



INSTITUTO DE HIGIENE E  
**MEDICINA TROPICAL**  
DESDE 1902

**Universidade Nova de Lisboa**

**Instituto de Higiene e Medicina Tropical**

Apresentação tardia ao diagnóstico de VIH: uma análise da  
definição de *late presenters*

**André Alexandre Correia Alemão**

DISSERTAÇÃO PARA A OBTENÇÃO DO GRAU DE MESTRE EM CIÊNCIAS BIOMÉDICAS

(março, 2024)





INSTITUTO DE HIGIENE E  
MEDICINA TROPICAL  
DESDE 1902

**Universidade Nova de Lisboa**

**Instituto de Higiene e Medicina Tropical**

Apresentação tardia ao diagnóstico de VIH: uma análise da  
definição de *late presenters*

**Autor:** André Alexandre Correia Alemão

**Orientador:** Mafalda Miranda

**Coorientador:** Ana Abecasis

Dissertação apresentada para cumprimento dos requisitos necessários à obtenção do  
grau de mestre em Ciências Biomédicas

(março ,2024)

## **Agradecimentos**

Ao concluir a dissertação de mestrado, é com o maior prazer, que agradeço a todos os que de alguma forma contribuíram para a sua realização.

Em primeiro lugar gostava de agradecer à Dra. Mafalda Miranda, pelo facto de ter aceitado orientar-me ao longo da realização desta dissertação, sendo a sua primeira vez a orientar um aluno a realizar uma dissertação de mestrado. Para além disso, gostava de agradecer pela excelente orientação, pela ajuda ao longo da realização do trabalho, de forma a ser possível que tudo se tornasse mais fácil de concretizar, pela permanente disponibilidade, pela paciência, pela dedicação e pelo conhecimento transmitido ao longo deste tempo.

Em segundo lugar, gostava de agradecer à minha coorientadora, a Professora Ana Abecassis por me ter recebido de braços abertos na sua equipa e pela disponibilidade sempre demonstrada.

À minha família, em especial aos meus pais por acreditarem em mim, por me incentivarem e apoiarem ao longo do meu percurso académico, pois sem eles nada disto seria possível. Agradeço ainda por toda a paciência, amor e disponibilidade que demonstraram sempre.

Aos meus amigos, que sempre estiveram presentes e sempre me apoiaram, um enorme obrigado por toda a paciência e carinho demonstrado, e entretida nos bons e maus momentos.

## Resumo

O diagnóstico tardio do VIH pode estar associado a um maior risco de transmissão do vírus e conseqüentemente levar a custos mais elevados de cuidados de saúde. Os programas de testagem do VIH são muito importantes para aumentar o diagnóstico precoce. Em 2010, o Grupo de Trabalho Europeu de Consenso sobre *late presenters*, definiu os apresentadores tardios como indivíduos com uma contagem de células CD4 inferior a 350 células/ $\mu$ L no momento do diagnóstico, ou um evento definidor de SIDA no momento do diagnóstico, independentemente da contagem de células CD4. Devido à abrupta queda inicial da contagem de CD4 que ocorre na fase aguda da infecção pelo VIH-1, a definição de apresentação tardia levou por vezes a uma classificação errada de indivíduos na fase aguda como apresentadores tardios. Em 2022, um segundo grupo de trabalho sugeriu uma atualização na definição de apresentação tardia, considerando outros critérios. Nesta definição, os *non-late presenters* são definidos como indivíduos que apresentam evidência laboratorial de infecção recente, último teste de VIH negativo nos últimos 12 meses ou evidência clínica de infecção aguda. Quando os indivíduos apresentam CD4 abaixo de 350 células/ $\mu$ L, temos de avaliar estes três parâmetros com o intuito de excluir um indivíduo classificado erradamente como *late presenter*. Os indivíduos diagnosticados durante o período de seroconversão não devem ser considerados como apresentadores tardios. No entanto, esta definição baseia-se em informações clínicas que podem não estar disponíveis em muitos casos. Assim, este estudo tem o objetivo de analisar a definição de *late presenters* através da utilização de outras metodologias para além da definição existente considerando os valores de células CD4 e o último teste de VIH negativo, com o intuito de evitar este viés de classificação.

Palavras-chave: *Late presenters*, *Non-Late Presenters* VIH, CD4

## **Abstract**

Late HIV diagnosis is thought to be associated with a higher risk of HIV transmission and higher health care costs. HIV testing programs are very important to increase early diagnosis. In 2010 the European Late Presenter Consensus Working Group, defined late presenters as individuals with a CD4 cell count lower than 350 cells/ $\mu$ L at diagnosis, or an AIDS-defining event at diagnosis, regardless of CD4 cell count. Due to the initial drop of CD4 count in the acute phase of HIV-1 infection, the definition of late presentation sometimes led to a misclassification of individuals in acute infection as late presenters. In 2022, Sara Croxford's team suggested an update to the definition of late presenters. In this definition, non-late presenters are defined as individuals who have laboratory evidence of a recent infection, a last negative HIV test in the last 12 months, or clinical evidence of an acute infection. Individuals who present a CD4 cell count below 350cells/ $\mu$ L need to be evaluated for these three parameters with the objective of excluding a misclassified individual as late presenter. However, this definition relies on previous information that may not be available in many cases. Therefore, this study aims to analyze the definition of late presenters through the use of other methodologies in addition to the existing definition considering CD4 cell count and the last negative HIV test, with the aim of avoiding this classification bias.

Keywords: *Late presenters, Non-Late Presenters* HIV, CD4

# Índice

1-Introdução.....	12
1.1- Epidemiologia e Contexto Histórico do VIH e SIDA.....	12
1.2- Origem VIH.....	13
1.4- Estrutura da Partícula Viral e Genoma do VIH-1 .....	15
1.5- Ciclo replicativo .....	17
1.6- Vias de transmissão .....	19
1.7- Fases Clínicas.....	19
1.8- Terapia Antirretroviral.....	21
1.9- Resistência à terapia Antirretroviral (ARV).....	24
1.10- <i>Late Presenters</i> .....	24
2- Objetivos.....	27
2.1- Objetivo Geral .....	27
2.2- Objetivos Específicos .....	27
3- Materiais e Métodos .....	28
3.1- População em estudo .....	28
3.2- Classificação com base no valor de células CD4 .....	28
3.3- Classificação com base na taxa de ambiguidade.....	28

3.4- Classificação com base na análise filogenética .....	29
3.5- Determinação dos subtipos do VIH-1 .....	29
3.6-Análise Estatística .....	30
4- Resultados.....	31
4.1- Dados sociodemográficos e clínicos dos indivíduos .....	31
4.2- Análise 1 .....	32
4.3- Análise 2 .....	33
4.4- Análise 3 .....	35
4.5- Análise 4 .....	36
5- Discussão e Conclusões.....	38
6- Bibliografia.....	43

## Abreviaturas

ADN	Ácido Desoxirribonucleico
ARN	Ácido Ribonucleico
ARV	Resistência à terapia Antirretroviral
CA	Cápside
CDC	<i>Center for Disease Control</i>
CRF	Formas Recombinantes Circulantes
ECDC	Centro Europeu de Prevenção e Controle de Doenças
EI	Inibidores de Entrada
FDA	<i>Food and Drug Administration</i>
gp41	Glicoproteína 41
gp120	Glicoproteína 120
GTR	<i>General Time Reversible</i>
HAART	Terapia antirretroviral altamente ativa
HSH	Homens que têm sexo com Homens
IVDA	Transmissão por seringas e instrumentos infetados (IVDA)
LP	<i>Late presenters</i>
LTR	Long Terminal Repeats
MA	Matriz

NC	Nucleocápside
NLP	<i>Non-Late Presenters</i>
nm	Nanômetro
NNRTI	Inibidores da Transcriptase Reversa não-análogos a nucleósidos
NRTI	Inibidores da Transcriptase Reversa análogos a nucleósidos
OMS	Organização Mundial da Saúde
PI	Inibidores da Protease
RT	Transcriptase Reversa
SIDA	síndrome da imunodeficiência adquirida
SIV	Vírus da imunodeficiência dos símios
SPSS	<i>Statistical Package for the Social Science</i>
UDIs	Transmissão por utilização de drogas injetáveis
URF	Formas Recombinantes Únicas
VIH	Vírus da Imunodeficiência Human

# 1-Introdução

## 1.1- Epidemiologia e Contexto Histórico do VIH e SIDA

A síndrome da imunodeficiência adquirida (SIDA) foi identificada em 1981 nos Estados Unidos da América pelo *Center for Disease Control* (CDC), devido ao aparecimento de inúmeras doenças oportunistas, tais como a pneumonia e o sarcoma de Kaposi, em homens que têm sexo com homens e pessoas que usavam drogas injetáveis (Greene, 2007; Pang *et al.*, 2018).

Em 1983, uma equipa francesa do Instituto Pasteur isolou pela primeira vez o agente etiológico causador de SIDA (Barré-Sinoussi, 2009). Esse agente etiológico ficou conhecido como vírus da imunodeficiência humana (VIH), agente responsável pela SIDA. O VIH pode-se transmitir através de transmissão vertical, via endovenosa ou via sexual (Barré-Sinoussi, 2009; De Cock, Jaffe and Curran, 2012).

Desde o início da pandemia já foram infetadas 85,6 milhões de pessoas, entre as quais 40,4 milhões morreram de SIDA. Estima-se que em 2022 existiam 39 milhões de pessoas a viverem com VIH em todo o mundo, onde 1,3 milhões tinham sido infetados em 2022 e 630 mil pessoas morreram nesse mesmo ano de SIDA. Dos 39 milhões de pessoas que vivem com VIH, 53% pertencem ao sexo feminino. Em 2022, cerca de 86% dos indivíduos que tinham VIH sabiam que estavam infetados. No entanto cerca de 5,5 milhões de indivíduos estavam infetados, mas não sabiam que estavam (UNAIDS, 2023).

Enquanto em 1995 foram infetadas 3,2 milhões de pessoas com VIH, em 2022 estas foram 1,3 milhões, tendo resultado num decréscimo de 59% nas novas infeções diagnosticadas anualmente. O sexo feminino representava 46% das novas infeções em 2022. Desde 2010 que as infeções recentes reduziram em 38% passando de 2,1 milhões de pessoas para 1,3 milhões (UNAIDS, 2023).

