20*, 30070. [CrossRef]
31. Euresist Data Analysis—Database. Available online: http://engine.euresist.org/database/ (accessed on 4 January 2021).
32. Lawyer, G.; Altmann, A.; Thielen, A.; Zazzi, M.; Sönnerborg, A.; Lengauer, T. HIV-1 mutational pathways under multidrug therapy. *AIDS Res. Ther.* **2011**, *8*, 26. [CrossRef] [PubMed]
33. Zazzi, M.; Incardona, F.; Rosen-Zvi, M.; Prosperi, M.; Lengauer, T.; Altmann, A.; Sönnerborg, A.; Lavee, T.; Schülter, E.; Kaiser, R. Predicting Response to Antiretroviral Treatment by Machine Learning: The EuResist Project. *Intervirology* **2012**, *55*, 123–127. [CrossRef]
34. Pineda-Peña, A.-C.; Faria, N.R.; Imbrechts, S.; Libin, P.; Abecasis, A.; Deforche, K.; Gómez-López, A.; Camacho, R.; de Oliveira, T.; Vandamme, A.-M. Automated subtyping of HIV-1 genetic sequences for clinical and surveillance purposes: Performance evaluation of the new REGA version 3 and seven other tools. *Infect. Genet. Evol.* **2013**, *19*, 337–348. [CrossRef] [PubMed]
35. Struck, D.; Lawyer, G.; Ternes, A.-M.; Schmit, J.-C.; Bercoff, D.P. COMET: Adaptive context-based modeling for ultrafast HIV-1 subtype identification. *Nucleic Acids Res.* **2014**, *42*, e144. [CrossRef] [PubMed]
36. Andersson, E.; Shao, W.; Bontell, I.; Cham, F.; Cuong, D.D.; Wondwossen, A.; Morris, L.; Hunt, G.; Sönnerborg, A.; Bertagnolio, S.; et al. Evaluation of sequence ambiguities of the HIV-1 pol gene as a method to identify recent HIV-1 infection in transmitted drug resistance surveys. *Infect. Genet. Evol.* **2013**, *18*, 125–131. [CrossRef] [PubMed]
37. Sesen, M.B.; Nicholson, A.E.; Banares-Alcantara, R.; Kadir, T.; Brady, M. Bayesian Networks for Clinical Decision Support in Lung Cancer Care. *PLoS ONE* **2013**, *8*, e82349. [CrossRef] [PubMed]
38. Kumar, Y.; Sahoo, G. Analysis of Parametric & Non Parametric Classifiers for Classification Technique using WEKA. *Int. J. Inf. Technol. Comput. Sci.* **2012**, *4*, 43–49. [CrossRef]

4.3. Manuscript III

frontiers | Frontiers in Microbiology

ORIGINAL RESEARCH
published: 13 April 2022
doi: 10.3389/fmicb.2022.846943

Trends of Transmitted and Acquired Drug Resistance in Europe From 1981 to 2019: A Comparison Between the Populations of Late Presenters and Non-late Presenters

Mafalda N. S. Miranda[1], Marta Pingarilho[1], Victor Pimentel[1], Maria do Rosário O. Martins[1], Rolf Kaiser[2], Carole Seguin-Devaux[3], Roger Paredes[4], Maurizio Zazzi[5], Francesca Incardona[6,7] and Ana B. Abecasis[1]*

[1] Global Health and Tropical Medicine (GHTM), Institute of Hygiene and Tropical Medicine, New University of Lisbon (IHMT/UNL), Lisbon, Portugal, [2] Institute of Virology, University of Cologne, Cologne, Germany, [3] Laboratory of Retrovirology, Department of Infection and Immunity, Luxembourg Institute of Health, Esch-sur-Alzette, Luxembourg, [4] Infectious Diseases Department and IrsiCaixa AIDS Research Institute, Hospital Universitari Germans Trias i Pujol, Badalona, Spain, [5] Department of Medical Biotechnologies, University of Siena, Siena, Italy, [6] IPRO—InformaPRO S.r.l., Rome, Italy, [7] EuResist Network, Rome, Italy

OPEN ACCESS

Edited by:
Kok Keng Tee,
University of Malaya, Malaysia

Reviewed by:
Lingjie Liao,
National Center for AIDS/STD Control and Prevention, Chinese Center for Disease Control and Prevention, China
Sunee Sirivichayakul,
Chulalongkorn University, Thailand
Emmanuel Ndashimye,
Western University, Canada

***Correspondence:**
Mafalda N. S. Miranda
mafalda_nsm@hotmail.com

Specialty section:
This article was submitted to Virology, a section of the journal Frontiers in Microbiology

Received: 31 December 2021
Accepted: 15 February 2022
Published: 13 April 2022

Citation:
Miranda MNS, Pingarilho M, Pimentel V, Martins MdRO, Kaiser R, Seguin-Devaux C, Paredes R, Zazzi M, Incardona F and Abecasis AB (2022) Trends of Transmitted and Acquired Drug Resistance in Europe From 1981 to 2019: A Comparison Between the Populations of Late Presenters and Non-late Presenters. Front. Microbiol. 13:846943. doi: 10.3389/fmicb.2022.846943

Background: The increased use of antiretroviral therapy (ART) has decreased mortality and morbidity of HIV-1 infected people but increasing levels of HIV drug resistance threatens the success of ART regimens. Conversely, late presentation can impact treatment outcomes, health costs, and potential transmission of HIV.

Objective: To describe the patterns of transmitted drug resistance (TDR) and acquired drug resistance (ADR) in HIV-1 infected patients followed in Europe, to compare its patterns in late presenters (LP) vs non-late presenters (NLP), and to analyze the most prevalent drug resistance mutations among HIV-1 subtypes.

Methods: Our study included clinical, socio-demographic, and genotypic information from 26,973 HIV-1 infected patients from the EuResist Integrated Database (EIDB) between 1981 and 2019.

Results: Among the 26,973 HIV-1 infected patients in the analysis, 11,581 (42.9%) were ART-naïve patients and 15,392 (57.1%) were ART-experienced. The median age was 37 (IQR: 27.0–45.0) years old and 72.6% were males. The main transmission route was through heterosexual contact (34.9%) and 81.7% of patients originated from Western Europe. 71.9% of patients were infected by subtype B and 54.8% of patients were classified as LP. The overall prevalence of TDR was 12.8% and presented an overall decreasing trend (p for trend < 0.001), the ADR prevalence was 68.5% also with a decreasing trend (p for trend < 0.001). For LP and NLP, the TDR prevalence was 12.3 and 12.6%, respectively, while for ADR, 69.9 and 68.2%, respectively. The most prevalent TDR drug resistance mutations, in both LP and NLP, were K103N/S, T215rev, T215FY, M184I/V, M41I/L, M46I/L, and L90M.

Conclusion: Our study showed that the overall TDR (12.8%) and ADR (68.5%) presented decreasing trends during the study time period. For LP, the overall TDR was slightly lower than for NLP (12.3 vs 12.6%, respectively); while this pattern was opposite for ADR (LP slightly higher than NLP). We suggest that these differences, in the case of TDR, can be related to the dynamics of fixation of drug resistance mutations; and in the case of ADR with the more frequent therapeutic failure in LPs.

Keywords: HIV-1 infection, transmitted drug resistance, acquired drug resistance, late presenters, non-late presenters

INTRODUCTION

In 2014, UNAIDS implemented the Fast-Track approach driven by the 95-95-95 targets. These targets have the aim to end the pandemic by 2030 by achieving 95% of diagnosis among people living with HIV, 95% of those receiving antiretroviral treatment and 95% of those reaching viral suppression (Joint United Nations Programme on Hiv/Aids (Unaids), 2015). In the meantime, UNAIDS has developed a set of targets for 2025 to help achieve the previous goals until 2030, which are people-centered and right-based (Unaids).

At the end of 2020, there were 37.7 million people living with HIV and at least 50% of the new diagnoses were related to late HIV infection [late presenters (LP)], with regional differences. LP are patients newly diagnosed with a baseline CD4 count lower than 350 cells/mm^3 or with an AIDS-defining event, regardless of CD4 cell count (Miranda et al., 2021). Between 2000 and 2020 the percentage of new HIV infections dropped by 49% and HIV-related deaths dropped by 55% due to antiretroviral therapy (ART; World Health Organization).

The advent of highly active ART has greatly improved the prognosis of HIV-1 infection and reduction of the risk of HIV transmission (Cdc). Today, 73% of people living with HIV have access to ART. Drug resistance could be acquired drug resistance (ADR), due to selective pressure of antiretrovirals (ARVs) in individuals, or transmitted drug resistance (TDR) due to an infection by HIV strains that harbor drug resistance mutations (DRMs; Clutter et al., 2016; Pingarilho et al., 2020).

Drug resistance testing is recommended for individuals with HIV infection who are newly diagnosed or ART-naïve patients, individuals on ART with a viral load higher than 200 copies/mL, individuals who did not achieve viral suppression, and individuals who interrupted ART with a non-nucleoside reverse transcriptase inhibitor (NNRTI; Günthard et al., 2019). For ART-naïve patients, genotypic drug-resistance testing involved testing for mutations in the reverse transcriptase (RT), protease (PR) and integrase (IN) genes. In ART-experienced patients, genotypic and phenotypic resistance testing is recommended in individuals suspect of multi drug-resistance mutations and virological failure (Nih).

The most common DRMs among ART-naïve and ART-experienced patients for nucleoside reverse transcriptase inhibitors (NRTIs) were M41L and M184V, respectively, and K103N for NNRTIs (Rossetti et al., 2018; Zou et al., 2020).

In 2016, the World Health Organization (WHO) recommended the following guidelines as a first-line ART regimen: the combination of two NRTIs, such as tenofovir (TDF) and lamivudine (3TC) or emtricitabine (FTC), plus an integrase strand inhibitor (INSTI), such as dolutegravir (DTG), or instead of DTG the combination with the NNRTI efavirenz (EFV). The recommendations for second-line regimens included the combination of two NRTIs plus one protease inhibitor (PI), like atazanavir (ATV) or lopinavir/ritonavir (LPV/RTV) or two NRTIs and DTG. Third-line regimens included the combination of one PI, such as darunavir (DRV), DTG, and one or two NRTIs (World Health Organization).

Resistance to ART could decrease the success of first line regimens and is a major threat to halt the UNAIDS targets, as well as late presentation. Resistance to antiretrovirals and late presentation are still existing problems that could delay the success of regimens and continue the onward transmission of HIV-1 infection. In this study, we aim to describe the patterns of TDR and ADR, as well as compare them in LP and non-late presenter (NLP) populations included in this study. We also analyzed the most prevalent drug resistance mutations and their prevalence in HIV-1 subtypes among LP and NLP HIV-1 infected patients followed in Europe.

METHODS

Study Group

Clinical, socio-demographic, and genomic information from 26,973 HIV-1 infected patients from the EuResist Integrated Database (EIDB) between 1981 and 2019 were included in this study. The EIDB is one of the largest existing datasets which integrate clinical, socio-demographic, and viral genotypic information from HIV-1 patients. It integrates longitudinal, periodically updated data mainly from Italy (ARCA database), Germany (AREVIR database) Spain (CoRIS and IRSICAIXA), Sweden, Belgium, Portugal, and Luxembourg (EuResist; Lawyer et al., 2011; Zazzi et al., 2012).

In this study, information from the ARCA, AREVIR, Luxembourg, IRSICAIXA, and Portugal databases were used.

Exclusion Criteria

Among the 89,851 HIV-1 infected patients included in the EuResist database, only 54,176 patients had sequence information for the RT and PR regions. Those patient sequences went through the quality control process. We calculated the ambiguity rate for each genomic sequence and only included those sequences that were larger than 500 nucleotides and with an ambiguity rate lower

than 2.5%, resulting in the elimination of 4,044 sequences. Our final study population included 26,973 HIV-1 infected patients, because of the 50,132 patients, only 26,973 had information regarding their date of first ARV therapy.

Institutional Review Board Statement

All procedures performed in this study were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki declaration. The database enrolled anonymized patients' information, including demographic, clinical, and genomic data from patients from the EuResist Integrated Database (Date of approval: January 15, 2021).

Drug Resistance Analysis and Subtyping

HIV pol sequences were derived from existing routine clinical genotypic resistance tests (Sanger method, e.g., Viroseq, Trugene and in house genotyping). The size of RT and PR fragments used for this analysis were between 500 and 1,000 nucleotides. Only the first HIV genomic sequence per patient was analyzed. TDR was defined as the presence of one or more surveillance drug resistance mutations in a sequence, according to the WHO 2009 surveillance list (Bennett et al., 2009). The sequences were submitted to the Calibrated Population Resistance tool version 8.0. Clinical resistance to ARV drugs was calculated through the Standford HIVdb version 9.0.

We analyzed TDR and ADR overall proportions between 1981 and 2019, although we only used the years 1995–2019, divided into three time periods (1995–2002; 2003–2010, and 2011–2019), to compute TDR and ADR trends, since the absolute number before 1995 was smaller than 10 patients per year. We also analyzed TDR and ADR proportions in countries of follow-up. For this analysis, we limited the analyses to the last 10 years divided into two time periods (2008–2012 and 2013–2019).

HIV-1 subtyping was performed using the consensus of the result obtained based on three different subtyping tools: Rega HIV Subtyping Tool version 3.46[1] (Pineda-Peña et al., 2013), COMET: adaptive context-based modeling for HIV-1[2] (Struck et al., 2014) and SCUEAL.[3]

Study Variables

New variables were created according to:

- Migrant/Native—based on country of origin and country of follow-up (if country of origin and country of follow-up is the same, then patient was classified as native; otherwise as migrant)
- Age at Drug Resistance Test—based on the difference between year of birth and date of the first drug resistance test;
- Region of Origin—based on country of origin;
- Treatment Status at Date of First Drug Resistance Test—based on the difference between sample collection date for first drug resistance test and start date of first therapy:

  ART-naïve → patients who had a sample collection date for first drug resistance test before the start date of first therapy
  ART-experienced → patients who had a sample collection date for first drug resistance test after the start date of first therapy

- Recentness of infection—based on ambiguity rate of genomic sequences. We defined as Chronic if the ambiguity rate was higher than 0.45% otherwise was defined as Recent infection, as previously described (Andersson et al., 2013).

LP vs NLP at HIV diagnosis- based on CD4 count, LP were defined as patients with a baseline CD4 count $\leq$ 350 cells/mm^3 and NLP were defined as patients with baseline CD4 count > 350 cells/mm^3 (Antinori et al., 2011).

Statistical Analysis

The proportion and median [interquartile range (IQR)] were calculated for every categorical and continuous variable, respectively. The treatment status variable was compared with the categorical variables with the Chi-square test and continuous variables with the Mann-Whitney U test. Also, we analyzed the trends over time for the overall TDR and ADR through logistic regression models. Data was analyzed using RStudio (Version 1.2.5033).

RESULTS

Characteristics of European Population

Among the 26,973 HIV-1 infected patients included in the analysis, 11,581 (42.9%) were ART-naïve patients and 15,392 (57.1%) were ART-experienced patients. Other socio-demographic characteristics of the population of patients has been analyzed and described in "Determinants of Determinants of HIV-1 Late Presentation in Patients Followed in Europe" (Miranda et al., 2021).

In the total population, the median age was 37 (IQR: 27.0–45.0) years old and 72.6% of HIV-1 infected patients were males. The main transmission route was through heterosexual contact (34.9%) and 81.7% were originated from Western Europe. The most prevalent subtype observed in this population was subtype B (71.9%). Most patients included in this study were native (77.4%) and as having chronic infection (63.6%) based on the ambiguity rate of the first genomic sequence. CD4 count at diagnosis and viral load at diagnosis (log10) presented a median of 318 cells/mm^3 (IQR 151–513) and log10 4.3 copies/mL (IQR 3.3–5.0), respectively.

54.8% of patients were classified as LP (CD4 < 350 cells/mm^3). In ART-naïve patients, 52.8% were LP, meanwhile in ART-experienced patients, 56.4% were LP at time of diagnosis (**Table 1**).

[1]https://www.genomedetective.com/app/typingtool/hiv
[2]https://comet.lih.lu
[3]http://classic.datamonkey.org/dataupload_scueal.php

TABLE 1 | Sociodemographic and clinical patient characteristics.

Patient characteristics	Total	ART-naive	ART-experienced	*p*-value
Total, *n* (%)	26973 (100)	11581 (42.9)	15392 (57.1)	
Gender, *n* (%)	26475 (98.2)	11458 (43.3)	15017 (56.7)	
Male	19224 (72.6)	8797 (76.8)	10427 (69.4)	$p < 0.001$
Female	7251 (27.4)	2661 (23.2)	4590 (30.6)	
Median age at resistance test in years IQR, *n* (%)	26973 (100)	11581 (42.9)	15392 (57.1)	$p < 0.001$
	37.0 (27.0–45.0)	37.0 (30.0–45.0)	37.0 (0.0–44.0)	
≤18	5047 (18.7)	761 (6.6)	4286 (27.8)	$p < 0.001$
19–30	3423 (12.7)	2468 (21.3)	955 (6.2)	
31–55	16707 (61.9)	7472 (64.5%)	9235 (60.0)	
≥56	1796 (6.7)	880 (7.6)	916 (6.0)	
Transmission route, *n* (%)	18118 (67.2)	8336 (46.0)	9782 (54.0)	
Heterosexual	6326 (34.9)	3130 (37.5)	3196 (32.7)	$p < 0.001$
MSM	6124 (33.8)	3863 (46.3)	2261 (23.1)	
IDU	4370 (24.1)	838 (10.1)	3532 (36.1)	
Other	1298 (7.2)	505 (6.1)	793 (8.1)	
Region of origin, *n* (%)	19881 (73.7)	9460 (47.6)	10421 (52.4)	
Western Europe	16249 (81.7)	7436 (78.6)	8813 (84.6)	$p < 0.001$
Eastern Europe	554 (2.8)	377 (4.0)	177 (1.7)	
Africa	2109 (10.6)	1051 (11.1)	1058 (10.2)	
South America	611 (3.1)	338 (3.6)	273 (2.6)	
Other	358 (1.8)	258 (2.7)	100 (1.0)	
Migrant status, *n* (%)	19881 (73.7)	9460 (47.6)	10421 (52.4)	
Migrant	4494 (22.6)	2616 (27.7)	1878 (18.0)	$p < 0.001$
Native	15387 (77.4)	6844 (72.3)	8543 (82.0)	
Recentness of infection, *n* (%)	26973 (100)	11581 (42.9)	15392 (57.1)	
Chronic	17151 (63.6)	6915 (59.7)	10236 (66.5)	$p < 0.001$
Recent	9822 (36.4)	4666 (40.3)	5156 (33.5)	
Subtype, *n* (%)	26973 (100)	11581 (42.9)	15392 (57.1)	
HIV-1 Subtype B	19387 (71.9)	8047 (69.5)	11340 (73.7)	$p < 0.001$
HIV-1 Subtype non-B	7586 (28.1)	3534 (30.5)	4052 (26.3)	
Median (IQR) CD4 count at diagnosis (cells/mL), *n* (%)	24442 (90.6)	10937 (44.7)	13505 (55.3)	$p < 0.001$
	318.0 (151.0–513.0)	332.0 (160.0–518.0)	306.0 (147.0–508.5)	
LP	13390 (54.8)	5776 (52.8)	7614 (56.4)	$p < 0.001$
NLP	11052 (45.2)	5161 (47.2)	5891 (43.6)	
Viral Load at diagnosis (log10 copies/mL), *n* (%), IQR	14005 (51.9)	4589 (32.8)	9416 (67.2)	$p < 0.001$
	4.3 (3.3–5.0)	4.6 (3.8–5.3)	4.1 (3.2–4.9)	
≤4.0	5814 (41.5)	1410 (30.7)	4404 (46.8)	$p < 0.001$
4.1–5.0	4573 (32.7)	1580 (34.4)	2993 (31.8)	
≥5.1	3618 (25.8)	1599 (34.8)	2019 (21.4)	

Transmitted and Acquired Drug Resistance

The overall prevalence of TDR was 12.8% (95%CI: 12.2–13.4%). NRTI, NNRTI and PI TDR were detected in 8.2% (95%CI: 7.7–8.7%), 5.6% (95%CI: 5.2–6.0%) and 3.7% (95%CI: 3.4–4.1%) of ART-naïve patients, respectively. 9.1% (95%CI: 8.6–9.7%) of these patients presented single class resistance, 2.8% (95%CI: 2.5–3.1%) presented dual class resistance and 0.9% (95%CI: 0.8–1.1%) presented triple class resistance (**Table 2**).

