UNSW SYDNEY | Australia's Global University

Thesis/Dissertation Sheet

Surname/Family Name	:	**PINTO**
Given Name/s	:	**ANGIE**
Abbreviation for degree as give in the University calendar	:	**PhD**
Faculty	:	**Medicine**
School	:	**The Kirby Institute**
Thesis Title	:	**Aspects of primary HIV-1 infection in New South Wales**

Abstract 350 words maximum: (PLEASE TYPE)

Background: Primary HIV infection (PHI) is a unique time to study clinical, virological and public health aspects of HIV. Data on long term outcomes in this cohort are limited. New South Wales is at a crossroads in HIV strategy and more detailed insight is needed into molecular epidemiology, drug resistance and the role of PHI in driving the epidemic.

Aims: 1) To determine mortality in a cohort of PHI in NSW 2) To establish a state-wide drug resistance database and evaluate rates of transmitted and acquired drug resistance (TDR & ADR) 3) To understand the role of PHI in molecular epidemiology of NSW

Methods: Chapter Two describes data linkage methods used to link PHI cohort to deaths notifications. It also describes data linkage used to link HIV sequence dataset with HIV notifications. Chapter Three uses univariate and multivariate analysis to assess factors associated with mortality in a PHI cohort. Chapter Four evaluates the rates of transmitted and acquired drug resistance over ten years. Chapter Five uses univariate and multivariate analysis to assess factors associated with TDR. Chapter Six uses maximum likelihood phylogenetic inference and univariate and multivariate analysis to determine factors associated with cluster membership.

Key findings: Treatment with antiretrovirals within one year of PHI is associated with decreased mortality. Overall rate of prevalent drug resistance is decreasing, most markedly with ADR, while TDR rates have been stable. People residing in non-metropolitan regions and those aged 19-29 years had the highest odds of pre-treatment drug resistance. Acquisition in Australia, diagnosis post- 2012, early infection stage (but not laboratory defined PHI) are factors associated with cluster membership in NSW.

Conclusion: Primary HIV cohorts can provide valuable insights into the natural history of HIV, and the role of therapies commenced during this period. Ongoing surveillance of drug resistance mutations combined with molecular epidemiology in NSW is critical in an era of immediate universal treatment and widespread community access to pre exposure prophylaxis to understand where to focus public health interventions. Future research should focus on establishment on national HIV resistance surveillance and integration of molecular epidemiology into public health strategies.

FOR OFFICE USE ONLY Date of completion of requirements for Award:

ORIGINALITY STATEMENT

'I hereby declare that this submission is my own work and to the best of my knowledge it contains no materials previously published or written by another person, or substantial proportions of material which have been accepted for the award of any other degree or diploma at UNSW or any other educational institution, except where due acknowledgement is made in the thesis. Any contribution made to the research by others, with whom I have worked at UNSW or elsewhere, is explicitly acknowledged in the thesis. I also declare that the intellectual content of this thesis is the product of my own work, except to the extent that assistance from others in the project's design and conception or in style, presentation and linguistic expression is acknowledged.'

INCLUSION OF PUBLICATIONS STATEMENT

UNSW is supportive of candidates publishing their research results during their candidature as detailed in the UNSW Thesis Examination Procedure.

Publications can be used in their thesis in lieu of a Chapter if:

- The student contributed greater than 50% of the content in the publication and is the "primary author", ie. the student was responsible primarily for the planning, execution and preparation of the work for publication
- The student has approval to include the publication in their thesis in lieu of a Chapter from their supervisor and Postgraduate Coordinator.
- The publication is not subject to any obligations or contractual agreements with a third party that would constrain its inclusion in the thesis

Please indicate whether this thesis contains published material or not.

☐ *This thesis contains no publications, either published or submitted for publication (if this box is checked, you may delete all the material on page 2)*

☒ *Some of the work described in this thesis has been published and it has been documented in the relevant Chapters with acknowledgement (if this box is checked, you may delete all the material on page 2)*

☐ *This thesis has publications (either published or submitted for publication) incorporated into it in lieu of a chapter and the details are presented below*

CANDIDATE'S DECLARATION

I declare that:

- I have complied with the Thesis Examination Procedure
- where I have used a publication in lieu of a Chapter, the listed publication(s) below meet(s) the requirements to be included in the thesis.

Name	Signature	Date (dd/mm/yy)

Aspects of primary HIV-1 infection in New South Wales

Angie N. Pinto

A thesis submitted for the degree of Doctor of Philosophy

The Kirby Institute
Faculty of Medicine

TABLE OF CONTENTS

ACKNOWLEDGEMENTS

This work has been a long road with many turns, traffic lights, highway cruising, speed limits and rest stops along the way. It is a culmination of seemingly endless ethics applications, patiently waiting for data before finally analyzing and presenting results.

It is the result of the incredible leadership of my supervisors who have provided unwavering support along the way. I am profoundly grateful to the late Professor David Cooper, who was always thinking big, and several steps ahead of the rest. His foresight and compassion was unsurpassed in the field of HIV, and it is fitting that this thesis is a continuation of work he established over three decades ago. My heartfelt thanks go to Professor Tony Kelleher, the ultimate clinician scientist, who shared with me his expertise ranging from molecules to populations. I would like to thank Associate Professor Kathy Petoumenos, who provided the roadmap that helped me navigate the statistical minefield.

I am grateful to the National Health and Medical Research Council (NHMRC) for providing research funding, and to the Kirby Institute (KI) for fostering a collaborative research environment. Two chapters of this thesis relied on data linked by the Centre for Health Record Linkage, and was securely maintained in a database managed by Ansari Shaik. One chapter relied on multidisciplinary efforts within an NHMRC Partnership grant. I am grateful for the phylogenetic expertise of Francesca Di Giallonardo and also the support of colleagues from the Australian Molecular Epidemiology Network.

Finally I am eternally indebted to my parents Douglas and Alda for their hard work and sacrifice that has enabled me to be in the position I am today. And to my husband Deven for all the *on the spot guidance*, I would like to finally say *you got a credit*.

LIST OF PUBLICATIONS DURING CANDIDATURE

This thesis incorporates the published works #1, #2, #4 as listed below:

1. Pinto, A.N., *et al.*, *HIV-1 subtype diversity, transmitted drug resistance and phylogenetics in Australia.* Future Virology, 2018. **13**(8). p. 1-10.

2. Pinto, A., *et al*., *Early Treatment of Primary HIV Infection Is Associated with Decreased Mortality.* AIDS Res Hum Retroviruses, 2018. 34(11). p.936-941.

3. Pinto, A., *et al*., *A10 Using the molecular epidemiology of HIV transmission in New South Wales to inform public health response: Assessing the representativeness of linked phylogenetic data.* Virus evolution, 2018. **4**(suppl_1): 4(1). pp. 010. 009.

4. Pinto, A., *et al*., *Evolution of HIV-1 surveillance drug resistance mutations over ten years in New South Wales, Australia.* AIDS Res Hum Retroviruses, 2017.p. 1-16.

5. Pinto A and Cooper D, *Antiretroviral therapy: research, rollout and resistance.* Microbiology Australia, 2014. **35**(2): p. 79-82.

6. Pinto, A and Cooper D, *The end of HIV: how do we get there?* Med J Aust, 2014. **201**(2): p. 77-78.

LIST OF PRESENTATIONS (POSTER AND ORAL) DURING CANDIDATURE

1. Pinto A, *et al*. *Characteristics associated with HIV-1 Transmision Clusters in New South Wales*. in *2018 Australasian HIV&AIDS Conference*. 2018. Sydney, Australia.

2. Pinto A. *The Levinia Crooks Emerging Leader Award Oration: Zero by 2020*. in *HIV & AIDS Annual Australasian Society for HIV Medicine Conference*. 2018. Sydney, Australia.

3. Pinto, A., *et al*., *Using the molecular epidemiology of HIV transmission in NSW to inform the Public Health Response: assessing the representativeness of linked phylogenetic data*, in *22nd International Bioinformatics Workshop on Virus Evolution and Molecular Epidemiology (VEME)*. 2017: Universidade Nova de Lisboa, NOVA Campus - Campolide, Lisboa, Portugal.

4. Pinto A, *et al*. *Rates of Transmitted Drug Resistant Mutations in Newly Diagnosed HIV in NSW 2004-2015*. in *Australasian HIV&AIDS Conference* 2017. Canberra, Australia.

5. Pinto, A., *et al*. *Changing mortality and cause of death in an Australian cohort of primary HIV infection*. in *ASID Annual Scientific Meeting* 2016. Launceston, Tasmania.

6. Pinto, A., *et al*. *Evolution of HIV-1 Drug Resistance Mutations in NSW Over 10 Years*. in *World STI & HIV Congress 2015 & Australasian HIV & AIDS Conference 2015*. 2015. Brisbane, Australia.

7. Pinto, A., *et al*., *Ten year prevalence of HIV-1 drug resistance mutations in New South Wales, Australia* in *8th IAS Conference on HIV Pathogenesis, Treatment and Prevention* 2015: Vancouver, Canada.

LIST OF ABBREVIATIONS

3TC	lamivudine
ABC	abacavir
ABS	Australian Bureau of Statistics
AHOD	Australian HIV observational database
AIDS	acquired immune deficiency
APDC	Admitted Patient Data Collection
ART	antiretroviral therapy
cART	combination antiretroviral therapy
CCR	cancer care registry
CDC	Centers for Disease Control and Prevention
CHeReL	Centre for Health Record Linkage
CMR	crude mortality rate
CNS	central nervous system
CoDe	Coding Causes of Death in HIV
DNA	deoxyribose nucleic acid
DoB	date of birth
EDDC	Emergency Department Data Collection
EDI	estimated date of infection
EDTA	Ethylenediaminetetraacetic acid
EFV	efavirenz
ELISA	enzyme linked immunosorbent assay
GART	genotypic antiretroviral resistance test
HAART	highly active antiretroviral treatment
HBV	hepatitis B virus
HCV	hepatitis C virus
HIV	human immunodeficiency virus
INSTI	integrase strand transfer
LCI	lower confidence interval
MLK	master linkage key
NCIMS	notifiable Conditions Information Management System
NNRTI	non-nucleoside analogue reverse
NRTI	nucleos(t)ide analogue reverse
NSW	New South Wales
PBS	pharmaceutical benefits scheme
PCUD	primary combined united database
PDR	pre-treatment drug resistance
PHI	primary HIV infection
PI	protease inhibitor
PPN	project person number
PrEP	Pre exposure prophylaxis
PYFU	person year follow up
RBDM	registry of Births, Deaths and Marriages
RNA	ribonucleic acid
RPA	Royal Prince Alfred

rPI	ritonavir-boosted protease inhibitor
SAPS	Sydney AIDS prospective cohort
SDRM	surveillance drug resistance mutation
START	Strategic Timing of Antiretroviral Treatment
STR	single tablet regimens
STV	St Vincent's
TAM	thymidine analogue mutations
TDF	tenofovir
TDR	transmitted drug resistance
UCI	upper confidence interval

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

LITERATURE REVIEW

This literature review provides a background for the three main themes of this thesis. Firstly an introduction to primary HIV infection (PHI), secondly the classification and epidemiology of drug resistance mutations and finally an explanation and rationale for the molecular epidemiological study of HIV in New South Wales. It outlines strategies and approaches to PHI management in New South Wales prior to the commencement of this thesis in 2012, and includes literature reviewed until the end of 2018.

1.1 PRIMARY HIV INFECTION

1.1.1 Description of primary HIV infection

Primary human immunodeficiency virus infection was first described by Cooper *et al* [1] and is defined as the time from initial infection to complete seroconversion [2-4]. In up to 90% of patients it is characterised by an acute retroviral syndrome or seroconversion illness that occurs between 6 to 56 days after exposure. The severity of seroconversion has been shown to relate directly to viral load and clinical progression [2, 3].

In addition to the clinical and immunological implications of PHI, its role in the public health management of HIV infection is an area of considerable interest. Given the high level of viraemia, it is postulated that PHI plays a critical role in the onward transmission of HIV and is essential to incorporate into any HIV prevention strategy. Emphasis on the diagnosis and management of PHI has shifted from one of understanding disease pathogenesis to one of immediate treatment initiation with the dual aims of maximising treatment benefit for the individual and the public health strategy of preventing onward transmission.

1.1.2 Stages and definitions

Due to the challenges in recognising the clinical syndrome of PHI, standard definitions for staging of PHI are needed for inclusion in clinical trials, diagnostic assay evaluation and public health surveillance. A general description of terms used in this thesis is summarised below in Table 1.1:

Table 1.1 Definitions of Terms

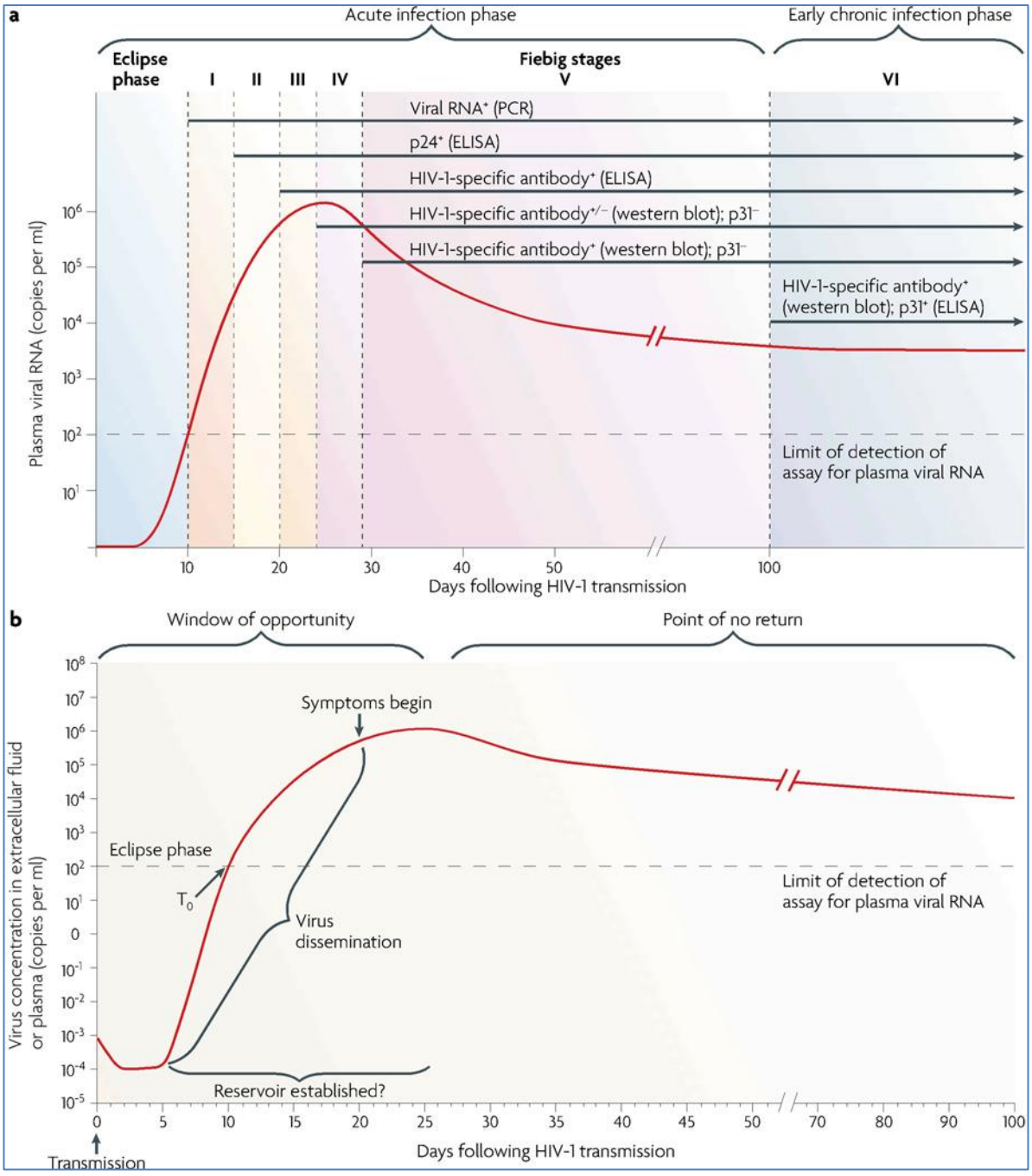
Term	Description	Duration	Definitions used in thesis
Acute HIV infection (AHI)	Time from human immunodeficiency virus (HIV) infection to the presence of HIV-specific antibodies; characterized by a rapid rise in the viremia	Approximately 24 days [4]	Either i) signs and symptoms of acute retroviral symptoms with the presence of positive p24 antigen/HIV-1 proviral DNA/HIV-1 RNA or ii) indeterminate or evolving Western blot with positive p24 antigen/HIV-1 proviral DNA/HIV-1 RNA
Early infection	Time from HIV infection to a viraemic steady state or set point; characterized by replication of the virus and the development of the immune response during which there is a decline in the CD4 T-cell counts	Approximately 52 days [4]	Positive ELISA and Western blot with a negative documented serological test within 6 months
Recent	State that begins when the biological process of HIV infection	First 2 years of HIV infection	NSW Ministry of Health uses the term "**Early stage infection**" for those with

Term	Description	Duration	Definitions used in thesis
infection	is first initiated; this term is used primarily in the context of measuring HIV incidence	[5, 6]	evidence of early stage infection/being infected in the 12 months prior to diagnosis: a seroconversion illness or negative or indeterminate HIV test within 12 months of diagnosis, irrespective of CD4 or an AIDS defining illness at diagnosis[7]
Chronic infection	For people not receiving antiretroviral therapy: fairly constant level of viremia (“set point”), steady declines in CD4 T-cell counts, and increasing rates of opportunistic infections For people receiving antiretroviral therapy: suppressed viral loads, increases in CD4 T-cell counts, and reducing rates of opportunistic	Approximately 10 years, but with a range of 2 to 20 years after infection for people not receiving antiretroviral therapy [8] or nearly lifelong for people receiving effective antiretroviral therapy [9]	Categories of CD4 T cell counts of 500+, 350-499, 200-349 cells/μl excluding acute, early and advanced stage categories.

Term	Description	Duration	Definitions used in thesis
	infections		
Late-stage infection	CD4 T-cell levels decline to very low levels, rapid rises in viremia, and death if infection is left untreated	Approximately 1 year if untreated [10]	NSW ministry of health uses the term uses the term "**Advanced stage"** for those with CD4 T cell counts count less than 200 cells/µl or an AIDS defining illness in absence of 'Early' criteria [7]

More stringent laboratory definitions for acute HIV infection have been devised based on a unique study in plasma donors by Fiebig *et al*, are are now widely known as Fiebig stages[4]. These are summarised in Figure 1.1 below.

Figure 1.1 Stages of acute HIV infection



a) Fiebig stages I-VI *b) Events in acute HIV-1 infection*

Given the importance in identifying recent infections, laboratory methods can be used to accurately diagnose recent infections in the absence of reliable clinical or epidemiological data points. Several assays have been evaluated to estimate the recency of infection and include a number of approaches: detuned assays use a lower cut-off and are used to identify recent infections as distinct from chronic infections; avidity index assesses the strength of antibody binding which increases with duration of infection; BED EIA calculates the proportion of HIV specific IgG compared to total IgG; antibody reactivity to an immunodominant epitope; isotype specific IgG3 antibody detection; antibody titre quantification and evaluation of the amount of viral genetic diversity, which increases with duration of infection [11]. However these approaches rely on specialised assays, and are not routinely available to most laboratories. A widely accepted and more practical approach to estimating the date of infection is based on the Fiebig stages outlined above and is the calculated midpoint between first positive and last negative HIV serology test result [12].

1.1.3 Clinical and immunological features

Seroconversion illness is a non-specific clinical syndrome characterised by a mononucleosis like syndrome consisting of a number of symptoms including: fevers, malaise, lethargy, anorexia, nausea, myalgia, arthralgia, headache pharyngitis, diarrhoea, lymphadenopathy and rash [1]. There may be a wide spectrum of clinical presentations affecting all organ systems, including neurological and haematological [13]. The presence of a seroconversion illness has been shown to be associated with adverse outcomes, with a “dose-response” relationship between severity and disease progression [14-16]. Awareness and education of the broad clinical presentation is essential to obtain early diagnosis and treatment as part of any HIV prevention program [17].

There are a number of immunopathological changes that have been described during PHI including an inverted CD4:CD8 ratio due to CD4 T cell depletion and expansion of cytotoxic CD8T lymphocytes [18-22]. During this time there is also immune activation, cytokine secretion [23] and immune escape mutants that results in inability to eradicate virus [20]. At this time there is establishment of reservoirs in peripheral blood and lymphoid tissues which remain some of the key barriers to achieving virological cure [24-27].

Studies in PHI are particularly valuable in gaining greater understanding of early viral and host characteristics and their relationship to long term outcomes. Studies of this nature are challenging to recruit, due to the non-specific and transient nature of

seroconversion symptoms. Cohorts of individuals identified and enrolled during PHI are a valuable resource to understand pathogenesis and disease progression from an early stage. There are caveats to recruiting cohorts of individuals identified with primary HIV infection. They may tend to be more representative of persons engaged in the health care system, seeking care for early symptoms that may otherwise be ignored in the general population. For example, those with shorter testing interval (time to negative test) were more likely to report symptoms of seroconversion [28]. They may therefore be more biased towards high risk populations such as men who have sex with men who have increased awareness about symptoms of seroconversion illness, compared to other risk groups, for example. There may also be a temporal bias, in that there was heightened awareness of symptoms of PHI around the initial time of its first description, and this recognition may have waned over time.The contributing studies of primary HIV infection cohorts whose data forms part of this thesis are summarised in Table 1.2

Table 1.2 Summary of PHI study cohorts included in thesis

Name	Inclusion	Enrolled (n)	Years enrolled	Design	Objective	Intervention	Treatment Interruption	Ref
SAPS: Sydney AIDS prospective cohort	MSM and bisexual men	1057	1984-85	prospective cohort	Immuno-epidemiological study	-	no	[1]
PHAEDRA: primary HIV and early disease research	HIV acquisition within 12 m	137	2002-2007	observational natural history cohort study • retrospectively included studies below • prospectively recruited all consenting seroconverters	Correlates of progression, biological specimen repository	-	no	[29]
VIRAX	ART within 6/12 PHI	35	1996-2001	RPC phase I/II	RPCDB trial of recombinant fowl pox IFG vaccine	PBO, FPV + HIV genes, FPV + HIV genes + IFG	no	[30]
SPARTAC	PHI	366	August 2003-July 2007	RCT	RCT of 0,12,48w ART	94% PI based	yes	[31]

PULSE	PHI	68	Jan 2000-Feb 2002	RCT	RCT of structured treatment interruptions	IDV, RIT, DDI, D4T/3TC +/- HU	yes	[32]
PHIIDO: primary HIV infection data observational	Newly diagnosed infection	16		observational	Immune and viral events in early HIV	-	no	[33]
AGOURON study 511	Not on Rx	297		phase III multicentre DBRPCT	to evaluate safety and activity of NFV 2 doses with cART	NFV 750/ NFV 500/PBO + ZDV/3TC	no	[34]
QUEST		148		RCT	addition of vaccine to ART --> viral suppression	ART 76wks then ART alone/ ART + ALVAC/ ART + ALVAC + remune	yes	[35]

Abbreviations: Rx treatment; RCT randomised controlled trial; RPCDB randomised placebo controlled double blind; ART antiretroviral treatment; IFG interferon gamma; PBO placebo; FPV fowlpox virus; IDV indinavir; RIT ritonavir; DDI didanosine; D4T stravudine; 3TC lamivudine; HU hydroxyurea; NFV nelfinavir; ZDV zidovudine; cART combination antiretroviral treatment

1.1.4 Treatment strategies for PHI

There have been changing treatment paradigms in PHI over the years, with very aggressive treatment during PHI in pre-highly active antiretroviral treatment (HAART) era, deferral of treatment in early post HAART era, followed by earlier treatment again as CD4 T cell count thresholds changed [36]. Structured treatment interruptions (STIs) have been evaluated to try to improve virological control after PHI while limiting long term toxicity from antiretroviral therapy. However it has been definitively shown that STIs did not confer a benefit in terms of risk of opportunistic infections or death, and did not reduce risk of toxicity from ART but and actually increased morbidity [37, 38]. Novel strategies such as adjunctive hydroxyurea and therapeutic vaccines were not shown to be beneficial in this setting [32, 35]. Cohort studies evaluating the role of early intervention gave inconsistent results and recommendations were based on expert opinion [19, 39] .Due to lack of robust clinical trials in PHI setting, uncertainty still existed about whether it was beneficial to commence ART during this time [40].