## 1.2- Origem VIH

A SIDA nos humanos, é causada pelo vírus da imunodeficiência humana que se pode classificar em dois tipos, o tipo 1 e o tipo 2 (VIH-1 e VIH-2). Ambos os tipos do vírus resultaram de infeções zoonóticas de humanos a partir de vírus da imunodeficiência dos símios (SIVs) que infetavam os primatas do continente africano. O VIH-1 pertence à família *Retroviridae* e ao género *Lentivirus* (Kirchhoff, 2016),

A transmissão cruzada entre os humanos e os símios ocorreu provavelmente devido ao processo de caça e de abate dos primatas, e consequentemente, seu ao comércio e domesticação. A origem do VIH-1 teve origem em pelo menos 4 transmissões zoonóticas distintas, originando os diferentes grupos (M, N, O e P) (Hemelaar, 2012).

Os grupos M e N de VIH-1 tiveram origem diretamente do SIVcpz encontrado nos chimpanzés (*Pan troglodytes troglodytes*) na África Centro-Occidental. Já os grupos O e P estão relacionados com o SIVgor encontrados em gorilas (*Gorilla gorilla*) nos Camarões (Hemelaar, 2012).

## 1.3. Diversidade Genética e Epidemiologia Molecular

O grupo M do VIH-1 é responsável pela pandemia global do VIH e é composto

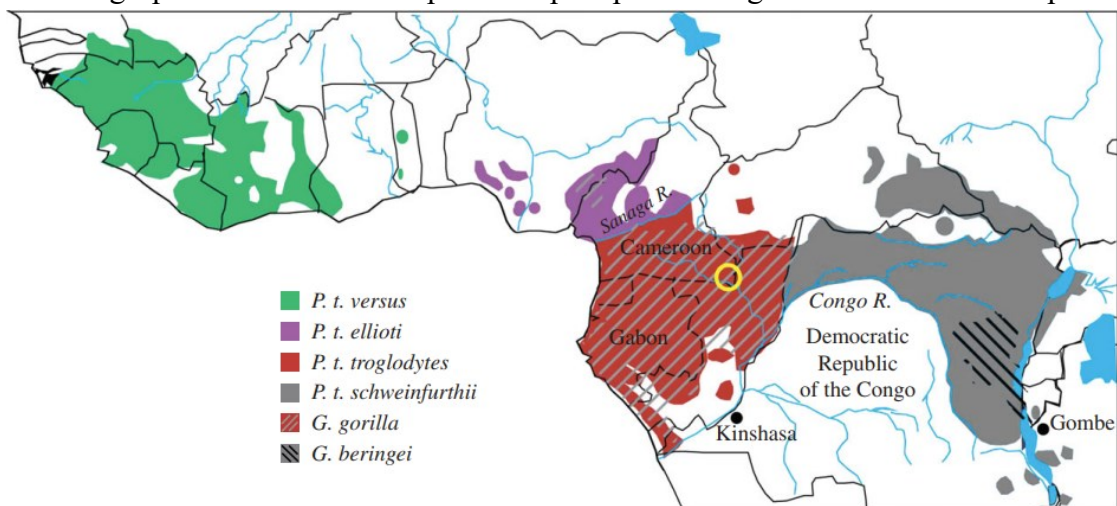


Figura 1- Mapa do continente africano com a representação da distribuição geográfica das subespécies de chimpanzés relacionadas com a origem do VIH-1 (Sharp and Hahn, 2010)

pelos subtipos A, B, C, D, F, G, H, J, K e L. Para além disso também existem 140 Formas Recombinantes Circulantes (CRF) e algumas Formas Recombinantes Únicas (URF),

representando 23% das estirpes de VIH em circulação (Cañada *et al.*, 2021; Mendes Da Silva *et al.*, 2021; Los Alamos Database, 2023).

Devido à replicação viral e aos altos níveis de mutação do VIH que ocorrem dentro da célula hospedeira, um indivíduo pode ser infetado por mais do que uma estirpe do vírus em simultâneo. Este fenómeno origina as CRFs (*Circulating Recombinant Form*) ou URFs (*Unique Recombinat Form*). A principal diferença entre as duas é que enquanto a CRF tem impacto epidemiológico, a URF não tem. (Hemelaar *et al.*, 2019; Cañada *et al.*, 2021; Tang *et al.*, 2021).

O subtipo C é responsável por cerca de 46,6% das infeções de VIH a nível mundial, sendo o subtipo com maior prevalência na região do sul de África e em países como a Etiópia e a Índia. O subtipo B, o mais disseminado a nível mundial, é responsável por 12,1% das infeções, sendo mais prevalente na América Latina, na América do Norte, na Europa Ocidental e Central. Já o subtipo A é responsável por 10,2% das infeções sendo mais prevalente na África Oriental, na Europa Oriental e na Ásia Central. O CRF02\_AG é responsável por 7,8% das infeções a nível mundial sendo mais prevalente na região da África central. O CRF01\_AE é responsável por 5,3% das infeções sendo mais prevalente no continente asiático. (Hemelaar *et al.*, 2019).

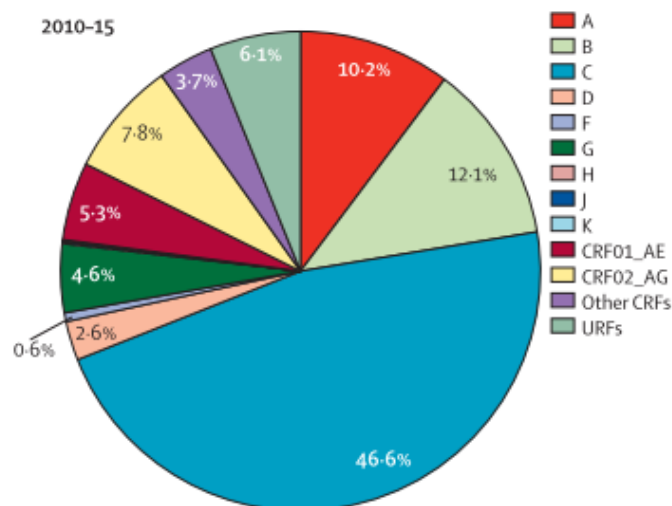


Figura 2- Distribuição mundial dos subtipos do VIH-1, CRFs e URFs entre 2010-2015 (Hemelaar *et al.*, 2019)

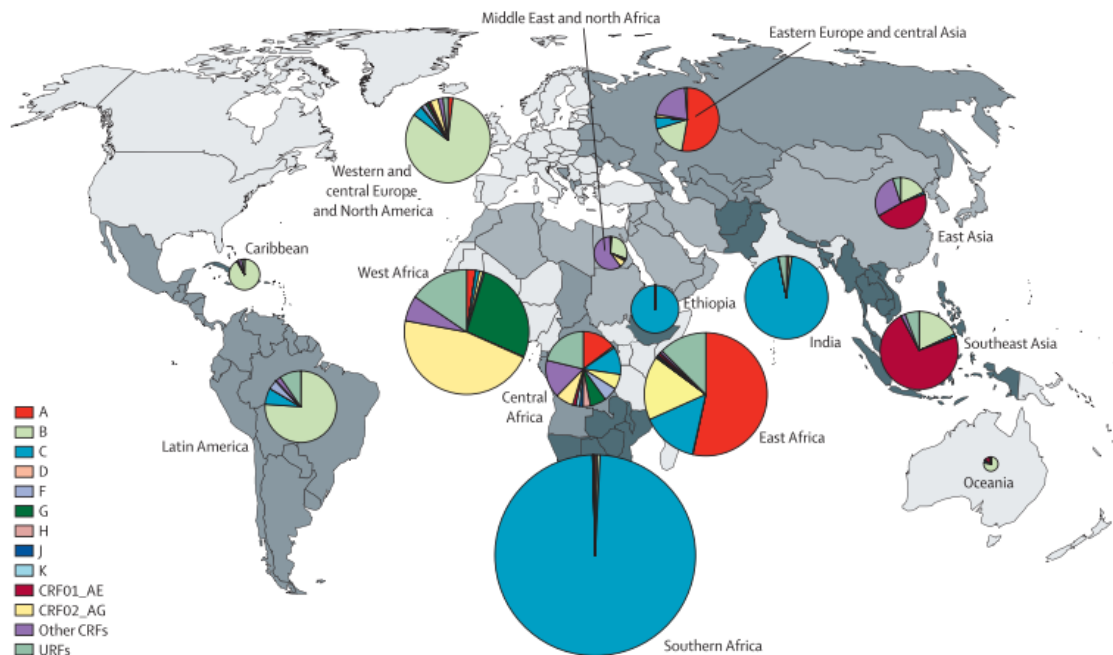


Figura 3- Distribuição geográfica dos subtipos do VIH-1, CRFs e URFs entre 2010-2015 (Hemelaar et al., 2019)

#### 1.4- Estrutura da Partícula Viral e Genoma do VIH-1

A partícula viral do VIH tem uma forma esférica com 100 nm de diâmetro e contém uma bicamada lipídica que é derivada da célula hospedeira, que rodeia uma cápside em forma de cone truncado. A bicamada lipídica está ligada à glicoproteína de superfície (gp120) e à glicoproteína transmembranar (gp41). Para além disso, a partícula viral também é constituída pela matriz (MA), pela cápside (CA), pela nucleocapsíde (NC), que estão localizadas no genoma do VIH, e ainda é constituído pelo ARN (Ácido Ribonucleico) genómico, pela protease viral, pela transcriptase reversa (RT) e pela integrase (Emanuele Fanales-Belasio, Mariangela Raimondo, 2010; Kirchhoff, 2016).