68.5% (95%CI: 67.8–69.2%) of experienced patients presented ADR, with higher drug resistance mutations for NRTI (59.1%; 95%CI: 58.3–59.8%), followed by NNRTI (42.2%; 95%CI: 41.4–43.0%) and by PI (24.2%; 95%CI: 23.5–24.9%). 23.5% (95%CI: 22.8–24.2%) of ART-experienced patients presented single class resistance, 33.0% (95%CI: 32.3–33.8%) presented dual class resistance and 12.0% (95%CI: 11.5–12.5%) presented triple class resistance (**Table 2**).

TDR presented an overall decreasing trend between 1995 and 2019 (p for trend < 0.001; **Table 2** and **Supplementary Data**). The same decreasing trend for TDR was observed for NRTIs, NNRTIs and PIs drug classes (p for trend < 0.001; **Table 2**). TDR between three time-periods (1995–2002; 2003–2010, and 2011–2019) was analyzed and it was observed that the overall TDR decreased from 20.0% to 13.3% to 10.7%. The same happened for every drug class, PIs (8.2% to 3.8% to 2.7% for the three time-periods, respectively), NRTIs (17.0% to 8.9% to 5.4% for the three time-periods, respectively) and for the NNRTIs (8.1% to 6.0% to 4.4% for the three time-periods, respectively). Moreover, between the 2003–2010 time-period, the overall TDR had a statistically significant decreasing trend (OR = 0.87; $p = 0.001$; **Figure 1A**).

TABLE 2 | Proportion of transmitted drug (TDR) and acquired drug resistance (ADR) between 1991 and 2019.

	Transmitted drug resistance (TDR)			Acquired drug resistance (ADR)		
	***n* (%)**	**95% CI**	***p* for trend**	***n* (%)**	**95% CI**	***p* for trend**
Total	11581 (100)			15392 (100)		
Any DRMs	1482 (12.8)	12.2–13.4	<0.001	10543 (68.5)	67.8–69.2	<0.001
NRTI resistance	944 (8.2)	7.7–8.7	<0.001	9089 (59.1)	58.3–59.8	<0.001
NNRTI resistance	644 (5.6)	5.2–6.0	<0.001	6499 (42.2)	41.4–43.0	<0.001
PI resistance	427 (3.7)	3.4–4.1	<0.001	3727 (24.2)	23.5–24.9	<0.001
Single class resistance	1056 (9.1)	8.6–9.7	0.049	3617 (23.5)	22.8–24.2	<0.001
Dual class resistance	319 (2.8)	2.5–3.1	<0.001	5080 (33.0)	32.3–33.8	<0.001
Triple class resistance	107 (0.9)	0.8–1.1	<0.001	1846 (12.0)	11.5–12.5	<0.001
PI + NRTI resistance	115 (1.0)	0.8–1.2	<0.001	1671 (10.9)	10.4–11.4	<0.001
PI + NNRTI resistance	13 (0.1)	0.07–0.2	0.452	63 (0.4)	0.3–0.5	0.179
NRTI + NNRTI resistance	191 (1.6)	1.4–1.9	<0.001	3346 (21.7)	21.1–22.4	<0.001

p-value for trend of TDR and ADR between 1995 and 2019. DRM, drug resistance mutations; NRTI, nucleotide reverse transcriptase inhibitors; NNRTI, non-nucleotide reverse transcriptase inhibitors; PI, protease inhibitors; CI, confidence interval.

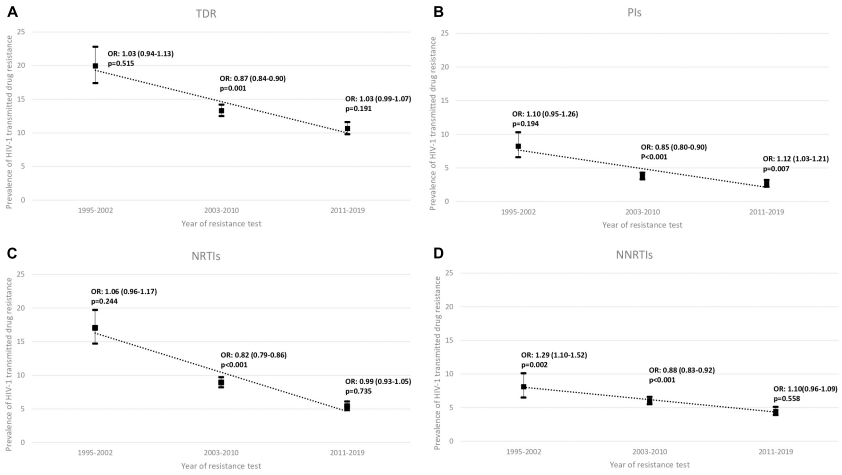

FIGURE 1 | Proportion of **(A)** overall transmitted drug resistance (TDR), **(B)** of protease inhibitors (PIs), **(C)** of nucleoside reverse transcriptase inhibitor (NRTIs) and **(D)** of non-nucleoside reverse transcriptase inhibitor (NNRTIs) in sequences from drug-naïve patients between three periods 1995–2002, 2003–2010, and 2011–2019. OR, Odds Ratio; *p*, *p*-value.

For the same time-period, the ARV drug classes also showed a decreasing trend, PI (OR = 0.85; $p < 0.001$), NNRTIs (OR = 0.82; $p < 0.001$) and NNRTIs (OR = 0.88; $p < 0.001$; **Figures 1A–D**).

Regarding the overall ADR trend, it has been decreasing over the three time-periods (80.0% to 70.7% to 44.5%) as well as in all drug classes studied except for NNRTIs (**Figure 1A**). PIs decreased from 36.3% to 24.8% to 5.9% and NRTIs decreased from 74.3% to 61.4% to 29.8%. Conversely, NNRTIs increased from 36.9% to 47.0% and then decreased to 31.4%. In the last time-period, 2011–2019, the overall ADR showed a decreasing trend (OR = 0.96; $p = 0.018$). The drug classes, in the same time-period, also showed a decreasing trend, but without being statistically significant PIs (OR = 0.94; $p = 0.092$), NRTIs (OR = 0.97; $p = 0.163$) and NNRTIs (OR = 0.98; $p = 366$; **Figure 2A–D**).

Differences in TDR and ADR prevalence between different countries included in this study were also analyzed between two time-periods (2008–2012 and 2013–2018). In our study population, in the first time-period (2008–2012), Luxembourg had the higher rate of TDR (16.8%). This scenario changed for TDR when the last time-period (2013–2018) was analyzed, since Germany (13.9%) presented the highest TDR rate. Comparing

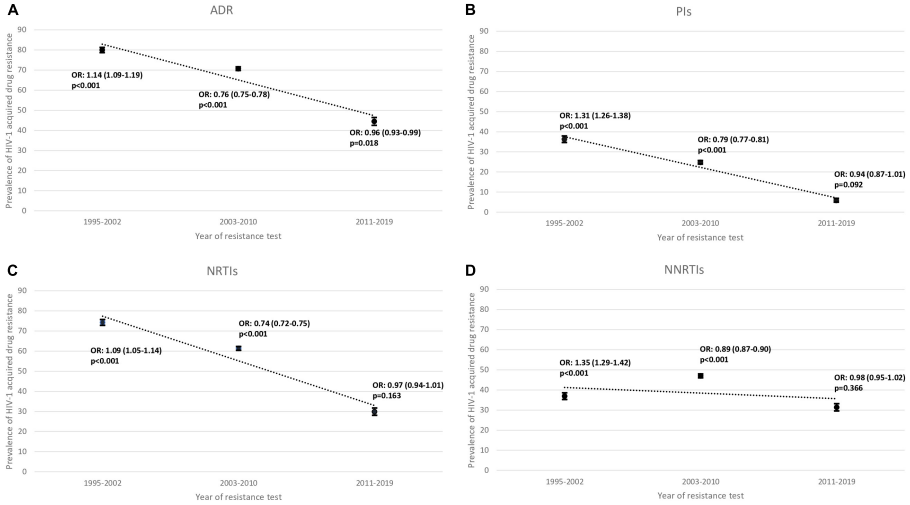

FIGURE 2 | Proportion of **(A)** overall acquired drug resistance (ADR), **(B)** of protease inhibitors (PIs), **(C)** of nucleoside reverse transcriptase inhibitor (NRTIs) and **(D)** of non-nucleoside reverse transcriptase inhibitor (NNRTIs) in sequences from drug-experienced patients between three periods 1995–2002, 2003–2010 and 2011–2019. OR, Odds Ratio; *p*, *p*-value.

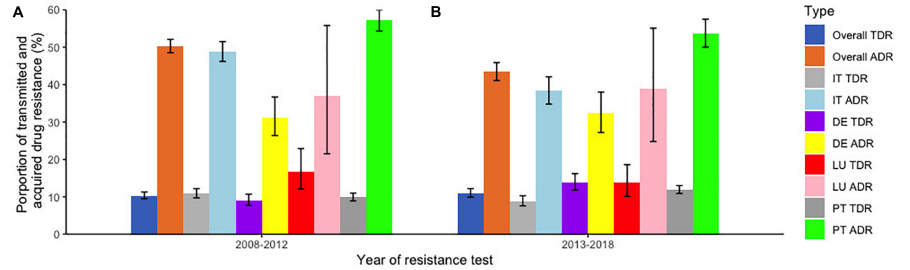

FIGURE 3 | Proportion of transmitted and acquired drug resistance per country of follow-up in two different time periods. **(A)** Between 2008 and 2012; **(B)** between 2013 and 2018. TDR, Transmitted drug resistance; ADR, Acquired drug resistance; IT, Italy; DE, Germany; LU, Luxembourg; PT, Portugal.

each country in those two time-periods, the TDR rate of Italy and Luxembourg decreased from one period to another (10.9% to 8.8%; 16.8% to 13.8%, respectively), while the rates of Germany and Portugal increased (9.9% to 11.9%; 9.1% to 13.9%, respectively).The ADR rates for the first time-period, indicated that all the countries, with the exception of Portugal (57.2%), presented a ADR lower than 50% (**Figure 3A**) and for the last time-period Portugal maintained the highest rate (53.7%; **Figure 3B**). Comparing the ADR rates between the same time-periods, the rate of Italy and Portugal decreased from one period to another (48.9% to 38.4%; 57.2% to 53.7%, respectively), while the rates of Germany and Luxembourg increased (31.3% to 32.4%; 37% to 38.9%, respectively; **Figure 3**).

Transmitted and Acquired Drug Resistance Among Late Presenters and Non-late Presenters

Focusing now on the LP and NLP population, we observed a TDR of 12.3% (95%CI: 11.5–13.2) for LP population and 12.6% (95%CI: 11.8–13.6) for NLP population. In relation to drug resistance classes, the rates of resistance were higher in

TABLE 3 | Proportion of transmitted drug (TDR) and acquired drug resistance (ADR) in Late Presenters (LP) and Non-Late Presenters (NLP) between 1991 and 2019.

Transmitted drug resistance (TDR)	Late presenters (LP)		Non-late presenters (NLP)	
	n (%)	95% CI	*n* (%)	95% CI
Total	5776 (100)		5161 (100)	
Any DRMs	710 (12.3)	11.5–13.2	652 (12.6)	11.8–13.6
NRTI resistance	446 (7.7)	7.1–8.4	428 (8.3)	7.6–9.1
NNRTI resistance	317 (5.5)	4.9–6.1	269 (5.2)	4.6–5.9
PI resistance	202 (3.5)	3.1–4.0	191 (3.7)	3.2–4.3
Acquired drug resistance (ADR)				
Total	7614 (100)		5891 (100)	
Any DRMs	5319 (69.9)	68.8–70.9	4016 (68.2)	67.0–69.3
NRTI resistance	4588 (60.3)	59.2–61.4	3538 (60.1)	58.6–61.1
NNRTI resistance	3354 (44.1)	42.9–45.2	2327 (39.5)	38.3–40.8
PI resistance	2047 (26.9)	25.9–27.9	1328 (22.5)	21.5–23.6

DRM, drug resistance mutations; NRTI, nucleotide reverse transcriptase inhibitors; NNRTI, non-nucleotide reverse transcriptase inhibitors; PI, protease inhibitors; CI, confidence interval.

the NLP when compared to LPs, except for the NNRTIs class. LP presented higher rates of ADR—69.9% (95%CI: 68.8–70.9)—when compared to NLP: 68.2% (95%CI: 67.0–69.3). Contrary to TDR, the rates of ADR were higher in LP when compared to NLP (**Table 3**).

In both LP and NLP populations, the NNRTIs class K103N/S mutation presented the highest prevalence (3.1%; **Figure 4**). For PIs, M46I/L was more prevalent (1.5% for both LP and NLP) followed by L90M (1.4% for LP and 1.2% for NLP). Futhermore, in the PIs class there were two mutations present in LP (I47VA and V32I, respectively), that were not present in NLP (**Figure 4**). In the NLP, for NRTIs, we observed that M41I/L (3.2%) was the mutation with highest prevalence, followed by T215 revertants (3.0%) and by D67N/G/E and M184I/V (2.5%). Conversely, in the LP population, T215 revertants were more prevalent (3.2%), followed by M41I/L (2.4%) and M184I/V (2.3%).

Drug resistance mutations in ART-experienced patients in both LP and NLP populations were also analysed and compared (**Figure 5**). The more prevalent mutations consistently presented higher prevalences in LPs than in NLPs. Similarly to ART-naïve patients, for NNRTIs drug class, K103N/S mutation presented the highest prevalence (21.0% in LP and 19.0%, in NLP; **Figure 5**). For NRTIs, M184I/V had the highest prevalence (42.5% for LP and 41.7% for NLP). In the PIs class, the mutations with higher prevalence were L90M (11.8% NLP and 14.3% LP) and M46I/L (9.4% for NLP and 12.4% for LP). Also, K238TN mutation from the NNRTIs class was present only in the LP population. The presence of these mutations could lead to reduced susceptibility to some specific ARV.

Analysis of Mutations Per Subtype Among Late Presenters and Non-late Presenters Patients

Finally, we compared mutations in LP and NLP, according to subtype B and non-B subtypes. As we can see in **Figure 6**, for subtype B ART-naïve patients, for both NRTIs and NNRTIs, most mutations—except T215rev—were more prevalent in NLP when compared to LP. K103N/S mutation was the one

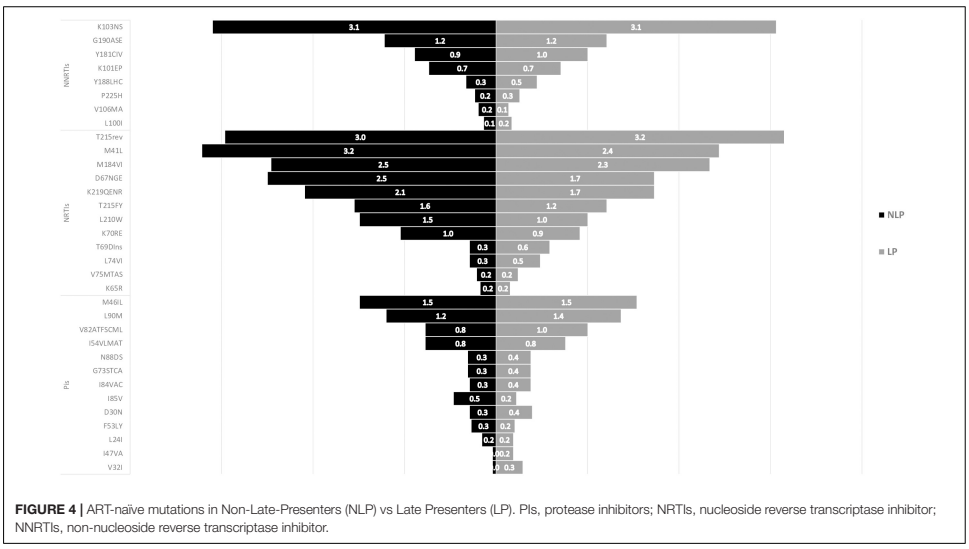

FIGURE 4 | ART-naïve mutations in Non-Late-Presenters (NLP) vs Late Presenters (LP). PIs, protease inhibitors; NRTIs, nucleoside reverse transcriptase inhibitor; NNRTIs, non-nucleoside reverse transcriptase inhibitor.

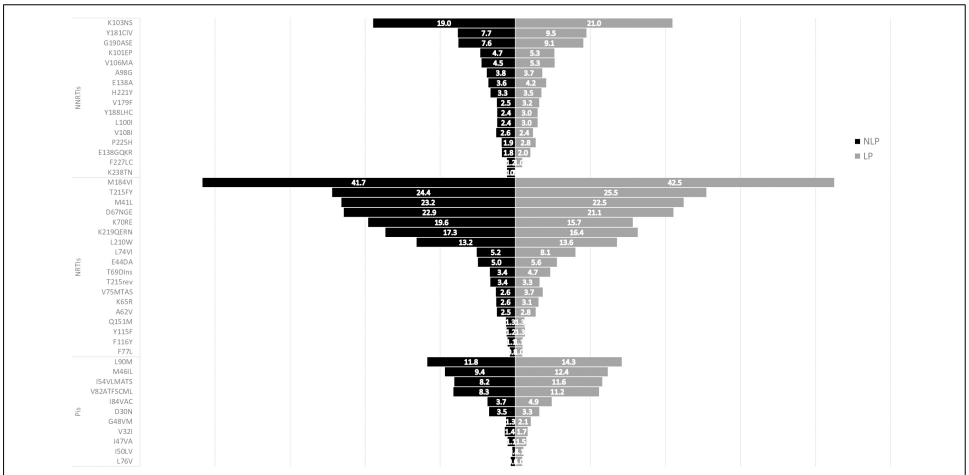

FIGURE 5 | ART-experienced mutations in Non-Late-Presenters (NLP) vs Late Presenters (LP). PIs, protease inhibitors; NRTIs, nucleoside reverse transcriptase inhibitor; NNRTIs, non-nucleoside reverse transcriptase inhibitor.