The arguments supporting early ART were based on improved immunological markers and delayed disease progression. A controlled study showed improved CD4 T cell counts at 24 weeks, but not sustained in the longer term [41]. Normalisation of CD4/CD8 ratio was also shown with ART commenced within 6 months [44]. Treatment within 6 months of seroconversion controls the viral reservoirs [41-44] and the proportion of latently infected cells compared to the chronically infected untreated population [45]. There is a decay in latently infected T cells after treatment, but this phenomenon plateaus out in a biphasic manner [45]. In terms of virological effects, there was a lower mean viral load and viral set point in primo-SHM trial with early treatment, but the arms converged after 2-3 years [46, 47]. Delayed viral rebound with early ART has also been demonstrated[48] with duration of viral control related to longer duration of ART during PHI [49]. There is also the added effect of ART in controlling viraemia with delayed need for restarting ART in the future [46, 50]. The Concerted Action on Seroconversion to AIDS and Death in Europe (CASCADE) cohort is a combination of 23 seroconverter studies from Europe, Australia and Canada[51]. This collaboration showed there were regional differences in time to initiation of ART in seroconverters with more initiating during PHI in sub-Saharan Africa compared with western countries in the with corresponding lower mortality in this group [52].

The Short Pulse Anti-Retroviral Therapy at Seroconversion (SPARTAC) was a randomised clinical trial aimed at evaluating the role of early ART in delaying disease progression [31]. Although this was shown to be the case the benefits did not extend

much beyond the duration of treatment [16, 31, 53-55].The CASCADE collaboration, bringing together multiple European primary infection cohorts, showed slower disease progression with treatment started at presentation in those with CD4 less than 500 cells/µl [55]. This was also the aim of the SETPOINT study, however this study was ceased prematurely as the delayed therapy group progressed to meet criteria for ART, supporting the argument that early treatment delays disease progression [56].

Despite this body of evidence supporting treatment during PHI, several studies cast doubt on these findings, creating uncertainty in this area. A meta-analysis showed improvements in CD4 count and lower viral loads, but this effect was not sustained beyond 12 months after treatment interruption [19]. In the ANRS PRIMO cohort, few patients maintained very low viral loads after treatment interuption [57]. Early treatment even within 10 days of PHI did not affect the establishment of a reservoir of infected T cells with integrated HIV DNA [26, 27]. The long half life of reservoir cells means that it is the duration of treatment rather than timing of treatment that has the greatest impact [58]. Following in this line of argument, there was also no additional benefit of an intensified 5 drug regimen compared to 3 drug regimen [59]. Studies have also shown better immunological responses in those commmenced on continuous treatment during chronic infection compared with PHI treated group who interrupted therapy, however this study's findings probably relate to the effects of treatment interruption, rather than starting during PHI itself [21].

There are numerous descriptions in the literature of varying degrees of clinical progression and control after infection with HIV, with a wide range of phenotypes ranging from elite controllers to rapid progressors [60]. Immune control has been demonstrated in patient with low viral loads after treatment of acute HIV [46, 61] and also in the setting of infants starting ART soon after birth [62-64]. Controllers have low HIV viral loads in peripheral blood monoculear cells and high HIV speciifc immunity[65]. Post treatment control is thought to be most effective when treatment is started in the acute PHI phase soon after seroconversion, most likely due to low amount of reservoir [66, 67].

1.1.5 Long term outcomes and mortality

Long term outcomes are thought to be related to the severity of seroconversion illness [15], and the degree of immunological and virological responses that occur during PHI. Higher viral loads are thought to be related to disease progression and long term outcomes, although there is wide variation in viral load before reaching a set point after host-virus interaction [68].

Survival overall in those with PHI has improved since introduction of HAART [69]. The CASCADE cohort showed mortality nearing that of the general population, but increases with duration of infection [70]. Analyses of 38 studies showed that age at seroconversion and time from seroconversion were major determinants of survival prior to ART [51]. Large observational cohort studies of chronic HIV in various geographic settings have showed a decrease in AIDS related deaths, with increase in non-AIDS conditions such as liver disease and overdose, but there are no comparable studies in PHI cohorts [71-75]. An Australian cohort also showed a decline in overall cancer rate; however this study mainly recruited chronically infected HIV individuals [76]. Overall mortality still remains higher in HIV infected patients compared to the general population [77]. Mortality studies specifically pertaining to PHI cohorts are lacking due to difficulty in recruiting these populations, and the long duration of follow up required.

1.1.6 Public health aspects

In recent years, the more compelling argument for early treatment is the public health approach to prevent onward transmission. There are several studies outlining the role of PHI in driving onward transmission [12, 78-82]. Although the case for early treatment has conclusively been made through the START study, this clinical trial included patients with a median time since diagnosis of one year, without any indication about time of acquisition, although median CD4 count was > 500 cells/µl at baseline [83]. With comprehensive published evidence for benefits of treatment in the individual, guidelines have also shifted towards early treatment for all, particularly for population based public health benefits [84, 85].

The Test and Treat strategy aims to reduce infectivity, which is particularly relevant during PHI as local studies have shown the impact of seroconverters on onward transmission of HIV infection [79, 86]. While it follows that the higher viral load during primary infection could lead to onward transmission, published estimates of this worldwide vary from between 4-50%, making this an area for further research [87, 88].

In Australia, the adoption of international guidelines advocating universal treatment was enabled in April 2014 with the removal of CD4 count criterion for subsided ART. This resulted in a reduction in median time to treatment initiation from 84 days to 19 days [89]. Australian guidelines have supported initiation of treatment during acute and recent infection with standard combination regimens [90].

1.1.7 Summary

PHI is a critical time to study the clinical, immunological and virological events that occur after initial HIV infection. It is challenging to identify and recruit individuals for these studies, and study cohorts containing these participants remain a unique resource to further understand PHI. In an era where universal treatment is recommended regardless of stage of infection, the question around treatment during PHI has added implications beyond disease progression and mortality, such as reduction of size of reservoir and improved immunological markers. The question surrounding the impact of starting treatment during PHI is still unanswered, in particular whether early treatment in this group improves long term clinical outcomes [91]?

1.2 HIV-1 DRUG RESISTANCE (HIVDR)

As the management of HIV in NSW has evolved towards immediate treatment with antiretrovirals, the need to recognise baseline drug resistance mutations is even more important. Section 1.2 and 1.3 outlines the definitions, epidemiology and evolution of drug resistant mutations in the context of treatment rollout in Australia. It includes sections that have been previously published as part of a review article “HIV-1 subtype diversity, transmitted drug resistance and phylogenetics in Australia” in Future Virology, 2018 [92].

1.2.1 Drug resistance mutations classification

HIV-1 is a rapidly evolving virus and is prone to errors during viral replication. The *polymerase* gene encodes the *protease*, *reverse transcriptase* and *integrase* enzymes. Mutations in this gene can impact on susceptibility to antiretroviral therapy and may be associated with treatment failure. However these mutations can come at a fitness cost, and may not be as transmissible to a new host. Antiretroviral mutations can lead to several ways in reducing overall viral fitness. These include reduced replication capacity, altered mutation rate and changes in host cell tropism. This first phase of loss of viral fitness can lead to a clinical phenotype of reduced viral load, despite the individual being on a failing regimen. However over time, compensatory mutations may develop such as amino acid changes and altered cleavage sites that could lead to increased viral replication capacity and rising viral load over time, which in turn could increase transmissibility to a new individual [93]. Therefore it is important to characterise the type of drug resistance mutations according to standard definitions.

Transmitted drug resistance (TDR)

According to the WHO Global action plan, transmitted HIVDR (TDR) is “detected in ARV drug naive people with no history of ARV drug exposure. TDR occurs when

previously uninfected individuals are infected with virus that has drug resistance mutations" [94].

Mutations occur at a fitness cost, and these viruses may be less transmissible; over time these revertant mutations may revert to wildtype. Other mutations (e.g. K103N, Y181, L90M) have similar fitness to wildtype and therefore survive longer within population, and can remain as a potential reservoir for onward transmission [95].

Acquired drug resistance (ADR)

Acquired HIVDR (ADR) "develops when HIV mutations emerge due to viral replication in individuals receiving ARV drugs", as defined in the WHO action plan [94]. Resistant variants may be selected out during therapy, and can remerge as the predominant population upon discontinuation of therapy. In the days of monotherapy, the emergence of resistance variants was what often led to treatment failures. This knowledge and further clinical studies eventually led to the development of three drug antiretroviral regimens as the gold standard. Over time, the overall efficacy of antiretroviral regimens has improved [96], and simultaneously with these better tolerated regimens, rates of acquired drug resistance has been decreasing [97, 98]. With newer regimens in the setting of immediate treatment, the risk of ADR has been estimated at between 2-3% [99] . It should be noted that the risk estimates of ADR vary between 2-3% depending on the timing of starting treatment which may vary between studies [99]. In developing regions, where access to viral load monitoring is limited there is wide variance (2-26%) in virological failure rates [100].

Thymidine associated mutations (TAMS) are mutations that are associated with older thymidine analogue reverse transcriptase inhibitors such as AZT and d4T. They include M41L, D67N, K70R, L210W, T215Y/F and K219Q/E[101]. They remain clinically relevant as may be associated virological failures to current first line NRTI regimens [102].

The *integrase* inhibitor class of antiretrovirals has a high genetic barrier to resistance. Acquired resistance to this class is generally low, with one case report in Australia [103], however has been reported to increasing in the United States up to 1.4% [104], while remains stable in the Swiss cohort despite widespread rollout [105].

Pre-treatment drug resistance (PDR)

Pre-treatment HIVDR (PDR) is "detected in ARV drug naive people initiating ART or people with prior ARV drug exposure(s) initiating or reinitiating first-line ART" [94]. This may include mothers started on antiretrovirals to prevent mother to child transmission,

and also in the setting of pre-exposure prophylaxis. One study did review resistance mutations that were common in clinical practice, and found a few key mutations that occur frequently are reflective of common prescribing patterns [106].

Surveillance drug resistance mutations (SDRM)

To provide standardisation of mutations across surveillance programs, a consensus list of 93 mutations has been recommended by the World Health Organisation [107]. This list has been developed to enable standardised comparison over temporal and geographic regions, to allow population based meta-analysis of data. The criteria for inclusion in this list includes: 1) mutation known to cause drug resistance; 2) mutation occurs in non-polymorphic positions ;3) mutation seen in the eight most common subtypes; 4) excludes rare mutations [107].The WHO SDRM list differs slightly from major mutations that have the potential to cause clinically significant virological failure. There are several lists of these mutations compiled by experts, such as the IAS-USA drug resistance mutation list and the Stanford HIVdb drug resistance interpretation algorithm [108, 109]. These have been developed using different criteria, and are suited to clinical decision making for an individual patient. They are more readily updated, and adapted with availability of newer antiretroviral drugs and classes, and are based upon published data. Criteria for inclusion of mutations in these lists includes: 1) in vitro passage studies; 2) laboratory antiviral susceptibility testing; 3) mutation demonstrated in individuals failing therapy; 4) association studies with baseline genotyping and development of mutation after treatment failure. A comparison of the WHO SDRM list to major mutations in the HIVdb list is shown in Table 1.3 [109].

Table 1.3 Comparison of WHO SDRM list and major drug resistance mutations from the HIVdb algorithm

NRTI			NNRTI			PI		
	SDRM	**HIVdb**		**SDRM**	**HIVdb**		**SDRM**	**HIVdb**
M41	L	L	**L100**	I	I	**L23**	I	
K65	R	K, R	**K101**	E, P	E, P	**L24**	I	
D67	N, G, E	N	**K103**	N, S	N, S	**D30**	N	N
T69	D, Ins	Ins	**V106**	M, A	M, A, I	**M33**		F
K70	R, E	R	**E138**		ASK Q	**V32**	I	
L74	V, I	V, I	**V179**	F		**M46**	I, L	I, L
V75	M, T, A, S		**Y181**	C, I, V	C, I, V	**I47**	V, A	V, A
F77	L		**Y188**	L, H, C	L, H, C	**G48**	V, M	V, M
Y115	F	F	G190	A, S, E	A, S, E	**I50**	V, L	V, L
F116	Y		**P225**	H		**F53**	L, Y	
Q151	M	M	**M230**	L	L	**I54**	V, L, M, A, T, S	V, L, M, A, T
M184	V, I	V, I				**G73**	S, T, C, A	
L210	W	W				**L76**	V	V
T215	Y, F, I, S, C, D, V, E	Y, F				**V82**	A, T, F, S, C, M, L	A, T, F, S, L
K219	Q, E, N, R	Q, E				**N83**	D	
						I84	V, A, C	
						I85	V	V
						N88	D, S	D, S
						L90	M	M

1.2.2 Epidemiology of HIVDR

Globally scale up of ART has resulted in increased rates of TDR, mainly observed in developed countries such as Australia, United Kingdom and United States. Australia has been noted to have the highest rates of TDR in a recent published meta-analysis as well as in the START clinical trial data [110-112]. Limitations of both these publications are the relatively small sample sizes, limited geographic sites sampled, as

well as older data which may be more reflective of historic ART regimens. There is a need for a more comprehensive, contemporaneous and inclusive study on drug resistance in Australia.

Table 1.4 summarises recently published key studies of drug resistance in different regions. Countries that have shown increasing rates of TDR include the United Kingdom, with the highest rates in NNRTIs [113]. There are differences within risk groups, with a decreasing trend in MSM, while rates remained stable in heterosexual transmission risk groups. Increasing trends in TDR were observed in Latin America and Caribbean countries from 6.0% (2000-2005) to 8.2% (2006-2015)[114]. The highest rates of TDR have also been noted in younger age groups [115].

Regions that have reported recent decreasing rates include Canada in the (people who use drugs (PWUD) population [116, 117] as well as Washington DC [118]. These observations may be due to more rigorous follow up and linkage to care in these subpopulations.

The prevalence of TDR in the recently infected population is particularly important to monitor as it is reflective of mutations that are circulating and being transmitted within a population, before being archived and potentially replaced by reversion or by treatment induced mutations in chronic infection. Rates can vary between different populations, and it is important to monitor this group within each jurisdiction. For example, in the recently infected PWUD population in Canada, TDR rates are decreasing [116] whereas in Spain, they are stable [119]. While in Africa, no TDR was observed in this group[120]. Recent data in the recently infected population in New South Wales has not been previously published.

Table 1.4 Summary of drug resistance studies

Author (Ref)	Year of Publication	Years sampled	Number of Sequences (individuals)	Site-Type (Risk-group)	Stage	Overall-TDR (%)	NRTI-TDR (%)	NNRTI-TDR (%)	PI-TDR (%)	ADR
Tostevin [113]	2017	2010-2013	16425	UK MSM Hetero	ART naïve	8.1→6.6 8.7 6.4				
Socias [117]	2017	1996-2015	(573)	PWUD	ART naïve	9.8	3.0	5.4	1.9	
Rocheleau [97]	2017	1996-2014	23271 (6543)	Canadian cohort	ART naïve	12→18				39→3
Onywera [120]	2017	2012	87	Kenya	AHI, RHI	9.2	4.6	6.9	1.2	
Margot [121]	2017	2000-2013	(6704)	Clinical studies UK Europe	ART naïve	5.2→11.4	3.1	0.3->7.1		

Aldous [104]	2017	1999-2014	(3411)	US	ART naïve & exp	20.5	15.0→5.5			40.5
Scherrer[98]	2016	1999-2013	(11084)	Swiss cohort	ART exp					57.0→37.1
Park [122]	2016	1999-2012	(928)	South Korea	ART naïve	4.8	2.2	2.7	0.3	
Mendoza [123]	2016	2007-2013	(717)	Panama	ART naïve	9.2			4.1	87.6
Rhee [112]	2015	2000-2013	(50870)	SSA SSEA Asian LA Europe N. America	ART naïve	2.8 2.9 5.6 7.6 9.4 11.5				

Abbreviations: PWUD people who use illicit drugs; AHI acute HIV infection; RHI recent HIV infection; SSA sub-Saharan Africa, SSEA south/southeast Asia; LA Latin America/Caribbean

1.2.3 Antiretroviral treatment in Australia

With treatment paradigms shifting towards earlier treatment of HIV, it is unknown what impact this will have on primary and secondary drug resistance, although some models predict an increase [124, 125]. National HIV guidelines currently recommend combination antiretroviral regimen inclusive of an integrase strand transfer inhibitor[90]. There are several reports of trends in drug resistance in the international literature, however comprehensive local data is lacking [126-130]. Globally, the overall trend in transmitted drug resistance has been reviewed by Frentz *et al* who found a decreasing rate of NRTI resistance in resource rich countries, with an increasing rate in Asia and Africa, reflecting stages of ART rollout and prescribing [131].

In 2017, an estimated 87% of those diagnosed are on antiretroviral treatment[132]. The median time to starting treatment has reduced from 84 days in 2012 to 19 days in 2016 in an Australian cohort [89]. Individuals will be receiving ARV for longer durations due to starting ARV treatment earlier in the course of illness along with longer life expectancies [133]. Given the shifts towards earlier HIV treatment, there is a need for a population based survey of drug resistance in Australia.

1.2.4 Genotypic antiretroviral resistance testing in Australia

Genotypic antiretroviral resistance testing (GART) has been recommended in HIV treatment guidelines since 2008 in Australia, and is performed in seven reference laboratories around Australia as shown in Table 1.5. GART is recommended in acute HIV infection, ART-naïve patients, virological failure of ART, suboptimal suppression of HIV plasma load, and in HIV-infected pregnant women before ART initiation [134]. All laboratories currently perform Sanger sequencing of the *polymerase* gene to obtain sequences of the protease or reverse transcriptase regions; *integrase* sequencing is routinely performed at two laboratories. All laboratories participate in external accredited quality assurance programs to monitor assay quality and accuracy. Newer and automated whole genome sequencing (WGS) assays are being assessed, and may offer earlier and more sensitive detection of low-level antiretroviral resistance mutations. The WGS approach may also improve the reliability of molecular epidemiological analyses. The genotypic resistance interpretation algorithm (Stanford HIVdb Program) has been shown to be adequate for the management of HIV infection compared to virtual phenotype approach, and is currently used by all laboratories [135, 136]. Phenotypic antiretroviral resistance testing is not performed routinely in Australia.

Table 1.5 Summary of laboratories performing GART in Australia

Laboratory	Method	Region
Western Australia	In house	polymerase (RT & PR) and integrase
South Australia	In house	polymerase (RT & PR)
Victoria	In house	polymerase (RT & PR) integrase (on request) envelope (CCR5) (on request)
Queensland	ViroSeq HIV-1 Genotyping System	polymerase (RT & PR)
ICPMR, Westmead, NSW Health Pathology	In house	polymerase (RT only) and integrase
RPA immunology laboratory, NSW Health Pathology	ViroSeq HIV-1 Genotyping System	polymerase (RT & PR)
NSW State reference laboratory for HIV/AIDS & Molecular Diagnostics	ViroSeq HIV-1 Genotyping System	polymerase (RT & PR) integrase (on request) envelope (CCR5)

Abbreviations: PR protease; RT reverse transcriptase

1.2.5 Transmitted drug resistance in Australia

Availability of HAART has significantly reduced the morbidity and mortality of patients infected with HIV. However, long-term treatment and associated adherence issues have facilitated the emergence of drug resistant viruses which can be directly transmitted to an antiretroviral naïve person at the time of their primary infection. Individuals with TDR may have a reduced genetic barrier for resistance which could lead to risk of virological failure, increased resistance to the drugs in their initial regimen and the possibility of cross-resistance [137].

The occurrence of TDR at the time of infection has been well documented, and in high-income countries where HAART has been in use for more than 20 years, its incidence has ranged from 10% in Europe [138] to upwards of 20% in certain parts of the US [118]. Globally, rates of TDR in Australia have been some of the highest reported in the world, with overall rates of between 15.5-17.5% [110, 111]; however, both studies had a median year of 2006 and 2007 respectively, and were very limited in their sampling, representing a single site with 100% MSM [111], and a clinical trial with 98 Australian participants in the other [110].

A summary of Australian TDR studies is presented in Table 1.6. Early studies on TDR rates in the acutely infected, predominantly MSM population between 1992 and 2001 in Sydney, NSW, reported mainly resistance to NRTIs (18.4%) [139], reflective of treatment options available prior to the introduction of HAART in Australia in 1996 [139, 140]. Similarly, in an 11-year study of TDR surveillance in Victorian patients recently infected between 1996 and 2006, an overall TDR prevalence of 16% mostly associated with NRTI resistance, was demonstrated [141]. More recent studies performed in both of these states have reported a decline in the prevalence TDR in similar populations [142, 143]. A Victorian study examining the effect of newer ARV medications on the prevalence of TDR to RTIs and PIs by comparing two time periods (2005-2010 vs 2011-2015), reported a decline in PI- and RTI- associated resistance from 11.4% to 7.9% respectively [142], whilst a recent study in NSW assessing TDR in treatment naïve individuals during 2009-2013, reported an overall prevalence of 13.6% as well as overall decreasing rates by calendar year [143]. This trend is in contrast to that reported in North America, which has seen increasing overall TDR rates in the context of treatment rollout [112]. Ongoing monitoring for TDR transmission is essential, with a single study reporting a case of triple class resistance [144].