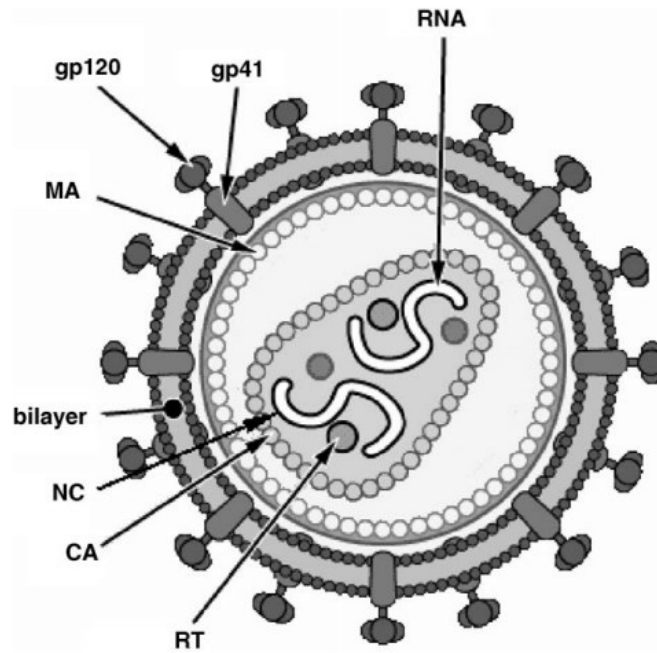


Figura 4- Estrutura da partícula viral do VIH-1 (Sierra, Kupfer and Kaiser, 2005)

O genoma viral é composto por duas moléculas de ARN, por 3 genes principais (*gag*, *pol* e *env*), por genes acessórios (*tat*, *env* e *rev*) e por genes reguladores (*vpr*, *vif* e *nef*). Para além disso, é flanqueado pelas LTR (*Long Terminal Repeats*), onde contém o promotor viral, sendo responsável pelo início da transcrição do genoma (Emanuele Fanales-Belasio, Mariangela Raimondo, 2010; Kirchhoff, 2016).

O gene *gag* codifica as proteínas estruturais do núcleo, entre as quais a cápside (p24) e a nucleocápside (p7) e ainda codifica a matriz (p17). O gene *env* codifica a glicoproteína de superfície (gp120) e a glicoproteína transmembranar (gp41), que reconhecem os recetores celulares. Por fim, o gene *pol* codifica as enzimas necessárias para que ocorra a replicação viral, sendo estas: a transcriptase reversa, que transforma Ácido Ribonucleico (ARN) viral em Ácido Desoxirribonucleico (ADN), a integrase que incorpora o ADN viral nos cromossomas da célula hospedeira e a protease que cliva os precursores proteicos dos genes *gag* e *pol* nos seus componentes (Emanuele Fanales-Belasio, Mariangela Raimondo, 2010; Kirchhoff, 2016).

Dentro dos genes reguladores/essenciais, temos o gene *tat*, o gene *env* e o gene *rev* e dentro dos genes não essenciais para o ciclo replicativo temos os genes *vpr*, *vif*, *nef* e *vpu* (Emanuele Fanales-Belasio, Mariangela Raimondo, 2010; Kirchhoff, 2016).

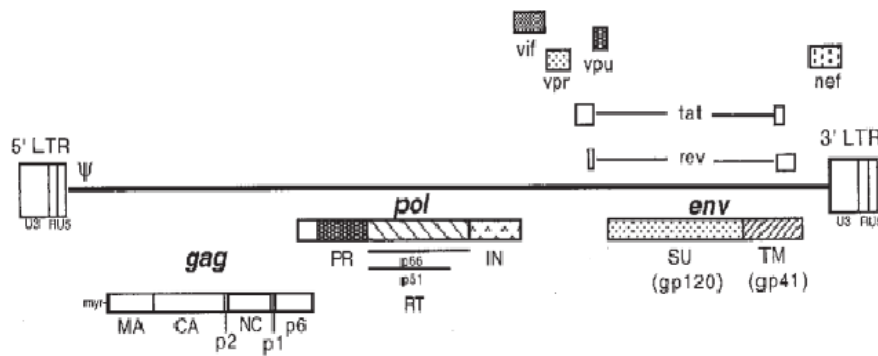


Figura 5- Estrutura do Genoma do VIH-1 (Freed, 2002)

### 1.5- Ciclo replicativo

A gp120 e a gp41 são responsáveis pelo reconhecimento do vírus levando assim à sua entrada na célula hospedeira. A gp41 contém um péptido hidrofóbico que é essencial para que ocorra a fusão da membrana do vírus. A gp120 liga-se às células CD4, que é expresso na superfície da célula, funcionando como coreceptor de membrana. Assim que a gp120 se liga às células CD4, o envelope viral sofre uma alteração conformacional, expondo o domínio específico da gp120 capaz de se ligar aos recetores de quimiocina na membrana da célula (Emanuele Fanales-Belasio, Mariangela Raimondo, 2010).

As interações dos coreceptores induzem uma alteração conformacional adicional permitindo a fusão da gp41 com a membrana celular, efetuando o primeiro contacto direto

entre o vírus e a célula alvo. A seguir à fusão da membrana, o material genético do vírus é libertado na célula hospedeira, libertando o material genético da partícula viral na célula hospedeira (Emanuele Fanales-Belasio, Mariangela Raimondo, 2010; Kirchhoff, 2016).

A conversão do ARN viral em ADN proviral ocorre devido à ação da transcriptase reversa e da integrase. A transcriptase reversa inicia a sua função no citoplasma da célula hospedeira, transformando uma molécula de ARN numa molécula de ADN de dupla hélice, que é integrada no genoma da célula hospedeira através da integrase (Emanuele Fanales-Belasio, Mariangela Raimondo, 2010).

Em seguida, ocorre a transcrição do ADN proviral em ARN mensageiro. O processo de transcrição resulta da síntese das proteínas reguladoras do VIH-1 como a Tat e a Rev., e da ação das enzimas da célula hospedeira. O ARN mensageiro migra para o citoplasma, onde as proteínas do novo virião são sintetizadas (Emanuele Fanales-Belasio, Mariangela Raimondo, 2010).

A formação da nova partícula viral ocorre em dois passos, as cadeias de ARN virais associam-se às enzimas de replicação, enquanto as proteínas do núcleo juntam-se a estas formando a cápside (Emanuele Fanales-Belasio, Mariangela Raimondo, 2010).

Esta partícula viral imatura vai para a superfície celular, ligando-se à membrana celular, adquirindo assim um novo envelope, sendo a nova partícula viral libertada para o espaço extracelular (Emanuele Fanales-Belasio, Mariangela Raimondo, 2010).

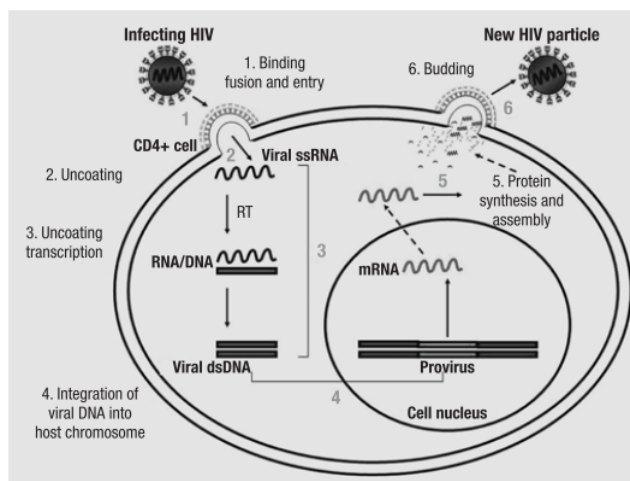


Figura 6- Ciclo replicativo do VIH (Emanuele Fanales-Belasio, Mariangela Raimondo, 2010)

### 1.6- Vias de transmissão

O VIH pode ser transmitido através de diversas vias. As principais vias de transmissão do VIH são a transmissão sexual, transmissão por utilização de drogas injetáveis (UDIs), transmissão vertical (mãe para filho) e transmissão por contacto com sangue ou produtos de sangue infetado (Galvin, Cohen and Hill, 2004).

O tipo de transmissão mais frequente varia de acordo com a área geográfica em estudo. Por exemplo, na África Subsaariana o tipo de transmissão mais frequente é através do contacto heterossexual, por outro lado, na Europa Oriental, Ásia Central, Norte de África e Médio Oriente é o uso de drogas injetáveis. Na América do Norte, na Europa Ocidental, e na Oceânia a forma mais recorrente de transmissão de VIH é por contacto sexual entre Homens que têm sexo com Homens (HSH) (Fettig *et al.*, 2014).

### 1.7- Fases Clínicas

Existem 3 fases de infeção pelo VIH, caracterizadas por sintomas clínicos e por marcadores biológicos, que são bastante importantes no diagnóstico e na monitorização do desenvolvimento clínico (Parekh *et al.*, 2019)

A primeira fase é a fase aguda, sendo caracterizada por uma rápida multiplicação e propagação do vírus no corpo humano, que pode durar entre 2 a 4 semanas após a infecção. Durante esta fase, ocorre um aumento do antígeno p24 na corrente sanguínea. A sintomatologia associada a esta fase é semelhante à de uma gripe comum, algumas pessoas apresentam dor de cabeça, febre e erupções cutâneas que podem durar várias semanas (Parekh *et al.*, 2019). É nesta fase que há um possível viés de classificação de indivíduos como *late presenters*, pois as células CD4 podem atingir níveis muito baixos, o que depois pode de alguma forma reverter (Croxford *et al.*, 2022).

A segunda fase é a fase crónica ou assintomática, em que a multiplicação viral continua a ocorrer, mas a um ritmo mais lento, e a maioria das pessoas infetadas não apresentam nenhum sintoma. Nesta fase, o sistema imunológico começa a produzir anticorpos, coincidindo com um decréscimo da carga viral (Parekh *et al.*, 2019).

Para além disso, os antígenos p24 detetáveis vão diminuindo, uma vez que o antígeno p24 liga-se aos anticorpos, formando um complexo anticorpo-antígeno p24. O período desde a infecção até o aparecimento dos anticorpos é conhecido como “período de seroconversão”. Se um indivíduo infetado permanecer sem tratamento, à medida que a replicação viral continua, as células CD4, que servem como alvo da replicação viral, vão sendo destruídas gradualmente, levando a um decréscimo da sua concentração no sangue (Parekh *et al.*, 2019).

Na terceira fase ou mais conhecida como SIDA, continua a ocorrer a replicação viral e depleção das células CD4, levando a um enfraquecimento do sistema imune, sendo caracterizada pelo aparecimento de doenças oportunistas, tais como a tuberculose, a pneumonia, a candidíase e a infecção pelo citomegalovírus. (Pang *et al.*, 2018; Parekh *et al.*, 2019).