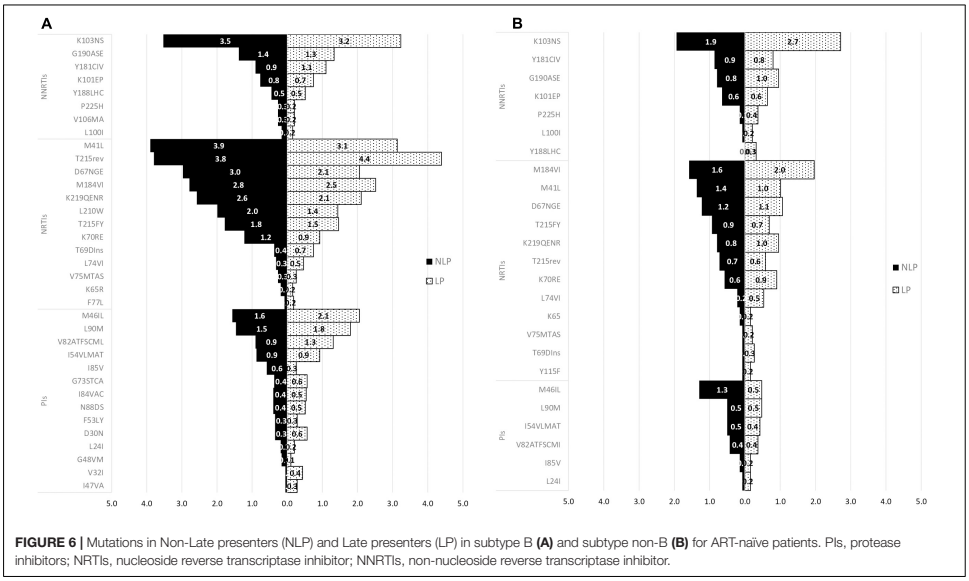

FIGURE 6 | Mutations in Non-Late presenters (NLP) and Late presenters (LP) in subtype B **(A)** and subtype non-B **(B)** for ART-naïve patients. PIs, protease inhibitors; NRTIs, nucleoside reverse transcriptase inhibitor; NNRTIs, non-nucleoside reverse transcriptase inhibitor.

with higher prevalence for NNRTIs (3.5% for NLP and 3.2% for LP). For NRTIs, M41L was the mutation with highest prevalence (3.9% for NLP vs 3.1% for LP), while for LP it was T215rev mutation (4.4% LP vs 3.8% NLP). For the PIs class, conversely, M46I/L and L90M were the mutations with the highest prevalence with higher prevalence in LP compared to

NLP (1.6 and 1.5% for NLP and 2.1 and 1.8% for LP, respectively; **Figures 6A,B**).

Regarding the non-B subtypes, K103N/S mutation was more prevalent in LP compared to NLP (2.7 vs 1.9%, respectively) which was the one with the highest prevalence. For NRTIs, M184V/I, M41L and D67NGE mutations (1.6, 1.4, and 1.2% for NLP and 2.0, 1.0, and 1.1 for LP, respectively) were the ones with higher prevalence. For PIs, M46I/L (1.3% for NLP and 0.5% for LP) was the one with the higher prevalence (**Figures 6A,B**). Comparing both populations regarding subtype non-B, opposite to what happens in subtype B, we observed that the LP population carried higher a prevalence of the most prevalent mutations (**Figures 6A,B**). Also, the K103N/S and the M184V/I were the mutations that were present in more non-B subtypes in the LP population, while the M46I/L was the one for the NLP populations. The most prevalent non-B subtype was subtype C (data not shown).

In ART-experienced patients, both in subtype B and in non-B subtypes, the most prevalent mutations occurred more frequently in LP than in NLP. For NNRTIs class K103N/S mutation had the highest prevalence in both NLP and LP (18.8 and 20.5%, respectively). For NRTIs the mutation with the highest prevalence was M184V/I mutation (43.6% for NLP and 43.9% for LP), and for PIs L90M and M46I/L were the mutations with the highest prevalence (12.7 and 10.4% for NLP and 16.7 and 14.2% for LP, respectively; **Figures 7A,B**).

Regarding the non-B subtypes, similiar to subtype B, K103N/S mutation (19.7% for NLP and 22.3% for LP) for NNRTIs, and M184I/V (33.1% for NLP and 38.8% LP) for NRTIs, were the ones with the highest prevalence. While in the PIs class, I54VLMATS (7.5% for NLP and 8.8% for LP) and L90M mutations (7.3% for NLP and 7.4% for LP) were the ones with the higher prevalence (**Figures 7A,B**). Also, M184V/I was the mutation that was present in the most diversity and proportion of non-B subtypes in both NLP and LP populations. The most prevalent non-B subtype was subtype G (data not shown).

DISCUSSION

There are no recent studies with updated information regarding TDR and ADR prevalence in Europe and the most recent study about this topic only includes TDR and is based on the median overall values from different studies (Rhee et al., 2020). In our study, we presented updated information of the prevalence of TDR and ADR in the overall population and compared its patterns between LP and NLP. Overall, TDR had a prevalence of 12.8% and ADR of 68.5%. The TDR and ADR prevalence from our study was slightly higher when compared to other studies and this could be explained by the fact that our timeline includes patients diagnosed between 1981 and 2019 (Tostevin et al., 2017; Zazzi et al., 2018). Regarding the overall trends, both TDR and ADR presented a decreasing trend, consistently with other studies in and outside of Europe (Schmidt et al., 2014; Rocheleau et al., 2018).

We also compared TDR and ADR for the countries of follow-up included in the database divided into two time periods (2008–2012 and 2013–2018). For Italy, TDR prevalence decreased within time-periods (2008–2012:10.9% and 2013–2018: 8.8%), which is in accordance with studies from that country and around the same timeline (Franzetti et al., 2018; Rossetti et al., 2018). The prevalence of ADR also decreased

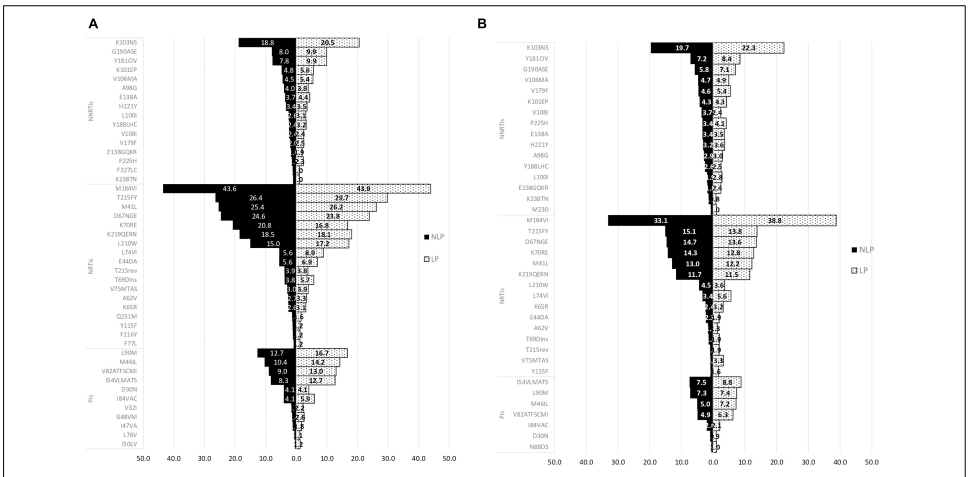

FIGURE 7 | Mutations in Non-Late presenters (NLP) and Late presenters (LP) in subtype B **(A)** and subtype non-B **(B)** for ART-experienced patients. PIs, protease inhibitors; NRTIs, nucleoside reverse transcriptase inhibitor; NNRTIs, non-nucleoside reverse transcriptase inhibitor.

in Italy (2008–2012: 48.9% and 2013–2018; 38.4%), and these results are slightly lower than those from a study from the Italian ARCA database. Moreover, the decrease in the last 5 years is in accordance with that study (Lombardi et al., 2021). For Germany, TDR prevalence was 9.1% and ADR prevalence was 31.3% between 2008 and 2012, and for a similar time-period, the TDR rate was around the same, but our ADR rate was lower than in another study reported in this country (Schmidt et al., 2014). For Luxembourg, the TDR prevalence was 16.8% and the ADR prevalence was 37% between 2008 and 2012, which is higher when compared to the values in Europe (Hofstra et al., 2016). For Portugal, TDR prevalence increased between time-periods (2008–2012:9.9% and 2013–2018: 11.9%), while ADR prevalence decreased between the same time-periods (2008–2012: 57.2% and 2013–2018: 53.7%). The TDR prevalence in the first time-period was closer to the one from a study conducted in Portugal between 2001 and 2017 and that same study indicated an increase trend for TDR. Our ADR prevalence for Portugal in the first time-period, had a lower value than the overall ADR prevalence from that study, although the decreasing trend was concordant (Pingarilho et al., 2020).

We also compared drug resistance in LP vs NLP, both in ART-naïve or ART-experienced patients. There were no major differences in the prevalence of drug resistance mutations in both LP and NLP from the ART-naïve population. However, LPs presented a lower prevalence of TDR than NLP, potentially suggesting a reversion of these mutations when patients are diagnosed late. The most prevalent mutations were the K103N/S, T215 revertants, the M184V/I, the M41I/L, the M46I/L and the L90M. However, in the LP, there were two mutations—I47V/A and V32I—that were not present in the NLP. Despite the lack of significance of these findings, we were not expecting to find mutations occurring specifically in late presenters, that could eventually indicate the irreversible fixation of these mutations in some cases, where they are not associated with a fitness cost (Winand et al., 2015; Nagaraja et al., 2016). In the ART-experienced population, there were also no significant differences between the LP and NLP populations, however, LPs presented a higher prevalence of ADR compared to NLP. The most prevalent mutations among LP and NLP were the K103N/S, the M184IV/I, the L90M and M46I/L. The K103N/S mutation presented similar prevalence in LP and NLP in ART-naïve, while ART-experienced LP had higher prevalence compared to NLP (Hiv Drug Resistance Database). T215rev in drug naïve patients was more prevalent in LP compared to NLP.The NRTIs T215rev mutants is associated with risk of virological failure to zidovidine (AZT) or stavudine (d4T). M41I/L impacts negatively virological response to regimens with abacavir (ABC), didanosine (ddl) or tenofovir (TDF). Together, these mutations confer high-level resistance to AZT and d4T. For the same drug class, M184V/I mutation reduces susceptibility to lamivudine (3TC) and emtricitabine (FTC; Hiv Drug Resistance Database). PI mutations were consistently more prevalent in LP compared to NLP, both in experienced and naïve patients, indicating a potential irreversible fixation of these mutations when they occur. The most prevalent were M46I/L which is associated with a reduction in the susceptibility to atazanavir (ATV), fosamprenavir (FPV), indinavir (IDV), lopinavir (LPV) and NFV, and L90M which is associated to reduced susceptibility to almost all PIs, except for tipranavir (TPV) and darunavir (DRV; Hiv Drug Resistance Database).

It is known that some mutations are closely related to specific subtypes and recombinant forms. As such, we conducted a final analysis distinguishing the patterns found in subtype B when compared to non-B subtypes. The most prevalent subtype was subtype B and the mutation with the highest prevalence in NLP ART-naïve patients was M41L from the NRTIs drug class. This result is in accordance with a study of mutations according to subtypes in Brazil (Westin et al., 2011).

In the LP and NLP patients, in the ART-experienced population, for both subtypes B and non-B, M184V/I mutation was the one with the higher prevalence.

This study was the first to analyze and compare transmitted and ADR in LP and NLP populations. Despite the lack of significant differences, we consistently found higher levels of TDR in NLP and higher levels of ADR in LP. We find this pattern consistent, except for non-B subtypes and the PIs class. This suggests different dynamics of reversion and irreversible fixation of mutations that should be further investigated in future studies.

Limitations

Our study had some limitations. For example, concerning the analysis time-period, the first years and the more recent ones can be a bias in the analysis, since the number of individuals of those years is low compared to other years of resistance test collection date. Also, our population is mainly from Western Europe, providing a certain imbalance when characterizing the population and the TDR and ADR origins regarding geographical distribution. Another limitation of our study is the definition of LP as there is lack of consensus as to whether this definition ("baseline CD4 count in newly diagnosed patient is lower than 350 cells/mm^3 or has an AIDS-defining event, regardless of CD4 cell count") is the correct one to characterize those who present late to diagnosis. Some discuss that the threshold should be CD4 count lower than 200 cells/mm^3, i.e., those characterized in LP with advanced disease.

CONCLUSION

In conclusion, our study showed that the overall TDR and ADR had a decreasing trend and the prevalence has been steady through the years. There were no significant differences in the TDR rate between the LP and NLP (around 12% in both), with slightly higher levels in the NLP. The mutation profile was also similar, again with most mutations presenting a higher prevalence of TDR in NLP and higher prevalence of ADR in LP. Late presentation for HIV remains a key unresolved challenge in HIV/AIDS with serious adverse consequences at the individual and societal levels. Our study highlights ADR and TDR patterns and drug resistance mutations, alone and according to subtypes in the LP population, when compared to NLP.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/**Supplementary Material**, further inquiries can be directed to the corresponding author/s.

AUTHOR CONTRIBUTIONS

MNSM, MP, and AA: conceptualization. MNSM, MP, VP, MdROM, and AA: methodology. MNSM, VP: software. MNSM, MP, FI, and AA: validation. MNSM, VP, MP, and AA: formal analysis. MNSM, MP, VP, and MdROM: investigation. CS-D, RP, RK, MZ, and FI: resources. CS-D, RP, RK, MZ, and FI: data curation. MNSM, MP, and AA: writing—original draft preparation. MNSM, MP, FI, and AA: writing—review, and editing. MNSM, MP, VP, MdROM, and AA: visualization. AA: supervision, project administration, and funding acquisition. All authors contributed to the article and approved the submitted version.

FUNDING

This study was financially supported by FCT through the following projects: GHTM (UID/04413/2020), INTEGRIV (PTDC/SAU-INF/31990/2017) and the scholarship (PD/BD/135714/2018), and Gilead Génese HIVLatePresenters.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmicb.2022.846943/full#supplementary-material

REFERENCES

Andersson, E., Shao, W., Bontell, I., Cham, F., Cuong, D. D., Wondwossen, A., et al. (2013). Evaluation of sequence ambiguities of the HIV-1 pol gene as a method to identify recent HIV-1 infection in transmitted drug resistance surveys. *Infect. Genet. Evol.* 18, 125–131. doi: 10.1016/j.meegid.2013.03.050

Antinori, A., Coenen, T., Costagiola, D., Dedes, N., Ellefson, M., Gatell, J., et al. (2011). Late presentation of HIV infection: a consensus definition. *HIV Med.* 12, 61–64. doi: 10.1111/j.1468-1293.2010.00857.x

Bennett, D. E., Camacho, R. J., Otelea, D., Kuritzkes, D. R., Fleury, H., Kiuchi, M., et al. (2009). Drug resistance mutations for surveillance of transmitted HIV-1 drug-resistance: 2009 update. *PLoS One* 4:e4724. doi: 10.1371/journal.pone.0004724

Cdc *Treatment | Living with HIV | HIV Basics | HIV/AIDS | CDC*. Available online at: https://www.cdc.gov/hiv/basics/livingwithhiv/treatment.html. [accessed on May 20, 2021].

Clutter, D. S., Jordan, M. R., Bertagnolio, S., and Shafer, R. W. (2016). HIV-1 drug resistance and resistance testing. *Infect. Genet. Evol.* 46, 292–307. doi: 10.1016/j.meegid.2016.08.031

EuResist *Euresist Data Analysis - database*. Available online at: http://engine.euresist.org/database/. [accessed on Jan 04, 2021].

Franzetti, M., De Luca, A., Ceccherini-Silberstein, F., Spagnuolo, V., Nicastri, E., Mussini, C., et al. (2018). Evolution of HIV-1 transmitted drug resistance in Italy in the 2007–2014 period: A weighted analysis. *J. Clin. Virol.* 106, 49–52. doi: 10.1016/j.jcv.2018.07.009

Günthard, H. F., Calvez, V., Paredes, R., Pillay, D., Shafer, R. W., Wensing, A. M., et al. (2019). Human Immunodeficiency Virus Drug Resistance: 2018 Recommendations of the International Antiviral Society–USA Panel. *Clin. Infect. Dis.* 68, 177–187. doi: 10.1093/cid/ciy463

Hiv Drug Resistance Database *NNRTI Resistance Comments - HIV Drug Resistance Database*. Available online at: https://hivdb.stanford.edu/dr-summary/comments/NNRTI/. [accessed on February 22, 2021].

Hofstra, L. M., Sauvageot, N., Albert, J., Alexiev, I., Garcia, F., Struck, D., et al. (2016). Transmission of HIV drug resistance and the predicted effect on current first-line regimens in Europe. *Clin. Infect. Dis.* 62, 655–663. doi: 10.1093/cid/civ963

Joint United Nations Programme on Hiv/Aids (Unaids) (2015). *Understanding Fast-Track Targets: accelerating action to end the AIDS epidemic by 2030*. (Geneva: UNAIDS).

Lawyer, G., Altmann, A., Thielen, A., Zazzi, M., Sönnerborg, A., and Lengauer, T. (2011). HIV-1 mutational pathways under multidrug therapy. *AIDS Res. Ther.* 8:26. doi: 10.1186/1742-6405-8-26

Lombardi, F., Giacomelli, A., Armenia, D., Lai, A., Dusina, A., Bezenchek, A., et al. (2021). Prevalence and factors associated with HIV-1 multi-drug resistance over the past two decades in the Italian ARCA database. *Int. J. Antimicrob. Agents* 57:106252. doi: 10.1016/j.ijantimicag.2020.106252

Miranda, M. N. S., Pingarilho, M., Pimentel, V., Martins, M. D. R. O., Vandamme, A.-M., Bobkova, M., et al. (2021). Determinants of HIV-1 Late Presentation in Patients Followed in Europe. *Pathogens* 10:835. doi: 10.3390/pathogens10070835

Nagaraja, P., Alexander, H. K., Bonhoeffer, S., and Dixit, N. M. (2016). Influence of recombination on acquisition and reversion of immune escape and compensatory mutations in HIV-1. *Epidemics* 14, 11–25. doi: 10.1016/j.epidem.2015.09.001

Nih *Drug-Resistance Testing | NIH*. Available online at: https://clinicalinfo.hiv.gov/en/guidelines/adult-and-adolescent-arv/drug-resistance-testing. [accessed on October 25, 2018].

Pineda-Peña, A. C., Faria, N. R., Imbrechts, S., Libin, P., Abecasis, A. B., Deforche, K., et al. (2013). Automated subtyping of HIV-1 genetic sequences for clinical and surveillance purposes: Performance evaluation of the new REGA version 3 and seven other tools. *Infect. Genet. Evol.* 19, 337–348. doi: 10.1016/j.meegid.2013.04.032

Pingarilho, M., Pimentel, V., Diogo, I., Fernandes, S., Miranda, M., Pineda-Pena, A., et al. (2020). Increasing Prevalence of HIV-1 Transmitted Drug Resistance in Portugal: Implications for First Line Treatment Recommendations. *Viruses* 12:1238. doi: 10.3390/v12111238

Rhee, S., Kassaye, S. G., Barrow, G., Sundaramurthi, J. C., Jordan, M. R., and Shafer, R. W. (2020). HIV-1 transmitted drug resistance surveillance: shifting trends in study design and prevalence estimates. *J. Int. AIDS Soc.* 23:e25611. doi: 10.1002/jia2.25611

Rocheleau, G., Brumme, C. J., Shoveller, J., Lima, V. D., and Harrigan, P. R. (2018). Longitudinal trends of HIV drug resistance in a large Canadian cohort, 1996–2016. *Clin. Microbiol. Infect.* 24, 185–191. doi: 10.1016/j.cmi.2017.06.014

Rossetti, B., Di Giambenedetto, S., Torti, C., Postorino, M. C., Punzi, G., Saladini, F., et al. (2018). Evolution of transmitted HIV-1 drug resistance and viral subtypes circulation in Italy from 2006 to 2016. *HIV Med.* 19, 619–628. doi: 10.1111/hiv.12640

Schmidt, D., Kollan, C., Fätkenheuer, G., Schülter, E., Stellbrink, H. J., Noah, C., et al. (2014). Estimating trends in the proportion of transmitted and acquired HIV drug resistance in a long term observational cohort in Germany. *PLoS One* 9:e104474. doi: 10.1371/journal.pone.0104474

Struck, D., Lawyer, G., Ternes, A. M., Schmit, J. C., and Bercoff, D. P. (2014). COMET: Adaptive context-based modeling for ultrafast HIV-1 subtype identification. *Nucleic Acids Res.* 42:e144. doi: 10.1093/nar/gku739

Tostevin, A., White, E., Dunn, D., Croxford, S., Delpech, V., Williams, I., et al. (2017). Recent trends and patterns in HIV-1 transmitted drug resistance in the United Kingdom. *HIV Med.* 18, 204–213. doi: 10.1111/hiv.12414

Unaids *2025 AIDS TARGETS - UNAIDS*. Available online at: https://aidstargets2025.unaids.org/#section-targets. [accessed on Aug 03, 2021].