Since 2015, integrase strand transfer inhibitors (INSTI) have been included in five of six recommended first-line regimens in Australian treatment guidelines, and are also used as part of a three drug post exposure prophylaxis regimen [90, 145]. In Australia, two state-based laboratories routinely perform baseline integrase resistance testing, whilst others genotype the *integrase* region upon request (Table 1.4). Based on current rates of TDR to various treatment regimens, genotypic drug-resistance testing of the protease and RT region is currently recommended in ARV-naive individuals, and integrase resistance genotyping is only performed when INSTI resistance is suspected. There is currently little evidence of TDR to INSTIs, but with increased use there is concern that emergence of TDR to this drug class will occur. One study investigating

the frequency of TDR to INSTIs in 461 newly infected individuals in Victoria between 2010 and 2015 reported one major INSTI mutation, where the impact on INSTI resistance was unclear, and several accessory mutations which another study suggests may be linked to prior antiretroviral therapy use [142, 146]. More recently, the first Australian report of INSTI resistance in a treatment naïve patient with no risk factors for INSTI resistance, further highlights the need for vigilance around detection of TDR to INSTIs [103].

The most important transmitted mutations involve single amino acid changes conferring high-level resistance to drugs contained within first-line regimens. Despite an overall decline in prevalence over the years, the most common drug resistance mutation identified in the local ARV-naïve population is the NRTI-associated codon change at amino acid T215, which in isolation, has little impact on first line treatment options [142, 143]. More importantly, but less frequently identified was M184V/I which confers high-level resistance to lamivudine and emtricitabine, both contained in first-line NRTI backbones. This mutation is also important in the context of pre exposure prophylaxis (PrEP) rollout, where infection with an M184V/I containing virus may compromise the activity of emtricitabine in this dual therapy regimen. Most prevalent NNRTI-associated mutations identified in the ARV-naïve population include K103N, Y181C and G190A, all of which confer high-level resistance to first generation NNRTIs [143]. While NNRTI-based regimens are no longer recommended first-line in Australia, efavirenz and nevirapine based regimens are recommended first-line in many low to middle income countries, which in the future could lead to further increases in TDR in these regions [112].

Table 1.6 Summary of HIV-1 drug resistance studies in Australia

Author (Ref)	Year of Publication	Median Sample Year	Years sampled	Number of Sequences	Site-Type (Risk-group)	Stage	Overall-TDR (%)	NRTI-TDR (%)	NNRTI-TDR (%)	PI-TDR (%)
Ammaranond [139]	2003	1993	1992-2001	185	NSW	Acute (PHI)		18.4	2.7	1.6
Ammaranond [140]	2003	1996	1992-2000	130	NSW	Acute (PHI)		43.9→19.1		0.8
Russell [141]	2009	2004	1996-2007	466	newly infected VIC	Naïve	16	8.6	4.1	1.3
Hawke [147, 148]	2013	2006	2000-2010	418	MSM 62%	Newly diagnosed	21.9	10.8	11.5	3.8
Pham [111]	2014	2006	1999-2013	368	MSM 100%	acute/recent	15.5			
Baxter [110]	2015	2007	2000-2013	98	AUST START study 2009	Naïve	17.5	9.3	4.1	5.2
D'Costa [142]	2017	2007	2005-2010	343	newly infected VIC	Naïve	11.4	8.5	5	0.6
Chibo [149]	2011	2008	2005-2010	209	HIVC (MSM 88%)	Acute infection (Seroconversion)	13.4	7.2	6.7	1.4
Pinto [143]	2017	2010	2009-2013	450	all (MSM ~80%)	Naïve	13.6	9.2	4.9	2.7
D'Costa [142]	2017	2013	2005-2015	772	newly infected VIC	Naïve	9.5	6.6	3.8	1.2

1.3 MOLECULAR EPIDEMIOLOGY of HIV-1

1.3.1 Application of molecular epidemiology for public health

Molecular epidemiology is defined as the use of gene sequence (genotyping) in combination with epidemiological data to study the distribution, dynamics and determinants of disease in populations. Pathogen genotyping can be used to understand disease transmission and pathogenesis, for clinical management and antimicrobial drug selection, and in understanding disease outbreaks of public health importance [150, 151]. Molecular epidemiology allows the characterisation of the relatedness of microorganisms within a species based on a unique genetic fingerprint. It is routinely used as part of the investigation and control of outbreaks such as MRSA, norovirus, *E. coli* and tuberculosis [152-156] . Together with contact tracing, in some settings these molecular and epidemiological tools can rapidly identify the source of an outbreak, risk factors for ongoing transmission and contribute to the prevention response.

In Australia, the recent establishment of a national alert system for detecting critical antimicrobial resistance has been able to monitor the spread of bacterial resistance genes identified through molecular mechanisms, yet none exists for monitoring of antiretroviral resistance[157]. Genotyping of the *polymerase* and *integrase* genes of human immunodeficiency virus (HIV) is now routinely performed to identify mutations associated with antiretroviral drug resistance. HIV is a virus with a high evolutionary rate, that rapidly adapts to the environment, and evolves mutations with drug pressure. There is large genetic diversity which can provide insights into the geographic and temporal origins of a virus within a population [158]. As such, with the advances in computational capacity, genetic evolutionary methodologies and ready availability of HIV sequences, the field of HIV molecular epidemiology has undergone rapid development [159-161]. As there is sufficient variability within the *polymerase* gene, which is routinely sequenced for HIV resistance testing, it may be used to infer phylogenetic trees and transmission networks. Integration of this routinely collected data into phylogenetic based surveillance systems has been used to identify disease outbreaks and inform public health responses [162, 163].

Australia has a low HIV prevalence of 0.13%, due to successful implementation of a public health response to HIV incorporating a combination of prevention strategies including condom use, needle syringe programs, post exposure prophylaxis (PEP) and more recently, treatment as prevention (TasP) [164]. There has been a steady increase in the number of newly diagnosed HIV infections since 1998, with 2012 seeing the

highest number in over twenty years[165]. Similarly, the number of national newly acquired infections also increased to 350 in 2013, a key indicator of ongoing transmission in Australia[165]. New South Wales (NSW) has the largest burden with an estimated 46% of people living with HIV in Australia[166].

In NSW, the rates of new diagnoses remained stable despite expanded testing, treatment and prevention strategies [164, 167]. While primarily a concentrated epidemic, predominantly affecting men who have sex with men, populations from high prevalence countries with heterosexual transmission are emerging as high priority groups for targeted prevention [168, 169].

In 2012, a new HIV strategy was launched in New South Wales, with the aims of reducing transmission of HIV by 60% in 2015, and virtual elimination in 2020 [166]. This strategy includes a multi-pronged approach including a wide range of behavioural and biomedical interventions. One of the newer prevention strategies was the large scale rollout of PrEP via implementation trials, with over 8000 enrolled participants across three states[170]. During this period, the first case report of multi-drug resistant virus acquisition while on PrEP highlighted the need for a coordinated approach to drug resistance surveillance [171, 172]. The new strategy focuses on reducing onwards HIV transmission by diagnosing people with HIV much earlier ("test more") and reducing their viral load to undetectable ("treat early") so that they become essentially non-infectious while their viral load remains undetectable[173]. This "test and treat" approach may be compromised by transmissions that occur very early in infection, before the person becomes aware of their infection. To enhance the success of this strategy, a greater understanding of HIV transmission was needed, using phylogenetic methods.

Recent international studies have used phylogenetic analysis to define the role of clinical, demographic and risk factors in rates of HIV transmission and transmitted drug resistance [174, 175]. Phylogenetic studies can provide insights on whether new strains are being introduced into a population or whether it is pre-existing circulating strains that are perpetuating the epidemic [176]. Research into HIV transmission networks using *de-identified* socio-demographic data can inform the prevention response in real time [159, 177-179]. Studies such as these can demonstrate the need for different prevention strategies and allow targeting of key affected populations [161, 169, 180].

In Australia, health authorities in Victoria and South Australia have used phylogenetic investigation to enhance surveillance through linking (then de-identifying) state-wide genotypic and HIV notifications databases to understand changes in transmission networks, subtype diversity and to inform public health initiatives targeting key populations [147, 181].

1.3.2 Phylogenetic methodologies

Phylogeny refers to the relationship between the genetic sequence of organisms [182]. Unless the entire population of interest is diagnosed and sampled, any attempt to reconstruct a phylogeny will be an approximation, and subject to limitations. There is no way of sampling members of the population who remain undiagnosed, not engaged in care, and as a result there will be limitations with any phylogenetic analysis.

Types of phylogenetic inference

There are several approaches to reconstructing a phylogenetic tree including neighbour-joining, maximum-likelihood and Bayesian methods [158, 183]. Neighbour-joining methods use a distance based approach, and is based upon the genetic distance calculated directly from sequence, and used to construct a single phylogenetic tree [184]. Maximum likelihood and Bayesian methods use models to assess the probability a particular topology is reflective of the sequence data [158] . A maximum likelihood tree is the “best tree” that is reconstructed after multiple resampling episodes. Bayesian methods use posterior probability to create multiple tree topologies and produce a consensus tree. Each branch within that tree has a probability based upon the number of trees in which that branch exists [158].

Bootstrap

Standard bootstrapping is used to assess the statistical likelihood of a particular branch within a phylogenetic tree. Ultrafast bootstrapping is a less computationally intensive method used [185]. A bootstrap value of between 70-90% is widely accepted as a reasonable cut-offs in the HIV literature to indicate adequate support for a grouping [158].

Genetic distance

The amount of genetic similarity between two sequences can be determined using one of two approaches. The first is a direct calculation of the number of nucleotides that differ between two sequences, and is known as the pairwise distance (p-distance). The other is a computed distance, and is based on a nucleotide substitution model that takes into account multiple back mutations that may have occurred [186].

Transmission cluster definition

A transmission cluster refers to a pair or series of genetically similar sequences that share a most recent common ancestor (MRCA). There is not a widely accepted standard definition for a HIV-1 transmission cluster [161, 187]. There have been a variety of definitions used for defining a transmission clusters, and the method chosen depends on the research question being asked, the population sampled and the methods used and number of sequences contained within a phylogeny. Bootstrap values between 70 to 98% combined with pairwise genetic distance between <1.5 to <4.8% have been used in various studies [187, 188]. The importance of including controls should be emphasised, as it affects the interpretation of a phylogenetic tree. Ideally the control samples should be obtained from a similar geographic or epidemiological population as the study population [189]. In partner studies that rely on sequencing based endpoints such as linked HIV transmissions, a multipronged approached can be used to define a transmission event, including a combination of sequencing regions and techniques, analysis methods and expert adjudicators, in lieu of a widely accepted definition of a transmission event [190].

It should be noted that transmission clusters are only inferred from phylogeny and cannot describe directionality. There are significant limitations to interpretation of directionality in the absence of robust epidemiological and behavioural data. A number of factors can influence the interpretation of directionality within a transmission cluster. For example, the timing of sampling from the transmission event, potential for in host evolution of virus once infected, the amount and similarity of background sequences, capacity to detect HIV variants, superinfections and recombination events [191]. With the availability of next generation sequencing, it was hoped that increasing the number of genetic variants sequenced could improve the quality of phylogenetic trees and provide more insights around directionality. However a recent study showed that even next generation sequencing methods are inconsistent in predicting directionality of transmission events[192].

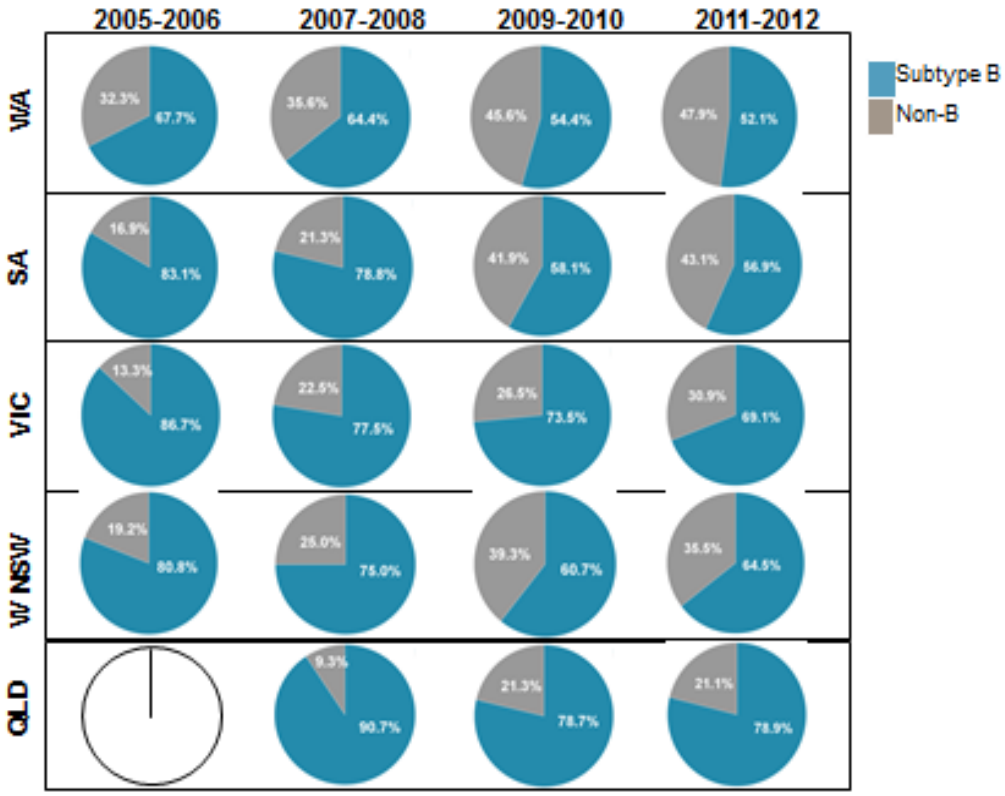
Nevertheless, there is still a role for phylogenetic studies into HIV transmission, particularly when supplemented with robust epidemiological data. For example in a study that supplemented phylogeny with partner notification data, 41% of partner notifications were not a part of a transmission cluster, and 59% of transmission clusters were not identified by partner notification, suggesting that phylogeny can provide greater insights than behavioural data alone [193].

1.3.3 Subtype diversity

Historically, HIV-1 subtypes and circulating recombinant forms (CRFs) have been associated with specific parts of the world and with particular modes of transmission [194]. However, the distribution of subtypes and CRFs is becoming more heterogeneous globally, related to population mobility, diversity of sexual contacts through travel and migration, and the impact of antiretroviral therapy [195, 196].

The HIV-1 epidemic in Australia is changing from predominantly subtype B infections in the Australian-born men who have sex with men (MSM) population to an increasing number of non-B infections within the heterosexual population, initially in those born overseas but increasingly with transmission to Australian-born persons [197, 198]. In a recent Australian study, just over one quarter of all infections were identified as non-B subtypes, which is slightly lower than observed in recent studies in North America, [199] and Europe, [200] [86]. The proportion of non-B infections has increased in all jurisdictions over time as shown in Figure 1.2; Western Australia had the highest proportion at baseline with 32.3% in 2005-2006, increasing to 48% in 2011-2012. South Australia had the steepest increase, from 16.9% during 2005-2006, to 43.1% in 2011-2012, including the emergence of a number of complex variants [86, 197]. A summary of studies examining subtype diversity in Australia is shown in Table 1.7.

Figure 1.2 HIV-1 subtype distribution over time



Abbreviations: WA Western Australia, SA South Australia, VIC Victoria, W NSW Western New South Wales, QLD Queensland.

Reproduced with permission from: Pinto *et al*. HIV-1 subtype diversity, transmitted drug resistance and phylogenetics in Australia. Future Virology. 2018; 13:8. 2018. Copyright © FUTURE MEDICINE LTD.

Table1.7 Summary of HIV-1 Subtype studies in Australia

Author (Ref)	Year of Publication	Title	Years sampled	Geographical area	Number of Sequences included	Region sequenced	B Subtype (%)
Chibo[149]	2012	Increasing diversity of Human Immunodeficiency Virus type 1 subtypes circulating in Australia.	2005-2010	Victoria	1056	*pol*	78
Hawke[147]	2013	HIV non-B subtype distribution: emerging trends and risk factors for imported and local infections newly diagnosed in South Australia.	2000-2010	South Australia	418	*pol*	76
Herring[201]	2003	Segregation of human immunodeficiency virus type 1 subtypes by risk factor in Australia	1993-2002	New South Wales	141	*env*	74
Castley[202]	2015	Longitudinal Trends in Western Australian HIV-1 Sequence Diversity and Viral Transmission Networks and Their Influence on Clinical Parameters	2000-2014	Western Australia	1021	*pol*	61
Castley[86]	2017	A national study of the molecular epidemiology of HIV-1 in Australia	2005-2012	National	4873	*pol*	75

Historically, people living in Australia infected with B or non-B virus subtypes represent highly distinct populations, with subtype B virus predominantly acquired in Australia by MSM [197]. Non-B strains were predominantly acquired overseas by people born overseas or people traveling to areas of high HIV prevalence. The most prevalent non-B strains in Australia are subtypes C and CRF01_AE, reflecting emerging patterns of migration from Africa, and overseas travel, especially to and from Asia [149]. Asian destinations make up seven of the top ten destinations for Australians travelling overseas. Non-B subtypes in Australia tend to be associated with heterosexual transmission and overseas acquisition [147, 201].

While CRF01_AE transmissions may relate to risk behaviours during travel to or migration from high-prevalence areas, there appears to be an increasing tendency for transmission to occur within Australia, and within groups not previously the target of relevant and focused HIV prevention campaigns [197]. It is also clear that over time the historical segregation between clades in terms of geography and risk group is becoming less distinct, with non-B infections occurring in the MSM population and subtype B infections occurring in overseas-born MSM males and Australian-born females, reflecting a growing global genetic variability [203-205]. There has also been an emergence of unique recombinant forms (URFs) acquired within Australia [197].

There is evidence that subtypes and CRFs differ clinically and phenotypically, by way of replication fitness, rate of disease progression, co-receptor utilization, transmission route, transmissibility, accuracy of current diagnostic assays, and response to therapy [198, 206, 207]. The accurate determination of antiretroviral drug resistance mutations in light of HIV-1 subtype-associated polymorphisms also requires careful consideration in the context of increasing non-B subtype diversity [208]. Subtype D infection, intersubtype recombinants and CXCR4 tropism [209] associated with CRF01_AE [210] have all been associated with rapid disease progression, [194], [207, 211], while a recent study found very fast disease progression for the newly identified 02_AG/A3 recombinant compared with subtype A virus, [212] which has been associated with slower disease progression [207]. Continued research into links between subtype and disease progression is needed to ensure currently recommended treatments continue to remain efficacious across all subtypes. Research into emerging subtypes can also have implications for molecular diagnostics, optimal first line therapies and future vaccine trials [207].

Reporting and interpretation of surveillance data can be problematic. Reporting of newly acquired infections does not necessarily mirror actual rates in the wider

community because HIV diagnoses represent only the subgroup of people who have been tested and had an HIV-positive result. These are people who have relatively easy access to health services and feel confident to access those services [213]. In Australia and elsewhere, immigrants, visa holders, and refugees face greater barriers when accessing health services for screening and treatment of HIV, arising from stigma, financial restrictions, limited support systems and English skills, and residence concern [213, 214]. Refugees in particular may be hard to reach because of traumatic life experiences before arrival in Australia. Programs that supply optimal ART to temporary residents unable to access government subsidized treatment have the potential to significantly reduce onward transmission, by an estimated 75% in an Australian study [214]. These types of concerns have prompted the United Nations to recognize migrants as one of the groups most vulnerable to HIV, and overseas-born people now comprise one-third of HIV notifications in Australia [215]. Problems with access to testing and the steady influx of new arrivals from low and middle income countries with high HIV prevalence are likely to lead to an underestimate of HIV infections in these populations, a possible increase in local transmission of non-B subtypes, and poor treatment adherence that could lead to drug resistance [198, 215].

The impact of the increasing number of non-B infections in the Australian population on prevention efforts and treatment outcomes is as yet unclear. Including molecular epidemiology data into current national surveillance is crucial for the development of appropriately targeted subtype-specific prevention and treatment options for populations most at risk [216]. Accurate identification of subtypes, CRFs and URFs globally and the linking of these with epidemiological data will assist greatly in understanding the course of the global HIV epidemic [197].

1.3.4 Phylogenetic studies of HIV in Australia

The HIV epidemic in Australia was previously regarded as predominantly affecting young MSM. State-based analyses show a shift from this perception and have highlighted differences in epidemics between regions within Australia. In Western Australia, the emergence of heterosexual transmission along with a wide age range at the time of infection was reported during a time of rapid socioeconomic change [202]. Other states have used phylogenetic analyses to investigate hypotheses generated from modelling data relating to onward transmission, suggesting that 19% of new diagnoses were acquired from those undergoing HIV seroconversion [82]. These analyses suggested that 30% of cases in Victoria were attributable to seroconverters, but these findings have not been replicated in other states [79]. Local phylogenetic

studies are required to inform locally relevant public health interventions, as there may be significant differences in the epidemic between states in Australia.

The most comprehensive retrospective study on molecular epidemiology of HIV-1 in Australia was performed from 2005 to 2012, with the aim of supporting rational, evidence based approaches to prevent, treat and monitor HIV-1 infection within Australia [86]. This study has demonstrated differences in transmission networks between B and non-B subtypes. As expected, more pairs and large networks were identified within male-only groups in the B subtype analysis, with higher proportions of male-and-female clusters in the C subtype analysis than in the CRF01_AE analysis. There was the observation of one large sub-epidemic B subtype cluster within the analysis [86].

Finally, a recent report from UNAIDS proposed an ambitious 2020 target of 90% of people living with HIV knowing their status, 90% of people diagnosed with HIV on treatment, and 90% of people on treatment with suppressed HIV-1 RNA. This target has recently been achieved in NSW, through enhanced surveillance methods and improved data collection [217]. Achieving this target Australia wide can be facilitated by knowledge gained through enhanced molecular surveillance of HIV-1, understanding changing subtype diversity, transmission networks and use of phylogenetic techniques to estimate population size.

1.3.5 Role of PHI in transmission of HIV

There are a wide range of estimates in the literature regarding the role of primary HIV in sustaining the epidemic. Estimates have ranged from as low as 2% to as high as 75%, illustrating the inconsistency between study results [218]. Reasons for this variation depend on the methods and models used to generate these estimates, and a number of variables including duration of infectivity, proportion of population diagnosed and sampled, and availability of behavioral and sexual network data. The proportion of infections that occur from people in very early infection can also be estimated through the use of molecular epidemiology techniques [219].