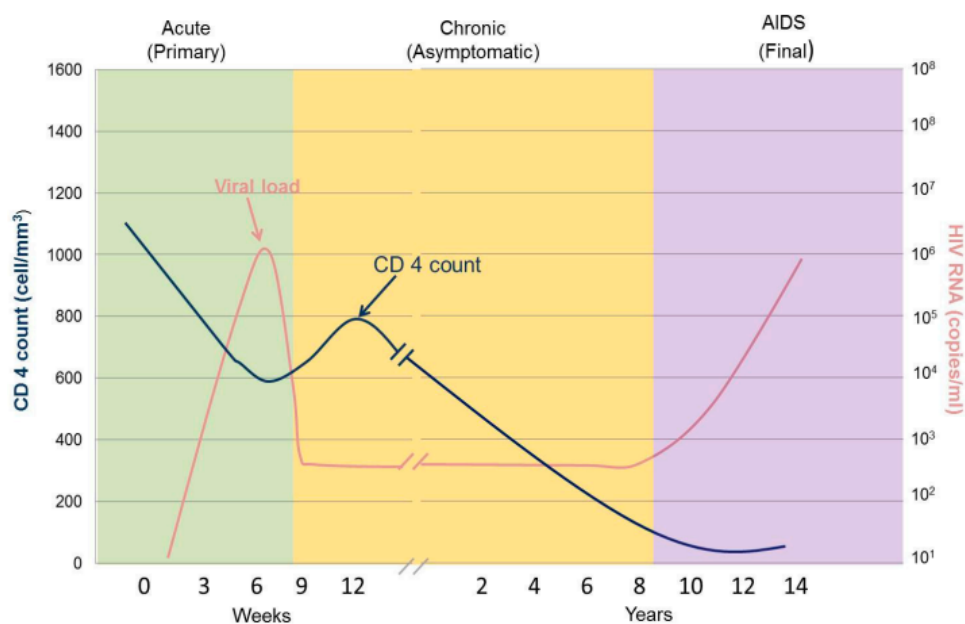


Figura 7- Fases clínicas da infecção pelo VIH-1 (Manoto, Lugongolo and Govender, 2018)

### 1.8- Terapia Antirretroviral

O avanço mais significativo no tratamento de indivíduos que vivem com VIH foi a introdução de medicamentos antirretrovirais, que suprimem a replicação do vírus a níveis indetectáveis o que leva a uma melhoria da qualidade de vida (Arts *et al.*, 2012).

Atualmente existem mais de 60 medicamentos aprovados pela *Food and Drug Administration* (FDA), disponíveis para tratar indivíduos que vivem com VIH (Arts *et al.*, 2012; Pau and George, 2015; U.S. Food and Drug Administration, 2019). Estes fármacos estão distribuídos em cinco classes diferentes, com base nos mecanismos moleculares, sendo eles: Inibidores da transcriptase reversa análogos de nucleósdeos (NRTIs), Inibidores da transcriptase reversa não-análogos de nucleósdeos (NNRTIs), Inibidores da integrase, Inibidores da protease (PI) e Inibidores de entrada (EI) (Arts *et*

*al.*, 2012). Para além dessas cinco classes, também existe uma nova classe que se trata do Inibidor seletivo da proteína da cápside (Dvory-sobol *et al.*, 2022).

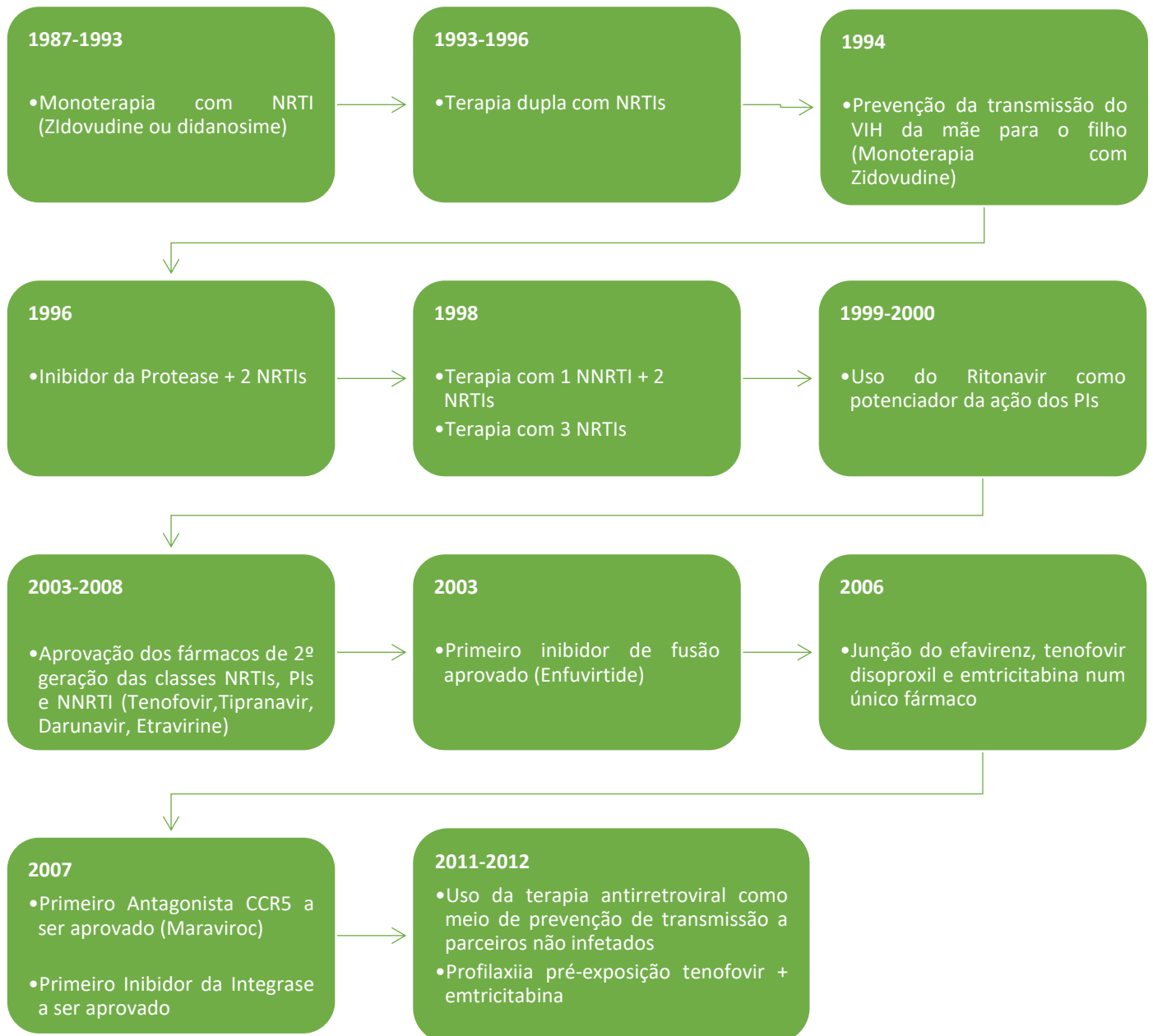


Figura 8 - Cronologia do Desenvolvimento da Terapia Antirretroviral (Pau and George, 2015)

O mecanismo de ação dos NRTIs passa pela inibição da ação enzimática da transcriptase reversa, por incorporação do nucleótido análogo, causando a terminação da cadeia de ADN ou competindo com o substrato natural do vírus. Assim, este processo

interrompe a transformação do ARN viral em ADN de fita dupla. Exemplos dos fármacos que se encontram nesta classe, são o abacavir, emtricitabina, tenofovir e zidovudina (Pau and George, 2015).

Os NNRTIs reduzem a ação da transcriptase reversa devido a uma alteração conformacional, alteração esta que irá inibir a capacidade que a transcriptase reverte de converter o ARN em ADN. Exemplos de NNRTIs são o efavirenz, etravirina, nevirapina e rilpivirina (Pau and George, 2015).

Os inibidores da integrase bloqueiam a ação desta enzima, impedindo a formação de ligações entre a célula hospedeira e o ADN viral. Isto, impede a incorporação do ADN viral no cromossoma da célula hospedeira. O dolutegravir, elvitegravir e raltegravir são exemplos de fármacos desta classe (Pau and George, 2015).

Os inibidores da protease exercem a sua função na fase tardia do ciclo replicativo do VIH. A sua função é ligarem-se às proteases do VIH, levando ao bloqueio da atividade enzimática, resultando na incapacidade de formação de viriões infecciosos. Exemplos destes antirretrovirais são o atazanavir, darunavir, fosamprenavir, lopinavir, saquinavir e tipranavir (Pau and George, 2015).

O maraviroc é antagonista de CCR5 aprovado para tratamento do VIH. Este liga-se seletivamente aos recetores CCR5 na membrana celular, bloqueando assim a interação entre a gp120 do VIH e o recetor CCR5 da célula hospedeira. Contudo, não impede a ligação do CXCR4 com a gp120 ou a gp41 (Pau and George, 2015).

O enfuvirtide é o único inibidor de fusão conhecido. O seu mecanismo de ação passa pela interferência no processo de fusão ao ligar-se à gp41, evitando assim mudanças conformacionais necessárias para a fusão do involucro viral com a membrana celular (Pau and George, 2015).

O lenacapavir é o único fármaco até agora que se encontra na classe dos inibidores seletivos da proteína da cápside. Ao ligar-se a duas subunidades vizinhas da proteína do capsídeo, este fármaco irá interferir em várias etapas do ciclo replicativo, incluindo a formação de viriões e a formação do núcleo do capsídeo (Dvory-sobol *et al.*, 2022).

### 1.9- Resistência à terapia Antirretroviral (ARV)

Em 1996, a terapia antirretroviral altamente ativa (HAART) foi introduzida nos países desenvolvidos, o que levou a uma diminuição tanto na mortalidade como na morbidade dos indivíduos que vivem com VIH. No entanto, as altas taxas de recombinação viral e de mutação do vírus levaram a uma elevada variabilidade genética. A alta variabilidade genética do VIH é um desafio significativo no tratamento com antirretrovirais. Essa diversidade resulta da rápida taxa de mutação do vírus durante a replicação. A algumas dessas variantes, que passaram por mutações, podem desenvolver resistência aos medicamentos antirretrovirais, quando isso acontece, a eficácia da terapia é comprometida. A resistência à terapia ARV ocorre quando o vírus desenvolve mutações que escapam às propriedades inibitórias destes mesmos fármacos (Pimentel *et al.*, 2020).