Westin, M. R., Biscione, F. M., Fonseca, M., Ordones, M., Rodrigues, M., Greco, D. B., et al. (2011). Resistance-Associated Mutation Prevalence According to Subtypes B and Non-B of HIV Type 1 in Antiretroviral-Experienced Patients in Minas Gerais, Brazil. *AIDS Res. Hum. Retroviruses* 27, 981–987. doi: 10.1089/aid.2010.0260

Winand, R., Theys, K., Eusébio, M., Aerts, J., Camacho, R. J., Gomes, P., et al. (2015). Assessing transmissibility of HIV-1 drug resistance mutations from treated and from drug-naive individuals. *AIDS* 29, 2045–2052. doi: 10.1097/QAD.0000000000000811

World Health Organization *Updated recommendations on first-line and second-line antiretroviral regimens and post-exposure prophylaxis and recommendations on early infant diagnosis of HIV*. (Geneva: World Health Organization).

World Health Organization *Global HIV Programme*. Available online at: https://www.who.int/teams/global-hiv-hepatitis-and-stis-programmes/hiv/strategic-information/hiv-data-and-statistics. [accessed on Aug 03, 2021].

Zazzi, M., Hu, H., and Prosperi, M. (2018). The global burden of HIV-1 drug resistance in the past 20 years,". *PeerJ* 6:e4848. doi: 10.7717/peerj.4848

Zazzi, M., Incardona, F., Rosen-Zvi, M., Prosperi, M., Lengauer, T., Altmann, A., et al. (2012). Predicting response to antiretroviral treatment by machine learning: The euresist project. *Intervirology* 55, 123–127. doi: 10.1159/000332008

Zou, X., He, J., Zheng, J., Malmgren, R., Li, W., Wei, X., et al. (2020). Prevalence of acquired drug resistance mutations in antiretroviral- experiencing subjects from 2012 to 2017 in Hunan Province of central South China. *Virol. J.* 17:38. doi: 10.1186/s12985-020-01311-3

Conflict of Interest: FI was employed by IPRO—InformaPRO S.r.l.

The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

The authors declare that this study received funding from Gilead Sciences. The funder was not involved in the study design, collection, analysis, interpretation of data, the writing of this article, or the decision to submit it for publication.

4.4. Manuscript IV

Title: The role of Late Presenters on HIV-1 transmission clusters in Europe

Runing title: HIV-1 TC in LP in Europe

Mafalda N. S. Miranda[1*], Marta Pingarilho[1], Victor Pimentel[1], Perpétua Gomes[2,3], Maria do Rosário O. Martins[1], Sofia G. Seabra[1], Rolf Kaiser[4,5], Michael Böhm[4,5], Carole Seguin-Devaux[6], Roger Paredes[7], Marina Bobkova[8], Maurizio Zazzi[9], Francesca Incardona[10,11] and Ana B. Abecasis[1]

Affiliations:

1- Global Health and Tropical Medicine (GHTM), Institute of Hygiene and Tropical Medicine, New University of Lisbon (IHMT/UNL), 1349-008 Lisbon, Portugal; martapingarilho@ihmt.unl.pt (M.P.); victor.pimentel@ihmt.unl.pt (V.P.); mrfom@ihmt.unl.pt (M.d.R.O.M.); sgseabra@ihmt.unl.pt (S.G.S); ana.abecasis@ihmt.unl.pt (A.A.)
2- Laboratório de Biologia Molecular (LMCBM, SPC, CHLO-HEM), Lisbon, Portugal; pcrsilva@chlo.min-saude.pt (P.G)
3- Centro de Investigação Interdisciplinar Egas Moniz (CiiEM), Instituto Universitário Egas Moniz, Costa da Caparica, Portugal.
4- Institute of Virology, Faculty of Medicine and University Hospital of Cologne, University of Cologne, Germany; rolf.Kaiser@uk-koeln.de (R.K.); michael.boehm@uk-koeln.de (M.B.)
5- DZIF Deutsches Zentrum für Infektionsforschung
6- Laboratory of Retrovirology, Department of Infection and Immunity, Luxembourg Institute of Health, Esch-sur-Alzette, Luxembourg; carole.devaux@lih.lu (C.S.-D.)
7- Infectious Diseases Department and IrsiCaixa AIDS Research Institute, Hospital Universitari Germans Trias i Pujol, Badalona, Spain; rparedes@irsicaixa.es (R.P.)
8- Gamaleya National Research Center of Epidemiology and Microbiology, Moscow, Russia; mrbobkova@mail.ru (M.B.)
9- Department of Medical Biotechnologies, University of Siena, Siena, Italy; maurizio.zazzi@unisi.it (M.Z.)
10- IPRO—InformaPRO S.r.l., Rome, Italy; f.incardona@informa.pro (F.I.)
11- EuResist Network, Rome, Italy

Correspondence:
Corresponding Author
Mafalda N.S. Miranda (MNSM)
mafalda_nsm@hotmail.com

Keywords: HIV-1 infection, Late presenters, Non-Late Presenters, Transmission clusters

Abstract

Background: Investigating the role of late presenters (LP) on HIV-1 transmission is important, as they can contribute to the onward spread of HIV-1 virus in the long period before diagnosis, when they are not aware of their HIV status

Objective: To describe the clinical and socio-demographic characteristics of HIV-1 infected individuals followed in Europe, to characterize patients in clusters and to compare transmission clusters (TC) in LP vs non-late presenters (NLP) populations.

Methods: Clinical, socio-demographic and genotypic information from 38531 HIV-1 infected patients was collected from the EuResist Integrated Database (EIDB) between 1981 and 2019. Sequences were aligned using VIRULIGN. Maximum likelihood (ML) phylogenies were constructed using FastTree. Putative transmission clusters were identified using ClusterPicker v1.332. Statistical analyses were performed using R.

Results: 32652 (84·7%) sequences were from subtype B, 3603 (9·4%) were from subtype G, and 2276 (5·9%) were from subtype A. The median age was 33 (IQR: 26·0-41·0) years old and 75·5% of patients were males. The main transmission route was through homosexual (MSM) contact (36·9%) and 86·4% were originated from Western Europe. Most patients were native (84·2%), 59·6% had a chronic infection, and 73·4% had acquired drug resistance (ADR). CD4 count and viral load at diagnosis (log10) presented a median of 341 cells/mm3, and of log10 4·3 copies/mL, respectively. 51·4% of patients were classified as LP and 21·6% patients were inside TCs. Most patients from subtype B (85·6%) were in clusters, compared to subtypes A (5·2%) and G (9·2%). Phylogenetic analyses showed consistent clustering of MSM individuals. In subtype A, patients in TCs were more frequently MSM patients and with a recent infection. For subtype B, patients in TCs were more frequently those with older age (≥ 56), MSM transmission route, originating from Western Europe, migrants, and with a recent infection. For subtype G, patients in TC were more frequently patients with recent infection and migrants. When analysing cluster size, we found that NLP more frequently belonged to large clusters (>8 patients) when compared to LP.

Conclusion: While late presentation is still a threat to HIV-1 transmission, LP individuals are more present either outside or in small clusters, indicating a limited role of late presentation to HIV-1 transmission.

Introduction

At the end of 2020, there were 37·7 million people living with HIV (1) and it is known that in HIV epidemics, certain risk groups contribute to the spread of HIV disproportionately more than others. This can be due to specific demographic, clinical or behavioral factors or to factors related to the infecting strain of the virus (2,3). On the one hand, the literature suggests that a recent infection, without diagnosis, could be disproportionately associated with transmission and spread of HIV-1 disproportionately (4). On the other hand, late presentation to diagnosis has been increasing over the years and in Europe, late presenters (LP) account for around 50% of HIV new diagnosis (5). Late presenters (LP), based on a definition consensus, are HIV-1 infected individuals defined by a baseline CD4 count lower than 350 cells/mm3 or with an AIDS-defining event, regardless of CD4 cell count (6). Late presentation is associated with high morbidity and mortality, at an individual level, and increased health costs (7). Besides those consequences, LP can also contribute to the onward spread of HIV-1 virus at the population level, as these individuals are not aware of their HIV status and could also spread the virus without knowing their infection status (8).

The use of powerful tools as phylogenetic trees and transmission clusters (TC) are essential to understand the dynamics of viral transmission and to identify groups of individuals connected to each other (2,9).

In this study, we aim to describe the clinical and socio-demographic characteristics of HIV-1 infected individuals followed in Europe according to subtype and to understand the determinants associated with clustering on each of the more prevalent subtypes. The analysis of transmission clusters for the most prevalent HIV-1 subtypes, B, A, and G, and comparison of the patterns of transmission clusters in late presenters (LP) vs non-late presenters (NLP) populations were included in this study.

Methods

Study Group

Clinical, socio-demographic and genomic information from 38531 HIV-1 infected patients from the EuResist Integrated Database (EIDB) between 1981 and 2019 were included in this study. The EuResist integrated database (EIDB) is one of the largest existing datasets which integrate clinical, socio-demographic and viral genotypic information from HIV-1 patients. It integrates longitudinal, periodically updated data mainly from Italy (ARCA database), Germany (AREVIR database) Spain (CoRIS and IRSICAIXA), Sweden, Belgium, Portugal, and Luxembourg databases (10).

In this study, information from the ARCA, AREVIR, Luxembourg, IRSICAIXA, Portugal and Russia databases were used.

Institutional Review Board Statement

All procedures performed in this study were in accordance with the ethical standards of the institutional and/or national research committees and with the 1964 Helsinki declaration. The database enrolled anonymized patients' information, including demographic, clinical and genomic data from patients from the EuResist Integrated Database (Date of approval: 15 January 2021).

Drug Resistance Analysis and Subtyping

HIV pol sequences were derived from existing routine clinical genotypic resistance tests (Sanger method). The size of RT and PR fragments used for this analysis was between 500 and 1000 nucleotides. Only the first HIV genomic sequence per patient was analyzed. TDR was defined as the presence of one or more surveillance drug resistance mutations

in a sequence, according to the WHO 2009 surveillance list (11). The sequences were submitted to the Calibrated Population Resistance tool version 8.0. Clinical resistance to ARV drugs was calculated through the Standford HIVdb version 9.0.
HIV-1 subtyping was performed using the consensus of the result obtained based on three different subtyping tools: Rega HIV Subtyping Tool version 3.46 (https://www.genomedetective.com/app/typingtool/hiv), COMET: adaptive context-based modeling for HIV-1 (https://comet.lih.lu) and SCUEAL ("http://classic.datamonkey.org/dataupload_scueal.php").

Transmission cluster (TC) identification

For the analysis of transmission clusters and construction of phylogenetic trees, the database was divided in three separate datasets, subtype B, A, and G. Control sequences were retrieved from the Los Alamos database and all HIV-1 pol subtype B, A, and G sequences from Europe, South America and Africa were included (http://www.hiv.lanl.gov) (12). We used as the outgroup reference three subtype B and C references retrieved from the Los Alamos database. For each subtype, sequences were aligned against the control sequences dataset using VIRULIGN (13). The HIV-1 K03455.1 (HXB2) pol nucleotide sequence (nt) was used as reference for codon correct alignment. The dataset was then manually edited to exclude sequences with low quality, duplicates and clones using MEGA7 software. The final datasets of subtypes B, A, and G consisted of 62543, 10122, and 5547 sequences, respectively, with a length of 948. Maximum likelihood (ML) phylogenies were constructed in FastTree with the generalized time reversible model. Statistical support of clades was assessed using the Shimodaira-Hasegawa-like test (SH-test). Putative transmission clusters were identified using ClusterPicker v1.332 (14) and defined a threshold that included a genetic distance of 0.030 and a branch support ≥0·90 aLRT. For analyses of cluster size, we defined clusters with 8 patients or more as large clusters and cluster with less than 8 patients as small clusters. The visual configuration of the phylogenetic trees was possible through the iTOL programme.

Study Variables

New variables were created according to:

- Migration Status- Based on Country of Origin and Country of Follow-up (if country of origin and country of follow-up is the same, then patient was classified as native; otherwise as migrant)
- Age at Resistance Test - Based on the difference between Year of Birth and Date of the first drug resistance test;
- Region of Origin- Based on Country of Origin;
- Treatment Status at date of first Drug Resistance Test – based on the difference between sample collection date for first drug resistance test and date of start of first therapy:
  ART-naïve→ patients who had a sample collection date for first drug resistance test before the date of start of first therapy
  ART-experienced→ patients who had a sample collection date for first drug resistance test after the date of start of first therapy
- Recentness of infection - Based on ambiguity rate of genomic sequences. We defined Chronic Infection if the ambiguity rate was higher than 0·45% otherwise Recent infection was defined, as previously described (15).

- LP vs NLP at HIV diagnosis- Based on CD4 count, LP were defined as patients with a baseline CD4 count =< 350 cells/mm3 and NLP were defined as patients with baseline CD4 count > 350 cells/mm3 (6).

Statistical analysis

The proportion and median (interquartile range, IQR) were calculated for every categorical and continuous variable, respectively. The treatment status variable was compared with the categorical variables with Chi-square test, and continuous variables with Mann-Whitney U test. Logistic regression was used to analyze the association between demographic, clinical factors, clustering status and the subtypes. First, we presented the logistic regression with the unadjusted odds ratios (uOR) and confidence intervals at 95% (95% CI), then we included only the variables with a p-value <0·05 in the final model. The final model was adjusted for sex, this variable was forced into the model regardless of its significance. Data was analyzed using RStudio (Version 1.2.5033).

Results

Characteristics of European Population

Among the 38531 HIV-1 infected patients from the EIDB included in the analysis, 32652 (84·7%) were from subtype B, 3603 (9·4%) were from subtype G and 2276 (5·9%) were from subtype A. The median age at resistance test was 40·0 (34·0-47·0) years old and 75·5% of HIV-1 infected patients were males. The main transmission route was through homosexual (MSM) contact (36·9%). For subtype B, the MSM route (40·7%) was also the most prevalent route whereas was the heterosexual route was the predominant route for subtype A and G (43·3% and 49·9%, respectively) (Data not shown). 86·4% of patients were originated from Western Europe and according to subtypes, South America was the most prevalent region of origin for subtype B, whereas subtype A-infected patients were mainly from Eastern Europe region and for G it was Africa region the most prevalent. Most patients included in this study were native (84·2%) and according to subtype, while in subtype B natives were more prevalent (94·8%), in subtype A and G were migrants (13·8% and 16·6%). Based on the ambiguity rate of the first genomic sequence, most patients were classified as presenting a chronic infection (59·6%). Most patients were ART-experienced (59%) and 73·4% had acquired drug resistance (ADR). CD4 count at diagnosis and viral load at diagnosis (log10) presented a median of 341 cells/mm3 (IQR 170-540) and log10 4·3 copies/mL (IQR 3·4-5·0), respectively. 21·6% of patients were represented within transmission clusters and 51·4% of patients were classified as LP (CD4<350 cells/mm3). Most patients from subtype B (85·6%) were located in clusters in contrast to subtypes A (5·2%) and G (9·2%) (Table S1).

Dynamics of subtype A HIV-1 epidemic in Europe

Based on the sequences from our database and the control sequences retrieved, we could observe that the majority of the subtype A population had its origin in Africa, and the major route of transmission was heterosexual. There were some individuals with IDU transmission. The phylogenetic analyses indicates that most EuResist patients cluster in two different parts of the tree, indicated with arrows A and B in figure 1, suggesting two parallel epidemics of subtype A in Europe. The first cluster was related to patients originating from Africa and Eastern Europe through heterosexual and IDU transmission (cluster A) and the other was linked to patients originating from Western Europe with MSM transmission (cluster B). LP individuals are mostly concentrated in cluster A, where the majority of individuals are also migrants (Figure 1).

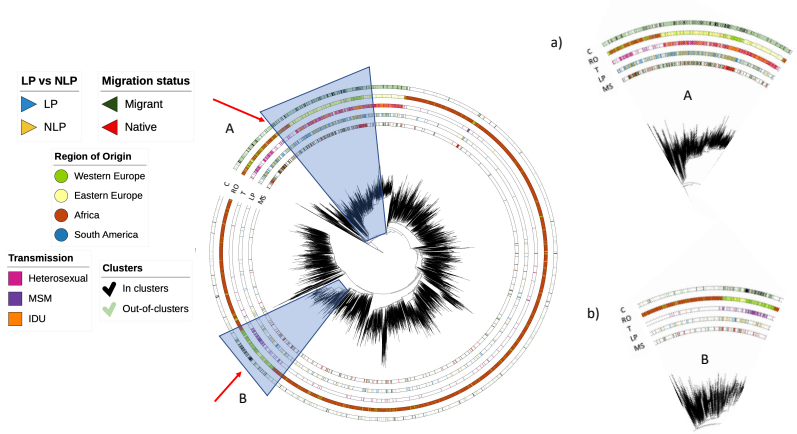

Figure 1. Phylogenetic tree for Subtype A. This image shows a visual phylogenetic tree of the subtype A population. The region A, figure 1.a), shows a clustering of individuals originating from Africa and Eastern Europe with a heterosexual and IDU transmission, the region B, figure 1.b), shows a clustering of individuals originating from Western Europe and MSM transmission. C- Clusters; RO- Region of origin; T- Transmission; LP- Late presenters vs non-late presenters; MS- Migration status

Dynamics of subtype B HIV-1 epidemic in Europe

Based on the sequences from our database and the control sequences retrieved, we could observe that most subtype B patients originated from Western Europe, are native and MSM. Individuals with IDU transmission originating from Western Europe dominate one cluster of the tree (indicated with an arrow). LP and NLP individuals are distributed evenly in the tree. Based on the configuration of the phylogenetic tree and apart from the cluster dominated by IDUs, there seems to be no major compartmentalization patterns in the subtype B epidemic in Europe (Figure 2).

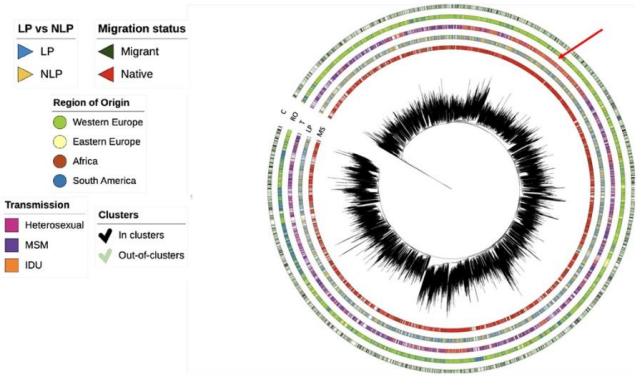

Figura 2. Phylogenetic tree for Subtype B. This image shows a visual phylogenetic tree of the subtype B population. The region highlighted with an arrow shows a clustering of individuals with an IDU transmission and originating from Western Europe. C- Clusters; RO- Region of origin; T- Transmission; LP- Late presenters vs non-late presenters; MS- Migration status

Dynamics of subtype G HIV-1 epidemic in Europe

Based on the sequences from our database and the control sequences retrieved, we could observe two major regions of origin - Western Europe and Africa – compose the subtype G epidemic of HIV-1 in Europe. These are largely divided in two major clusters indicated with arrows A and B in the Figure 3. Most individuals are heterosexuals and LP dominate. The tree configuration indicates lack of compartmentalization of subtype G epidemic in Europe and suggests frequent importations of subtype G.