It has been hypothesised that acute infection contributes significantly toward onward transmission. The Rakai study was first to illustrate this through a population based cohort that showed a very high rate of transmission per coital act [220]. This has also been shown in a neighbouring state in Australia, with up to one third of new infections attributed to seroconverters [79]. However there is uncertainty about timing of transmission events and the accurate classification of acute versus chronic infections

[221]. The difficulty with identifying acute infections clinically may lead to errors in assigning cluster membership, and an over estimate of the contribution of acute infection in onward transmission. Also as the period of high viraemia is transient, the transmission event may occur at a date distant from this high viraemia, casting further doubts on the contribution of PHI to onward transmission [12, 78].

In the setting where PHI has been identified as a significant contributor to onward transmission targeted public health campaigns have resulted in reduction in acute infections, which may contribute to a reduction in overall incidence [222]. It is also important to identify mixed stage clusters as this can have different implications for public health messages. For example where acute infections are the main driver, the emphasis may be on testing with rollout out of rapid testing and self-testing facilities to at risk populations. If chronic infections are identified as main drivers for onward transmission, then the emphasis will need to be on increasing engagement with care, access to treatment and viral load suppression [88].

1.3.6 Legal and ethical issues surrounding HIV phylogenetic research

Phylogenetic research poses unique ethical considerations relating to privacy, confidentiality and data security. A number of considerations such as community consultation needed to be addressed before proceeding with research in this area [223].

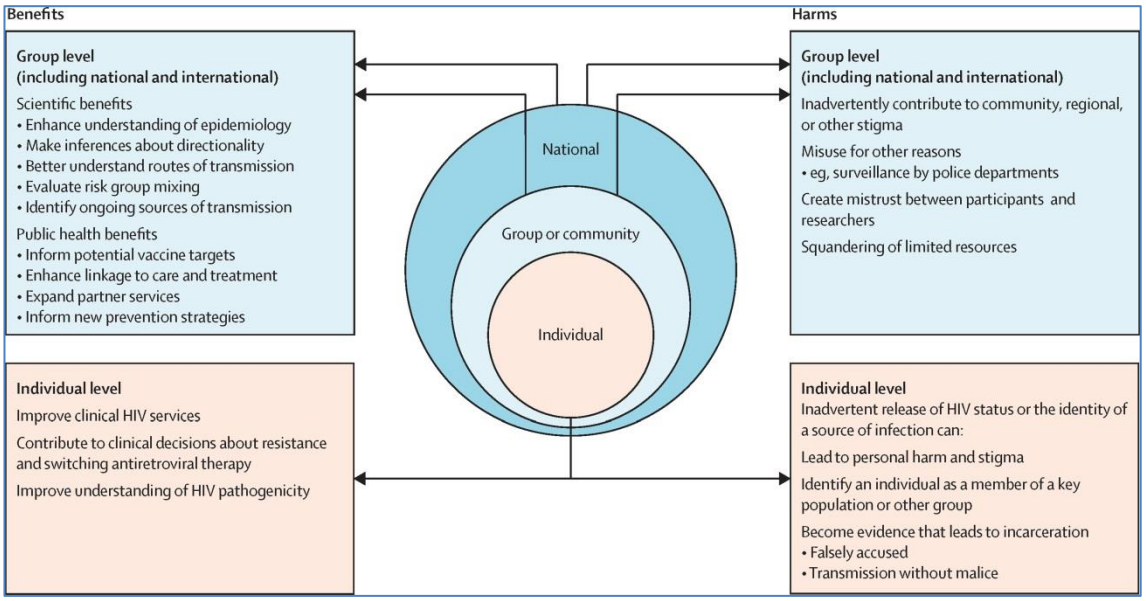
Australia's legal landscape impacts upon the development of research in HIV molecular epidemiology as there is potential for accessing and using the HIV genotype of an individual for evidence in court proceedings. Even though there are several limitations in the use of phylogenetic analyses such as inability to infer the direction of individual transmission events, inability to definitively prove a linked transmission due to third parties that have not been sampled, and limited community control samples from which to build phylogenetic trees, this type of evidence has been admitted and examined in the legal proceedings of three HIV transmission cases [224, 225]. This potential for forensic applications has constrained the ability to use routinely collected genotypic data for research and surveillance purposes.

In NSW at the time of commencement of this thesis, it was an offence under the *Public Health Act 2010* for a person with HIV not to disclose this prior to sexual intercourse. This has subsequently been amended on 18th October 2017 such that a person with HIV must take reasonable precautions against spreading the disease [226].

Nationally there is also potential for criminal law offences relating to intentional or reckless HIV transmission. In Australia there have been 38 criminal prosecutions for transmission of HIV in Australia [224, 227]. Five of those prosecutions have been in NSW, four resulting in convictions related to HIV transmission. For this and other privacy-related reasons there are serious concerns about the potential implications for use of this type of data if any personal identifiers remain attached to the combined dataset.

Research which links HIV sequence data with HIV register data containing clinical and demographic information is challenging due to confidentiality and privacy considerations. Although HIV notifications in Australia (with the exception of two jurisdictions) are currently provided in a de-identified format containing only a two by two name code and birthdate, additional precautions need to be undertaken to ensure the privacy and confidentiality of people living with HIV. For example, uncommon HIV subtypes that may have origins from particular geographical regions require cautious analysis, due to potential policy implications for particular cultural and migrant groups. A risk benefits analysis such as one outlined below should be undertaken to outline the potential of the research at the individual and community level [223]. An ideal research framework should involve engaged and sustainable partnerships with between researchers, clinicians, service providers, community and government.

Figure 1.3 Potential benefits and harms associated with HIV phylogenetic analysis



Reproduced with permission from: Coltart *et al* [223]. Ethical considerations in global HIV phylogenetic research. The Lancet HIV. 2018. Nov; 5(11):e656-e666. Copyright © Elsevier.

1.4 Thesis rationale and objectives

Primary HIV infection is a unique time to study several aspects of the virus, and its effects. It also affords a critical time point for interventions that can impact immunological and virological aspects of disease, as well as impact at a broader level on transmission and drug resistance. Although the benefits of early treatment are clear from the START study, the question of "how early" as applied to the natural history of HIV infection has not been answered. Chapter 3 investigates the factors associated with mortality in a primary HIV infection cohort.

Australia has moved towards a "test and treat" paradigm with rapid treatment initiation being promoted for both individual and public health benefit. In this context, the surveillance for transmitted drug resistance is urgently needed to inform first line treatment guidelines, and identify emerging drug resistance mutations in New South Wales, Australia's most populous state. Knowledge about common mutations can inform guidelines on preferred treatment regimens, and the appropriateness of regimens used in prevention and prophylaxis strategies. Chapter 4 describes the first analysis of state-wide HIV resistance mutations, and is essential to inform future treatment guidelines and public health interventions. Chapter 5 builds on this analysis using a dataset linked to HIV notifications to understand the demographic factors associated with transmitted drug resistance.

The landscape of HIV in Australia is changing in the context of expanded testing, treatment and prevention strategies. Concurrently, the molecular epidemiology of HIV is also shifting with greater subtype diversity and previously unrecognized transmission networks in different geographical settings. There is also an ongoing role for phylogenetic research using HIV genotype linked to HIV notification data. The third objective of this thesis is to understand the role of PHI in onward transmission of HIV, so that public health campaigns and resources may be better targeted to the highest risk groups. Chapter 6 describes a molecular epidemiological study using linked data to evaluate the factors associated with cluster membership in NSW.

CHAPTER 2

DATA LINKAGE METHODOLOGY

The analyses contained within three chapters of this thesis relied upon two separate data linkages to routinely collected health datasets which enhanced the quality of the research datasets. Chapter 3 used data linkage to identify additional deaths that may have missed from loss to follow up in a primary infection cohort that spanned three decades. Chapter 4 describes the establishment of a state-wide HIV sequence database to analyse prevalent drug resistance mutations. Chapters 5 and 6 describe further analysis of this dataset that is enhanced with data linkage to HIV notifications to combine clinical, epidemiological and phylogenetic data. This chapter provides an overview of data linkage methodologies and results of data linkages that were used for analyses that are described in detail in Chapters 3, 5 and 6.

2.1 Data linkage for Chapter 3

To determine the long term health outcomes of an established cohort of primary HIV infection, data linkage was required to capture additional endpoints that may have been missed from loss to follow up. In addition to mortality outcome that is described in Chapter 3, the primary infection cohort was linked to several other routinely collected health databases to determine validity of linkage.

2.1.1 The Primary Combined united database

The Primary HIV combined united database (PCUD) is an amalgamation of the early Sydney AIDS prospective cohort (SAPS) and the Phaedra/Core01 database. It is the largest database (608 participants from NSW and 54 from Melbourne) of patients with Primary HIV Infection in Australia. It records data on people from HIV seroconversion with the earliest information gained from results in 1983.

PCUD was established from a cohort of patients with PHI from several randomised clinical trials and prospective observational cohorts. Participants were enrolled through primary health clinics and hospitals in New South Wales and Victoria from 1984 to 2009, with active follow up until 31 December 2013. As is standard practice for trials of this sort, a name code comprising first two letters of the surname and first two letters of the given name is used instead of the patient's full name. PHI and estimated date of infection (EDI) was defined using strict laboratory criteria, details of these trials and eligibility criteria are published elsewhere: SAPS [228], VIRAX [30, 229], PULSE [29, 32], QUEST [35], AGOURON [34] CORE01 [230].

Ethics

Written informed consent was obtained for original study participants, which included the use of data for cross checking with state, national and Commonwealth databases. In addition, a waiver of consent was obtained specifically for data linkage. A waiver of consent was requested as it was impracticable to obtain consent for the following reasons:

1. Our cohort dates back to 1983 in the pre-HAART era, and unfortunately a large proportion of patients have died since entry into the study. As one of our outcomes measures is death, it would not be suitable to exclude these patients.
2. Similarly, due to significant length of follow up of 30 years, a large proportion may have moved since entry to the study, and it would be almost impossible to locate them.
3. Additionaly, it was feared that there would be further threats to privacy and in particular stigma associated with HIV/AIDS, by locating and contacting these patients for consent.
4. Given that the follow up period extends to over 30 years, there was also a risk of inflicting harm by contacting patients via public means where no relationship currently exists.

Ethics approval was granted by the NSW Population and Health Services Research Ethics Committee (NPHSREC: HREC/13/CIPHS/60).

Definitions

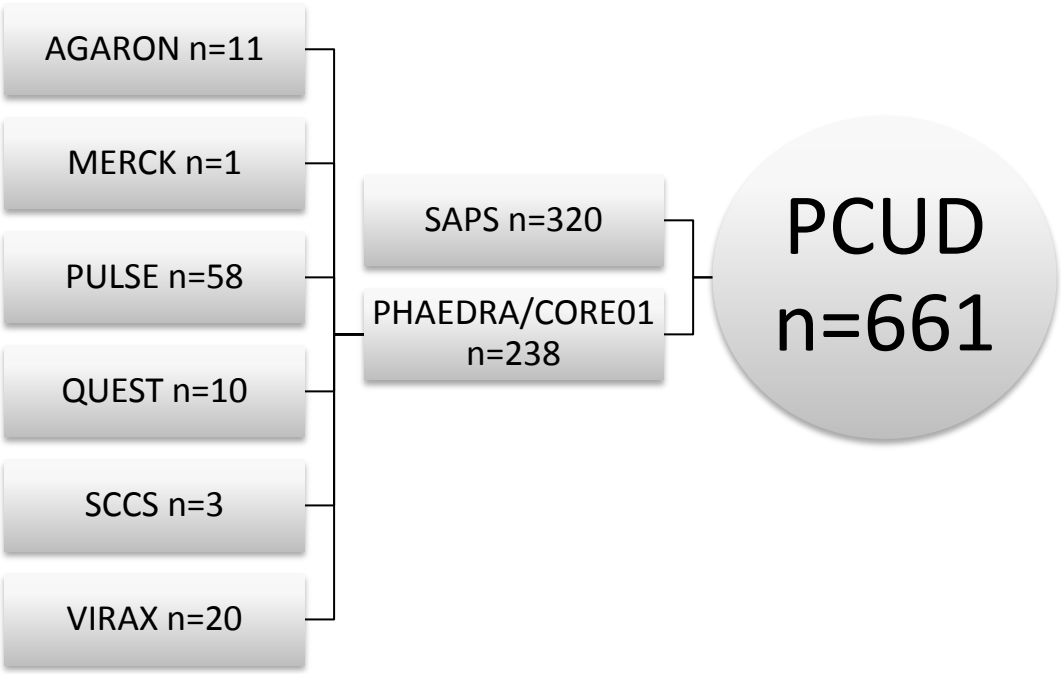
Acute infection was defined in study protocols as either:

i) signs and symptoms of acute retroviral symptoms with the presence of positive p24 antigen/HIV-1 proviral DNA/HIV-1 RNA or
ii) Indeterminate or evolving Western blot with positive p24 antigen/HIV-1 proviral DNA/HIV-1 RNA.

Early stage infection was defined as positive ELISA and Western blot with a negative documented serological test within 6 months.

Data collection was via clinical trials case record forms, centralised into the primary combined united database (PCUD) as shown in Figure 2.1. Consent for usage of data for future studies was obtained from original study protocols. Quality checks were performed including an accuracy check of 10% of records, completeness of data and manual review of records with duplicate name codes.

Figure 2.1 Data sources for PCUD database



2.1.2 Administrative datasets

The following descriptions have been provided in the Data Linkage Report by the Centre for Health record linkage [231].

Project: Long term health outcomes in people with HIV followed since seroconversion. Primary HIV infection linkage study. Date: 09 October 2014

- **NSW Admitted Patient Data Collection**
  The NSW Admitted Patient Data Collection (APDC) includes records for all hospital separations (discharges, transfers and deaths) from all NSW public and private hospitals and day procedure centres. APDC records include a range of demographic data items (e.g. date of birth, residential address, language spoken at home and country of birth), administrative items (e.g. admission and separation dates) and coded information (e.g. reason for admission, significant co-morbidities and complications and procedures performed during the admission).

- **NSW Emergency Department Data Collection**
  The Emergency Department Data Collection (EDDC) is maintained by the Demand and Performance Evaluation Branch of the NSW Health Department and provides information about presentations to the Emergency Departments of public hospitals in NSW. The data items included are demographic information, primary diagnosis and other clinical information.

- **NSW Central Cancer Registry**

  The Central Cancer Registry (CCR) maintains records of all cases of cancer diagnosed in NSW residents. The CCR records include a range of demographic data items (e.g. date of birth, residential address), staging information, year of diagnosis, and plus coded information (e.g. reason for death, ICD10 code, Morphology code etc.).

- **Notifiable Conditions Information Management System (NCIMS)**

  The Notifiable Conditions Information Management System (NCIMS) is a statutory population based surveillance system, to manage data and public health action on a number of the Scheduled Medical Conditions notifiable under the Public Health Act 1991. This is the source of information on notifiable infectious diseases in NSW.

- **HIV Administrative Dataset**

  Human Immunodeficiency Virus (HIV) infection notifications from the NCIMS dataset were obtained from NSW Health. To protect patient confidentiality, a name code comprising first two letters of the surname and first two letters of the given name is used instead of the patient's full name. This was consistent with the name code identifiers collected in the PCUD database. Other demographic fields provided included sex, date of birth and postcode.

- **NSW Registry of Births, Deaths and Marriages**

  The Registrar of the NSW Registry of Births, Deaths and Marriages (RBDM) is required to register all births and deaths in NSW. When a person dies, the Medical Certificate of Cause of Death is forwarded to the Registry which transcribes the information onto a computer database in un-coded format. A death registration number is assigned to each death. The Registry uses the details from the birth registration record to produce a NSW Birth Certificate.

- **ABS Mortality Data**

  The RBDM forwards the identifiable un-coded text information from the Medical Certificate of Cause of Death to the Australian Bureau of Statistics (ABS), which codes all the data. Cause of death is coded according to the International Statistical Classification of Diseases and Related Problems (ICD-9-CM or ICD-10-AM). Once the mortality data for a given year is complete and 'clean', ABS removes personal identifiers (name and street address) from the dataset. The death registration number is retained.

Subsets and external datasets

The NCIMS data custodian provided a subset of the NCIMS which matched the conditions: Hepatitis A, B and C, syphilis, gonorrhoea, tuberculosis, HIV and AIDS.

The PCUD custodian provided data with the following identifiers:

- 2-character codes for given name and surname
- Date of birth
- Gender
- Country of birth

The NCIMS custodian provided the HIV administrative data for individuals diagnosed with HIV in NSW with the following identifiers:

- 2-character codes for given name and surname
- Date of birth
- Gender
- Country of birth
- Post code
- Data Linkage

2.1.3 Data linkage

To capture any missed deaths due to loss of follow up, data linkage was performed with several health datasets including the NSW Registry of Births, Deaths and Marriages (RBDM) and Australian Bureau of Statistics (ABS) Mortality Data by the NSW Centre for Health Record Linkage (CHeReL) [232, 233]. ABS mortality database only extended until 2007 to minimise ascertainment bias, the RBDM was also used to capture deaths that may have occurred until census date of 31 Dec 2013. Identifiers used for linkage were two by two name codes, date of birth and sex.

The following has been provided from the Data Linkage Report by the Centre for Health record linkage [231].

Master Linkage Key

Identifying information such as name, address, date of birth and gender for each dataset is included in the Master Linkage Key (MLK). No health/content data are used in this process.

The MLK is being constructed by the Centre for Health Record Linkage (CHeReL)[231] using probabilistic record linkage methods and *ChoiceMaker* software[234].

ChoiceMaker uses 'blocking' and 'scoring' to identify definite and possible matches. During blocking, *ChoiceMaker* searches the target datasets for records which are possible matches to each other. There are two types of blocking. The exact blocking algorithm requires records to have the same set of valid fields and the same values for these fields. The automated blocking algorithm builds a set of conditions that are used to find as many as possible records that potentially match each other.

Scoring employs a combination of a probabilistic decision, which is computed using a machine learning technique, and absolute rules, which include upper and lower probability cut-offs, to determine the final decision as to whether each potential match denotes or possibly denotes the same person. Upper and lower probability cut-offs of 0.75 and 0.25 were used for the linkage. Record groups with probabilities in between the cut-offs were subject to clerical review.

Linkage of PCUD and HIV administrative data to APDC, EDDC, CCR, NCIMS records from the CHeReL MLK.

The APDC, EDDC, CCR, RBDM Deaths and NCIMS records were extracted from the Master Linkage Key (version 2014_08) for the periods shown in Table 2.1.

The PCUD and HIV Administrative records were linked to the MLK extract using deterministic passes, matching date of birth, country of birth, postcode (where available), sex and first two letters of given name and surname.

This deterministic linkage of the PCUD data to the MLK extract was carried out using a two-step procedure:

I. First pass: exact match on first 2 characters of given name, first 2 characters of surname, sex, date of birth and country of birth.
II. Second pass: exact match on first 2 characters of given name, first 2 characters of surname, sex and date of birth (excluding the records that were matched in first pass).

This deterministic linkage of the HIV data to the MLK extract was carried out using a three-step procedure:

I. First pass: exact match on first 2 characters of given name, first 2 characters of surname, sex, date of birth, country of birth and post code.
II. Second pass: exact match on first 2 characters of given name, first 2 characters of surname, sex, date of birth, and postcode (excluding the records that were matched in first pass).

III. Third pass: exact match on first 2 characters of given name, first 2 characters of surname and date of birth (excluding the records that were matched in first pass and second pass).

Deterministic linkage of RBDM death records and ABS Mortality records

The RBDM death records were deterministically linked to the ABS mortality records. This linkage was carried out using a five-step procedure:

I. First pass: year of registration, encrypted registration number and exact date of death.
II. Second pass: year of registration, encrypted registration number and either:
    a. one day difference in date of death or same year of death or
    b. Date of birth.
III. Third pass: year of registration, date of death, sex, postcode and date of birth
IV. Fourth pass: year of registration, date of death, sex, date of birth.
V. Fifth pass: date of death, sex and date of birth.

Once the linkage was finalised, the CHeReL created a Project Person Number (PPN) for each Person ID and assigned the PPN to all records in the linked datasets. The PPN and the record ID or patients ID from each source database were returned to the respective data custodians. Each data custodian will send to the study investigators a dataset comprising the approved information from the source database plus the PPNs. The investigators can merge the datasets using the PPN.

2.1.4 Results

The following has been provided from the Data Linkage Report by the Centre for Health record linkage [231].

Table 2.1 shows the number of records from each data source.

Table 2.2 summarises the outcome of linking the PCUD and HIV data to the APDC, EDDC, CCR and NCIMS MLK extract.

The linkage completed by CHeReL was checked for false positive linkages. A random sample of 1000 Person IDs from the linkage was selected and reviewed. The results of the review are: False positive rate: 5/1,000 Person IDs (0.5%)

Table 2.1: Data sources and record types

Data Source	Description	Number
Primary HIV combined united database (PCUD)	Primary HIV Notifications Study Cohort	605 records
HIV Administrative dataset (HIV)	Individuals diagnosed with HIV who were resident in NSW: Notification date: 1 Jan 1998 to 31 Dec 2013	8,436 records
NSW Admitted Patient Data Collection (APDC)	Episodes of hospital care: Admission date: 1 Jul 2000 – 31 Dec 2013 Separation date up to 31 March 2014	32,110,945 records
NSW Emergency Department Data Collection (EDDC)	Emergency Department admissions: Admission date: 1 Jan 2005 – 31 Dec 2013	18,610,882 records
NSW Central Cancer Registry (CCR)	Cancer notifications: Notification date: 1 Jan 1994 – 31 Dec 2009	540,832 records
Registry for Births, Deaths & Marriages (RBDM)	Death Notifications: Death date: 1 Jan 1985 – 31 Dec 2013	1,319,715 records
Notifiable Conditions Information Management System (NCIMS)	Notifiable Conditions notifications: Notification date: 1 Jan 1993 – 31 Dec 2013 Subset for specified conditions (Hepatitis A, B, C, syphilis, gonorrhoea, TB, HIV, AIDS)	695,335 records 228,169 records
Australian Bureau of Statistics (ABS) mortality data	ABS Mortality Data in NSW: Death date: 1 Jan 1985 – Dec 2007	1,020,798 records

Table 2.2: Summary of records returned to Study Investigators

Data Source	Record type	Number
Primary HIV combined united database (PCUD)	PCUD records linked to other datasets	466 records (465 persons)
	PCUD records not linked to other datasets	139 records (139 persons)
	Total PCUD records	605 records (604 persons)
APDC	APDC records linked to the PCUD records Admission date: 1 Jul 2000 – 31 Dec 2013 Separation date up to 31 March 2014	2,358 records (349 persons) 57.78%
EDDC	EDDC records linked to the PCUD records Admission date: 1 Jan 2005 - 31 Dec 2013	1,186 records (291 persons) 48.18%
RBDM Deaths	RBDM death records linked to PCUD records Death date: 1 Jan 1985 – 31 Dec 2013	70 records* (69 persons) 11.42%
CCR	CCR records (excl. *in situ* & skin cancers) linked to the PCUD records Notification date: 1 Jan 1994 – 31 Dec 2009	52 records (47 persons) 7.78%
NCIMS	NCIMS records linked to the PCUD records, specified conditions Notification date: 1 Jan 1993 – 31 Dec 2013	592 records (259 persons) 42.88%
HIV	HIV records linked to the PCUD records Notification date: 1 Jan 1998 – 31 Dec 2013	208 records (208 persons) 34.44%
ABS	ABS Mortality linked to PCUD records Mortality date: 1 Jan 1985 – 31 Dec 2007	57 records (57 persons) 9.44%
Total records to be returned to Study Investigators: Total Project Person Numbers (PPN):		5,128 records (604 persons)
*13 RBDM Deaths records (13 persons) had dates of death after December 2007, and would not have corresponding ABS Mortality records		

There were 661 participants in the PCUD cohort, 608 were NSW residents, 605 were eligible for data linkage (complete and unambiguous name codes) and 602 were included in the final analyses (Figure 2.2). Reasons for exclusions were incomplete data (baseline dates of infection), incorrect linkages (discrepancies in gender, duplicate records, discrepancies in death dates and hospital admissions, records that linked to multiple conflicting datasets), as illustrated in in Table 2.3.