Os indivíduos que vivem com VIH podem adquirir resistência aos ARV de duas maneiras: a primeira, resistência transmitida, é quando o indivíduo é infetado com uma estirpe que já tem alguma mutação que confere resistência ao ARV, e a segunda, resistência adquirida, é quando ocorre pressão seletiva dos antirretrovirais nos indivíduos infetados, onde as cargas virais não são totalmente suprimidas durante o tratamento (Baesi *et al.*, 2014).

### 1.10- *Late Presenters*

O diagnóstico tardio do VIH pensa-se estar associado a um maior risco de transmissão do vírus e por conseguinte a um maior custo de cuidados de saúde. Assim, os programas de teste de VIH são bastante importantes para evitar o diagnóstico tardio (Croxford *et al.*, 2022).

Os *lates presenters* têm não só uma maior progressão clínica como também apresentam maior taxa de mortalidade. Além de que, quando são diagnosticados têm uma recuperação imunológica mais lenta após o início da terapia antirretroviral, quando comparado com os *non-late presenters* (Jeong *et al.*, 2016).

Os indivíduos de apresentação tardia são usualmente homens heterossexuais, utilizadores de drogas injetáveis, migrantes, com mais de 50 anos ou indivíduos com pouco acesso à testagem de VIH (Zangerle *et al.*, 2014; Wójcik-Cichy *et al.*, 2018).

O número de indivíduos classificados como *late presenters* depende da definição usada. Por exemplo, na Europa, estima-se que entre 40 a 60% sejam *late presenters*. Na Ásia esta prevalência está entre os 72 e 83%, enquanto em África esta prevalência encontra-se entre os 35 e os 89% (Miranda *et al.*, 2021).

Devido à inconsistência das definições apresentadas para *late presenters*, que dificultavam a comparação entre vários estudos relacionados com apresentação tardia, o Centro Europeu de Prevenção e Controle de Doenças (ECDC) e a Organização Mundial da Saúde (OMS), em 2010, apoiaram a decisão do *European Late Presenter Consensus Working Group* em unificar a definição de *late presenters*. A definição preconiza que os indivíduos que vivem com VIH e apresentam uma contagem de células CD4 inferior a 350 cel/ $\mu$ L, ou, independentemente da contagem de células CD4, apresentam um evento definidor de SIDA, são considerados *late presenters* (Mukolo and Villegas, 2013; Croxford *et al.*, 2022).

A definição supracitada veio facilitar a padronização dos diferentes estudos realizados, permitindo uma comparação mais efetiva, evitando assim viés nas classificações de *late presenters*. No entanto, com esta definição de apresentação tardia, muitos indivíduos eram classificados como *late presenters* incorretamente. Uma das principais razões deve-se ao facto de alguns indivíduos que se encontravam na fase aguda apresentarem baixos níveis de células CD4 (Croxford *et al.*, 2022).

Assim, em 2022 a equipa de Sara Croxford sugeriu uma atualização da definição de *late presenters*. Desta forma, para evitar a classificação incorreta de LP, parâmetros de evidência de infeção recente passam a ser considerados para reclassificar ou não um LP em NLP. Ou seja, parâmetros como evidências laboratoriais de uma infeção recente (presença de antígeno p24, presente no sangue de indivíduos recentemente infetados), a existência de um teste negativo para o VIH nos últimos 12 meses ou uma evidência clínica

de infecção aguda indicam que o indivíduo se encontra em fase aguda e desta forma será reclassificado como NLP (Croxford *et al.*, 2022).

Muitas vezes a informação relativa ao último teste negativo é escassa, então é necessário encontrar outros parâmetros que sejam mais acessíveis, para caracterizar esta população, evitando o viés de classificação entre *late presenters* e indivíduos que se encontrem em fase aguda.

Neste estudo, para permitir uma melhor definição de *late presenter*, para além de utilizarmos a definição de apresentação tardia baseada nos valores de células CD4, recorreremos ao uso da taxa de ambiguidade da sequência genómica e ao possível tempo de infecção calculado através do ancestral comum mais recente com base em análises filogenéticas, para analisar a definição de apresentação tardia. Para além disso, o nosso objetivo foi também avaliar se estes parâmetros serão ou não fidedignos para serem utilizados para a classificação de *late presenters*.

## 2- Objetivos

### 2.1- Objetivo Geral

Analisar a precisão da definição de *late presenters* através da utilização de outros indicadores e considerando os valores de células CD4 e o último teste de VIH negativo.

### 2.2- Objetivos Específicos

1. Classificar em *late presenters vs. non late presenters* com base nos valores de células CD4;

2. Classificar em Infecção Crónica ou Infecção Recente com base na taxa de ambiguidade das sequências genómicas;

3. Utilizar a reconstrução de árvores filogenéticas para estimar o possível tempo de infecção dos indivíduos com classificação Infecção Crónica ou Infecção Recente;

4. Comparar os diferentes resultados obtidos através das diferentes ferramentas usadas.

### 3- Materiais e Métodos

#### 3.1- População em estudo

Para realizar este estudo utilizou-se a base de dados da *EuResist* como base para a nossa população em estudo. Esta base é uma das maiores bases de dados que existe e que integra dados clínicos, sociodemográficos e genómicos de indivíduos que vivem com VIH-1 (EuResist, 2023).

A partir desta base de dados selecionou-se as seguintes informações: ano de nascimento, sexo, continente de origem, via de transmissão, data do primeiro teste positivo, data da amostra, a sequência genética, contagem das células CD4 e a carga viral.

Posto isto, foram excluídos todos os indivíduos que não tinham 18 anos até à data de diagnóstico, e indivíduos que não tinham informações como a data de primeiro teste positivo, ou a sequência genética.

#### 3.2- Classificação com base no valor de células CD4

Para classificar os indivíduos em *late presenters* (LP) ou em *non-late presenters* (NLP), baseamo-nos no valor das células de CD4, os indivíduos com um valor de CD4 <350 células/ $\mu$ L foram classificados como LP, em contrapartida, todos os outros foram classificados como NLP (Croxford *et al.*, 2022).

#### 3.3- Classificação com base na taxa de ambiguidade

As sequências nucleotídicas obtidas na base de dados foram alinhadas contra a sequência de referência HXB2 utilizando o *software* Virulign (Libin *et al.*, 2019). A taxa de ambiguidade das sequências foi calculada através de um *script* do RStudio (versão 2023.03.1+446). Este *script* calcula o comprimento e remove os caracteres ambíguos de cada sequência definidos pelo padrão (RYKMSWBDHVN). Além do mais, calcula a

ambiguidade de cada sequência subtraindo o comprimento após a remoção de ambiguidades do comprimento original e calcula a razão da ambiguidade dividindo a ambiguidade pelo comprimento original de cada sequência. Para a definição de infecção recente ou crônica usamos um cut-off de 0.47%, sendo que a taxa de ambiguidade das sequências  $<0,47\%$  trata-se de uma infecção recente e quando a taxa de ambiguidade  $\geq 0,47\%$  trata-se de infecção crônica (Emmi Andersson et al., 2014). Além do mais, para controle de qualidade exclui-se sequências com tamanho inferior a 500 nucleótidos e uma taxa de ambiguidade superior a 2,5% (Emmi Andersson et al., 2014).

#### 3.4- Classificação com base na análise filogenética

Construiu-se uma árvore filogenética de máxima verosimilhança através do software FastTree, utilizando o modelo de substituição de nucleótidos *General Time Reversible* (GTR). Posteriormente, visualizou-se a árvore filogenética utilizando o software FigTree e eliminou-se os *outliers*, ou seja, os ramos da árvore que estavam desfasados (Price, Dehal and Arkin, 2010; Dearlove et al., 2021).

Em seguida, utilizamos o TreeTime com o objetivo de analisar o possível tempo de infecção. Este *software* é utilizado na área da bioinformática com o intuito de inferir sobre sequências ancestrais, e ao estimar o comprimento dos ramos consegue estudar evolução molecular ao longo do tempo. A data de infecção foi estimada através da diferença entre o ano de infecção do ancestral comum mais recente da árvore filogenética e o ano da amostra da sequência em estudo (Sagulenko, Puller and Neher, 2018).

Para os indivíduos com 0-1 anos de possível tempo de infecção classificou-se como infecção recente, os restantes indivíduos classificaram-se como infecção crônica.

#### 3.5- Determinação dos subtipos do VIH-1

Submeteu-se as sequências nucleotídicas a diferentes algoritmos de genotipagem de modo a poder definir o subtipo de cada sequência. Os algoritmos usados foram o Rega V3, o COMET e o Geno2pheno. Para chegar ao consenso dos subtipos utilizou-se o Rega

V3 (Pineda-Peña *et al.*, 2013) e o COMET (Struck *et al.*, 2014). Caso o subtipo obtido por estes dois *softwares* não fosse consensual, utilizaríamos o Geno2pheno para desempatar, e tornar a classificação consensual (Lengauer and Sing, 2006).

### 3.6-Análise Estatística

A análise estatística foi realizada recorrendo ao *software* estatístico SPSS versão 29.0.1.0. As tabelas de frequências foram feitas para as variáveis qualitativas, para descrever os dados sociodemográficos dos indivíduos. Para além disso, utilizou-se o teste do qui-quadrado para independência, para ver se existiam diferenças significativas entre as diferentes variáveis qualitativas que caracterizavam o tipo de infeção, dando especial atenção aquelas que são estatisticamente significativas (p-value <0,05).

## 4- Resultados

### 4.1- Dados sociodemográficos e clínicos dos indivíduos

Dos 2233 indivíduos, a maioria eram do sexo masculino (81.8%), sendo a faixa etária mais frequente entre os 31-55 anos (64.2%). O tipo de transmissão mais comum foi o contacto sexual homossexual/bissexual (56%), e 50.5% dos indivíduos eram originários da Europa Ocidental. O subtipo mais frequente foi o subtipo B (67.4%) (Tabela 1).