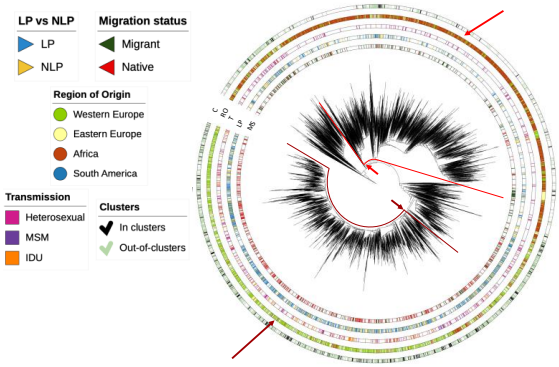

Figure 3. Phylogenetic tree for subtype G. This image shows a visual phylogenetic tree of the subtype G population. The regions highlighted with arrows shows a clustering of individuals originating from Western Europe and Africa. C- Clusters; RO- Region of origin; T- Transmission; LP- Late presenters vs non-late presenters; MS- Migration status

Determinants associated with transmission clusters of HIV-1 in Europe for different subtypes

In the first unadjusted logistic regression model for Subtype A, the variables associated with a patient being in clusters from subtype A were male individuals ($p<0·001$), MSM ($p<0·001$) route of transmission, having a recent infection ($p<0·001$) and being NLP ($p=0·004$) (Table S2).

In the final logistic regression model for subtype A, we adjusted the model to the variable sex, and individuals with a MSM transmission route were more likely to be in clusters when compared to heterosexual route (OR:2·65, $p=0·001$). Patients with a recent infection were more likely to be in clusters when compared to individuals with a chronic infection (OR:2·70, $p<0·001$) (Table 1.).

In the subtype B unadjusted logistic regression model, the variables associated with a patient being in clusters were male individuals ($p<0·001$), individuals with an age at resistance test between 19-30 years ($p<0·001$), route of transmission of MSM ($p<0·001$), patients originating from Eastern Europe ($p<0·001$), being migrant ($p<0·001$) and having a recent infection ($p<0·001$). Also, not having ADR ($p<0·001$), being NLP ($p<0·001$) and higher levels of viral load ($p<0·001$) were also associated with being in clusters from subtype B (Table S2).

The final logistic regression model was adjusted to sex and the determinants associated with a patient being in clusters from subtype B were males (OR:1·18, $p=0·037$), age at resistance testing, individuals with an age between 19-30 years were more likely to be within clusters (OR:1·49, $p=0·002$) when compared to older age (>56 years old). Individuals with a MSM transmission route were more likely to be in clusters when compared to heterosexual route (OR:1·74, $p<0·001$), while individuals with a IDU transmission route were less likely to be in clusters when compared to heterosexuals (OR:0·58, $p<0·001$). Patients originated from South America had a lower probability of being in clusters when compared to patients originated from Western Europe (OR:0·30, $p<0·001$). Native individuals were less likely to be in clusters when compared to migrants (OR:0·60, $p<0·001$) and individuals with a recent infection were more likely to be in clusters when compared to individuals with a chronic infection (OR:1·88, $p<0·001$) (Table 1.).

In the subtype G unadjusted logistic regression model, the variables associated with a patient being in clusters were female individuals ($p=0·012$), being native ($p=0·015$), having a recent infection ($p<0·001$), not having ADR ($p<0·001$) and being NLP ($p=0·037$) (Table S2).

The final logistic regression model was adjusted to sex and individuals from subtype G and in clusters were more likely to be native when compared to migrants (OR:1·55, $p=0·021$). Other factor associated with a patient being in clusters from subtype G was to have a recent infection when compared to those individuals with a chronic infection (OR: 1·99, $p<0·001$) (Table 1.).

Table 1. Determinants associated with belonging to a transmission cluster according to Subtype A, B and G.

		Subtype A		Subtype B		Subtype G	
In clusters/Out-of-clusters		**Final Model (Stepwise)**		**Final Model (Stepwise)**		**Final Model (Stepwise)**	
		aOR (95%CI)	**p-value**	**aOR (95%CI)**	**p-value**	**aOR (95% CI)**	**p-value**
Sex	**Female**	Ref	Ref	Ref	Ref	Ref	Ref
	Male	1·44 (0·95-2·18)	0·083	1·18 (1·01-1·38)	**0·037**	0·80 (0·56-1·16)	0·237
Age at resistance test	**<18**			1·32 (0·70-2·50)	0·391		
	19-30			1·49 (1·20-1·86)	**<0·001**		
	31-55			1·05 (0·87-1·28)	0·617		
	>56			Ref	Ref		
Transmission Route	**Heterosexual**	Ref	Ref	Ref	Ref		
	MSM	2·65 (1·50-4·69)	**0·001**	1·74 (1·52-2·00)	**<0·001**		
	IDU	0·79 (0·50-1·25)	0·308	0·58 (0·49-0·68)	**<0·001**		
	Other	0·91 (0·51-1·61)	0·747	0·73 (0·53-1·01)	0·057		
Region of Origin	**Western Europe**			Ref	Ref		
	Eastern Europe			1·00 (0·69-1·46)	0·997		
	Africa			0·75 (0·46-1·22)	0·249		
	South America			0·30 (0·21-0·44)	**<0·001**		
	Other			0·61 (0·39-0·95)	**0·028**		
Migration Status	**Migrant**			Ref	Ref	Ref	Ref
	Native			0·60 (0·49-0·72)	**<0·001**	1·55 (1·07-2·25)	**0·021**
Recentness of Infection	**Chronic**	Ref	Ref	Ref	Ref	Ref	Ref
	Recent	2·70 (1·88-3·88)	**<0·001**	1·88 (1·70-2·07)	**<0·001**	1·99 (1·37-2·87)	**<0·001**
ADR	**Yes**						
	No						
LP/NLP	**LP**						
	NLP						
Viral load groups	**<4.0**						
	4.1-5.0						
	>5.1						

Transmission Clusters analysis:

Of the 38531 patients, 8335 were in clusters (21·6%). The minimum clusters size was 2 and the maximum clusters size was 52 (data not shown).

The proportion of late presenters and non-late presenters in clusters was analyzed and LP were more in small clusters (95·7%) than NLP (92·4%). Also, LP within the small clusters were more in dual clusters (65·6%) (2 patients per clusters) (Figure 4.A).

We also analyzed the proportion of migrants vs natives in clusters, and migrants were more in large clusters (6·6%) than natives (4·8%) (Figure 4.A).

According to subtypes, subtype A had a higher proportion of patients in dual clusters (61·9%), subtype B had a higher proportion of patients in clusters of 9 or more (6·1%) and subtype G has a higher proportion of patients in clusters between 3-5 and 6-8 clusters (36·4% and 8·1%, respectively) (Figure 4.B).

A

	Small Clusters		Large Clusters	p-value
General population	2 patients	< 8 patients	≥ 8 patients	
	4614 (55.4%)	7765 (93.2%)	570 (6.8%)	
LP vs NLP				<0.001
LP (n= 2200)	1443 (65.6%)	2106 (95.7%)	94 (4.3%)	
NLP (n=2384)	1240 (52.0%)	2204 (92.4%)	180 (7.6%)	
Migrant vs Native				0.092
Migrant (n=532)	327 (61.5%)	497 (93.4%)	35 (6.6%)	
Native (n=2498)	1477 (59.1%)	2378 (95.2%)	120 (4.8%)	

B

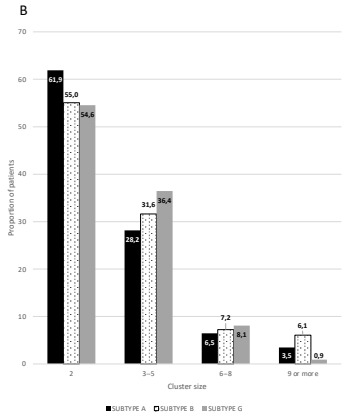

Figure 4. General clusters size characterization (A) and clusters size according to subtypes (B).

Transmission Clusters in LP vs NLP:

Here we studied the associated characteristics to LP vs NLP in clusters. Although, within the clinical and socio-demographic characteristics, there were specific characteristics of LP in clusters and NLP in clusters. LP were mainly out-of-clusters in all subtypes. In subtype A the variables associated with being LP vs NLP in clusters were age at resistance test (p=0·029) and recentness of infection (p=0·017), where LP individuals in clusters were older than 31yo (p=0·001; p=0·030), while NLP in clusters were males (p=0·002), with an age between 19-30yo (p=0·027) and a recent infection (p<0·001).

LP vs NLP with subtype B in clusters were associated with age at resistance testing (p<0·001), treatment status (p<0·001), transmission route (p=0·033), recentness of infection (p<0·001), viral load (p<0·001) and ADR (p=0·010). Where LP in clusters were older than 31yo (p<0·001; p<0·001), ART-experienced (p<0·001), originating from Eastern Europe (p<0·001), viral load higher than 5.1 copies/mL (p<0·001) and having ADR (p<0·001). While NLP in clusters were males (p<0·001), younger than 30yo (p<0·001;p<0·001), ART-naïves (p<0·001), with a MSM and IDU route of transmission (p<0·001 and p=0·047, respectively), from Western Europe (p<0·001), with a recent infection (p<0·001) and a viral load lower than 4.0 copies/mL (p<0·001).

LP vs NLP with subtype G in clusters were associated with the variables treatment status (p=0·015), recentness of infection (p=0·001), and transmission route (p=0·035). LP in clusters were mainly females (p=0·002), ART-experienced (p<0·001), with a chronic infection (p<0·001) and an IDU transmission route (p=0·011). For this subtype, MSM route in clusters was 100% related to NLP, but without being statistically significant (Table 2.)

346 **Table 2. Characteristics of LP and NLP in clusters according to subtypes**

	In clusters			
	LP	**NLP**	**p-value**	**p-value compared**
Subtype A				
Sex			0·052	
Male (n=109)	43 (39·4%)	66 (60·6%)		**0·002**
Female (n=64)	35 (54·7%)	29 (45·3%)		0·288
Age at resistance test			**0·029**	
<18 (n=4)	1 (25%)	3 (75%)		0·157
19-30 (n=33)	12 (36·4%)	21 (63·6%)		**0·027**
31-55 (n= 62)	40 (64·5%)	22 (35·5%)		**0·001**
>56 (n= 12)	8 (66·7%)	4 (33·3%)		**0·030**
Recentness of infection			**0·017**	
Chronic (n=84)	45 (53·6%)	39 (46·4%)		0·351
Recent (n=95)	34 (35·8%)	61 (64·2%)		**<0·001**
Subtype B				
Sex			0·879	
Male (n=3421)	1632 (47·7%)	1789 (52·3%)		**<0·001**
Female (n=629)	298 (47·4%)	331 (52·6%)		0·065
Age at resistance test			**<0·001**	
<18 (n=16)	2 (12·5%)	14 (87·5%)		**<0·001**
19-30 (n=608)	216 (35·5%)	392 (64·5%)		**<0·001**
31-55 (n=2307)	1215 (52·7%)	1092 (47·3%)		**<0·001**
>56 (n=245)	154 (62·9%)	91 (37·1%)		**<0·001**
Treatment Status			**<0·001**	
ART-naive (n=2121)	1014 (47·8%)	1107 (52·2%)		**0·004**
ART-experienced (n=1284)	707 (55·1%)	577 (44·9%)		**<0·001**
Transmission Route			**0.033**	
Heterosexual (n=712)	373 (52·4%)	339 (47·6%)		0·070
MSM (n=1626)	760 (46·7%)	866 (53·3%)		**<0·001**
IDU (n=427)	199 (46·6%)	228 (53·4%)		**0·047**
Region of Origin			0·256	
Western Europe (n=2795)	1332 (47·7%)	1463 (52·3%)		**<0·001**
Eastern Europe (n=81)	47 (58%)	34 (42%)		**0·042**
Africa (n=53)	26 (49·1%)	27 (50·9%)		0·853
South America (n=88)	46 (52·3%)	42 (47·7%)		0·542
Recentness of infection			**<0·001**	
Chronic (n=1942)	1126 (58%)	816 (42%)		**<0·001**
Recent (n=2181)	834 (38·2%)	1347 (61·8%)		**<0·001**
Viral load (log10)			**<0·001**	
≤ 4.0 (n=598)	201 (33·6%)	397 (66·4%)		**<0·001**
4.1-5.0 (n=601)	265 (44·1%)	336 (55·9%)		**<0·001**
≥ 5.1 (n=537)	342 (63·7%)	195 (36·3%)		**<0·001**
ADR			**0·010**	
Yes (n=240)	296 (61·3%)	187 (38·7%)		**<0·001**
No (n=483)	123 (51·2%)	117 (48·8%)		0·456
Subtype G				
Sex			0·460	
Male (n=142)	78 (54·9%)	64 (45·1%)		0·099
Female (n=140)	83 (59·3%)	57 (40·7%)		**0·002**
Treatment Status			**0.015**	
ART-naive (n=73)	39 (53.4%)	34 (46.6%)		0.075

	In clusters			
	LP	**NLP**	**p-value**	**p-value compared**
ART-experienced (n=92)	66 (71.7%)	26 (28.3%)		**<0.001**
Recentness of infection			**0·003**	
Chronic (n=152)	99 (65·1%)	53 (34·9%)		**<0·001**
Recent (n=130)	62 (47·7%)	68 (52·3%)		0·458
Transmission Route			**0·035**	
Heterosexual (n=42)	22 (52·4%)	20 (47·6%)		0·660
MSM (n=4)	0	4 (100%)		0·516
IDU (n=25)	17 (68%)	8 (32%)		**0·011**

Discussion:

In the current work, the use of the genomic sequences from the EuResist database combined with genomic sequences collected from public databases provides a comprehensive sample to study and characterize HIV-1 transmission clusters in Europe. On the one hand, HIV-1 transmission investigations are possible through reconstruction of transmission clusters. Information collected through such studies in a large scale can be highly useful for public health purposes, to fine-grain transmission patterns with higher resolution compared to classical epidemiology.

We reconstructed transmission clusters of HIV-1 in Europe with the specific objective of understanding the role of LP on transmission of infection. Specifically, we aimed to understand HIV-1 transmission clusters and determinants associated to transmission in clusters, taking into account the independent pandemics of the most prevalent subtypes in our population, A, B and C and to understand clustering patterns of LP and NLP in each subtype.

In our population the majority of patients were from subtype B, males, with MSM route of transmission and the region of origin of was Western Europe. These results are in accordance with a previous study conducted in Europe to analyse the distribution of subtypes (18).

We decided to study the transmission patterns of HIV-1 mainly according to subtypes, since there has been some discussion regarding the biological differences between them and mostly because of the higher prevalence of subtype B among Western Europe individuals (16). There were more patients outside TCs (78·4%) compared to those inside TCs (21·6%), in agreement with our study population based on sequences isolated at the first resistance test (17).

For subtype B, it was expected that one of the factors identified as associated with being inside a cluster was indeed the MSM transmission route (ref). For subtype A, we also found MSM transmission route as a factor associated with being inside clusters. The fact that the MSM route of transmission is being associated with clustering in other non-B subtypes is in accordance with some studies that report an increase of non-B subtypes associated with MSM route (19–21). Nevertheless, we expected an association of IDU route of transmission and subtype A since both subtype and route of transmission are highly prevalent in Eastern Europe where this type of transmission route is also prevalent (22). On the other hand, in subtype G, being inside a cluster was associated with heterosexual transmission, as expected (23,24). We also found that, for subtype B, the age at resistance test was associated with being in cluster, with a higher probability among individuals with younger age. This is in accordance with some other studies (17,25). Migration status was also associated with being in clusters: migrants infected with subtype B were mainly in a cluster, and migrants infected with subtype G were less likely to be in a cluster. These results were in accordance with a recent study focusing on

migration and HIV-1 in Portugal (26). These results could be explained by the region of origin of migrants. Subtype B had higher prevalence of migrants from Brazil, and Brazil has a concentrated HIV epidemic among MSM population (27,28).
Regarding the potential association between transmission clusters and late presentation, we found that both LP and NLP were mainly outside clusters. As for the differences between the populations of LP and NLP inside clusters, these patterns were consistent between subtypes A and B: concerning sex, there were more NLP males inside cluster; concerning age, there were older LP, with a higher and growing proportion as age increases. Subtype G had the most different patterns of all. Furthermore, for subtypes B and G individuals, there were more treated patients among LP than among NLP inside clusters. As for transmission route, for subtype G, we found more LP with an IDU transmission route inside clusters. Finally, for subtype B, it was interesting to observe that LP located inside clusters had higher viral loads than NLP. These results could not be compared by LP and NLP populations, nevertheless our results are overall in accordance with some studies (17,26,28).
Finally, we found that LP were more frequently present in small clusters or outside clusters compared to NLP which can indicate a limited role of this population on HIV-1 transmission, given the less frequent presence of these patients in TCs. However higher viral loads were observed in LP located inside clusters that can indicate higher transmissibility of infection within individuals from the TCs.
Finally, there is still scarce to none information regarding transmission clusters and late presentation. We studied here the association of transmission clusters according to subtypes in LP and NLP, and our results showed that the patterns of LP vs NLP in TCs presented similar characteristics in subtypes A and B, but not in subtype Gdominated by LP.

Limitations:
In our study we did not used the time and place of the most recent common ancester (tMERCA), instead we used a total number of sequences from a specific region. This methodology can cause some sampling bias since sequences can artificially be in cluster due to their shared region of origin.

Conclusion:
In conclusion, our study presented an updated description of the socio-demographic and clinical characteristics of HIV-1 infected individuals followed in Europe according to subtype. Our study also highlights the patterns of transmission clusters in LP vs NLP populations selected in the european dataset of EuResist. We conclude that late presentation could have a limited role on HIV-1 transmission. However further investigation should be considered to exclude LP classification bias, and to better estimate the time of infection based on phylogenetic trees reconstruction and molecular clock analysis.

Author Contributions: Conceptualization, M.N.S.M., M.P. and A.A.; Methodology, M.N.S.M., M.P., V.P., M.d.R.O.M. and A.A.; Software, M.N.S.M., V.P., S.G.S.; Validation, M.N.S.M., M.P., F.I. and A.A.; Formal Analysis, M.N.S.M., V.P., S.G.S., M.P. and A.A.; Investigation, M.N.S.M., M.P., V.P., and M.d.R.O.M.; Resources, C.S.-D., R.P., R.K., M.B. (Marina Bobkova),M.B (Michael Böhm), M.Z., P.G and F.I.; Data Curation, C.S.-D., R.P., R.K., M.B. (Marina Bobkova),M.B (Michael Böhm), M.Z., P.G and F.I.; Writing—Original Draft Preparation, M.N.S.M., M.P. and A.A.; Writing—

Review and Editing, M.N.S.M., M.P., C.S.-D., F.I. and A.A.; Visualization, M.N.S.M., M.P., V.P., M.d.R.O.M. and A.A.; Supervision, A.A.; Project Administration, A.A.; Funding Acquisition, A.A.

Funding: This study was financed by FCT through the following projects: GHTM-UID/04413/2020, INTEGRIV (PTDC/SAU-INF/31990/2017) and the scholarship PD/BD/135714/2018 and COVID/BD/152613/2022 and Gilead Génese HIVLatePresenters.