Figure 2.2 Flow chart of included persons

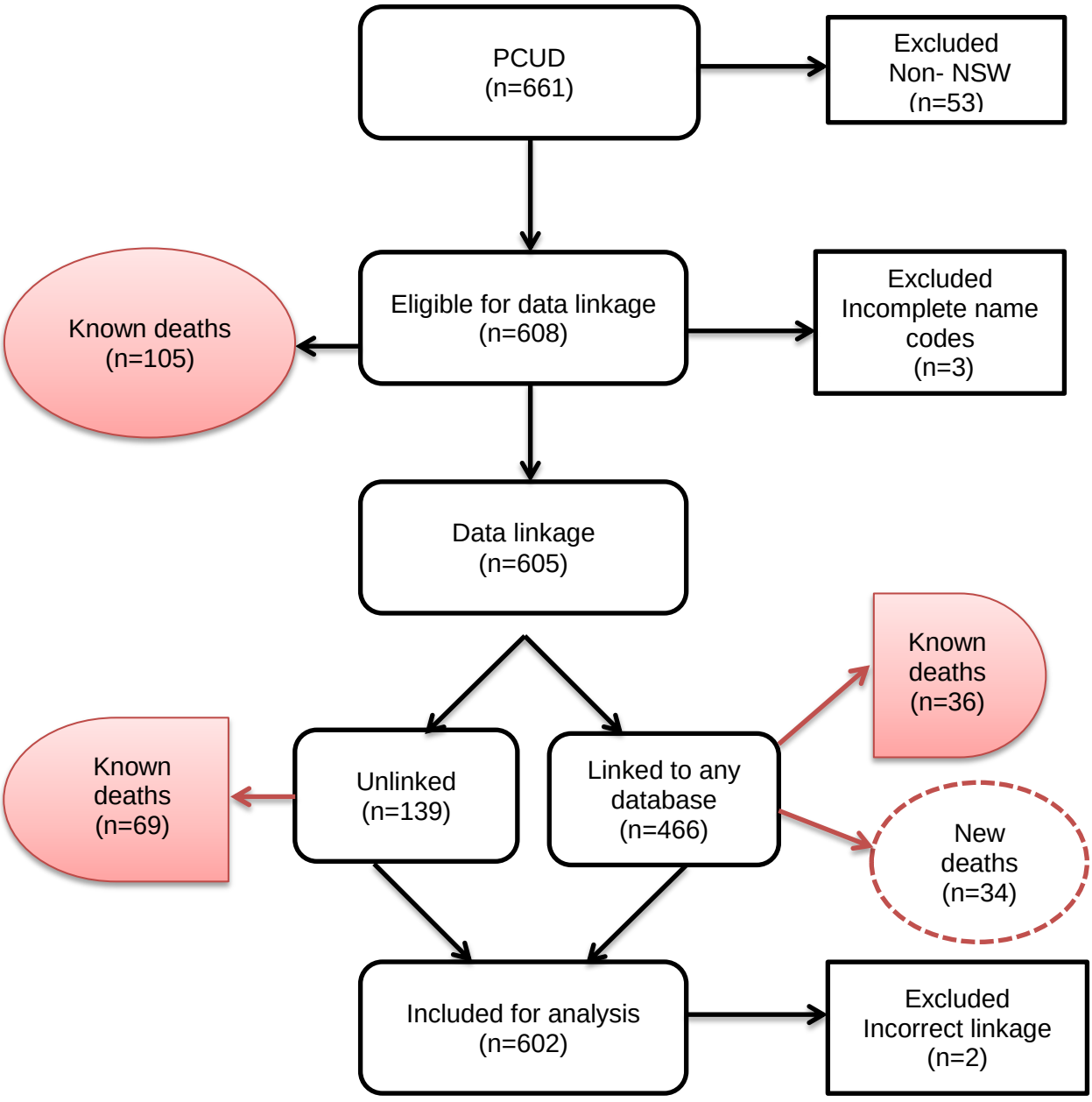
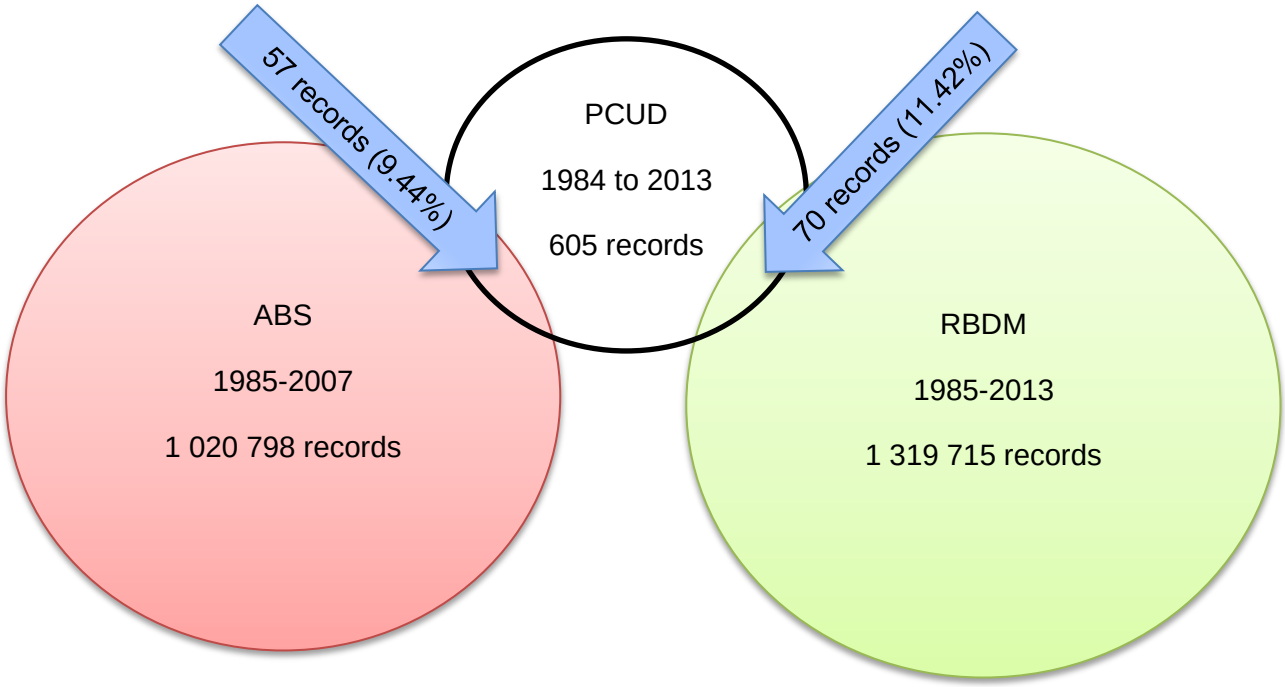


Table 2.3 Examples of incorrect linkages

Example	Date of Death	Date of admission	Comments
Example 1	2-Nov-88	1-Aug-11	wrong gender
Example 2	10-Mar-93	1-Apr-09	no information
Example 3	13-Dec-95	1-Jun-13	no HIV ICD codes listed
Example 4	1-May-06	1-May-13	only 1 HIV ICD listed among multiple admissions
Example 5	22-Jun-92	1-Aug-06	no HIV ICD codes listed
Example 6	1-Aug-90	1-May-13	no information

Figure 2.3 Summary of data linkage to deaths datasets



Sensitivity analysis

The proportion of PCUD records that linked to any dataset was 77% (466/605 records). For the purpose of this analysis, the proportion that linked to deaths databases was 11.42% for RBDM and 9.44% for ABS as shown in Figure 2.3. A sensitivity analysis for deaths was 34% (36 linked /105 known deaths). Known deaths were already identified from the PCUD cohort prior to data linkage.

139 of 605 individuals did not link to any database. Of these, 69 were known deaths (identified via PCUD database). Potential reasons for this discrepancy were investigated further, as any major errors in linkage could affect research outcomes [235]. To ensure the two by two name code was correctly submitted to CHeReL, original records were reviewed and checked, with 4 found to contain minor discrepancies in identifiers that may have cause incorrect linkage. On further review of records, it was found that three deaths took place interstate, whereas the RBDM only captures deaths within NSW. The incremental increased yield of linking to the national deaths registry was thought not to be sufficiently high enough to pursue further.

Overall, the proportion of records that linked to any health database was 77% (465/604) and was comparable to other studies (71%) that used name code based linkage of HIV positive individuals [236], which was considered sufficient to proceed with analysis.

2.2 Data linkage for Chapters 5 and 6

To enhance the analysis of the state-wide HIV sequence database (described in Chapter 4) data linkage was performed to the HIV notifications database.

The following has been provided from the Data Linkage Report by the Centre for Health record linkage [231].

Project: The HIV prevention revolution: Using molecular epidemiology of HIV transmission in NSW to inform the Public Health Response to HIV prevention.

2.2.1 The NSW HIV sequence database

- **GART Laboratory Database**

  Data collected by three reference laboratories in NSW (NSW State Reference Laboratory for HIV/AIDS & Molecular Diagnostics (STV), RPA and PathWest), which perform all genotypic antiretroviral resistance tests (GART). The following identifiers provided:

  - 2 letter first name code
  - 2 letter surname code
  - Gender
  - Postcode
  - Date of birth
  - Country of Birth* only supplied by PathWest and available on 283 (19%) of the 1,469 records

2.2.2 HIV notifications dataset

In NSW when a new HIV diagnosis is made, the treating clinician is required to complete a HIV notification form. This contains patient identifiers, laboratory information, CD4 and viral load, notification information, risk exposure and patient management (see Appendix Figure S2.1).

- **NSW Notifiable Conditions Information Management System – HIV data (NCIMS)**

  The Notifiable Conditions Information Management System (NCIMS) is a statutory population based surveillance system, to manage data and public health action on a number of the Scheduled Medical Conditions notifiable under the Public Health Act 1991. This is the source of information on notifiable infectious diseases in NSW.

HIV data was provided by NCIMS data custodian for individuals diagnosed with HIV in NSW from 1998 with following identifiers.

- 2-letter first name code
- 2-letter surname code
- Gender
- Date of birth
- Post code
- Country of birth

Ethics

A waiver of informed consent was sought for data linkage between these two datasets on the following basis:

1. The size of the population
2. The proportion of individuals who were likely to have moved or died since the health information was originally collected
3. The risk of creating additional threats to privacy by having to link information in order to locate and contact individuals
4. The risk of inflicting psychological, social or other harm by contacting individuals with particular conditions
5. The difficulty contacting individuals indirectly through public means such as advertisements or notices

Ethical approval was granted by the NSW Population & Health Services Research Ethics Committee (AU RED Reference: HREC/15/CIHS/38 & Cancer Institute NSW reference number: 2015/08/605) as well as the AIDS council of NSW (RERC Reference Number 2015/21).

2.2.3 Data linkage

Deterministic linkage of NCIMS and GART Laboratory records

The GART data and **NCIMS** records were linked using deterministic passes, matching date of birth, postcode (where available), sex and first two letters of given name and surname.

This deterministic linkage of the GART Labs data to the HIV **NCIMS** data was carried out using four deterministic passes:

I. First pass: Exact match on first 2 characters of given name, first 2 characters of surname, sex, date of birth and postcode.
II. Second pass: Exact match on first 2 characters of given name, first 2 characters of surname, sex and date of birth (excluding the records that were matched in first pass).

III. Third pass: Exact match on first 2 characters of given name, first 2 characters of surname, date of birth and postcode (excluding the records that were matched in first pass and second pass).

IV. Fourth pass: Exact match on first 2 characters of given name, first 2 characters of surname and date of birth (excluding the records that were matched in first pass, second pass and third pass).

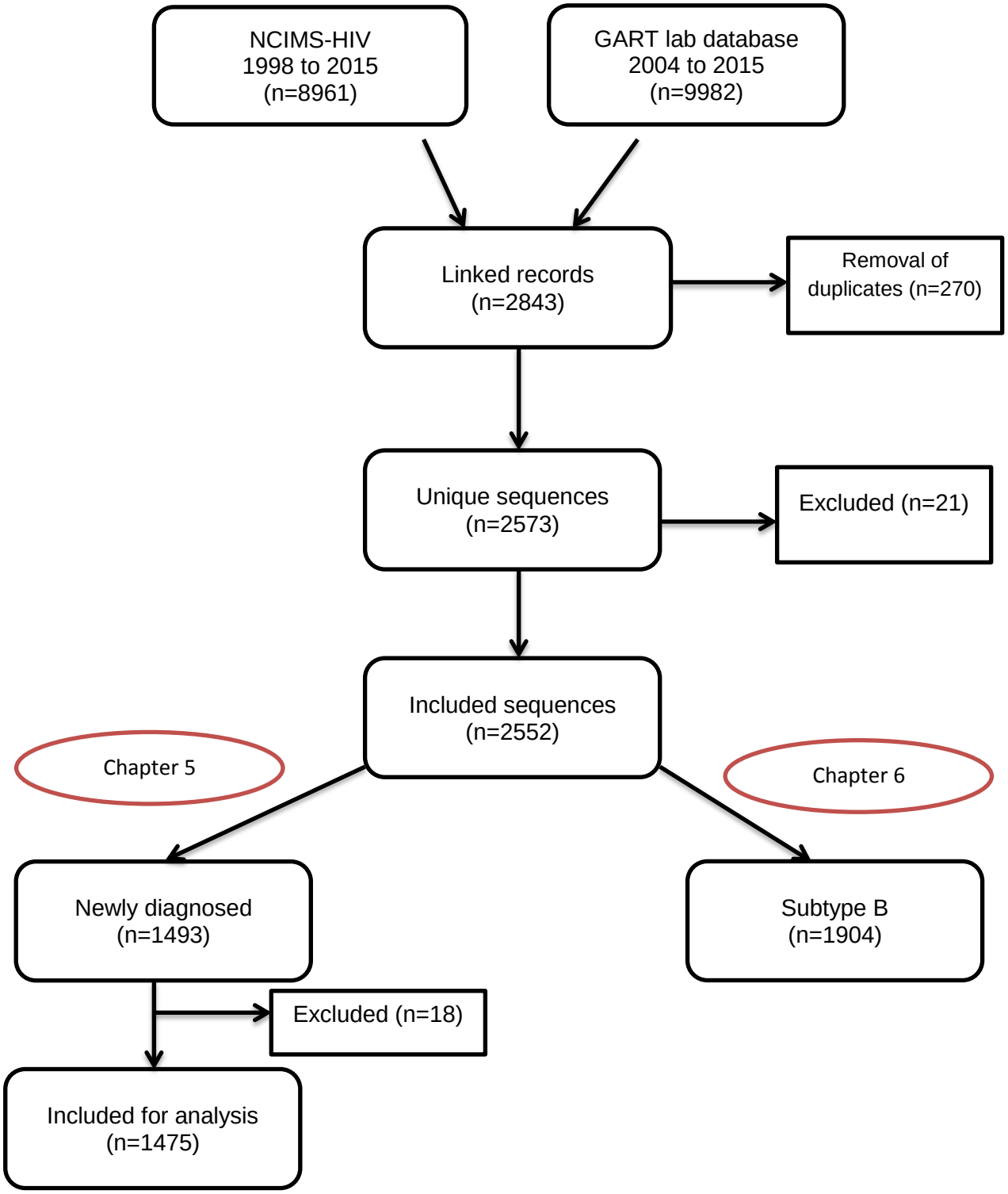
Once the linkage was finalised, the CHeReL created a Project Person Number (PPN) for each GART cohort Person ID and assigned the PPN to all records in the HIV dataset. Where a single GART person matched to multiple HIV records a single PPN was assigned to the linked records.

The PPN and the encrypted record number from each source database were returned to the respective data custodians. Each data custodian will send to the study investigators a dataset comprising the approved information from the source database plus the PPNs. The investigators can merge the datasets using the PPN.

2.2.4 Results

There were 2843 GART laboratory records that linked to NCIMS-HIV notifications database with a linkage rate of 28% (2843/9982), as summarised in Figure 2.4

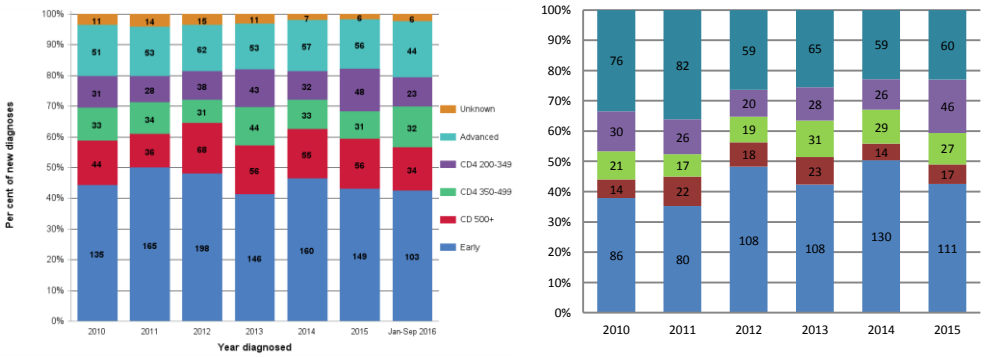
Figure 2.4 Flow chart of records and sequences included for analysis



Representativeness assessment

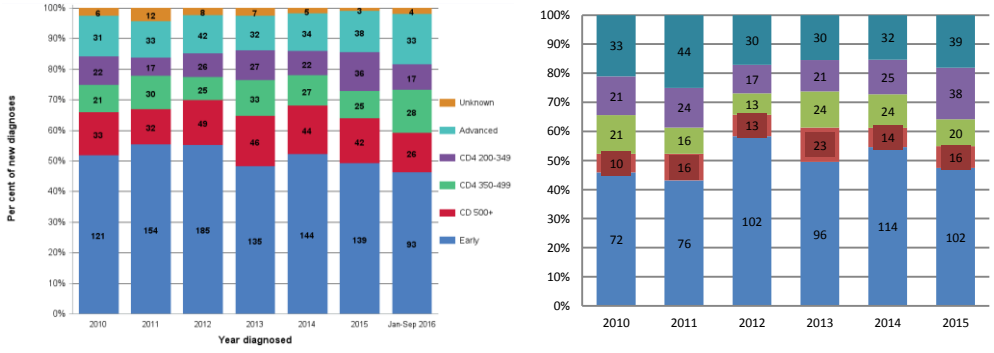
To assess the adequacy of records that linked, a comparative assessment of demographic characteristics was performed to determine whether the linked dataset was representative of the population diagnosed with HIV in NSW. Linked data was compared to the NSW HIV strategy data report released by the NSW Ministry of Health, and is summarised as Figures 2.5-2.9 below [237]. The demographic characteristics of the linked dataset were similar in proportion of subcategories to the population diagnosed with HIV in NSW in terms of stage of infection, testing history, age at diagnosis and risk exposure.