*Tabela 1- Caracterização sociodemográfica e clínica dos indivíduos em estudo*

<b>Características dos Indivíduos</b>	<b>Total</b>
<b>Total</b>	2233 (100%)
<b>Sexo, n (%)</b>	2233 (100%)
Masculino	1827 (81.8%)
Feminino	403 (18.1%)
Desconhecido	3 (0.1%)
<b>Idade ao diagnóstico, n (%)</b>	2233 (100%)
<18	13 (0.6%)
18-30	665 (29.8%)
31-55	1433(64.2%)
>56	122 (5.5%)
<b>Tipo de transmissão, n (%)</b>	2233 (100%)
Heterossexual	465 (20.8%)
Homossexual/Bissexual	1251 (56.0%)
Transmissão por seringas e instrumentos infetados (IVDA)	267 (12%)
Desconhecido	244 (10.9%)
Transmissão vertical	6 (0.3%)
<b>Região de Origem, n (%)</b>	2233 (100%)
África	219 (9.8%)
Europa Ocidental	1127 (50.5%)
Europa Oriental	571 (25.6%)
Outros	316 (14.2%)
<b>Subtipo, n (%)</b>	2233 (100%)
CRF 01_AE	176 (7.9%)
A1	62 (2.8%)
A6	173 (7.7%)
B	1505 (67.4%)
C	141 (6.3%)
G	33 (1.5%)
Outros	22 (1%)
Recombinantes	121 (5.4%)
<b>Carga Viral (log<sub>10</sub> cópias/mL), n (%)</b>	2233 (100%)
≤ 4.0	540 (24.2%)
4.1-5.0	872 (39.1%)

>5.0	808 (36.2%)
Desconhecido	13 (0.6%)

#### 4.2- Análise 1

Nesta primeira análise comparou-se a classificação obtida através do valor de células CD4 e a classificação obtida através dos valores da taxa de ambiguidade.

Através dos valores de CD4, 1195 dos indivíduos (53.5%) eram *late presenters*, e 1038 (46.5%) eram *non-late presenters*. Através dos valores da taxa de ambiguidade, 1113 indivíduos (49.8%) apresentavam uma infeção crónica, ou seja, apresentavam taxa de ambiguidade  $\geq 0,47\%$ , enquanto 1120 (50.2%) apresentavam uma infeção recente, ou seja, apresentavam uma taxa de ambiguidade  $< 0,47\%$  (Tabela 2).

Tabela 2 - Descrição dos tipos de classificação usando o valor de células CD4 e a Taxa de Ambiguidade

Classificações	LP vs NLP (CD4)		Tipo de Infeção (Taxa de ambiguidade)	p-value
<b>Total</b>	2233 (100%)		2233 (100%)	<0.001
LP	1195 (53.5%)	Infeção Crónica	1113 (49.8%)	
NLP	1038 (46.5%)	Infeção Recente	1120 (50.2%)	

Analisamos em seguida os resultados coerentes entre as duas técnicas: 1) classificação como LP com base no valor das células CD4 abaixo de 350 células/ $\mu\text{L}$  e classificação como infeção crónica com base na taxa de ambiguidades; 2) classificação como NLP com base no valor das células CD4 acima de 350 células/ $\mu\text{L}$  e como infeção aguda com base na taxa de ambiguidades.

Assim, observou-se que 65.2% dos resultados eram coerentes entre as duas classificações, 34.3% dos indivíduos eram classificados com LP e Infeção Crónica e 30.9% eram classificados como NLP e como Infeção Recente (Tabela 3).

Tabela 3- Comparação entre os tipos de classificação usando o valor de células CD4 e a Taxa de Ambiguidade

Classificações		LP vs NLP (CD4)		Total
		LP	NLP	
Tipo de Infecção (Taxa de ambiguidade)	Infecção Crônica	765 (34.3%)	348 (15.6)	1113 (49.8%)
	Infecção Recente	430 (19.3%)	690 (30.9%)	1120 (50.2%)
Total		1195 (53.5%)	1038 (46.5%)	2233 (100%)
Concordante		1455 (65.2%)		
Não Concordante		778 (34.8%)		

#### 4.3- Análise 2

Nesta segunda análise comparou-se a classificação obtida através do valor de células CD4 e a através do possível tempo de infecção obtida com base nos resultados da árvore do TreeTime baseado no ancestral comum mais recente.

Através dos valores de CD4, 1195 dos indivíduos (53.5%) eram *late presenters*, e 1038 (46.5%) eram *non-late presenters*. Através da análise do possível tempo de infecção, 1764 indivíduos (79.0%) apresentavam uma infecção crônica, ou seja, o tempo possível de infecção estimado, baseado no ancestral comum mais recente, era superior a 1 ano, enquanto 292 dos indivíduos (13,1%) apresentavam uma infecção recente, ou seja, o tempo possível de infecção estimado, baseado no ancestral comum mais recente, era inferior a 1 ano (Tabela 4).

Tabela 4- Descrição dos tipos de classificação usando o valor de células CD4 e o tipo de infecção baseado no ancestral comum mais recente

Classificações	LP vs NLP (CD4)		Tipo de Infecção (Ancestral comum mais recente)	p-value
<b>Total</b>	2233 (100%)		2233 (100%)	<0.001
LP	1195 (53.5%)	Infeção Crónica	1764 (79.0%)	
NLP	1038 (46.5%)	Infeção Recente	292 (13.1%)	
Inválidos	0 (0%)	Inválidos	177 (7.9%)	

Neste caso, considerámos como coerentes os casos em que os indivíduos que foram classificados como LP apresentam um tipo de infecção Crónica, e o caso em que os indivíduos classificados como NLP, apresentam infecção Recente.

Observou-se que 53.2% dos resultados usando estas duas técnicas eram concordantes, ou seja, 44.4% dos indivíduos eram classificados com LP (baseando-nos nos valores de CD4) ou como Infeção Crónica (baseando-nos no possível tempo de infecção, baseado no ancestral comum mais recente) e 8.8% eram classificados como NLP (baseando-nos nos valores de CD4) ou como Infeção Recente (baseando-nos no possível tempo de infecção, baseado no ancestral comum mais recente) (Tabela 5).

Tabela 5- Comparação entre os tipos de classificação usando o valor de células CD4 e o tipo de infecção baseado no ancestral comum mais recente

Classificações		LP vs NLP (CD4)		Total
		LP	NLP	
<b>Tipo de Infecção (Ancestral comum mais recente)</b>	Infeção Crónica	992 (44.4%)	772 (34.6%)	1764 (79.0%)
	Infeção Recente	96 (4.3%)	196 (8.8%)	292 (13.1%)
	Inválidos	107 (4.8%)	70 (3.1%)	177 (7.9%)
<b>Total</b>		1195 (53.5%)	1038 (46.5%)	2233 (100%)
<b>Concordante</b>		1188 (53.2%)		
<b>Não Concordante</b>		868 (38.9%)		
<b>Inválido</b>		177 (7.9%)		

#### 4.4- Análise 3

Nesta análise comparou-se a classificação obtida através da taxa de ambiguidade e através do possível tempo de infecção, baseado no ancestral comum mais recente.

Através da taxa de ambiguidade, 1113 dos indivíduos (49.8%) apresentavam infecção crónica, e 1120 (50.2%) apresentavam uma infecção recente. Através do possível tempo de infecção, baseado no ancestral comum mais recente, 1764 indivíduos (79.0%) apresentavam uma infecção crónica, enquanto 292 dos indivíduos (13,1%) apresentavam uma infecção recente, baseado no ancestral comum mais recente (Tabela 6).

*Tabela 6 Descrição dos tipos de classificação usando a Taxa de ambiguidade e o tipo de infecção baseado no ancestral comum mais recente*

<b>Classificações</b>	<b>Tipo de Infecção (Taxa de ambiguidade)</b>		<b>Tipo de Infecção (Ancestral comum mais recente)</b>	<b>p-value</b>
<b>Total</b>	2233 (100%)		2233 (100%)	<b>&lt;0.001</b>
Infeção Crónica	1113 (49.8%)	Infeção Crónica	1764 (79.0%)	
Infeção Recente	1120 (50.2%)	Infeção Recente	292 (13.1%)	
Inválidos	0 (0%)	Inválidos	177 (7.9%)	

Para além disso, observou-se que 50% dos resultados usando estas duas técnicas eram concordantes, ou seja, 41.3% dos indivíduos eram classificados como Infecção Crónica e 8.7% eram classificados como como Infecção Recente e 7,9% foram considerados como inválidos (Tabela 7).

Tabela 7 Comparação entre os tipos de classificação usando a Taxa de ambiguidade e o

tipo de infecção baseado no ancestral comum mais recente

Classificações		Tipo de Infecção (Taxa de ambiguidade)		Total
		Infeção Crónica	Infeção Recente	
Tipo de Infecção (Ancestral comum mais recente)	Infeção Crónica	923 (41.3%)	841 (37.7%)	1764 (79.0%)
	Infeção Recente	97 (4.3%)	195(8.7%)	292 (13.1%)
	Inválidos	93 (4.2%)	84 (3.8%)	177 (8.0%)
<b>Total</b>		1113 (49.8%)	1120 (50.2%)	2233 (100%)
<b>Concordante</b>		1118 (50%)		
<b>Não Concordante</b>		938 (42%)		
<b>Inválido</b>		177 (7.9%)		

#### 4.5- Análise 4

Nesta última análise, comparou-se a classificação com base no possível tempo de infecção, baseado no ancestral comum mais recente e os valores de carga viral de cada indivíduo.

Ao observarmos a Tabela 8 podemos constatar que os valores de carga viral estão distribuídos quase homogeneamente, tendo uma maior prevalência os indivíduos com carga viral entre 4.1-5.0 cópias/mL. Relativamente à classificação dos indivíduos baseada nos valores obtidos no tempo possível de infecção, baseado no ancestral comum mais recente, podemos observar que 79% tinham uma infecção crónica e 13% uma infecção recente.

Tabela 8- Descrição do tipo de classificação usando o tipo de infecção baseado no ancestral comum mais recente

e os valores de Carga Viral

Classificações	Carga Viral		Tipo de Infecção (Ancestral comum mais recente)	p-value
<b>Total</b>	2233 (100%)		2233 (100%)	

≤4.0	540 (24.2%)	Infeção Crónica	1764 (79.0%)	<0.001
4.1-5.0	872 (39.1%)	Infeção Recente	292 (13.1%)	
>5.0	808 (36.2%)	Inválidos	177 (7.9%)	
Desconhecido	13 (0,6%)	-----	-----	

No entanto, ao comparar, podemos observar que em ambos o tipo de infeção a maior prevalência de carga viral continua a ser entre as 4.1-5.0 cópias/mL (Tabela 9).