Conflicts of Interest: Author Francesca Incardona is employed by InformaPRO S.r.l. The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

References:

1. Global HIV & AIDS statistics — Fact sheet | UNAIDS [Internet]. [cited 2022 Jan 25]. Available from: https://www.unaids.org/en/resources/fact-sheet
2. Grabowski MK, Herbeck JT, Poon AFY. Genetic Cluster Analysis for HIV Prevention. Curr HIV/AIDS Rep [Internet]. 2018 Apr 19;15(2):182–9. Available from: https://doi.org/10.1007/s11904-018-0384-1
3. Abecasis AB, Wensing AMJ, Paraskevis D, Vercauteren J, Theys K, Van de Vijver DAMC, et al. HIV-1 subtype distribution and its demographic determinants in newly diagnosed patients in Europe suggest highly compartmentalized epidemics. Retrovirology [Internet]. 2013 Dec 14;10(1):7. Available from: https://retrovirology.biomedcentral.com/articles/10.1186/1742-4690-10-7
4. Brenner BG, Roger M, Routy J, Moisi D, Ntemgwa M, Matte C, et al. High Rates of Forward Transmission Events after Acute/Early HIV-1 Infection. J Infect Dis [Internet]. 2007 Apr;195(7):951–9. Available from: https://academic.oup.com/jid/article-lookup/doi/10.1086/512088
5. Miranda MNS, Pingarilho M, Pimentel V, Martins MDRO, Vandamme A-M, Bobkova M, et al. Determinants of HIV-1 Late Presentation in Patients Followed in Europe. Pathogens [Internet]. 2021 Jul 2;10(7):835. Available from: https://www.mdpi.com/2076-0817/10/7/835
6. Antinori A, Coenen T, Costagiola D, Dedes N, Ellefson M, Gatell J, et al. Late presentation of HIV infection: a consensus definition. HIV Med [Internet]. 2011 Jan;12(1):61–4. Available from: https://onlinelibrary.wiley.com/doi/10.1111/j.1468-1293.2010.00857.x
7. Late Presentation Working Groups in EuroSIDA and COHERE. Estimating the burden of HIV late presentation and its attributable morbidity and mortality across Europe 2010-2016. BMC Infect Dis [Internet]. 2020 Oct 7;20(1):728. Available from: https://bmcinfectdis.biomedcentral.com/articles/10.1186/s12879-020-05261-7
8. Rava M, Domínguez-Domínguez L, Bisbal O, López-Cortés LF, Busca C, Antela A, et al. Late presentation for HIV remains a major health issue in Spain: Results from a multicenter cohort study, 2004–2018. Andrei G, editor. PLoS One [Internet]. 2021 Apr 21;16(4):e0249864. Available from: https://dx.plos.org/10.1371/journal.pone.0249864
9. Pineda-Peña A-C, Pingarilho M, Li G, Vrancken B, Libin P, Gomes P, et al. Drivers of HIV-1 transmission: The Portuguese case. Blackard J, editor. PLoS One [Internet]. 2019 Sep 30;14(9):e0218226. Available from: https://dx.plos.org/10.1371/journal.pone.0218226
10. Euresist Data Analysis - database [Internet]. [cited 2021 Jan 4]. Available from:

http://engine.euresist.org/database/

11. Bennett DE, Camacho RJ, Otelea D, Kuritzkes DR, Fleury H, Kiuchi M, et al. Drug Resistance Mutations for Surveillance of Transmitted HIV-1 Drug-Resistance: 2009 Update. Nixon DF, editor. PLoS One [Internet]. 2009 Mar 6;4(3):e4724. Available from: https://dx.plos.org/10.1371/journal.pone.0004724
12. Stoesser G, Griffith M, Griffith OL. HIV Sequence Database. In: Dictionary of Bioinformatics and Computational Biology [Internet]. Chichester, UK: John Wiley & Sons, Ltd; 2004. p. 52–61. Available from: https://onlinelibrary.wiley.com/doi/10.1002/9780471650126.dob0322.pub2
13. Libin PJK, Deforche K, Abecasis AB, Theys K. VIRULIGN: fast codon-correct alignment and annotation of viral genomes. Hancock J, editor. Bioinformatics [Internet]. 2019 May 15;35(10):1763–5. Available from: https://academic.oup.com/bioinformatics/article/35/10/1763/5123354
14. Ragonnet-Cronin M, Hodcroft E, Hué S, Fearnhill E, Delpech V, Brown AJL, et al. Automated analysis of phylogenetic clusters. BMC Bioinformatics [Internet]. 2013 Dec 6;14(1):317. Available from: https://bmcbioinformatics.biomedcentral.com/articles/10.1186/1471-2105-14-317
15. Andersson E, Shao W, Bontell I, Cham F, Cuong DD, Wondwossen A, et al. Evaluation of sequence ambiguities of the HIV-1 pol gene as a method to identify recent HIV-1 infection in transmitted drug resistance surveys. Infect Genet Evol [Internet]. 2013 Aug;18(1):125–31. Available from: https://linkinghub.elsevier.com/retrieve/pii/S156713481300141X
16. Kanki PJ, Hamel DJ, Sankalé J, Hsieh C, Thior I, Barin F, et al. Human Immunodeficiency Virus Type 1 Subtypes Differ in Disease Progression. J Infect Dis [Internet]. 1999 Jan;179(1):68–73. Available from: https://academic.oup.com/jid/article-lookup/doi/10.1086/314557
17. Paraskevis D, Beloukas A, Stasinos K, Pantazis N, de Mendoza C, Bannert N, et al. HIV-1 molecular transmission clusters in nine European countries and Canada: association with demographic and clinical factors. BMC Med [Internet]. 2019 Dec 8;17(1):4. Available from: https://bmcmedicine.biomedcentral.com/articles/10.1186/s12916-018-1241-1
18. Abecasis AB, Wensing AM, Paraskevis D, Vercauteren J, Theys K, Van de Vijver DA, et al. HIV-1 subtype distribution and its demographic determinants in newly diagnosed patients in Europe suggest highly compartmentalized epidemics. Retrovirology [Internet]. 2013 Dec 14;10(1):7. Available from: http://retrovirology.biomedcentral.com/articles/10.1186/1742-4690-10-7
19. Delgado E, Benito S, Montero V, Cuevas MT, Fernández-García A, Sánchez-Martínez M, et al. Diverse Large HIV-1 Non-subtype B Clusters Are Spreading Among Men Who Have Sex With Men in Spain. Front Microbiol [Internet]. 2019 Apr 3;10(APR):1–19. Available from: https://www.frontiersin.org/article/10.3389/fmicb.2019.00655/full
20. Petersen A, Cowan SA, Nielsen J, Fischer TK, Fonager J. Characterisation of HIV-1 transmission clusters and drug-resistant mutations in Denmark, 2004 to 2016. Eurosurveillance [Internet]. 2018 Nov 1;23(44):1–9. Available from: https://www.eurosurveillance.org/content/10.2807/1560-7917.ES.2018.23.44.1700633
21. Hanke K, Faria NR, Kühnert D, Yousef KP, Hauser A, Meixenberger K, et al. Reconstruction of the Genetic History and the Current Spread of HIV-1 Subtype A in Germany. Kirchhoff F, editor. J Virol [Internet]. 2019 Jun 15;93(12).

Available from: https://journals.asm.org/doi/10.1128/JVI.02238-18
22. Lai A, Bozzi G, Franzetti M, Binda F, Simonetti FR, De Luca A, et al. HIV-1 A1 Subtype Epidemic in Italy Originated from Africa and Eastern Europe and Shows a High Frequency of Transmission Chains Involving Intravenous Drug Users. Tornesello ML, editor. PLoS One [Internet]. 2016 Jan 11;11(1):e0146097. Available from: https://dx.plos.org/10.1371/journal.pone.0146097
23. Yebra G, Holguín Á, Pillay D, Hué S. Phylogenetic and demographic characterization of HIV-1 transmission in Madrid, Spain. Infect Genet Evol [Internet]. 2013;14(1):232–9. Available from: http://dx.doi.org/10.1016/j.meegid.2012.12.006
24. Lorenzin G, Gargiulo F, Caruso A, Caccuri F, Focà E, Celotti A, et al. Prevalence of Non-B HIV-1 Subtypes in North Italy and Analysis of Transmission Clusters Based on Sequence Data Analysis. Microorganisms [Internet]. 2019 Dec 23;8(1):36. Available from: https://www.mdpi.com/2076-2607/8/1/36
25. Pingarilho M, Pimentel V, Miranda MNS, Silva AR, Diniz A, Ascenção BB, et al. HIV-1-Transmitted Drug Resistance and Transmission Clusters in Newly Diagnosed Patients in Portugal Between 2014 and 2019. Front Microbiol [Internet]. 2022;13. Available from: https://www.frontiersin.org/article/10.3389/fmicb.2022.823208
26. Pimentel VF, Pingarilho M, Sole G, Alves D, Miranda M, Diogo I, et al. Differential patterns of postmigration HIV-1 infection acquisition among Portuguese immigrants of different geographical origins. AIDS [Internet]. 2022 Jun 1;36(7):997–1005. Available from: https://journals.lww.com/aidsonline/Fulltext/9000/Differential_patterns_of_post_migration_HIV_1.96205.aspx
27. MacCarthy S, Brignol S, Reddy M, Nunn A, Dourado I. Late presentation to HIV/AIDS care in Brazil among men who self-identify as heterosexual. Rev Saude Publica [Internet]. 2016;50:54. Available from: http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0034-89102016000100233&lng=en&tlng=en
28. Kostaki EG, Gova M, Adamis G, Xylomenos G, Chini M, Mangafas N, et al. A Nationwide Study about the Dispersal Patterns of the Predominant HIV-1 Subtypes A1 and B in Greece: Inference of the Molecular Transmission Clusters. Viruses [Internet]. 2020 Oct 19;12(10):1183. Available from: https://www.mdpi.com/1999-4915/12/10/1183

575 **Table Supplementary 1**. Patients socio-demographic and clinical characteristics
576

Patient Characteristics	Total	Subtype A	Subtype B	Subtype G	p-value
Total	38531 (100)	2276 (5·9)	32652 (84·7)	3603 (9·4)	
Sex, n (%)	36699 (95·2)	2011 (5·5)	31200 (85·0)	3488 (9·5)	<0·001
Male	27715 (75·5)	1156 (4·2)	24570 (88·6)	1989 (7·2)	
Female	8984 (24·5)	855 (9·5)	6630 (73·8)	1499 (16·7)	
Median age at resistance teste in years IQR, n (%)	18421 (47·8)	631 (3·4)	17149 (93·1)	641 (3·5)	<0·001
	40·0 (34·0-47·0)	37·0 (30·0-45·0)	41·0 (35·0-47·0)	34·0 (28·0-41·0)	
≤ 18	233 (1·3)	32 (13·7)	143 (61·4)	58 (24·9)	<0·001
19-30	2529 (13·7)	150 (5·9)	2209 (87·4)	170 (6·7)	
31-55	14199 (77·1)	390 (2·7)	13430 (94·6)	379 (2·7)	
≥ 56	1460 (7·9)	59 (4·0)	1367 (93·6)	34 (2·4)	
Transmission Route, n (%)	18140 (47·1)	1230 (6·8)	16129 (88·9)	781 (4·3)	<0·001
Heterosexual	5568 (30·7)	532 (9·6)	4646 (83·4)	390 (7·0)	
MSM	6692 (36·9)	97 (1·4)	6566 (98·1)	29 (0·4)	
IDU	4883 (26·9)	433 (8·9)	4269 (87·4)	181 (3·7)	
Other	997 (5·5)	168 (16·9)	648 (65·0)	1818 (18·2)	
Region of origin, n (%)	23647 (61·4)	1277 (5·4)	20595 (87·1)	1775 (7·5)	<0·001
Western Europe	20440 (86·4)	482 (2·4)	18836 (92·2)	1122 (5·5)	
Eastern Europe	933 (3·9)	468 (50·2)	365 (39·1)	100 (10·7)	
Africa	1162 (4·9)	294 (25·3)	343 (29·5)	525 (45·2)	
South America	787 (3·3)	11 (1·4)	754 (95·8)	22 (2·8)	
Other	325 (1·4)	22 (6·8)	297 (91·4)	6 (1·8)	
Migration Status	16101 (41·8)	574 (3·6)	14629 (90·9)	898 (5·6)	<0·001
Migrant	2543 (15·8)	351 (13·8)	1771 (69·6)	421 (16·6)	
Native	13558 (84·2)	223 (1·6)	12858 (94·8)	477 (3·6)	
Clusters	38531 (100)	2276 (5·9)	32652 (84·7)	3603 (9·4)	0·006
In Clusters	8335 (21·6)	433 (5·2)	7136 (85·6)	766 (9·2)	
Out-of-Clusters	30196 (78·4)	1843 (6·1)	25516 (84·5)	2837 (9·4)	
Treatment Status	21687 (56·3)	709 (3·3)	19387 (89·4)	1591 (7·3)	<0·001
ART-naive	8887 (41·0)	443 (5·0)	8046 (90·5)	398 (4·5)	
ART-experienced	12800 (59·0)	266 (2·1)	11341 (88·6)	1193 (9·3)	
Recentness of infection	38531 (100)	2276 (5·9)	32652 (84·7)	3603 (9·4)	<0·001
Recent	15571 (40·4)	929 (6·0)	13322 (85·6)	1320 (8·5)	
Chronic	22960 (59·6)	1347 (5·9)	19330 (84·2)	2283 (9·9)	
TDR	7727 (20·1)	393 (5·1)	7035 (91·0)	299 (3·9)	<0·001
Yes	967 (12·5)	17 (1·8)	918 (94·9)	32 (3·3)	
No	6760 (87·5)	376 (5·6)	6117 (90·5)	267 (3·9)	
ADR	6184 (16·0)	133 (2·2)	5320 (86·0)	731 (11·8)	<0·001
Yes	4542 (73·4)	73 (1·6)	3958 (87·1)	511 (11·3)	
No	1642 (26·6)	60 (3·7)	1362 (82·9)	220 (13·4)	
Median CD4 count at diagnosis (cells/mL) IQR, n (%)	24321 (63·1)	1011 (4·2)	21583 (88·7)	1727 (7·1)	<0·001
	341·0 (170·0-540·0)	328·0 (174·0-510·0)	349·0 (172·0-547·0)	273·0 (139·0-445·0)	
LP	12501 (51·4)	545 (4·4)	10875 (87·0)	1081 (8·6)	<0·001
NLP	11820 (48·6)	466 (3·9)	10708 (90·6)	646 (5·5)	
Viral Load at diagnosis (log_{10} copies/mL) IQR, n (%)	15670 (40·7)	614 (3·9)	13105 (83·6)	1951 (12·5)	<0·001
	4·3 (3·4-5·0)	4·3 (3·4-5·1)	4·3 (3·3-5·0)	4·5 (3·7-5·1)	
≤ 4.0	6210 (39·6)	246 (4·0)	5315 (85·6)	649 (10·5)	<0·001
4.1-5.0	5361 (34·2)	197 (3·7)	4436 (82·7)	728 (13·6)	
≥ 5.1	4099 (26·2)	171 (4·2)	3354 (81·8)	574 (14·0)	

577

578 **Table Supplementary 2**. Unadjusted analysis for determinants associated with belonging to a transmission
579 cluster according to Subtype A, B and G

In clusters/Out-of-clusters		Subtype A Unadjusted Model		Subtype B Unadjusted Model		Subtype G Unadjusted Model	
		uOR (95%CI)	**p-value**	**uOR (95%CI)**	**p-value**	**uOR (95% CI)**	**p-value**
Sex	**Female**	Ref	Ref	Ref	Ref	Ref	Ref
	Male	1·84 (1·46-2·32)	**<0·001**	1·49 (1·38-1·59)	**<0·001**	0·81 (0·69-0·96)	**0·012**
Age at resistance test	**<18**	0·56 (0·16-1·90)	0·353	0·72 (0·45-1·17)	0·187	1·85 (0·35-9·71)	0·469
	19-30	1·19 (0·57-2·49)	0·641	1·72 (1·47-2·03)	**<0·001**	2·89 (0·65-12·80)	0·162
	31-55	0·76 (0·38-1·50)	0·423	0·91 (0·80-1·06)	0·223	2·16 (0·50-9·30)	0·303
	>56	Ref	Ref	Ref	Ref	Ref	Ref
Transmission Route	**Heterosexual**	Ref	Ref	Ref	Ref	Ref	Ref
	MSM	3·46 (2·12-5·65)	**<0·001**	1·84 (1·67-2·02)	**<0·001**	1·51 (0·62-3·67)	0·367
	IDU	0·76 (0·51-1·13)	0·178	0·60 (0·53-0·68)	**<0·001**	1·52 (0·99-2·34)	0·055
	Other	1·06 (0·64-1·77)	0·810	0·76 (0·60-0·97)	**0·025**	0·76 (0·46-1·25)	0·276
Region of Origin	**Western Europe**	Ref	Ref	Ref	Ref	Ref	Ref
	Eastern Europe	0·63 (0·46-0·87)	**0·005**	1·58 (1·26-1·99)	**<0·001**	0·67 (0·40-1·14)	0·141
	Africa	0·55 (0·37-0·80)	**0·002**	1·11 (0·86-1·43)	0·424	0·57 (0·43-0·75)	**<0·001**
	South America	1·20 (0·31-4·59)	0·793	0·94 (0·78-1·12)	0·478	2·12 (0·90-5·02)	0·087
	Other	1·20 (0·46-3·13)	0·714	1·02 (0·77-1·34)	0·917	0·61 (0·07-5·27)	0·656
Migration Status	**Migrant**	Ref	Ref	Ref	Ref	Ref	Ref
	Native	1·48 (0·97-2·27)	0·070	0·72 (0·64-0·81)	**<0·001**	1·57 (1·09-2·26)	**0·015**
Recentness of Infection	**Chronic**	Ref	Ref	Ref	Ref	Ref	Ref
	Recent	2·70 (2.18-3.35)	**<0·001**	2·22 (2·10-2·34)	**<0·001**	2·31 (1·96-2·72)	**<0·001**
ADR	**Yes**	Ref	Ref	Ref	Ref	Ref	Ref
	No	2·01 (0·87-4·64)	0·102	1·54 (1·30-1·82)	**<0·001**	2·35 (1·47-3·74)	**<0·001**
LP/NLP	**LP**	Ref	Ref	Ref	Ref	1	1
	NLP	1·61 (1·16-2·23)	**0·004**	1·15 (1·08-1·23)	**<0·001**	1·32 (1·02-1·71)	**0·037**
Viral load groups	**<4.0**	Ref	Ref	Ref	Ref	1	1
	4.1-5.0	1·15 (0·67-1·98)	0·603	1·27 (1·14-1·42)	**<0·001**	1·16 (0·88-1·53)	0·297
	>5.1	1·66 (0·98-2·82)	0·060	1·55 (1·38-1·73)	**<0·001**	1·24 (0·93-1·66)	0·144

580

5. General Discussion and Conclusions

5.1. Discussion

There are some studies regarding LP worldwide, but none regarding recent years. The most recent in Europe was published in 2020 and included two major European cohorts of HIV-1 infected individuals, although the timelines of the study were between 2010-2013 (the COHERE) and between 2001-2016 (the EuroSIDA), and in that study LP accounted for 48.4% of the study population (1).

This thesis presents up to date information regarding HIV-1 and late presentation in a timeline between 1981 and 2019, based on the in-depth analyses of the EIDB, which included information combining ARCA (Italy), AREVIR (Germany), CoRIS and IRSICAIXA (Spain), Portugal, United Kingdom, Russia and Luxembourg databases from treated and naïve patients followed up in these countries. Since it has been hypothesized that LP could sustain the HIV epidemic and contribute to the onward spread of the virus, this work had the objective of explaining the determinants of late presentation for HIV-1 infection in Europe, give an overview of TDR and ADR in the overall European population and in LP and NLP and to construct transmission clusters of HIV-1 infection and identify and describe the role of LP and NLP populations in such clusters.