Figure 2.5 Stage of infection of newly diagnosed HIV



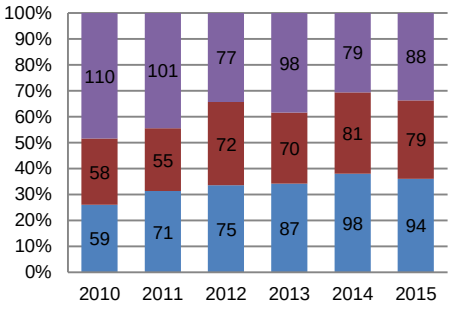
a) NSW **b) Linked dataset**

Figure 2.6 Stage of infection of newly diagnosed HIV in men who have sex with men



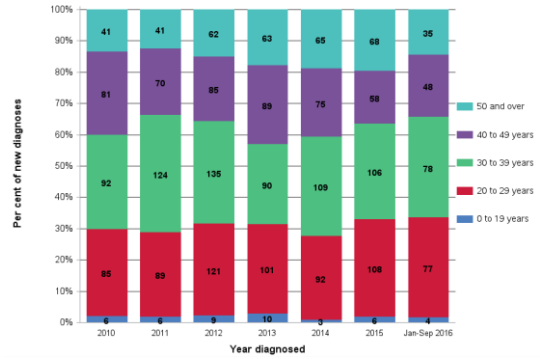
a) NSW **b) Linked dataset**

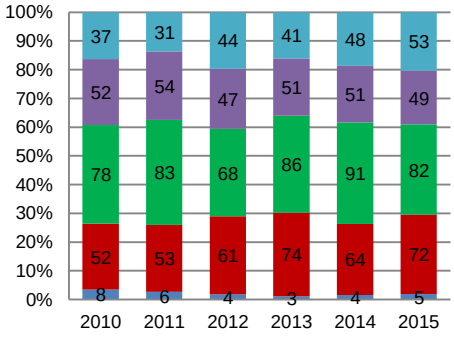
Figure 2.7 Testing history of newly diagnosed HIV in men who have sex with men



a) NSW **b) Linked dataset**

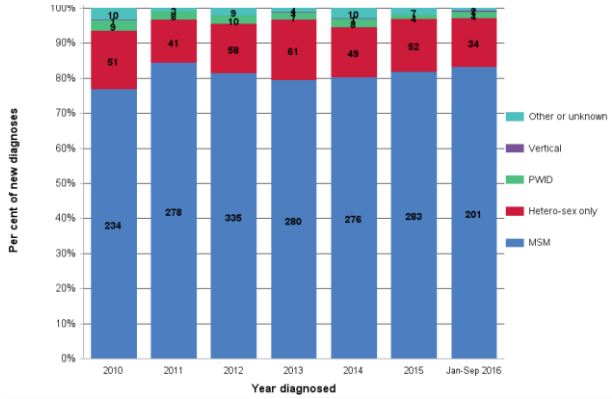
Figure 2.8 Age at diagnosis of newly diagnosed HIV

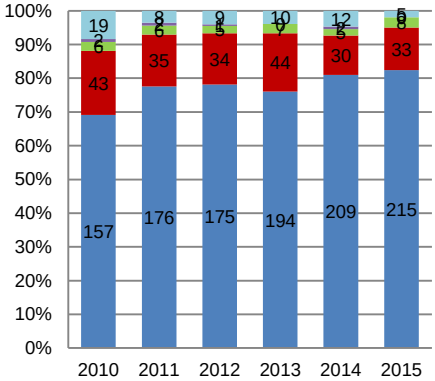




a) NSW **b) Linked dataset**

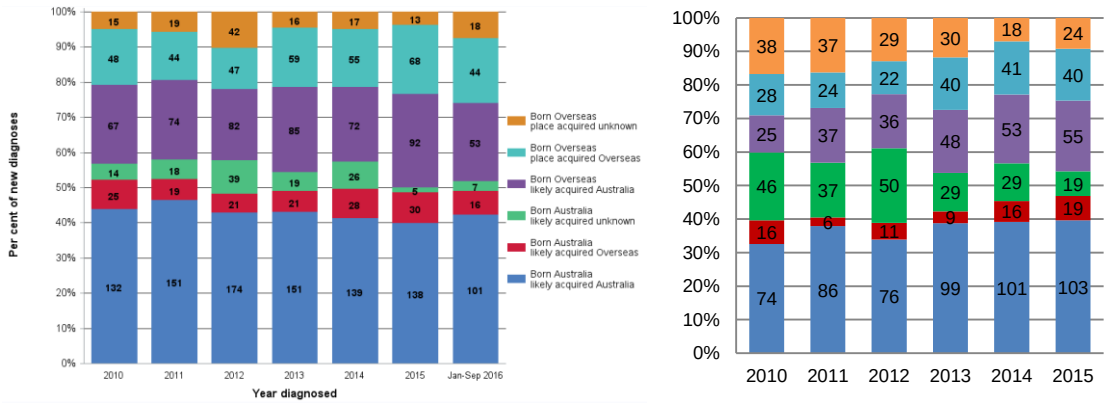
Figure 2.9 Risk exposure of newly diagnosed HIV





a) NSW **b) Linked dataset**

Figure 2.10 Place of birth and place of acquisition of newly diagnosed HIV



a) NSW **b) Linked dataset**

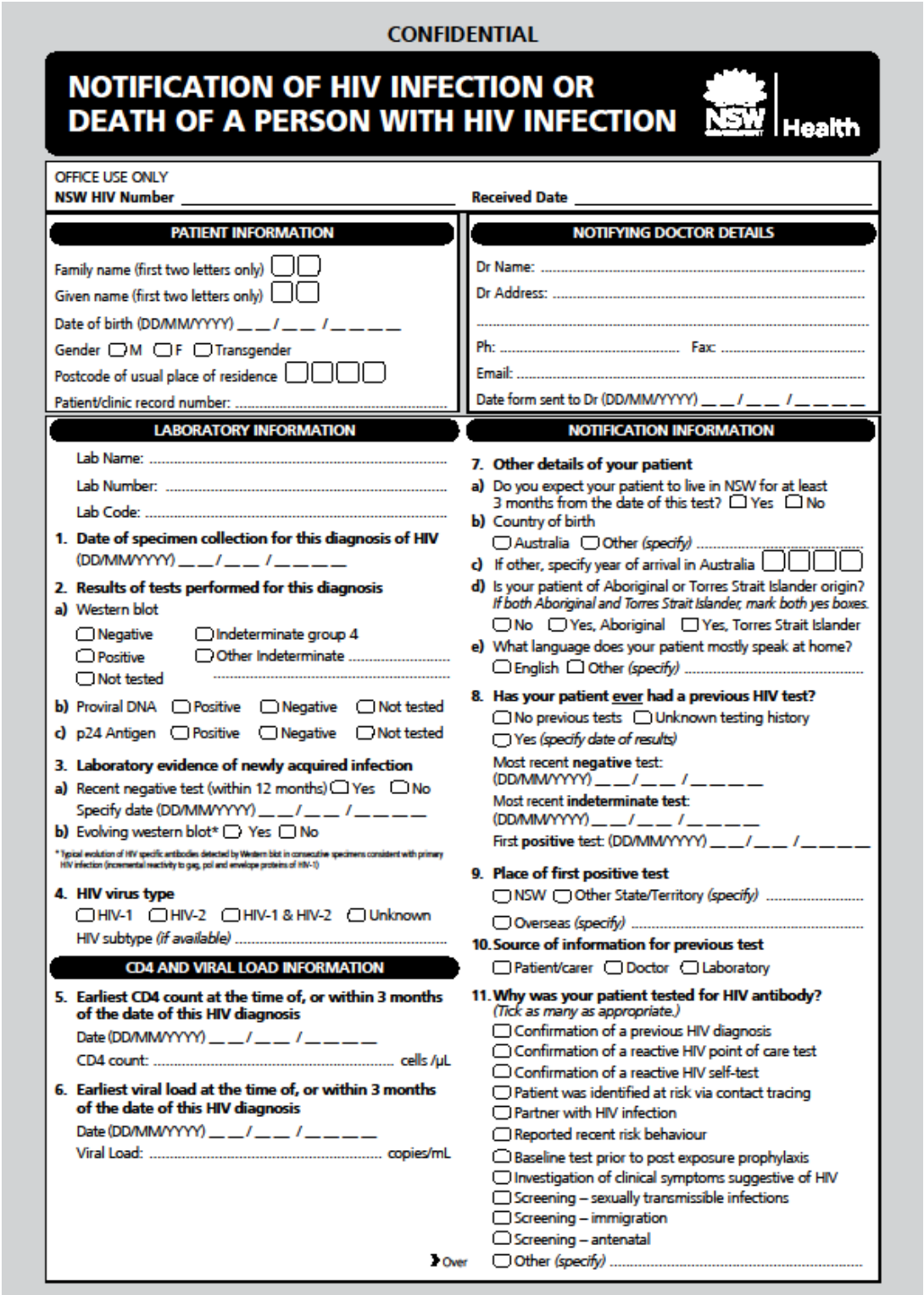
2.3 Conclusion

Data linkage was performed to enhance research datasets with routinely collected health data. The PCUD cohort was linked to identify any additional deaths that may not have been captured during the study period. Of the 605 PCUD participants, 466 linked to any database within NSW, and included 34 previously unidentified deaths. Incorrect linkages were identified and excluded, leaving a final linked dataset of 602 individuals for analysis in Chapter 3.

The state wide HIV laboratory sequence dataset was linked to HIV notifications dataset to enhance the final sequence with demographic and epidemiological data. Although only 2843 (28%) laboratory records linked to the HIV notifications dataset, the final dataset was representative of the population diagnosed with HIV in NSW, in terms of baseline demographics, stage of infection, testing history and risk exposure. This dataset was used for analysis in Chapters 5 and 6.

2.4 Appendix

Figure S2.1 NSW HIV Notification form



CONFIDENTIAL

NOTIFICATION OF HIV INFECTION OR DEATH OF A PERSON WITH HIV INFECTION

NSW Health

OFFICE USE ONLY

NSW HIV Number _____ **Received Date** _____

PATIENT INFORMATION

Family name (first two letters only) ☐☐

Given name (first two letters only) ☐☐

Date of birth (DD/MM/YYYY) __ __ / __ __ / __ __ __ __

Gender ☐ M ☐ F ☐ Transgender

Postcode of usual place of residence ☐☐☐☐

Patient/clinic record number: .....

NOTIFYING DOCTOR DETAILS

Dr Name: .....

Dr Address: .....

.....

Ph: ..... Fax: .....

Email: .....

Date form sent to Dr (DD/MM/YYYY) __ __ / __ __ / __ __ __ __

LABORATORY INFORMATION

Lab Name: .....

Lab Number: .....

Lab Code: .....

1. Date of specimen collection for this diagnosis of HIV
(DD/MM/YYYY) __ __ / __ __ / __ __ __ __

2. Results of tests performed for this diagnosis

a) Western blot

☐ Negative ☐ Indeterminate group 4

☐ Positive ☐ Other Indeterminate .....

☐ Not tested .....

b) Proviral DNA ☐ Positive ☐ Negative ☐ Not tested

c) p24 Antigen ☐ Positive ☐ Negative ☐ Not tested

3. Laboratory evidence of newly acquired infection

a) Recent negative test (within 12 months) ☐ Yes ☐ No
Specify date (DD/MM/YYYY) __ __ / __ __ / __ __ __ __

b) Evolving western blot* ☐ Yes ☐ No

* Typical evolution of HIV specific antibodies detected by Western blot in consecutive specimens consistent with primary HIV infection (incremental reactivity to gag, pol and envelope proteins of HIV-1)

4. HIV virus type

☐ HIV-1 ☐ HIV-2 ☐ HIV-1 & HIV-2 ☐ Unknown

HIV subtype *(if available)* .....

CD4 AND VIRAL LOAD INFORMATION

5. Earliest CD4 count at the time of, or within 3 months of the date of this HIV diagnosis

Date (DD/MM/YYYY) __ __ / __ __ / __ __ __ __

CD4 count: ..... cells /µL

6. Earliest viral load at the time of, or within 3 months of the date of this HIV diagnosis

Date (DD/MM/YYYY) __ __ / __ __ / __ __ __ __

Viral Load: ..... copies/mL

Over

NOTIFICATION INFORMATION

7. Other details of your patient

a) Do you expect your patient to live in NSW for at least 3 months from the date of this test? ☐ Yes ☐ No

b) Country of birth

☐ Australia ☐ Other *(specify)* .....

c) If other, specify year of arrival in Australia ☐☐☐☐

d) Is your patient of Aboriginal or Torres Strait Islander origin?
If both Aboriginal and Torres Strait Islander, mark both yes boxes.

☐ No ☐ Yes, Aboriginal ☐ Yes, Torres Strait Islander

e) What language does your patient mostly speak at home?

☐ English ☐ Other *(specify)* .....

8. Has your patient ever had a previous HIV test?

☐ No previous tests ☐ Unknown testing history

☐ Yes *(specify date of results)*

Most recent **negative** test:
(DD/MM/YYYY) __ __ / __ __ / __ __ __ __

Most recent **indeterminate test**:
(DD/MM/YYYY) __ __ / __ __ / __ __ __ __

First **positive** test: (DD/MM/YYYY) __ __ / __ __ / __ __ __ __

9. Place of first positive test

☐ NSW ☐ Other State/Territory *(specify)* .....

☐ Overseas *(specify)* .....

10. Source of information for previous test

☐ Patient/carer ☐ Doctor ☐ Laboratory

11. Why was your patient tested for HIV antibody?
(Tick as many as appropriate.)

☐ Confirmation of a previous HIV diagnosis
☐ Confirmation of a reactive HIV point of care test
☐ Confirmation of a reactive HIV self-test
☐ Patient was identified at risk via contact tracing
☐ Partner with HIV infection
☐ Reported recent risk behaviour
☐ Baseline test prior to post exposure prophylaxis
☐ Investigation of clinical symptoms suggestive of HIV
☐ Screening – sexually transmissible infections
☐ Screening – immigration
☐ Screening – antenatal
☐ Other *(specify)* .....

CHAPTER 3

EARLY TREATMENT OF PRIMARY HIV INFECTION IS ASSOCIATED WITH DECREASED MORTALITY

Data linkage

To determine the long-term health outcomes of an established cohort of primary HIV infection, data linkage was required to capture additional endpoints that may have been missed from loss to follow up. In addition to the mortality outcome that is described in this chapter, the primary infection cohort was linked to several other routinely collected health databases as described in Chapter 2.

Work from this chapter has been published as :

Pinto, A., et al., Early Treatment of Primary HIV Infection Is Associated with Decreased Mortality. AIDS Res Hum Retroviruses, 2018. 34(11). p.936-941.

3.1 Abstract

Background: The aim of this study was to understand factors associated with increased mortality in a cohort of primary HIV infection (PHI) in New South Wales over three decades.

Methods: Six hundred and two patients with PHI were enrolled from 1984 to 2009. Probabilistic data linkage was performed to NSW Registry of births deaths and marriages, and Australian Bureau of Statistics mortality database. Mortality was measured by crude death rate. The pre highly active antiretroviral therapy (pre-HAART) era was defined as prior to 1 January 1997. A Cox proportional hazard model was used to identify factors associated with death.

Results: One hundred and thirty-eight deaths occurred during 6223 person years follow up (PYFU). Overall crude death rate was 2.2 per 100 PY [95%CI, 1.9-2.6]; 3.6 [95%CI, 3.1-4.3] in pre-HAART era, and 0.20 [95%CI, 0.08-0.47] in post-HAART era. AIDS was the most frequent cause of death (52%, 72/138), all occurring in the pre-HAART era. Of non-AIDS deaths, the leading known cause was Non-AIDS cancer 8% (11/138) followed by suicide 4% (6/138). On multivariate analysis, estimated date of infection in pre-HAART era and time to commencement of ART greater than one year post diagnosis were more likely to be associated with death ($p<0.05$).

Conclusions: Mortality in PHI has decreased significantly in the post-HAART era. Non AIDS deaths due to malignancy and suicide are emerging as leading causes in this population in the post-HAART era. Time to starting ART greater than one year was associated with increased mortality.

3.2 Introduction

Primary HIV infection (PHI) is the earliest clinical stage following acquisition of HIV and may be characterised by a seroconversion illness[1]. Severity of symptoms during PHI can be associated with disease progression and death [15, 68]. It is a critical time for the establishment of the viral reservoir and host immune response. Cohort studies of PHI can provide insights into the natural history of HIV; however diagnosing patients at this early stage is challenging as seroconversion may be transient and have a nonspecific clinical presentation.

Some cohort studies of PHI have demonstrated the impact of early treatment in delaying disease progression [56, 238]. Other studies have only demonstrated short term effects that did not extend beyond the duration of treatment [31].

It should be noted there are limitations to recruiting cohorts of individuals identified with primary HIV infection. They may tend to be more representative of persons engaged in the health care system, seeking care for early symptoms that may otherwise be ignored in the general population. They may therefore be more biased towards high risk populations such as men who have sex with men who have increased awareness about symptoms of seroconversion illness, compared to other risk groups, for example. There may also be a temporal bias, in that there was heightened awareness of symptoms of PHI around the initial time of its first description, and this recognition may have waned over time [239]. Some studies have demonstrated higher incidence of PHI in the sexual health clinic setting compared to general HIV clinics [240].

Mortality rates were highest before the availability of HAART in 1997. In this pre-HAART era, there were no effective treatment options to prevent the progression to AIDS and death. The early HAART era from 1997-1999 involved treatment with mono or dual therapies that may have led to the development of resistance. Only in the late HAART era, were less toxic combination regimens widely available.

Large international cohorts have shown decreases in mortality (1.33-1.75 per 100 PY) and changing causes of death in the post-HAART era [241, 242]. Although life expectancy in Australia has increased for people living with HIV, mortality in the post-HAART era was still ten times greater than the general population [243, 244]. Mortality in PHI within the first five years of acquisition approximates the general population, but is increased as the duration of infection increases [70]. As the population diagnosed with PHI is also shifting, with increased community based testing, understanding of disease outcomes in this unique group will be more challenging to determine.

We evaluated mortality rates and causes of death within a NSW based Australian cohort of PHI over three decades spanning several eras of HIV treatment.

3.3 Methods

We established a cohort of patients with PHI from several clinical trials and prospective observational cohorts. Participants were enrolled through primary health clinics and hospitals in New South Wales and Victoria from 1984 to 2009, with active follow up until 31 December 2013. PHI and estimated date of infection (EDI) was defined using strict laboratory criteria. Details of these studies and eligibility criteria are published elsewhere: SAPS [228], VIRAX [30, 229], SPARTAC [31], PULSE [29, 32], QUEST [35], AGOURON [34] CORE01 [230]. All participants in clinical trials were asked to roll over into a long term follow up cohort (Core01) at the end of the trial specified follow-up. Acute infection was defined in study protocols as either: i) signs and symptoms of acute retroviral syndrome with the presence of positive p24 antigen/HIV-1 proviral DNA/HIV-1 RNA or ii) indeterminate or evolving Western blot with positive p24 antigen/HIV-1 proviral DNA/HIV-1 RNA. Early stage infection was defined as positive ELISA and Western blot with a negative documented serological test within 6 months. Data collection was via clinical studies case record forms, centralised into the primary combined united database (PCUD). Data collection included twice yearly CD4 T cell count(CD4), viral load, treatment regimen, treatment start and stop dates, AIDS diagnoses and death. Consent for usage of data for future studies was obtained from original study protocols.

The main outcome measures were death and cause of death. Deaths were categorised according to the causes of death in HIV (CoDe) protocol [245]; AIDS defining conditions assigned according to CDC [246]. Deaths were recorded via original study protocols, with both written and verbal confirmation from treating physicians. To capture any missed deaths due to loss of follow up, probabilistic data linkage was performed with the NSW Registry of Births, Deaths and Marriages (RBDM) and Australian Bureau of Statistics (ABS) Mortality Data by the NSW Centre for Health Record Linkage [232]. ABS mortality database only extended until 2007 hence to minimise ascertainment bias, the RBDM was also used to capture deaths that may have occurred until census date of 31 Dec 2013. Identifiers used for linkage were two by two name codes, date of birth and sex.

Crude death rate was calculated using the Poisson exact method with 95% confidence intervals (CI) reported. Where lower CI was zero, one sided 97.5% CI was used.

Person-years of follow up were estimated from date of infection (EDI) until death or last recorded date of follow up.

A Cox proportional hazard model was used to identify factors associated with death. Univariate analysis was performed on the following categorical variables: sex, time period of enrolment (1986-1996 and 1997-2006), stage of infection (Acute or Early as defined above), number of PHI symptoms, time to start treatment, first regimen; and continuous variables: age at PHI, CD4 count at PHI (expressed per 20 cells/µl), percent CD4 change per year (CD4 PCPY). Post-HAART era was defined as EDI after 1 January 1997. Number and duration of PHI symptoms were included as continuous variables. Covariates were included in a multivariate model if they had a p-value of <0.05 in the univariate analysis. The multivariate model was determined using a step wise approach with a two sided statistical significance ($p<0.05$). Kaplan-Meier method was used for survival analyses. Analyses were performed using STATA (version 14; StataCorp LP, College Station, TX, USA) and SAS (version 9.3; SAS Institute INC., Cary, NC, USA). Written informed consent was obtained for original study participants. A waiver of consent was obtained for data linkage. Ethics approval was granted by the NSW Population and Health Services Research Ethics Committee (NPHSREC: HREC/13/CIPHS/60).

3.4 Results

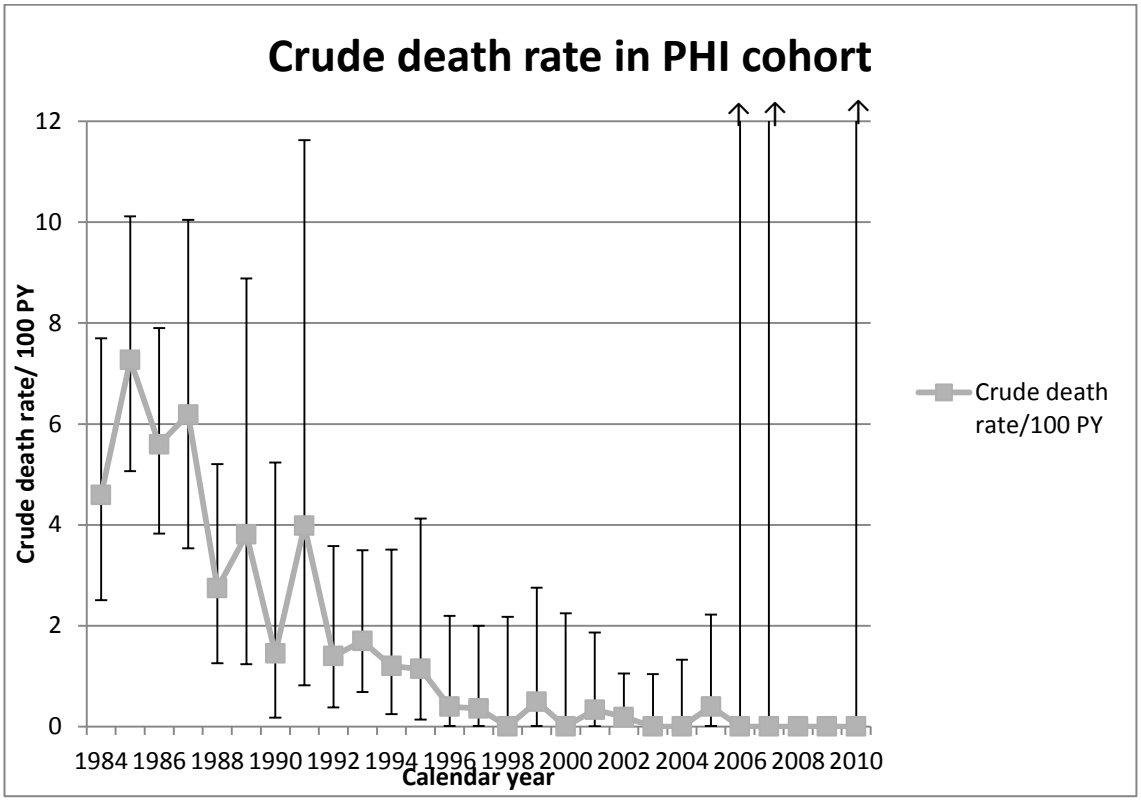
There were 667 participants in the PCUD cohort, 602 were NSW residents eligible for data linkage, and included in these analyses. There was a total 6222 PYFU (3667 in the pre-HAART and 2555 in post HAART period). The median age at PHI diagnosis was 33 years (Interquartile range (IQR) 28-39), with a median 9 (IQR 5-13) person-years of follow up (PYFU). The median baseline T cell CD4 count was 560 cells/µl (IQR 396-744), 99% of participants were male, and 81.7% received antiretroviral therapy at any time. Prior to linkage, there were 105 known deaths.