*Tabela 9- Comparação entre o tipo de infeção baseado no ancestral comum mais recente e a Carga Viral*

Classificações		Carga Viral				Total
		≤4.0	4.1-5.0	>5.0	Inválido	
<b>Tipo de Infeção (Ancestral comum mais recente)</b>	Infeção Crónica	461 (20.6%)	683 (30.6%)	608 (27.2%)	12 (0.5%)	1764 (79.0%)
	Infeção Recente	50 (2.2%)	128 (5.7%)	114 (5.1%)	0 (0%)	292 (13.1%)
	Inválidos	29 (1.3%)	61 (2.7%)	86 (3.9%)	1 (0%)	177 (7.9%)
<b>Total</b>		540 (24.2%)	872 (39.1%)	808 (36.2%)	13 (0.6%)	2233 (100%)

## 5- Discussão e Conclusões

O objetivo deste estudo teve como base analisar a definição atual de *late presenters*, e tentar encontrar outras metodologias alternativas para evitar um viés na classificação entre LP vs NLP, isto porque alguns indivíduos que se encontram em fase aguda podem apresentar uma baixa concentração de células CD4, sendo por isso mal classificados como LP.

Tendo como base as metodologias estudadas, entre as quais se encontram a classificação em LP vs NLP, com base na contagem de células CD4, classificação em Infecção Crónica vs Infecção Recente tendo como base a taxa de ambiguidade e o tempo possível de infecção, baseado no ancestral comum mais recente recorrendo ao *software* TreeTime para construir a árvore filogenética. As metodologias que apresentaram resultados mais concordantes foram a combinação da classificação em LP vs NLP baseado no valor das células CD4, e a classificação em Infecção Crónica e Infecção Recente, tendo como base a taxa da ambiguidade. Ao cruzar estas duas metodologias obteve-se uma taxa de concordância de 65.2%. Existem duas equipas que foram referenciadas nesta dissertação, a equipa da Sara Croxford que classificou os indivíduos em LP vs NLP tendo como base a contagem de células CD4 (Croxford *et al.*, 2022) e a equipa da Emmi Andersson que classificou os indivíduos tendo como base o tipo de infecção (Infecção crónica vs Infecção recente) (Emmi Andersson *et al.*, 2014).

Inicialmente, foi feita uma descrição sociodemográfica e clínica da população em estudo. Em seguida efetuaram-se quatro análises distintas: 1) comparou-se a classificação em LP vs NLP tendo como base a contagem de células CD4 e a classificação em Infecção Crónica vs Infecção Recente, tendo como base a taxa de ambiguidade; 2) comparou-se a classificação em LP vs NLP tendo como base a contagem de células CD4 e a classificação em Infecção Crónica vs Infecção Recente, tendo como base o possível tempo de infecção, baseado no ancestral comum mais recente; 3) comparou-se a classificação em Infecção Crónica vs Infecção Recente, tendo como base a taxa de ambiguidade e a classificação em Infecção Crónica vs Infecção Recente, tendo como base o possível tempo de infecção, baseado no ancestral comum mais recente; e por último 4) comparou-se a Infecção Crónica

vs Infecção Recente, tendo como base possível tempo de infecção, baseado no ancestral comum mais recente com a carga viral.

Ao analisar os dados sociodemográficos, pode-se concluir que o gênero predominante era o sexo masculino (81.8%), a idade ao diagnóstico mais frequente é entre os 31-55 anos (64.2%), e o tipo de transmissão mais comum foi o contacto homossexual/bissexual (56%). Os resultados obtidos neste estudo estão de acordo com dados existentes na literatura que descrevem as características sociodemográficas de indivíduos que vivem com VIH na Europa (Wójcik-Cichy *et al.*, 2018; Serra *et al.*, 2020; Miranda *et al.*, 2021). A região de origem mais prevalente foi a Europa Ocidental (50.5%) e os subtipos mais frequentes foram o subtipo B (67.4%), o subtipo A (9.5%), e CRF 01\_AE (7.9%). Apesar da elevada proporção de indivíduos do subtipo B, podemos aqui observar que os subtipos não-B estão cada vez mais a aumentar na Europa, possivelmente devido às migrações de indivíduos provenientes de outros continentes com uma prévia infecção de VIH de outro subtipo que não o B. Os resultados deste estudo relativamente à distribuição de subtipos na população, verificam-se noutros estudos, como por exemplo, no estudo de Miranda *et al.* (2021) 64.4% dos indivíduos eram do subtipo B, 13.5% do subtipo A e 2.4% apresentavam CRF 01\_AE. Relativamente à carga viral, 39.1% dos indivíduos apresentaram um valor de carga viral entre as 4.1-5.0 cópias/mL, estando de acordo com dados do mesmo estudo acima mencionado, que descreve uma população de indivíduos que vivem com VIH na Europa, onde 34.4% dos indivíduos apresentavam a carga viral entre 4.1-5.0 cópias/mL (Miranda *et al.*, 2021).

Na comparação entre os resultados obtidos no nosso estudo e os estudos conduzidos por Sara Croxford e Emmi Andersson, observamos algumas variações notáveis nas características da população e nas taxas de apresentação tardia (LP) e categorização de infecção recente vs infecção crônica.

O estudo de Sara Croxford, que se concentrou exclusivamente em indivíduos do Reino Unido, identificou que 42% dos indivíduos eram LP. Em contraste, o nosso estudo, que incluiu uma amostra mais diversificada com uma proporção significativa de indivíduos provenientes da África Sub-Sahariana, revelou uma taxa ligeiramente mais alta de 53.5% de LP. Essa discrepância pode ser atribuída à distribuição geográfica

variada da nossa amostra, sugerindo que as diferenças na apresentação tardia podem ser influenciadas pela origem dos indivíduos. Quanto à categorização em LP e não LP, nossos resultados (53.5% LP e 46.5% não LP) alinham-se com estudos recentes na Europa, onde as taxas de LP variavam entre 50% e 63%, incluindo referências como Miranda et al. (2022), Chkhartishvili et al. (2017), e Petrakis et al. (2020).

No estudo de Emmi Andersson, 48% dos participantes foram classificados como tendo infecção recente, enquanto 52% tinham infecção crônica, com base nos valores da taxa de ambiguidades. No nosso estudo, observou-se que 50.2% dos indivíduos apresentam infecção recente e 49.8% infecção crônica, sendo estes resultados concordantes com a literatura (Emmi Andersson et al., 2014). No entanto, ao compararmos esses resultados com o estudo de Miranda et al. (2022), observamos discrepâncias, possivelmente devido à aplicação de diferentes *cutoffs*, enquanto Miranda et al. (2022) utilizou um *cutoff* de 0.45%, o nosso estudo e o de Emmi Andersson et al. (2014) utilizou-se um *cutoff* de 0.47%.

Na avaliação dos resultados obtidos através de duas metodologias distintas, podemos concluir que 65.2% dos resultados obtidos (classificação em LP vs NLP e classificação em infecção crônica vs infecção recente) eram concordantes. Isto demonstra que apesar da maioria ser concordante, ainda existe uma ligeira discrepância nos resultados, o que pode indicar uma mais-valia a utilização de mais do que uma metodologia na classificação de LP vs NLP.

Na etapa subsequente da análise, classificamos os indivíduos em LP vs NLP com base na contagem das células CD4 e em infecção crônica vs infecção recente considerando o possível tempo de infecção com base ancestral comum mais recente. Os resultados revelaram que 79% dos indivíduos apresentavam uma infecção crônica, enquanto 13.1% tinham uma infecção recente. A escassez de estudos que utilizam métodos semelhantes dificulta a comparação direta dos nossos resultados com a literatura existente. No entanto, é importante destacar que 53.2% dos resultados eram concordantes ao utilizar essas duas metodologias. Essa concordância sugere uma consistência razoável entre as classificações de LP vs NLP e infecção crônica vs infecção recente com base na contagem de células CD4 e no tempo de infecção, baseado no ancestral comum mais recente. Contudo,

reconhecemos a necessidade potencial de incluir uma terceira metodologia para fortalecer ainda mais o consenso entre essas abordagens.

Ao avaliar os resultados, torna-se evidente que a introdução de uma terceira metodologia pode ser benéfica para a obtenção de uma perspectiva mais abrangente e robusta sobre a classificação dos indivíduos. Esta consideração é particularmente relevante diante da complexidade e heterogeneidade que pode existir nos dados de infecção pelo VIH. Em resumo, embora tenhamos alcançado uma significativa concordância entre as metodologias utilizadas, reconhecemos a importância de explorar abordagens adicionais para solidificar as nossas conclusões e contribuir para uma compreensão mais completa das características da infecção pelo VIH na nossa amostra.

No contexto da classificação entre Infecção Recente vs Infecção Crónica, utilizando a taxa de ambiguidade e o possível tempo de infecção baseado no ancestral comum mais recente, como critérios, observamos uma concordância de 50% entre essas duas metodologias. Esta observação reforça a sugestão anterior de que a combinação de múltiplas abordagens pode ser essencial para alcançar uma homogeneidade nos resultados obtidos.

A comparação dos indivíduos com base nos valores de carga viral e na classificação em Infecção Crónica vs Infecção Recente, considerando o possível tempo de infecção baseado no ancestral comum mais recente, revelou que tanto os indivíduos com Infecção Crónica como os com Infecção Recente apresentavam uma carga viral superior a 4 cópias/mL.

Ao discutir estes resultados, é importante considerar que a incorporação de uma terceira metodologia, como mencionado anteriormente, pode ser valiosa para reforçar a consistência e a confiabilidade das conclusões, especialmente quando há uma complexidade considerável nos dados.

A análise comparativa entre estas diferentes metodologias destaca a necessidade contínua de refinamento e aprofundamento na abordagem metodológica, enfatizando a importância de uma discussão mais aprofundada sobre as implicações e limitações dessas técnicas na interpretação dos resultados.

Ao analisar as diferentes combinações de metodologias utilizadas neste estudo, identificamos que a classificação em LP vs NLP com base na contagem das células CD4, juntamente com a categorização em Infecção Crônica vs Infecção Recente usando os valores da taxa de ambiguidade, apresentou a maior percentagem de valores concordantes. Este resultado sugere que essa combinação pode ser uma das alternativas para minimizar o viés na definição de *late presenters*. No entanto, é reconhecido que a adição de uma terceira metodologia pode conferir ainda mais rigor à classificação.