The first study (Manuscript I) described late presentation and its determinants in Portugal, based on a combined laboratory and clinical database from Egas Moniz hospital in Lisbon. The study showed that LP accounted for half of the infected individuals in this database. LPs in this study were mainly males, self-reported heterosexual transmission route, aged between 31-55 years and the main region of origin was Portugal and, for migrants, was Sub-Saharan Africa. The results we achieved were in accordance with studies from other European countries, where for example, migrants from Sub-Saharan Africa had a higher prevalence among LP in a study in Spain, Germany and Switzerland (2–4). In a study in United Kingdom, late presentation was also associated with older age (5) and in a study from Germany heterosexual transmission was also higher among LP (3). Also, given that in this study we had information regarding stage of infection, we could use this dataset of patients as a partial validation for our classification of LP and

NLP populations: logistic regression indicated a significant association between stage of infection and late presentation status, as expected. Given these findings and that our results were in accordance with other European countries studies, we extended our analyses in order to understand the determinants associated with late presentation for the rest of the Europe and its evolution through time.

For that purpose, we used the EuResist database for the other studies. The second study (Manuscript II) was the most recent update on the HIV epidemic of late presentation in Europe, since the latest published study analyzed data collected until 2016 (1). Herein, we did not have information concerning stage of infection. As such late presentation was defined based on the CD4 count thresholds established by the European Late Presenter Consensus working group, which states LP as an individual presenting a TCD4+ count lower than 350 cells/mm3 or an AIDS-defining event, regardless of TCD4+ cell count (6). To overcome that bias, we created a new variable, based on the level of ambiguity on the genomic sequences of the patients. As such, to validate our classification of LP, we studied the association between CD4 count and the ambiguity rate. Results showed that there was a negative correlation, that is, when CD4 count was higher, the ambiguity rate was lower, consistently indicating a more recent infection (R^2= 0.023). On the other hand, our logistic regression model indicated the ambiguity rate as a determinant associated with LP, i.e., that LP had a higher probability of having a higher ambiguity rate. This definition of chronic vs recent infection according to ambiguity rate is well documented throughout Manuscript II and in a study from Andersson et.al regarding ambiguity rate (7). The fact that LP are associated with a more chronic infection indicates consistency of the two definitions, one based on the ambiguity rate of the genomic sequence (chronic vs recent), and the other based on the CD4 count of the patients (LP vs NLP). The results found herein were also in accordance with our first study made in Portugal (Manuscript I) and where we had information regarding stage of infection. In future work, we expect to test and validate our approach on datasets that included genomic sequences generated through next-generation sequencing, in which the genetic diversity of the quasispecies can be measured and be more accurate to establish time since infection than ambiguity rate on its own (8,9).

As in our Portuguese study (Manuscript I), late presenters in Europe accounted for half of the HIV-1 infections (50.4%). Compared to other studies, the population of LP in our study presented similar characteristics: there were mainly males, with heterosexual transmission route and with Western Europe as main region of origin, as found in our Portuguese study and some other from Europe (10–12). The determinants associated with being LP in Europe, besides having a chronic infection, were Africa origin, older age, heterosexual contact and a higher viral load at diagnosis. These were again consistent with those found in our Portuguese study (Manuscript I) and in other studies regarding late presentation (13–16). The final step of our analyses - Bayesian network (BN) analyses – allowed to correct for potential dependencies between variables considered as independent for the multivariate logistic regression. Bayesian network analyses results strengthen the results obtained through logistic regression, where the variables directly associated with LP were the same as in the logistic regression model. An important achievement of this study was that we could demonstrate the evolution of LP through the years, and it was shown that late presentation was constant and always with a value around 50%. Considering these results, we could hypothesize that the present strategies for HIV prevention and early diagnosis have not been efficient for LP populations. These findings could help future stakeholders and health authorities to elaborate prevention strategies directed at this population. At the end of this section, we elaborate on potential strategies for these populations.

Studies 3 and 4 (Manuscripts III and IV) had an innovative perspective towards LP, since they focus on drug resistance and transmission clusters among the populations of LP and NLP. It is known that the high error rate of RT, recombination, replication rate and population sizes during infection lead to high levels of HIV-1 genetic diversity that allows for development of drug resistance, which is recognized as a major problem globally. As such, continuous surveillance and research on HIV-1 genetic diversity and drug resistance and of its evolution throughout the years remains mandatory (17). We acknowledge drug resistance and late presentation as two important problems that may contribute to the failure of the UNAIDS targets for ending the pandemic by 2030. In study 3 (Manuscript III), we had the objective of analyzing HIV-1 TDR and ADR in the populations of LP and NLP using the same database used in Manuscript II. Since information regarding this topic in this specific population is still scarce, we decided to

adopt a comparative perspective of the TDR and ADR patterns among those two populations. Furthermore, this study also compares the most prevalent drug resistance mutations between subtypes: subtype B vs non-B subtypes.

There are many studies about the prevalence of TDR and ADR among European countries individually (18–20). However, there are no recent studies regarding TDR and ADR prevalence in Europe with various countries included in the analysis. The most recent study about this topic includes only TDR and is a median of overall TDR values from different studies according to different countries (21). Herein, we presented an overview of the proportions and trends of TDR and ADR among the HIV-1 infected population, and then to compare its patterns among LP and NLP populations. Overall, TDR had a prevalence of 12.8%, which was higher when compared to individual studies from other countries in and outside Europe (20,22–26) and to a study including 26 European countries between 2008-2010 where TDR prevalence was 8.3% (27). Regarding ADR, the overall prevalence was 68.5% which was also higher when compared to other studies of ADR prevalence (28–30). The fact that our results might be higher when compared to other studies could be due to the timeline included in our study being so broad (patients diagnosed between 1981 and 2019) or it could also be related to the low genetic barrier of previous generation of ARV drugs (31). We also estimated the overall trends for TDR and ADR and both presented a decreasing trend over the years, which is consistent with other studies in and outside of Europe (29,32) These decreasing trends could also be related to the hypothesis mentioned above. In this study (Manuscript III), we also compared the prevalence of TDR and ADR according to country of follow-up in two time-periods (between 2008-2012 and 2013-2018), contrarily to other studies in which results are not stratified per individual countries. All countries studied (Italy, Germany, Luxembourg and Portugal) had prevalence of TDR and ADR concordant with what has been reported in other studies from their countries (20,26,32–34), except for Luxembourg, where both TDR and ADR prevalence was higher, although we could not compare to a study made in the country and instead the values were compared to a previous study in Europe (27).

On the other hand, we compared levels of drug resistance in LP vs NLP and we did not find differences in the prevalence in neither of those population both in TDR or ADR.

The prevalence of TDR and ADR for LP and NLP found in our study was similar to the prevalence of the overall population. We could not compare our results of TDR and ADR in LP and NLP with other studies since such studies are scarce or inexistent. On the other hand, we compared patterns of drug resistance between LP and NLP. LP presented PI DR mutations that were not found in NLP. However, we were expecting the opposite result - that NLP would present mutations that were not present in LP - since the infection for LP population is older and, without the selective pressure of ARV drugs, some transmitted DR mutations can revert to the wild type (35). In some specific cases, it could perhaps indicate the irreversible fixation of mutations considering that they are not associated with a fitness cost (36,37).

In this study it was also shown that the most prevalent mutations whether in ART-naïve or ART-experienced individuals were the K103N/S (NNRTIs), T215 revertants and T215FY, M184V/I and the M41I/L (NRTIs), and M46I/L and L90M (PIs). These results are in accordance with several studies regarding DRMs in Europe (38–40) and reflect the long standing use of NNRTIs as first-line with low genetic barrier to resistance.

The importance of the study of DRMs is that their presence can impact treatment outcomes and lead to first-line treatment failure for ART-naïve individuals initiating ART and virological failure or an increased burden of treatment for ART-experienced individuals (41). Therefore, the most prevalent DRMs found in this study could lead to some of those consequences mentioned above, for example in NRTIs class the presence of T215rev is associated with high risk of virological failure to AZT or d4T, the presence of M41I/L could have a negative impact in the virological response to regimens which include ABC, ddl or TDF and the presence of M184V/I could reduce susceptibility to regimens including 3TC and FTC (42). In NNRTIs class the mutation K103N/S can reduce susceptibility to NVP and EFV (43). In PIs class, M46I/L mutation is associated with a reduction of susceptibility to ATV, FPV, IDV, LPV and NFV, while L90M is associated to a reduction in susceptibility to almost all PIs, except for TPV and DRV (44). Furthermore, we performed an analysis distinguishing the patterns found in subtype B when compared to non-B subtypes, since it is well known that some mutations occur more frequently in specific subtypes and recombinant forms (45). In our study, the most

prevalent subtype was subtype B compared to non-B and the mutations with the highest prevalence in both LP and NLP were similar to those mentioned above, and these results were in accordance with other studies about mutations per subtypes (26,40,46).

The third study (Manuscript III) was a continuation of the second study (Manuscript II) but with a different aim, in the first we described LP and NLP populations in Europe and the determinants associated and in the second we presented trends and patterns of TDR, ADR and DRMs in a comparative perspective for the two populations. Therefore, we provided important information that can suggest different dynamics of reversion and/or irreversible fixation of mutations in LP and NLP populations, and with that it is possible to consider further research in these populations and new strategies and methodologies for future studies.

HIV-1 transmission investigations are possible through reconstruction of transmission clusters. The footprint of the viral strain infecting each patient is so accurate that this approach is used even to support or reject the hypothesis of potential transmission of infection in the context of HIV-1 court cases (47).

On the other hand, information collected through such studies in a large scale can be highly useful for public health purposes, to fine-grain transmission patterns with higher resolution compared to classical epidemiology. The accuracy of such reconstructed transmission clusters is obviously affected by sampling issues and by the methods used for reconstruction of phylogenetic trees. In this context, the use of the genomic sequences from the EuResist database combined with genomic sequences collected from public databases provides a comprehensive sample to study and characterize HIV-1 transmission clusters in Europe. In Manuscript IV, we reconstructed transmission clusters of HIV-1 in Europe with the specific objective of understanding the role of LP on transmission of infection. Specifically, we aimed to understand HIV-1 transmission clusters and determinants associated to transmission in clusters, taking into account the independent pandemics of the most prevalent subtypes A, B and C and to understand differential clustering of LP and NLP in each subtype.

For subtypes A, B and C datasets of 10122, 62543 and 5547 patients were used that included sequences from EuResist database and control sequences. In our results, there were more patients outside TCs (78.4%) compared to those inside TCs (21.6%), which is a result expected considering our study population (48). For subtype B, it was not a surprise that one of the factors identified as associated with being inside cluster was the MSM transmission route. For subtype A, we also found MSM transmission route as a factor associated with being inside clusters. The fact that in other non-B subtypes the MSM route of transmission is being associated with clustering is in accordance with some studies that report an increase of non-B subtypes associated with MSM route (49–51). Nevertheless, we expected an association of IDU route of transmission and subtype A since this subtype is highly prevalent in Eastern Europe where this type of transmission route is also prevalent (52). On the other hand, in subtype G, being inside cluster was associated with heterosexual transmission (53,54). We also found that, for subtype B, age at resistance test was associated with being in cluster and the probability of clustering was higher among younger individuals. This is in accordance with some other studies (48,55). Migration status was also associated with being in clusters: migrants infected with subtype B were mainly in cluster, and migrants infected with subtype G were less likely to be in cluster, in accordance with a recent study about migration and HIV-1 in Portugal (56). These results could be explained by the region of origin of migrants. Subtype B had higher prevalence of migrants from Brazil, and Brazil has a concentrated HIV epidemic among MSM population (57,58).

Regarding the potential association between transmission clusters and late presentation, we found that both LP and NLP were mainly outside clusters. As for the differences between the populations of LP and NLP inside clusters, these patterns were consistent between subtypes A and B: concerning sex, there were more NLP males inside cluster; concerning age, there were older LP, with a higher and growing proportion as age increases. Furthermore, for subtypes B and G, there were more treated patients among LP than among NLP that were inside clusters. As for transmission route, for subtype G, we found more LP with an IDU transmission route inside clusters. Finally, for subtype B, it was interesting to observe that LP that were inside clusters had higher viral loads than

NLP. These results could not be compared by LP and NLP populations, but according to an overall, our results are in accordance with some studies (48,56,58).

Finally, we found that LP were more frequently present in small clusters or outside clusters compared to NLP which can indicate a limited role of this population on HIV-1 transmission, given the less frequent presence of these patients in TCs. However, on the other hand, the finding of higher viral loads in LP that are inside clusters can indicate higher transmissibility of infection for those that are present in TCs.

Our results also showed that migrants were more frequently inside clusters for subtype B, while the opposite pattern was found for subtype G, where migrants were more frequently outside clusters. These findings, together with the findings of (Manuscript II), where we found that migrants had a higher probability of being LP, pinpoint HIV-1 infected migrants as a particularly vulnerable population. This can be explained by various factors, that can be related to access to health care, laws and regulations concerning for example HIV testing in the host country, conditions of unemployment, poverty, household conditions and stigma (59,60).

In this thesis, we performed the first in-depth analyses of the HIV-1 LP population in Europe. It is clearly important to target this population in the context of UNAIDS goals for ending the pandemic by 2030 and for the 95-95-95 target. This can be done by designing prevention measures specific for that population, and by reinforcing the follow-up and retention in care. Also, we highlighted that there is a need to increased HIV testing and include screening programs as a prevention strategy, since the rate of LPs throughout the years in Portugal and broadly in Europe has never been lower than 45%. With the implementation of new strategies and the attention of policymakers and health workforce to these specific population it might be possible to reduce late presentation and achieve better health outcomes and earlier HIV diagnosis.

5.2. Conclusion

In this thesis we performed an analysis regarding the late presenters and non-late presenters populations that was lacking in the literature.

We provided information not only regarding the determinants of late presentation, but also their drug resistance patterns, their role on transmission of HIV-1 and the characteristics of LP contributing to transmission. This information can help to overcome adverse results in therapy and help to guide strategies for reaching these populations.

This thesis could also impact the view and integration of this population in the community and health care systems, since we provided information on the socio-demographics of this population. Many of LP in our studies were migrants, and the migrant population is also a vulnerable population and at risk of contracting HIV and other diseases.

Our final objective with this work was to provide knowledge so that healthcare professionals and policymakers could design directed and evidence-based informed strategies regarding prevention strategies and therapy regimens for this population.

5.3. Future research

The results of this thesis pinpointed some important points that should be addressed in the field of HIV research. An update of the surveillance list of SDRMs is crucial for future DR and transmission studies, since this list was last updated more than 10 years ago, in 2009 (61). With the rapid evolutionary rate of HIV and the shift in prevalence and a higher proportion of non-B subtypes under treatment, new mutations have been emerging that could threaten the success of current ART regimens, the E138A mutation is one example (62).

As our work focused mainly on the LP vs NLP populations, we also faced problems in using the current consensus definition of LP. For that reason, we believe that the definition of late presentation should also be updated to prevent bias on the classification. It is known that when an individual is acutely infected, the CD4 cell count drops before becoming higher again, which means that if he is diagnosed in that time period, i.e. acute stage, the patient can be misdiagnosed as late presenter. To overcome this bias in our studies, we combined ambiguity rate of genomic sequences with CD4 cell count definition. So, for future research, we propose the use of a new definition of late presentation that can combine information from CD4 and ambiguity rate of the genomic sequence, furthermore, the information about estimated time of infection based on phylogenetic trees reconstruction and molecular clock analysis could also be useful for that purpose.

5.4. Policy Implications

Considering the consistent high prevalence of LP within HIV-1 infected patients and the characteristics of that population found in our studies, we propose here some strategies to reach this population.

Since LP are mainly males with heterosexual transmission, migrants and individuals aged above 50 years old, these hard-to-reach subpopulations need to be targeted. Early testing strategies are highly necessary, but until now those strategies were proven ineffective.

Two approaches that could be relevant in the context of this population are stigma reduction and health education, which could have effects on the long term. One important strategy is HIV-1 health education, lectures about HIV and the importance of HIV prevention and testing could be implemented in schools and faculties. These lectures could be important to reach young individuals and their relatives and to reduce stigma. The lectures should be given by health specialists in HIV research and clinical area and should aim to highlight the main reasons why it is important to be regularly tested at a sexually active age. All the risk clinical and behavioural aspects should be mentioned in order to raise awareness to the importance of testing for this chronic disease.

The individuals aged above 50 years old are hard-to-reach subpopulations in terms of early diagnosis. The gap report of 2014 by UNAIDS (63), states that individuals of this age group exhibit the same HIV risk behaviours that are common among young individuals. However, according to the report, older individuals suffer more psychological consequences due to discrimination and stigma and could have less satisfactory treatment outcomes due to the presence of other chronic diseases. To reach this population, it is necessary to raise awareness regarding the possibility of acquiring HIV even with older ages. For that reason, awareness campaigns should be implemented for HIV prevention, involving big and small companies and among public and private sectors that could be more misinformed.

For the migrants subpopulation, in Portugal there has been some strategies for HIV testing in checkpoints close to the regions of residency. However, the migrant population is very hard to reach, since they are mainly from Sub-Saharan Africa and there is still a lot of stigma towards HIV infection. There is also a difficulty in access to health, which makes it harder for them to be tested for HIV.