Mortality

In total, there were 138 deaths confirmed via study protocol or data linkage. Overall crude death rates were 2.2 per 100 person-years [95%CI, 1.9-2.6]. For those that were diagnosed with PHI in the pre-HAART era, mortality was 3.6 per 100 PY [95%CI, 3.1-4.3] compared with 0.20 per 100 PY [96%CI, 0.08-0.47] in the post-HAART era. After five years of follow up, there were 38 deaths, with 481 patients remaining at risk. The cumulative hazard was 0.07 [lower confidence interval (LCI)-upper CI (UCI) 0.0533-0.099]. At nine years follow up, there were 99 deaths, with 305 patients remaining at risk. The cumulative hazard was 0.22 [LCI-UCI 0.18-0.27]. There has been a decrease

in death rate over time, with only 5 deaths occurring in the post-HAART era group (Figure 3.1).

Figure 3.1 Crude death rates per 100 person years (PY)



95% confidence intervals (CI) are shown. Where lower CI was zero, one sided 97.5% CI was used with arrows indicating extension beyond axis.

On univariate analysis, those enrolled into a study cohort between the years of 1984-1998 had a significantly increased risk of death. Time to starting treatment beyond 1 year also had a significantly increased risk of death. Factors associated with reduced hazards for death include acquiring PHI in the post-HAART era, presence of PHI symptoms and starting a protease inhibitor containing regime compared with mono or dual NRTI/NNRTI based treatment as first regimen (Table 3.1). There was no significant difference in death rate based on gender ($p>0.20$), baseline CD4 count or rate of change of CD4 count per year.

Table 3.1 Cox proportional hazard model of risk of death in primary HIV infection studies in NSW

		UNIVARIATE ANALYSIS					MULTIVARIATE ANALYSIS				
Variable	n	Hazard ratio	p value	Lower CI	Upper CI	Overall p	Hazard ratio	p value	Lower CI	Upper CI	Overall p
Time period						<0.001					0.002
1986-1996	306										
1997-2006	296	0.044	<0.001	0.0181	0.109		0.192	0.002	0.0671	0.550	
Stage of infection						<0.001					0.04
ACUTE	230										
EARLY	149	1.9	0.07	0.94	3.9		1.26	0.5	0.609	2.59	
UNKNOWN	223	7.6	<0.001	4.4	13.4		1.99	0.03	1.08	3.65	
Number of symptoms	Freq.					<0.001					
No symptoms	149										
1-5 symptoms	220	0.47	<0.001	0.32	0.70						
6-10 symptoms	219	0.43	<0.001	0.28	0.65						
>10 symptoms	11	0.37	0.2	0.090	1.51						
missing	3	<0.001	1	<0.001	.						
Time to start treatment						<0.001					0.002
<= 1 year	281										
1 to 2 years	51	5.6	<0.001	2.8	11.1		2.9	0.003	1.42	5.9	
> 2 years	158	7.7	<0.001	4.5	13.3		2.6	0.001	1.48	4.7	
missing	112	4.2	<0.001	2.2	7.9		<0.001	<0.001	<0.001	<0.001	
First regimen						<0.001					0.001
Mono or dual	160										
Protease inhibitor [a]	227	0.088	<0.001	0.048	0.162		0.51	0.06	0.251	1.04	
Other [b]	105	0.127	<0.001	0.061	0.261		0.243	<0.001	0.116	0.51	
No treatment	35	<0.001	1	<0.001	.		<0.001	1	<0.001	.	
Missing	75	0.59	0.02	0.38	0.92		.	.	.	.	
Age at PHI	602	0.98	0.09	0.96	1.00	<0.20					
CD4 at PHI	454	1.01	0.4	0.99	1.02	0.4					
CD4 PCPY [c]	251	1.05	0.2	0.97	1.14	0.2					

Footnote: a) Protease inhibitor included as part of HAART; b) includes other combination regimens containing integrase inhibitors, non-nucleoside reverse transcriptase inhibitors

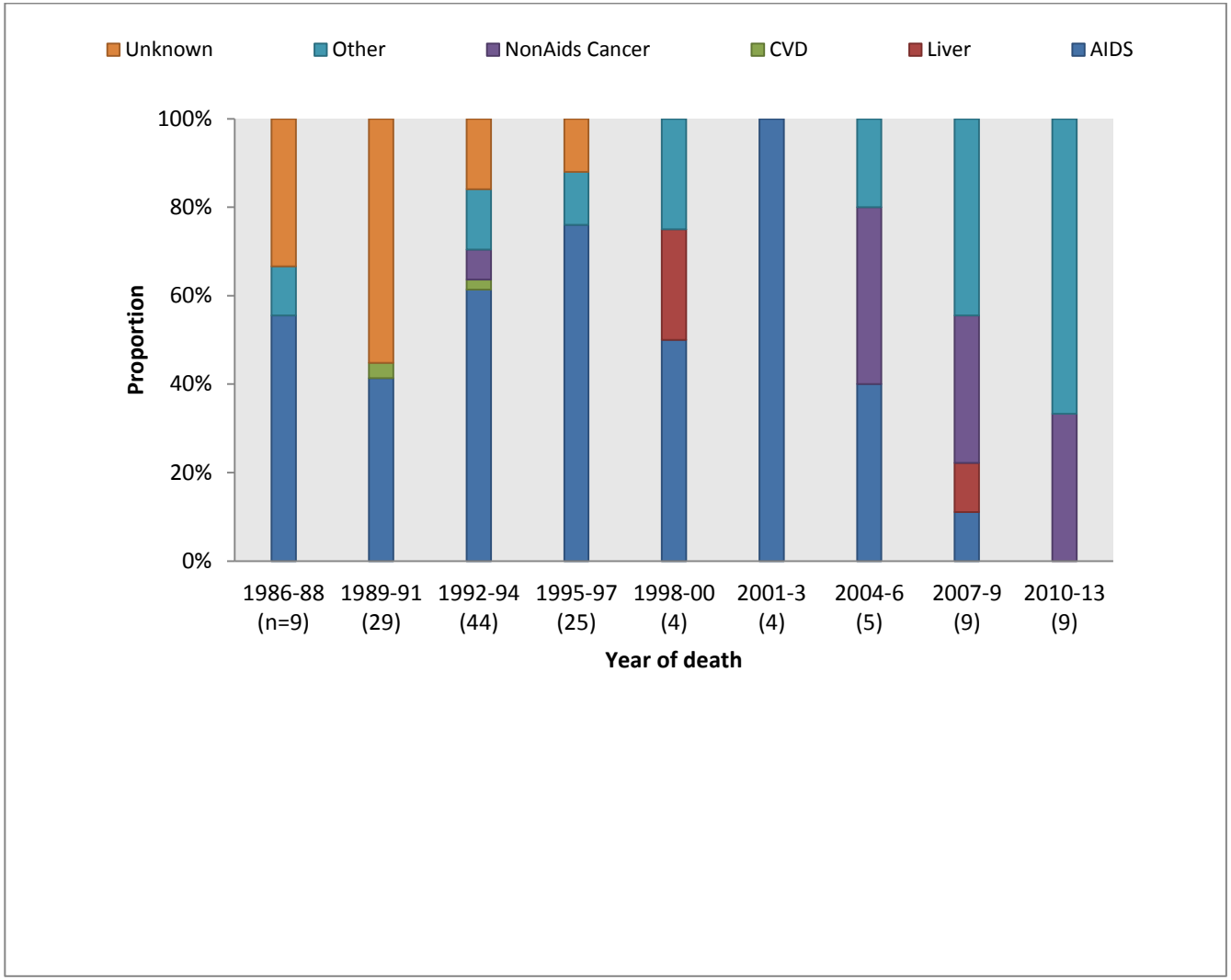
c) CD4 change expressed as percent per year (PCPY)

On multivariate analysis, acquiring PHI in the post-HAART era had a significantly decreased risk of death (hazard ratio (HR) 0.192, 95%CI (0.0671-0.550), p = 0.002). Stage of infection approached significance and was associated with increased risk of death for early or unknown compared to acute stage of infection (HR 1.26, 95%CI(0.609-2.59) and HR 1.99, 95%CI(1.08-3.65) respectively, overall p= 0.040). Time to starting anti-retroviral therapy beyond one year conferred an increased risk of death (HR 2.9, 95%CI (1.42-5.9), p=0.003), as did a monotherapy or dual nucleoside only containing first regimen.

Cause of death

The most frequent causes of death were AIDS (n=72, 52%), malignancy (n=11, 8%) and suicide (6, 4%). 30 deaths (22%) were classified as having an unknown cause. Of the remaining 18 deaths, causes included accident (n=4, 2.9%), substance abuse (n=3, 2.17%), ischaemic heart disease and digestive system disease (n=2, 1.45% each), unclassifiable, respiratory disease, renal failure, liver failure, infection, central nervous system (CNS) disease and chronic viral hepatitis (n=1, 0.72% each). In the pre-HAART era, the majority of deaths were AIDS related, while in the post-HAART era, the number of AIDS related deaths decreased, non AIDS cancer and other causes predominated (Figure 3.2).

Figure 3.2 Cause of death over time



3.5 Discussion

This is the first study to report long term mortality outcomes in an Australian cohort recruited at the time of primary HIV infection, and is unlikely to be replicated. Due to the challenging nature of recruiting at the time of PHI, together with the current “test and treat” landscape, establishing and maintaining PHI cohorts such as this with long term follow up would not be feasible. Overall, mortality in this cohort has decreased over three decades, in keeping with mortality rates in comparable observational cohorts [243]. The highest death rates were seen in those recruited prior to 1998 which coincides with the pre-HAART era and also in the SAPS study, which recruited almost exclusively during this time period.

In addition, time to starting antiretroviral treatment was an important predictor of death, with delays beyond one year resulting in a greater than five times risk of death compared to starting treatment within a year of acquiring HIV. In the setting of PHI, this is a notable finding, as it suggests that the impact of early treatment is related to more than just the CD4 T cell count, a surrogate of immune deficiency. This study suggests that a longer duration of untreated viral replication at the earliest stage of infection has an impact on all-cause mortality, regardless of immune function. Even though large international randomised trials now also conclusively recommend early treatment, these studies were conducted on participants with a median time of 1 year from diagnosis and compared starting immediately versus when the CD4 count fell below 350 cells/µl [83]. This data supports the current WHO recommendations of Test and Start, and presents an even more compelling argument for early treatment as it is from a cohort of primary HIV infection[84]. In addition, the advantages of early treatment during primary HIV infection, a time of peak viraemia, bolsters the public health argument of prevention of transmission, thus also aligning with the UNAIDS 90-90-90 targets for epidemic control[247].

In this cohort of PHI, recruited at the earliest time point in the natural history of infection, CD4 counts are relatively preserved, suggesting that treatment has other effects beyond simply preservation of immune function. Although long term longitudinal CD4 data was not complete in this cohort, analysis of CD4 change within the first year was performed, and not found to be significantly associated with death. Acute stage of infection was associated with decreased death compared to early stage, and may be related to a greater proportion of earlier treatment in this group (71% vs 46% started treatment within one year). This variable was confounded by a high proportion of unknown stage, which when investigated further was found to contain a greater

proportion of earlier years of infection, who delayed treatment initiation beyond two years and hence had a higher hazard ratio for death. To date, the long-term effects of treatment started during primary HIV have been conflicting, with some studies demonstrating immunological benefits, while others suggesting that toxicity of long term treatment outweighs the benefit of starting early [25, 26, 41-44, 57, 58]. This study supports current clinical and public health directives for earlier treatment in the primary infection cohort.

A number of factors were significantly associated with a decreased risk of death including the type of regimen (protease inhibitor containing compared nucleoside only), and the number and severity of symptoms during PHI. The later observation is in contrast to that reported in Vanhems *et al* in a similar but smaller cohort, which suggested a dose response relationship between severity of PHI infection and a faster rate of disease progression. The median duration of follow up in that study was only 3 years, and was performed in the pre-HAART era [15]. This study was much larger, and included a broader range of treatment eras, with three times longer median follow up. Reporting bias due could have resulted in a greater proportion without symptoms in the SAPS study compared to the Phaedra/Core01 study as SAPS relied on a self-administered questionnaire. In addition, the confounding may have contributed to this finding, as those with more symptoms started treatment earlier.

Cause of death in PHI has changed from predominantly AIDS related to non-AIDS related with malignancy and suicide as leading causes of death in the post-HAART era. This is consistent with findings in other large cohorts that show non-AIDS related malignancies and suicides as prominent causes of death [72, 77, 248], but this is the first to our knowledge to demonstrate this in the context of people identified at primary infection. A large proportion of causes of death were listed as unknown. This occurred exclusively in those with PHI from 1985-1991 and was due to these participants not linking to any deaths registry. Possible reasons for not linking may be due to acquiring PHI at a time of intense fear and social stigma around HIV and AIDS. As a result, incorrect identifiers (name code, DOB, postcode) may have been provided to researchers, which restricted the ability to link to official death registries. Alternate reasons could be that despite living in NSW, these participants may have been from interstate, and when faced with deteriorating health and no available treatment options at that time, they may have returned to their state of origin before death. When these unknown deaths were analysed, all but 1 of 30 had a diagnosis of AIDS prior to death,

making it probable that in an era of no effective treatment for HIV, that AIDS was the cause of death.

The availability of antiretroviral treatments over time has had a direct impact on survival in HIV. Initially regimens were limited to monotherapy or dual therapy. Delays in starting treatment in the pre-HAART era may have been due to the practice of delaying therapy until there was clear disease progression. However our cohort also included clinical trials where early antiretroviral regimens were included in the protocol. Univariate analysis was performed by study type, and this was not found to be significant in relation to mortality.

Another limitation of this study was the entire treatment history such as treatment interruptions was incomplete, and therefore only the first treatment regimen was included for this analysis. Viral load suppression at one year post treatment was analysed as a surrogate measure for treatment adherence, and not found to be associated with death. Treatment interruptions are known to be inferior to continuously administered HAART however interruptions after commencing first regimen were not analysed in this study [37].

Other considerations in interpretation of this retrospective cohort study are potential survivor bias from those that remained alive until the availability of combination antiretroviral therapy. This may account for the marked decreased mortality in the post-HAART era. Coinfection such as Hepatitis C was not analysed, as the serostatus for this variable within the cohort was incomplete. Other important behavioural factors related to mortality were not a part of this study, such as cigarette smoking and intravenous drug usage. As mentioned previously, PHI cohorts may recruit a more health literate population, as this condition is frequently under recognised. This may select out a more educated population with more engagement in their health care over time, resulting in potential biases with lower rates of morbidity and mortality. Within these limitations, the study findings are consistent with international cohort studies on HIV mortality and causes of death [71-75]. PHI provides a cohort which is more generalizable than other HIV observational cohorts as all participants are enrolled at the earliest time of HIV acquisition, allowing accurate measurement of time of infection.

In summary, death rates in PHI have decreased significantly over time. Delaying antiretroviral treatment beyond one year was associated with increased risk of death. AIDS related mortality was a leading cause of death in pre-HAART era, replaced more recently by malignancy and suicide as leading causes of death in this cohort.

Acknowledgements

The contribution of the many investigators of contributing studies of this cohort is gratefully acknowledged.

Funding

The funding provided by those listed below is gratefully acknowledged:

1. Personnel support: NHMRC postgraduate research scholarship (APP1074467) and an NHMRC Practitioner Fellowship.
2. CORE01 was supported by an NIH grant (UOI AI52403-04).
3. SAPS, Phaedra and this linkage study were supported though funding to the Kirby Institute from the Australian Federal Department of Health and an NHMRC Program Grant.

CHAPTER 4

EVOLUTION OF HIV-1 SURVEILLANCE DRUG RESISTANCE MUTATIONS FROM 2004 TO 2013 IN NEW SOUTH WALES, AUSTRALIA

4.1 Abstract

New South Wales (NSW) has the greatest burden of HIV in Australia, with 2012 and 2013 recording the highest rates of new diagnoses in over twenty years. Concurrently, there have been significant changes in antiretroviral treatments and testing paradigms. We compiled a state-wide resistance database to characterize changes in HIV-1 resistance mutations over a ten year period.

Genotypic antiretroviral resistance testing (GART) was performed on request at three reference laboratories using commercial and in house methods. 7629 HIV-1 polymerase sequences obtained from GART from 2004-2013 were retrospectively collated, reformatted, de-identified and analysed using Stanford HIVdb program 7.0 and the 2009 WHO Surveillance Drug Resistance Mutations (SDRM). Analyses were performed on subgroups of known treatment naïve, treatment experienced and seroconverters.

There has been a decrease in overall rates of prevalent drug resistance mutations from 57.8% in 2004 to 21% in 2013 ($p<0.0001$). Dual and triple class resistance mutations have decreased from 32.7% in 2004 to 5.8% in 2013 and 16.4% to 1.2% respectively. In treatment naïve individuals (n=450), the frequency of protease inhibitor (PI) mutations remains low at 2.7%. In treatment experienced, rates remained stable with 36.0%, 18.9%, 29.1% and 6.4% in overall, NNRTI, NRTI and PI mutations. The most common mutations in treatment experienced patients occurred at position M184V (NRTI), K103N (NNRTI); M46I (PI). In seroconverters, the overall rate of major mutations was 17.4% (49/281), with 1.4% PI, 5.7% NRTI and 11.7% NNRTI. Apparent decreases in prevalent SDRMs can be attributed to changes in GART testing indications over time.

4.2 Introduction

Australia is a high income country with an estimated 27150 (24630-30310) people living with HIV, 19097 of whom are men who have sex with men (MSM) [249, 250]. In NSW, there were 347 new diagnoses on average over the period 2009 to 2014, [251].

82% of those newly diagnosed reported MSM risk exposure [251]. Combination antiretroviral therapy has been universally available since 1997 via the publicly funded Federal Government pharmaceutical benefits scheme (PBS). Post-exposure prophylaxis is freely available and recommended in national guidelines for high risk exposures[145]. Although pre-exposure prophylaxis (PrEP) has not been approved for inclusion on the PBS, implementation trials commenced in 2016 in the most populous state, New South Wales (NSW). The state is in the midst of rolling out a multidimensional strategy to virtually eliminate HIV transmission in NSW by 2020 [252]. Rates of TDR in Australia are reported to be the highest world-wide [15.5% (12.8–18.3%)][111]; however, populations sampled for these studies may not be representative as were from single centres, and included data collected during the pre-Highly active antiretroviral therapy (HAART) era. Rates as high as 23.2% were reported in a NSW-based study from 1992-2001 but this was from a single centre in the pre-HAART era and included only MSM [139]. Other states in Australia have reported TDR rates of 16% in Victoria [141] and 21.9% in South Australia [253]. In some states, there has been an increase in the number of non-B subtypes, mainly acquired heterosexually [147, 254]. More recent clinical trial data from the Strategic Timing of Antiretroviral Treatment (START) study also suggests that Australia has the highest overall TDR rate in world (17.5%) however small sample size limits the interpretation[110]. Currently there is no national approach to surveillance of transmitted drug resistance in Australia. In the context of evolving treatment paradigms, and concurrent rollout of PrEP in NSW, the aim of this study was to describe prevalent HIV drug resistance mutations in NSW over the ten years prior to implementation of current HIV strategies.

4.3 Methods

Indications for resistance testing

In NSW, all genotypic antiretroviral resistance testing (GART) including RT and protease is currently performed at three reference laboratories (Lab 1, 2, 3). In the early years of this study, testing was restricted to treatment experienced patients with virological failure; however testing indications have since changed. Local and international guidelines since at least 2007 have recommended GART at diagnosis or prior to initiation of ART [255]. In 2011, testing was made universally available and up to two tests per year were funded through the Medicare benefits schedule (MBS) when performed at presentation, before antiretroviral therapy or at treatment failure, provided HIV load is greater than 1000 copies per mL plasma[256].

Ethical approvals

This low to negligible risk study was approved by the Human Research and Development Committee of St Vincent's Hospital with site specific approvals obtained for Westmead Hospital (Pathology West) and Royal Prince Alfred Hospital.

Data sources

HIV polymerase sequences were obtained from three reference laboratories. Additional data including date of birth (DoB), sex, date of infection, seroconverter status and ART treatment status was extracted from the electronic health record of individuals attending hospital based HIV clinics at laboratories 1 and 2. This information was not available for those that were referred from other centres such as outpatient sexual health clinics.

Data storage

All study information will be maintained in strictest confidence in accordance with the Kirby Institute Privacy and Confidentiality guidelines and the NHMRC National Statement on Ethical Conduct in Research Involving Humans (2007).
Once the linked data file without names and addresses is received by the researchers at the Kirby Institute, all study data will be transferred to a secure network protected by a firewall and individual files will also be password-protected. Only authorised staff will have access to study information, in either paper or electronic format. All results will be published in a form that will not allow individuals to be identified, that is, in tabular, aggregate form only; no individual results will be disclosed.

Inclusion criteria

All GART performed from 1st January 2004 until 31st December 2013 was included for analysis. A single sequence containing protease and reverse transcriptase genes was included for a single patient per calendar year.

Exclusion criteria

Duplicate sequences per patient within calendar year were excluded. Duplicate individuals were flagged by identical DoB (2006-2013) and duplicate sequences were identified by phylogenetic distance (2004-5) where DoB was not available. Poor quality sequences containing stop codons or frameshift mutations were excluded. *Integrase* mutations were not included, as integrase region sequencing was not routinely performed during the study timeframe.

Definitions

Seroconverter status was recorded between 2004-2013 at a single site (Lab 1) using the following definitions: i) laboratory confirmed HIV with documented negative serology within previous six months; ii) positive detuned BED ELISA; iii) <= 3 bands on western blot with positive p24 antigen OR <= 3 bands on western blot with detectable proviral HIV-1 DNA or <=3 bands on western blot with plasma HIV-1 RNA> 10 000 c/mL.

Treatment status could only be determined from 2009 onward. Treatment naïve status was assigned when there was no known previous antiretroviral regimen and included all seroconverter as well as chronic infections. Treatment experienced status was assigned when there was evidence of a current or previous antiretroviral regimen, regardless of the duration of infection.

HIV sequencing

Plasma samples were collected from whole blood collected into EDTA containing tubes and transported to reference laboratories. Plasma was separated and stored at between -65°C to -80°C and batched for HIV-1 RNA extraction. A minimum viral load of 2000 copies/ml is required.

Population based Sanger method sequencing of reverse transcriptase and protease was performed on purified amplicons. For laboratories 1 and 2, the Trugene® HIV-1 genotypic kit was used, where as an in-house assay was used by laboratory 3. Internal quality controls included a negative and positive HIV control included in each run. All laboratories participate in the NRL HIV genotyping quality assessment programme twice a year.