A sugestão de explorar diferentes valores da taxa de ambiguidade é válida, pois pode contribuir para encontrar uma maior concordância entre as técnicas utilizadas. Isso destaca a importância de ajustar e otimizar os parâmetros das metodologias para melhorar a consistência e a confiabilidade dos resultados. Entretanto, é importante reconhecer as limitações deste estudo. A amostra reduzida é um ponto a ser considerado, pois pode impactar a generalização dos resultados. A aplicação limitada da metodologia do tempo possível de infecção também representa uma limitação, especialmente devido à diversidade de métodos para criar árvores temporais e à utilização de diferentes sequências de controlo. Essa variabilidade pode levar a incertezas na classificação entre Infecção Crônica vs Infecção Recente com base no possível tempo de infecção. Além disso, ao classificar os indivíduos em LP vs NLP, apenas a contagem de células CD4 foi considerada, o que pode ter introduzido um viés na classificação. A escolha do *cut-off* para a classificação em infecção recente vs infecção através da taxa de ambiguidade também pode ser uma fonte potencial de viés.

Em resumo, apesar de encontrar combinações de metodologias que apresentam maior concordância, é fundamental reconhecer e discutir as limitações do estudo. A consideração dessas limitações pode orientar futuras pesquisas por abordagens mais robustas e confiáveis na classificação de LP vs NLP e Infecção Crônica vs Infecção Recente.

## 6- Bibliografia

Andersson, E. *et al.* (2014) 'Evaluation of Sequence Ambiguities of the HIV-1 pol gene as a Method to Identify Recent HIV-1 Infection in Transmitted Drug Resistance Surveys', *Bone*, 23(1), pp. 1–7. doi: 10.1016/j.meegid.2013.03.050.Evaluation.

Arts, E. J. *et al.* (2012) 'HIV-1 Antiretroviral Drug Therapy BASIC PRINCIPLES OF ANTIRETROVIRAL ir'.

Baesi K, Ravanshad M, Ghanbarisafari M, Saberfar E, SeyedAlinaghi S, V. J. (2014) 'Associated With Antiviral Therapy', *Antiviral Therapy*, 55(November 2005), pp. 52–55. doi: 10.1002/jmv.

Barré-Sinoussi, F. (2009) 'HIV: A discovery opening the road to novel scientific

knowledge and global health improvement (Nobel Lecture)', *Angewandte Chemie - International Edition*, 48(32), pp. 5809–5814. doi: 10.1002/anie.200901918.

Cañada, J. E. *et al.* (2021) 'Identification of a new hiv-1 bc intersubtype circulating recombinant form (Crf108\_bc) in Spain', *Viruses*, 13(1), pp. 1–13. doi: 10.3390/v13010093.

Chkhartishvili, N. *et al.* (2017) 'Late presentation of HIV infection in the country of Georgia: 2012-2015', *PLoS ONE*, 12(10), pp. 2012–2015. doi: 10.1371/journal.pone.0186835.

De Cock, K. M., Jaffe, H. W. and Curran, J. W. (2012) 'The evolving epidemiology of HIV/AIDS', *Aids*, 26(10), pp. 1205–1213. doi: 10.1097/QAD.0b013e328354622a.

Croxford, S. *et al.* (2022) 'Late diagnosis of HIV : An updated consensus definition ', *HIV Medicine*, (October), pp. 1–7. doi: 10.1111/hiv.13425.

Database, L. A. (2023) 'HIV Circulating Recombinant Forms (CRFs)'. Available at: <https://www.hiv.lanl.gov/components/sequence/HIV/crfdb/crfs.comp>.

Dearlove, B. *et al.* (2021) 'Factors influencing estimates of HIV-1 infection timing using BEAST', *PLoS Computational Biology*, 17(2), pp. 4–19. doi: 10.1371/JOURNAL.PCBI.1008537.

Dvory-sobol, H. *et al.* (2022) 'Lenacapavir : a first-in-class HIV-1 capsid inhibitor', 0, pp. 4–10. doi: 10.1097/COH.0000000000000713.

Emanuele Fanales-Belasio, Mariangela Raimondo, B. S. and S. B. (2010) 'HIV virology and pathogenetic mechanisms of infection : a brief overview', pp. 5–14. doi: 10.4415/ANN.

EuResist (2023) 'EuresistDataAnalysis-database'.

Fettig, J. *et al.* (2014) 'Global Epidemiology of HIV HIV Global health Epidemiology Antiretroviral therapy Risk behavior', *Infectious Disease Clinics of NA*, 28(3), pp. 323–

337. doi: 10.1016/j.idc.2014.05.001.

Galvin, S. R., Cohen, M. S. and Hill, C. (2004) 'THE ROLE OF SEXUALLY TRANSMITTED DISEASES IN HIV TRANSMISSION', 2(January). doi: 10.1038/nrmicro794.

Greene, W. C. (2007) 'A history of AIDS: Looking back to see ahead', *European Journal of Immunology*, 37(SUPPL. 1), pp. 94–102. doi: 10.1002/eji.200737441.

Hemelaar, J. (2012) 'The origin and diversity of the HIV-1 pandemic', *Trends in Molecular Medicine*, 18(3), pp. 182–192. doi: 10.1016/j.molmed.2011.12.001.

Hemelaar, J. *et al.* (2019) 'Global and regional molecular epidemiology of HIV-1, 1990–2015: a systematic review, global survey, and trend analysis', *The Lancet Infectious Diseases*, 19(2), pp. 143–155. doi: 10.1016/S1473-3099(18)30647-9.

Jeong, S. J. *et al.* (2016) 'Late presentation into care of HIV disease and its associated factors in Asia : Results of TAHOD', pp. 1–25.

Kirchhoff, F. (2016) 'Encyclopedia of AIDS', *Encyclopedia of AIDS*, (January 2013). doi: 10.1007/978-1-4614-9610-6.

Lengauer, T. and Sing, T. (2006) 'Bioinformatics-assisted anti-HIV therapy', *Nature Reviews Microbiology*, 4(10), pp. 790–797. doi: 10.1038/nrmicro1477.

Libin, P. J. K. *et al.* (2019) 'VIRULIGN: Fast codon-correct alignment and annotation of viral genomes', *Bioinformatics*, 35(10), pp. 1763–1765. doi: 10.1093/bioinformatics/bty851.

Mendes Da Silva, R. K. *et al.* (2021) 'Genetic characterization of a new hiv-1 sub-subtype a in cabo verde, denominated a8', *Viruses*, 13(6), pp. 2–10. doi: 10.3390/v13061093.

Miranda, A. C. *et al.* (2021) 'Determinants of HIV-1 Late Presentation in Patients Followed in Europe', *AIDS Research and Human Retroviruses*, 37(11), pp. 846–851. doi: 10.1089/aid.2020.0175.

- Miranda, M. N. S. et al. (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', 13(April), pp. 1–12. doi: 10.3389/fmicb.2022.846943.
- Mukolo, A. and Villegas, R. (2013) 'Predictors of Late Presentation for HIV Diagnosis : A Literature Review and Suggested Way Forward', pp. 5–30. doi: 10.1007/s10461-011-0097-6.
- Pang, W. et al. (2018) 'Prevalence of opportunistic infections and causes of death among hospitalized HIV-infected patients in Sichuan, China', *Tohoku Journal of Experimental Medicine*, 244(3), pp. 231–242. doi: 10.1620/tjem.244.231.
- Parekh, B. S. et al. (2019) 'crossm Diagnosis of Human Immunodeficiency Virus Infection', 32(1), pp. 1–55.
- Pau, A. K. and George, J. M. (2015) 'Antiretroviral Therapy : Current Drugs', 28(3), pp. 371–402. doi: 10.1016/j.idc.2014.06.001.Antiretroviral.
- Petrakis, V. et al. (2020) 'Late Presenters of HIV Infection in an HIV Unit of a Tertiary University Hospital in a Rural Region of Greece', *AIDS Research and Human Retroviruses*, 36(7), pp. 601–605. doi: 10.1089/aid.2019.0246.
- Pimentel, V. et al. (2020) 'Molecular epidemiology of hiv-1 infected migrants followed up in Portugal: Trends between 2001-2017', *Viruses*, 12(3), pp. 1–15. doi: 10.3390/v12030268.
- Pineda-Peña, A. C. 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', *Infection, Genetics and Evolution*, 19(100), pp. 337–348. doi: 10.1016/j.meegid.2013.04.032.
- Price, M. N., Dehal, P. S. and Arkin, A. P. (2010) 'FastTree 2 - Approximately maximum-likelihood trees for large alignments', *PLoS ONE*, 5(3). doi: 10.1371/journal.pone.0009490.

Sagulenko, P., Puller, V. and Neher, R. A. (2018) 'TreeTime: Maximum-likelihood phylodynamic analysis', *Virus Evolution*, 4(1), pp. 1–9. doi: 10.1093/ve/vex042.

Serra, M. A. A. O. *et al.* (2020) 'Sociodemographic and Behavioral Factors Associated with HIV Vulnerability according to Sexual Orientation', *AIDS Research and Treatment*, 2020. doi: 10.1155/2020/5619315.

Struck, D. *et al.* (2014) 'COMET: Adaptive context-based modeling for ultrafast HIV-1 subtype identification', *Nucleic Acids Research*, 42(18), pp. 1–11. doi: 10.1093/nar/gku739.

Tang, R. *et al.* (2021) 'Genetic source completeness of HIV-1 circulating recombinant forms (crfs) predicted by multi-label learning', *Bioinformatics*, 37(6), pp. 750–758. doi: 10.1093/bioinformatics/btaa887.

U.S. Food and Drug Administration, U. S. D. of H. and H. S. (2019) 'HIV and AIDS: Medicines to Help You', *Silver Spring*. Available at: <https://www.fda.gov/consumers/free-publications-women/hiv-and-aids-medicines-help-you>.

Wójcik-Cichy, K. *et al.* (2018) '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 - Psychological and Socio-Medical Aspects of AIDS/HIV*, 30(12), pp. 1538–1541. doi: 10.1080/09540121.2018.1470306.

Zangerle, R. *et al.* (2014) 'Delayed HIV diagnosis and initiation of antiretroviral therapy: Inequalities by educational level, COHERE in EuroCoord', *Aids*, 28(15), pp. 2297–2306. doi: 10.1097/QAD.0000000000000410.