5.5. General discussion and conclusion references

1. Late Presentation Working Groups in EuroSIDA and COHERE. Estimating the burden of HIV late presentation and its attributable morbidity and mortality across Europe 2010-2016. BMC Infect Dis [Internet]. 2020 Oct 7;20(1):728. Available from: https://bmcinfectdis.biomedcentral.com/articles/10.1186/s12879-020-05261-7
2. Conway AS, Esteve A, Fernández-Quevedo M, Casabona J. Determinants and Outcomes of Late Presentation of HIV Infection in Migrants in Catalonia, Spain: PISCIS Cohort 2004–2016. J Immigr Minor Heal [Internet]. 2019 Oct 30;21(5):920–30. Available from: http://dx.doi.org/10.1007/s10903-018-0834-2
3. Zoufaly A, an der Heiden M, Marcus U, Hoffmann C, Stellbrink H, Voss L, et al. Late presentation for HIV diagnosis and care in Germany. HIV Med [Internet]. 2012 Mar;13(3):172–81. Available from: https://onlinelibrary.wiley.com/doi/10.1111/j.1468-1293.2011.00958.x
4. Hachfeld A, Ledergerber B, Darling K, Weber R, Calmy A, Battegay M, et al. Reasons for late presentation to HIV care in Switzerland. J Int AIDS Soc [Internet]. 2015 Jan;18(1):20317. Available from: http://doi.wiley.com/10.7448/IAS.18.1.20317
5. Iwuji CC, Churchill D, Gilleece Y, Weiss HA, Fisher M. Older HIV-infected individuals present late and have a higher mortality: Brighton, UK cohort study. BMC Public Health [Internet]. 2013 Dec 26;13(1):397. Available from: http://bmcpublichealth.biomedcentral.com/articles/10.1186/1471-2458-13-397
6. Antinori A, Coenen T, Costagiola D, Dedes N, Ellefson M, Gatell J, et al. Late presentation of HIV infection: a consensus definition. HIV Med [Internet]. 2011 Jan;12(1):61–4. Available from: https://onlinelibrary.wiley.com/doi/10.1111/j.1468-1293.2010.00857.x
7. Andersson E, Shao W, Bontell I, Cham F, Cuong DD, Wondwossen A, et al. Evaluation of sequence ambiguities of the HIV-1 pol gene as a method to identify recent HIV-1 infection in transmitted drug resistance surveys. Infect Genet Evol [Internet]. 2013 Aug;18(1):125–31. Available from: https://linkinghub.elsevier.com/retrieve/pii/S156713481300141X
8. Puller V, Neher R, Albert J. Estimating time of HIV-1 infection from next-generation sequence diversity. Wilke CO, editor. PLOS Comput Biol [Internet]. 2017 Oct 2;13(10):e1005775. Available from: https://dx.plos.org/10.1371/journal.pcbi.1005775
9. Liu S, Rodrigo AG, Shankarappa R, Learn GH, Hsu L, Davidov O, et al. HIV Quasispecies and Resampling. Science (80-) [Internet]. 1996 Jul 26;273(5274):415–6. Available from: https://www.science.org/doi/10.1126/science.273.5274.415
10. EuroCoord T late presenters working group in C in. Late presentation for HIV care across Europe: update from the Collaboration of Observational HIV Epidemiological Research Europe (COHERE) study, 2010 to 2013. Eurosurveillance [Internet]. 2015 Nov 26;20(47). Available from: https://www.eurosurveillance.org/content/10.2807/1560-7917.ES.2015.20.47.30070
11. Chkhartishvili N, Chokoshvili O, Bolokadze N, Tsintsadze M, Sharvadze L, Gabunia P, et al. Late presentation of HIV infection in the country of Georgia: 2012-2015. Nikolopoulos GK, editor. PLoS One [Internet]. 2017 Oct 30;12(10):e0186835. Available from: https://dx.plos.org/10.1371/journal.pone.0186835
12. Darling K, Hachfeld A, Cavassini M, Kirk O, Furrer H, Wandeler G. Late presentation to HIV care despite good access to health services: current epidemiological

trends and how to do better. Swiss Med Wkly [Internet]. 2016 Aug 21;146(August):w14348. Available from: http://doi.emh.ch/smw.2016.14348
13. Wilton J, Light L, Gardner S, Rachlis B, Conway T, Cooper C, et al. Late diagnosis, delayed presentation and late presentation among persons enrolled in a clinical HIV cohort in Ontario, Canada (1999-2013). HIV Med [Internet]. 2019 Feb;20(2):110–20. Available from: https://onlinelibrary.wiley.com/doi/10.1111/hiv.12686
14. Wójcik-cichy K, Jabłonowska O, Piekarska A. The high incidence of late presenters for HIV / AIDS infection in the Lodz province , Poland in the years 2009 – 2016 : we are still far from the UNAIDS 90 % target. AIDS Care [Internet]. 2018;0(0):1–4. Available from: https://doi.org/10.1080/09540121.2018.1470306
15. Girardi E, Sabin CA, Monforte ADA. Late Diagnosis of HIV Infection: Epidemiological Features, Consequences and Strategies to Encourage Earlier Testing. JAIDS J Acquir Immune Defic Syndr [Internet]. 2007 Sep 1;46(Suppl 1):S3–8. Available from: https://journals.lww.com/00126334-200709011-00002
16. Darcis G, Lambert I, Sauvage A, Frippiat F, Meuris C, Uurlings F, et al. Factors associated with late presentation for HIV care in a single Belgian reference center: 2006–2017. Sci Rep [Internet]. 2018 Dec 5;8(1):8594. Available from: http://www.nature.com/articles/s41598-018-26852-0
17. Clutter DS, Jordan MR, Bertagnolio S, Shafer RW. HIV-1 drug resistance and resistance testing. Infect Genet Evol [Internet]. 2016 Dec;46:292–307. Available from: http://dx.doi.org/10.1016/j.meegid.2016.08.031
18. Tostevin A, White E, Dunn D, Croxford S, Delpech V, Williams I, et al. Recent trends and patterns in HIV-1 transmitted drug resistance in the United Kingdom. HIV Med [Internet]. 2017 Mar;18(3):204–13. Available from: https://onlinelibrary.wiley.com/doi/10.1111/hiv.12414
19. Pineda-Peña A-C, Schrooten Y, Vinken L, Ferreira F, Li G, Trovão NS, et al. Trends and Predictors of Transmitted Drug Resistance (TDR) and Clusters with TDR in a Local Belgian HIV-1 Epidemic. López-Galíndez C, editor. PLoS One [Internet]. 2014 Jul 8;9(7):e101738. Available from: https://dx.plos.org/10.1371/journal.pone.0101738
20. Pingarilho M, Pimentel V, Diogo I, Fernandes S, Miranda M, Pineda-Pena A, et al. Increasing Prevalence of HIV-1 Transmitted Drug Resistance in Portugal: Implications for First Line Treatment Recommendations. Viruses [Internet]. 2020 Oct 30;12(11):1238. Available from: https://www.mdpi.com/1999-4915/12/3/268
21. Rhee S, Kassaye SG, Barrow G, Sundaramurthi JC, Jordan MR, Shafer RW. HIV-1 transmitted drug resistance surveillance: shifting trends in study design and prevalence estimates. J Int AIDS Soc [Internet]. 2020 Sep 16;23(9):1–12. Available from: https://onlinelibrary.wiley.com/doi/10.1002/jia2.25611
22. Coelho LPO, Matsuda EM, Nogueira RS, de Moraes MJ, Jamal LF, Madruga JVR, et al. Prevalence of HIV-1 transmitted drug resistance and viral suppression among recently diagnosed adults in São Paulo, Brazil. Arch Virol [Internet]. 2019 Mar 20;164(3):699–706. Available from: http://link.springer.com/10.1007/s00705-018-04122-8
23. Weng Y-W, Chen I-T, Tsai H-C, Wu K-S, Tseng Y-T, Sy C-L, et al. Trend of HIV transmitted drug resistance before and after implementation of HAART regimen restriction in the treatment of HIV-1 infected patients in southern Taiwan. BMC Infect Dis [Internet]. 2019 Dec 23;19(1):741. Available from: https://bmcinfectdis.biomedcentral.com/articles/10.1186/s12879-019-4389-1
24. Kantzanou M, Karalexi MA, Papachristou H, Vasilakis A, Rokka C, Katsoulidou

A. Transmitted drug resistance among HIV-1 drug-naïve patients in Greece. Int J Infect Dis [Internet]. 2021 Apr;105:42–8. Available from: https://doi.org/10.1016/j.ijid.2021.02.043
25. Lunar MM, Židovec Lepej S, Tomažič J, Vovko TD, Pečavar B, Turel G, et al. HIV-1 transmitted drug resistance in Slovenia and its impact on predicted treatment effectiveness: 2011–2016 update. Zhang C, editor. PLoS One [Internet]. 2018 Apr 26;13(4):e0196670. Available from: https://dx.plos.org/10.1371/journal.pone.0196670
26. Rossetti B, Di Giambenedetto S, Torti C, Postorino M, Punzi G, Saladini F, et al. Evolution of transmitted HIV-1 drug resistance and viral subtypes circulation in Italy from 2006 to 2016. HIV Med [Internet]. 2018 Oct;19(9):619–28. Available from: https://onlinelibrary.wiley.com/doi/10.1111/hiv.12640
27. Hofstra LM, Sauvageot N, Albert J, Alexiev I, Garcia F, Struck D, et al. Transmission of HIV Drug Resistance and the Predicted Effect on Current First-line Regimens in Europe. Clin Infect Dis [Internet]. 2016 Mar 1;62(5):655–63. Available from: https://academic.oup.com/cid/article-lookup/doi/10.1093/cid/civ963
28. Lai A, Franzetti M, Bergna A, Saladini F, Bruzzone B, Di Giambenedetto S, et al. Marked decrease in acquired resistance to antiretrovirals in latest years in Italy. Clin Microbiol Infect [Internet]. 2021 Jul;27(7):1038.e1-1038.e6. Available from: https://doi.org/10.1016/j.cmi.2020.09.028
29. Rocheleau G, Brumme CJ, Shoveller J, Lima VD, Harrigan PR. Longitudinal trends of HIV drug resistance in a large Canadian cohort, 1996–2016. Clin Microbiol Infect [Internet]. 2018 Feb;24(2):185–91. Available from: https://doi.org/10.1016/j.cmi.2017.06.014
30. Abela IA, Scherrer AU, Böni J, Yerly S, Klimkait T, Perreau M, et al. Emergence of Drug Resistance in the Swiss HIV Cohort Study Under Potent Antiretroviral Therapy Is Observed in Socially Disadvantaged Patients. Clin Infect Dis [Internet]. 2020 Jan 2;70(2):297–303. Available from: https://academic.oup.com/cid/article/70/2/297/5370584
31. Luber AD. Genetic Barriers to Resistance and Impact on Clinical Response. J Int AIDS Soc [Internet]. 2005 Feb;7(1):69–69. Available from: http://doi.wiley.com/10.1186/1758-2652-7-3-69
32. Schmidt D, Kollan C, Fätkenheuer G, Schülter E, Stellbrink H-J, Noah C, et al. Estimating Trends in the Proportion of Transmitted and Acquired HIV Drug Resistance in a Long Term Observational Cohort in Germany. Harrigan PR, editor. PLoS One [Internet]. 2014 Aug 22;9(8):e104474. Available from: https://dx.plos.org/10.1371/journal.pone.0104474
33. Franzetti M, De Luca A, Ceccherini-Silberstein F, Spagnuolo V, Nicastri E, Mussini C, et al. Evolution of HIV-1 transmitted drug resistance in Italy in the 2007–2014 period: A weighted analysis. J Clin Virol [Internet]. 2018;106(October 2017):49–52. Available from: https://doi.org/10.1016/j.jcv.2018.07.009
34. Lombardi F, Giacomelli A, Armenia D, Lai A, Dusina A, Bezenchek A, et al. Prevalence and factors associated with HIV-1 multi-drug resistance over the past two decades in the Italian ARCA database. Int J Antimicrob Agents [Internet]. 2021 Feb;57(2):106252. Available from: https://linkinghub.elsevier.com/retrieve/pii/S0924857920304726
35. Gandhi RT, Wurcel A, Rosenberg ES, Johnston MN, Hellmann N, Bates M, et al. Progressive Reversion of Human Immunodeficiency Virus Type 1 Resistance Mutations In Vivo after Transmission of a Multiply Drug-Resistant Virus. Clin Infect Dis [Internet].

2003 Dec 15;37(12):1693–8. Available from: https://academic.oup.com/cid/article-lookup/doi/10.1086/379773
36. Winand R, Theys K, Eusébio M, Aerts J, Camacho RJ, Gomes P, et al. Assessing transmissibility of HIV-1 drug resistance mutations from treated and from drug-naive individuals. AIDS [Internet]. 2015 Sep 24;29(15):2045–52. Available from: http://content.wkhealth.com/linkback/openurl?sid=WKPTLP:landingpage&an=00002030-201509240-00017
37. Nagaraja P, Alexander HK, Bonhoeffer S, Dixit NM. Influence of recombination on acquisition and reversion of immune escape and compensatory mutations in HIV-1. Epidemics [Internet]. 2016 Mar;14:11–25. Available from: http://dx.doi.org/10.1016/j.epidem.2015.09.001
38. Rahim S, Fredrick LM, da Silva BA, Bernstein B, King MS. Geographic and Temporal Trends of Transmitted HIV-1 Drug Resistance Among Antiretroviral-Naïve Subjects Screening for Two Clinical Trials in North America and Western Europe. HIV Clin Trials [Internet]. 2009 Apr 6;10(2):94–103. Available from: http://www.tandfonline.com/doi/full/10.1310/hct1002-94
39. Rhee S-Y, Jordan MR, Raizes E, Chua A, Parkin N, Kantor R, et al. HIV-1 Drug Resistance Mutations: Potential Applications for Point-of-Care Genotypic Resistance Testing. Harrigan PR, editor. PLoS One [Internet]. 2015 Dec 30;10(12):e0145772. Available from: https://dx.plos.org/10.1371/journal.pone.0145772
40. Parczewski M, Leszczyszyn-Pynka M, Witak-Jędra M, Maciejewska K, Rymer W, Szymczak A, et al. Transmitted HIV drug resistance in antiretroviral-treatment-naive patients from Poland differs by transmission category and subtype. J Antimicrob Chemother [Internet]. 2015 Jan 1;70(1):233–42. Available from: https://academic.oup.com/jac/article-lookup/doi/10.1093/jac/dku372
41. Zou X, He J, Zheng J, Malmgren R, Li W, Wei X, et al. Prevalence of acquired drug resistance mutations in antiretroviral- experiencing subjects from 2012 to 2017 in Hunan Province of central South China. Virol J [Internet]. 2020 Dec 17;17(1):38. Available from: https://virologyj.biomedcentral.com/articles/10.1186/s12985-020-01311-3
42. NRTI Resistance Notes - HIV Drug Resistance Database [Internet]. [cited 2022 Mar 24]. Available from: https://hivdb.stanford.edu/dr-summary/resistance-notes/NRTI/#thymidine.analog.mutations.tams.
43. NNRTI Resistance Notes - HIV Drug Resistance Database [Internet]. [cited 2022 Mar 24]. Available from: https://hivdb.stanford.edu/dr-summary/resistance-notes/NNRTI/
44. PI Resistance Notes - HIV Drug Resistance Database [Internet]. [cited 2022 Apr 13]. Available from: https://hivdb.stanford.edu/dr-summary/resistance-notes/PI/
45. Westin MR, Biscione FM, Fonseca M, Ordones M, Rodrigues M, Greco DB, et al. Resistance-Associated Mutation Prevalence According to Subtypes B and Non-B of HIV Type 1 in Antiretroviral-Experienced Patients in Minas Gerais, Brazil. AIDS Res Hum Retroviruses [Internet]. 2011 Sep;27(9):981–7. Available from: http://www.liebertpub.com/doi/10.1089/aid.2010.0260
46. Abecasis AB, Deforche K, Bacheler LT, McKenna P, Carvalho AP, Gomes P, et al. Investigation of baseline susceptibility to protease inhibitors in HIV-1 subtypes C, F, G and CRF02_AG. Antivir Ther [Internet]. 2006;11(5):581–9. Available from: http://www.ncbi.nlm.nih.gov/pubmed/16964826
47. Abecasis AB, Pingarilho M, Vandamme A-M. Phylogenetic analysis as a forensic

tool in HIV transmission investigations. AIDS [Internet]. 2018 Mar 13;32(5):543–54. Available from: https://journals.lww.com/00002030-201803130-00002
48. Paraskevis D, Beloukas A, Stasinos K, Pantazis N, de Mendoza C, Bannert N, et al. HIV-1 molecular transmission clusters in nine European countries and Canada: association with demographic and clinical factors. BMC Med [Internet]. 2019 Dec 8;17(1):4. Available from: https://bmcmedicine.biomedcentral.com/articles/10.1186/s12916-018-1241-1
49. Delgado E, Benito S, Montero V, Cuevas MT, Fernández-García A, Sánchez-Martínez M, et al. Diverse Large HIV-1 Non-subtype B Clusters Are Spreading Among Men Who Have Sex With Men in Spain. Front Microbiol [Internet]. 2019 Apr 3;10(APR):1–19. Available from: https://www.frontiersin.org/article/10.3389/fmicb.2019.00655/full
50. Petersen A, Cowan SA, Nielsen J, Fischer TK, Fonager J. Characterisation of HIV-1 transmission clusters and drug-resistant mutations in Denmark, 2004 to 2016. Eurosurveillance [Internet]. 2018 Nov 1;23(44):1–9. Available from: https://www.eurosurveillance.org/content/10.2807/1560-7917.ES.2018.23.44.1700633
51. Hanke K, Faria NR, Kühnert D, Yousef KP, Hauser A, Meixenberger K, et al. Reconstruction of the Genetic History and the Current Spread of HIV-1 Subtype A in Germany. Kirchhoff F, editor. J Virol [Internet]. 2019 Jun 15;93(12). Available from: https://journals.asm.org/doi/10.1128/JVI.02238-18
52. Lai A, Bozzi G, Franzetti M, Binda F, Simonetti FR, De Luca A, et al. HIV-1 A1 Subtype Epidemic in Italy Originated from Africa and Eastern Europe and Shows a High Frequency of Transmission Chains Involving Intravenous Drug Users. Tornesello ML, editor. PLoS One [Internet]. 2016 Jan 11;11(1):e0146097. Available from: https://dx.plos.org/10.1371/journal.pone.0146097
53. Yebra G, Holguín Á, Pillay D, Hué S. Phylogenetic and demographic characterization of HIV-1 transmission in Madrid, Spain. Infect Genet Evol [Internet]. 2013;14(1):232–9. Available from: http://dx.doi.org/10.1016/j.meegid.2012.12.006
54. Lorenzin G, Gargiulo F, Caruso A, Caccuri F, Focà E, Celotti A, et al. Prevalence of Non-B HIV-1 Subtypes in North Italy and Analysis of Transmission Clusters Based on Sequence Data Analysis. Microorganisms [Internet]. 2019 Dec 23;8(1):36. Available from: https://www.mdpi.com/2076-2607/8/1/36
55. Pingarilho M, Pimentel V, Miranda MNS, Silva AR, Diniz A, Ascenção BB, et al. HIV-1-Transmitted Drug Resistance and Transmission Clusters in Newly Diagnosed Patients in Portugal Between 2014 and 2019. Front Microbiol [Internet]. 2022;13. Available from: https://www.frontiersin.org/article/10.3389/fmicb.2022.823208
56. Pimentel VF, Pingarilho M, Sole G, Alves D, Miranda M, Diogo I, et al. Differential patterns of postmigration HIV-1 infection acquisition among Portuguese immigrants of different geographical origins. AIDS [Internet]. 2022 Jun 1;36(7):997–1005. Available from: https://journals.lww.com/aidsonline/Fulltext/9000/Differential_patterns_of_post_migration_HIV_1.96205.aspx
57. MacCarthy S, Brignol S, Reddy M, Nunn A, Dourado I. Late presentation to HIV/AIDS care in Brazil among men who self-identify as heterosexual. Rev Saude Publica [Internet]. 2016;50:54. Available from: http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0034-89102016000100233&lng=en&tlng=en
58. Kostaki EG, Gova M, Adamis G, Xylomenos G, Chini M, Mangafas N, et al. A

Nationwide Study about the Dispersal Patterns of the Predominant HIV-1 Subtypes A1 and B in Greece: Inference of the Molecular Transmission Clusters. Viruses [Internet]. 2020 Oct 19;12(10):1183. Available from: https://www.mdpi.com/1999-4915/12/10/1183
59. Fakoya I, Álvarez-del Arco D, Copas AJ, Teixeira B, Block K, Gennotte A-F, et al. Factors Associated With Access to HIV Testing and Primary Care Among Migrants Living in Europe: Cross-Sectional Survey. JMIR Public Heal Surveill [Internet]. 2017 Nov 6;3(4):e84. Available from: http://publichealth.jmir.org/2017/4/e84/
60. Migrants | UNAIDS [Internet]. [cited 2022 Jul 21]. Available from: https://www.unaids.org/en/resources/documents/2014/Migrants
61. Bennett DE, Camacho RJ, Otelea D, Kuritzkes DR, Fleury H, Kiuchi M, et al. Drug Resistance Mutations for Surveillance of Transmitted HIV-1 Drug-Resistance: 2009 Update. Nixon DF, editor. PLoS One [Internet]. 2009 Mar 6;4(3):e4724. Available from: https://dx.plos.org/10.1371/journal.pone.0004724
62. Sluis-Cremer N, Jordan MR, Huber K, Wallis CL, Bertagnolio S, Mellors JW, et al. E138A in HIV-1 reverse transcriptase is more common in subtype C than B: Implications for rilpivirine use in resource-limited settings. Antiviral Res [Internet]. 2014 Jul;107(1):31–4. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3624763/pdf/nihms412728.pdf
63. People aged 50 years and older | UNAIDS [Internet]. [cited 2022 Jul 21]. Available from: https://www.unaids.org/en/resources/documents/2014/Peopleaged50yearsandolder