Analysis of sequencing data

Sequences were assembled and proofread with GeneObjects (Lab. 1) Seqscape® (Lab. 2) and Sequencher® (Lab. 3). A segment containing less than 100 bases was considered invalid; at least 6 valid segments are required for analysis. Electrophoretic signals were read and analysed and consensus sequences are read by two independent operators.

Sequence files were saved in FASTA format using laboratory generated accession number as the file name. Individual FASTA files were merged into a multiFASTA file using DNA baser®, and then converted to text format. All text file sequences were submitted to Stanford HIV Drug Resistance Database version 7.0 for quality analysis [136, 148]. Sequences containing stop codons, frameshift mutations were excluded. Resistance mutations were defined according to WHO 2009 Surveillance drug

resistance mutation (SDRM) criteria [107]. Subtype was determined for *reverse transcriptase* gene according to the Stanford HIV Drug Resistance Database v7.0[148].

Statistical analysis

The main outcome was the frequency of resistance mutations for *reverse transcriptase* and *protease* genes defined according to WHO SDRM list [107]. Baseline characteristics including age, sex and subtype were compared across sites. Frequency of SDRMs was determined overall and for non-nucleoside reverse transcription inhibitors (NNRTI), NRTI and protease inhibitor (PI) classes using population analysis with Calibrated Population Resistance tool version 6.0[257]. Stata 12.0 (StataCorp LP, TX, USA) was used for chi-square statistical analysis and for the trend of proportions (ptrend).

4.4 Results

Out of a total of 7639 GART performed in NSW from 2004-2013, 6901 (90%) contained unique sequences within a calendar year. 738 duplicates within a calendar year were excluded. Treatment history was available for 799 (12%) sequences from 2009-2013 only, as summarized in Table 4.1. A mean 690 (SD 76) GART were performed each year, with 550 in 2004 and a maximum of 788 in 2010.

Table 4.1 Number of genotypic antiretroviral resistance tests performed

Year	Site of testing			Treatment status		Newly diagnosed [a]	Total
	Lab 1	Lab 2	Lab 3	Naïve	Experienced		
2004	191	285	74				550
2005	211	352	46				609
2006	180	368	94			397	642
2007	304	371	78			386	753
2008	303	263	92			326	658
2009	343	348	68	121	62	339	759
2010	457	253	78	123	64	309	788
2011	469	202	71	83	53	331	742
2012	468	165	96	54	80	407	729
2013	416	175	80	69	90	354	671
Total	3,342	2,782	777	450	349	2849	6,901

a) Data from Table 2, page 27 of Kirby Institute Annual Surveillance Report 2016 [250]

Table 4.2 summarises characteristics of those that had demographics available, 5028 (90%) were male with a median age of 41 years (IQR 33-48).

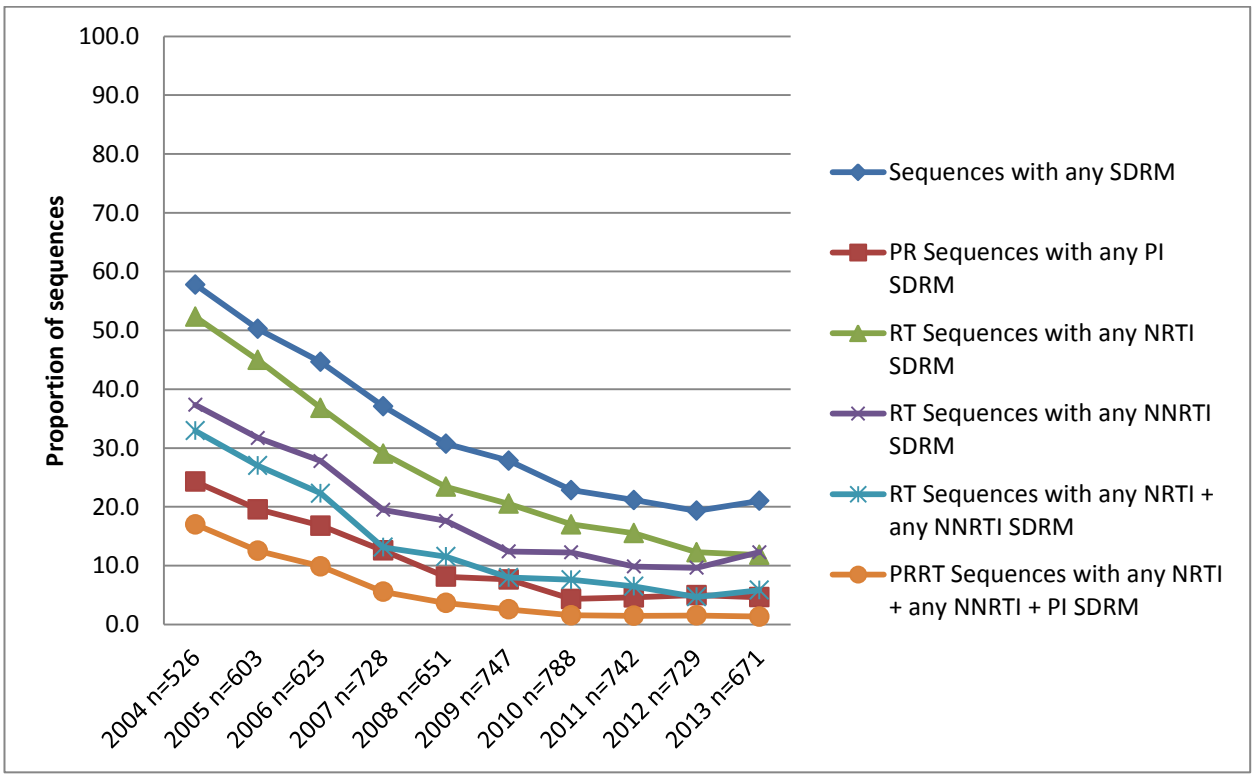
Table 4.2 Baseline characteristics of all individuals sequenced

	Lab 1	Lab 2	Lab 3
Number of sequences, n (%)	3342 (48.4)	2782 (40.3)	777 (11.3)
Mean age (SD)	40.9 (11.1)	42.7 (10.9)	38.2 (12.1)
Male (%)	2967 (94.2)	1494 (89.8)	567 (74.6)
ART history available, n (%)	617 (18.5)	305 (11.0)	0 (0)
Subtype, n (%)			
B	2915 (87.2)	2419 (87.0)	561 (72.2)
C	100 (3.0)	109 (3.9)	77 (9.9)
CRF01_AE	260 (7.8)	198 (7.1)	68 (8.8)
Other	67 (2.0)	56 (2.0)	71 (9.1)

Surveillance drug resistance mutations

There has been an overall decrease in frequency of prevalent SDRM from 57.8% in 2004 to 21% in 2013, with corresponding declines in PI (24% to 4.6%), NRTI (52.3 to 11.8%) and NNRTI (37.5% to 12.3%) resistance mutations over this time (Figure 4.1).

Figure 4.1 Frequency (%) of SDRMs in all sequences 2004-2013



From 2009-2013 the period for which there was treatment status data available, there was a significant decrease in overall frequency of SDRMs from 54.8% to 30%

($p<0.0001$) for treatment experienced patients (Table 4.3). There were corresponding declines in PI, NRTI and NNRTI classes in treatment experienced group, whereas SDRM rates in treatment naïve individuals have been stable with overall rate of 13.6%.

Table 4.3 Frequency (%) of SDRMs for treatment experienced and naive

Resistance in treatment experienced	2009 N=62 %	2010 N=64 %	2011 N=53 %	2012 N=80 %	2013 N=90 %	No. analysed	No. SDRM	Overall %
Sequences with any SDRM	54.8	45.3	32.1	30.0	30.0	347	125	36
PR Sequences with any PI SDRM	10.0	12.5	2.0	3.8	4.4	344	22	6.4
RT Sequences with any NRTI SDRM	43.5	32.3	24.5	23.8	24.4	344	100	29.1
RT Sequences with any NNRTI SDRM	25.8	19.4	13.2	18.8	16.7	344	65	18.9
Resistance in treatment naive	**2009 N=121 %**	**2010 N=123 %**	**2011 N=83 %**	**2012 N=54 %**	**2013 N=69 %**	**No. analysed**	**No. SDRM**	**Overall %**
Sequences with any SDRM	17.4	15.4	8.4	18.5	11.6	449	61	13.6
PR Sequences with any PI SDRM	3.3	2.4	3.6	1.9	1.4	448	12	2.7
RT Sequences with any NRTI SDRM	14.0	9.8	4.8	9.6	4.3	446	41	9.2
RT Sequences with any NNRTI SDRM	4.1	5.7	2.4	3.8	8.7	446	22	4.9

The most frequent mutation in the treatment naïve group was T215 which has shown a decline since 2009, while M184VI also showed a declining trend until 2012 in treatment experienced group (Figure 4.3). T215 was seen most frequently in treatment naïve (3.6% overall), whereas M184 was the most frequent mutation (14% overall) in treatment experienced individuals. In the treatment experienced population 85% (295/347) isolates retained susceptibility to tenofovir, with 1.1% (4/349) containing K65R mutation and 3.4% (12/349) with K70R mutation.

The most common NRTI mutations differed between treatment experienced (M184V) and naïve groups (M41L). For PI mutations, the most common in treatment experienced was M46I and L90M in treatment naïve. For NRTI mutations K103N was the most common across both groups. The only one of these mutations with a significant decline in frequency in both treatment experienced and naïve groups was M41L (p test for trend <0.05). T215 revertant mutations in treatment naïve group also had a significant decline in frequency over time. A summary of comparative mutations is shown in Figure 4.2.

There were 10 sequences containing thymidine associated mutations in treatment naïve group, and 34 in the treatment experienced group, with no significant change in the proportion of TAMs over time in either group. The number of TAMs identified within each sequence were more frequent in the treatment experienced group (up to five TAMs identified per sequence), while in the naïve group, the maximum number within a single sequence was three.

Figure 4.2 Comparison of most frequent major mutations between treatment experienced and naïve groups

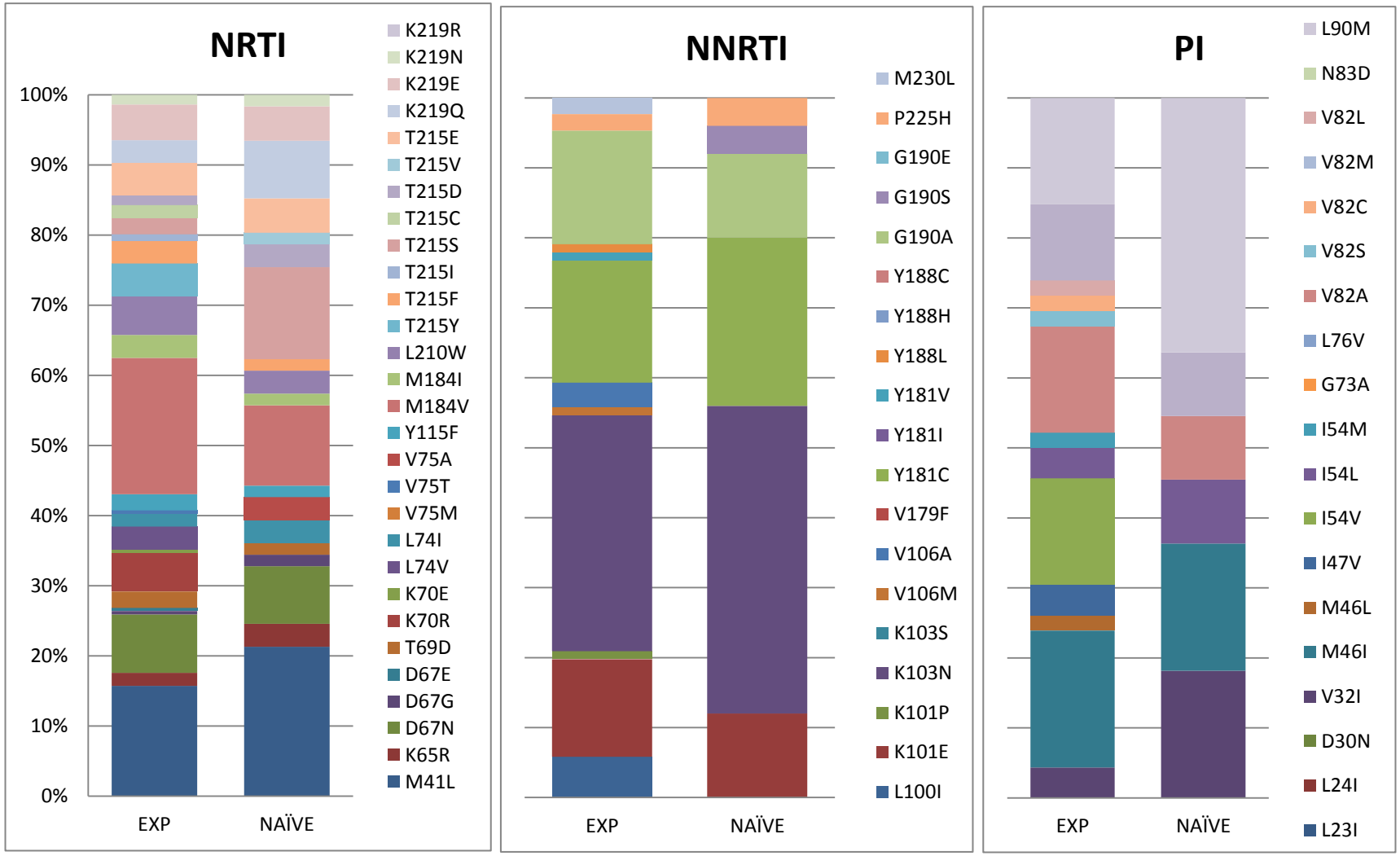
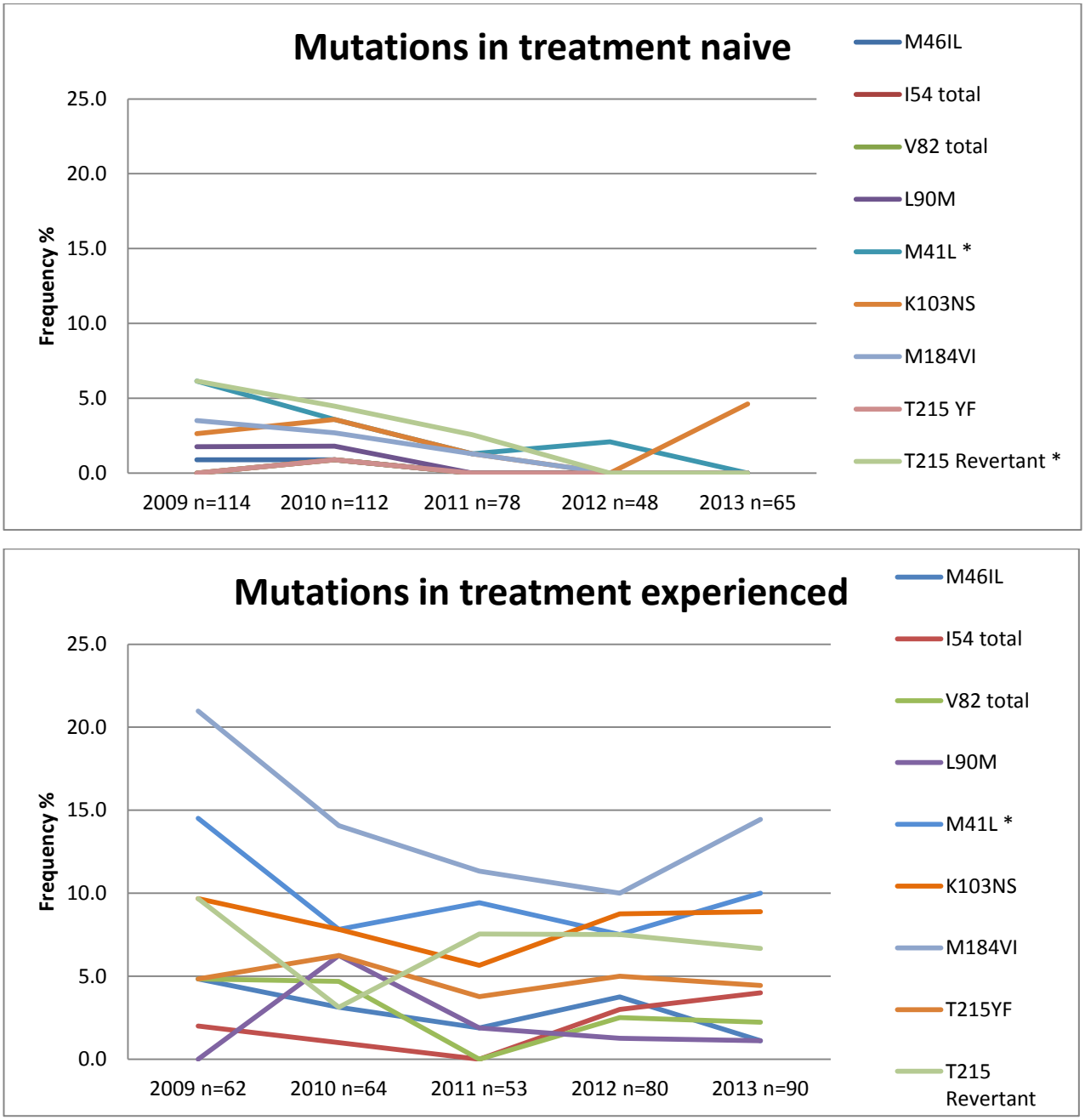


Figure 4.3 Frequency of most common mutations over time 2009-2013



* Significant p test for trend

In a subpopulation of seroconverters at a single laboratory between 2004 and 2013 (n=281) the overall rate of sequences containing a major mutation was 17.4% (49/281), with 1.4% PI, 5.7% NRTI and 11.7% NNRTI mutations. The complete list of mutations is shown in Table 4.4.

Table 4.4 Major mutations identified in seroconverters 2004 to 2013

	Freq.	%
PI		
G48DGV	1	0.4
I54ST	1	0.4
L90M	1	0.4
M46IM	1	0.4
None	277	98.6
NRTI		
D67N,K219Q	1	0.4
D67N,T215V,K219Q	1	0.4
K219R	1	0.4
M41L	1	0.4
M41L,T215C	3	1.1
M41L,T215CS	1	0.4
M41L,T215S	4	1.4
M41LM	1	0.4
T215E	1	0.4
T215L	1	0.4
T69N	1	0.4
None	265	94.3
NNRTI		
E138A	4	1.4
E138AE	1	0.4
K101E	1	0.4
K101E,E138A,G190A	1	0.4
K103N	4	1.4
K103T	1	0.4
V106I	5	1.8
V179D,P225H	1	0.4
V179E	3	1.1
V90I	9	3.2
V90IV	1	0.4
Y181SY,H221HY	1	0.4
Y188L	1	0.4
None	248	88.3
Total	**281**	**100.0**

4.5 Discussion

This is the first analysis of all GART performed in NSW over a ten year period. There has been a decline in frequency of SDRM in treatment experienced populations, reflecting the availability of better tolerated regimens and improved adherence. Rates in treatment naïve patients have been stable. Due to the referral pathway for resistance testing, the true rate of TDR could not be determined for all sequences in this study, as testing indication (recently infected, newly diagnosed, antiretroviral naive versus treatment failure) is not routinely available to the reference laboratories. SDRM rates in a subpopulation of recently infected seroconverters are an alternate method for approximating the rate of TDR, and has the advantage of lower rates of reversion mutations[258].

Based on data from the Australian HIV observational database (AHOD), the proportion of patients currently receiving antiretroviral treatment in NSW has increased from 88.8% (539/607) in 2004 to 96% (923/961) in 2013 [259, 260]. Despite the increases in treatment uptake seen over the ten-year period, SDRMs in the treatment experienced population has declined. Treatment naïve populations have also seen a decline in M41L mutations and T215 revertant mutations. The latter may reflect a shift in treatment regimens away from older zidovudine based regimens to newer agents, which are associated with improved adherence, and as a result less circulating mutations with transmission potential in the infective population. Contemporaneous knowledge of the prevalent rate of drug resistant mutations at a state-wide level is important to inform and update local antiretroviral guidelines. This is of particular relevance to NSW, which is participating in a large scale pre-exposure prophylaxis implementation trial which has enrolled over 4000 high risk individuals [170]. This study has coverage of an estimated 83% (6901/8291) of those on treatment in NSW, providing a background rate of HIV-1 drug resistance prevalence prior to implementation of this strategy, allowing comparisons with future surveillance studies. Despite universal access to ART and availability of subsidised GART in Australia, there is currently no national antiretroviral resistance surveillance programme.

During the study period (2004-2013), there has been a shift in the resistance testing paradigm, initially only available to those failing a combination antiretroviral regimen. Laboratories will proceed with HIV resistance testing if there is a plasma viral load of at least 2000 copies/mL. Since 1 July 2011, laboratories received reimbursement for testing when performed at presentation, before ART and with treatment failure [256]. This change has altered the population being tested in NSW, and needs to be considered in interpretation of study results. The apparent decrease in prevalence of

SDRMs can be attributed to an increased proportion in treatment naïve patients undergoing testing in the years leading up to 2011. One limitation of this study is the ability of the reference laboratory to identify prior ART exposure. This was possible for some of the sequences; however this review only included hospital-based records. Thus treatment that was commenced at an earlier time in the community would not have been captured. This may have led to an overestimation in treatment naïve individuals, partly explaining the higher than expected prevalence of M184 and T215 mutations in this group. Given the absence of treatment status data at Lab 3, which also had the highest proportion of non-B subtypes, comparison of differences in mutations between subtypes was not performed as this would be limited by small numbers and selection bias.

This study differs from most resistance surveillance reports which focus solely on TDR. Due to format and availability of sequence data collected, this study reflects the prevalent rates of mutations at a population level. We could not directly compare the proportion of viral load suppression at a population level in New South Wales for this study period. There has however been a steady upward trend in proportion of patients with an undetectable viral load, increasing from 62-84% in 2013 to 89-97% in 2018 [261]. The inclusion of acquired drug mutations also can provide insights into the longevity and tolerability of various regimens. For those known to be treatment experienced, there was a decrease in the rate of acquired drug resistance, reflecting newer better tolerated antiretroviral therapy and a shift towards single tablet regimens (STR). The most common regimen from 2008 to 2014 in the AHOD cohort was FTC+ TDF+ EFV which has been available as a STR since 1 January 2010[259]. Importantly, the frequency of mutations specific to EFV (P225H) was low in both naïve and experienced populations, although it is now recommended as an alternative regimen in Australian guidelines. Rates of TDF mutations are low in comparison to global rates, possibly reflecting the availability of other ART-based regimens such as ABC/3TC, which is also listed as a recommended regimen in national guidelines [90, 262] In addition, this may be due to the relatively high levels of adherence and education in this population, as evidence by the stable rates of M184V mutations in the treatment experienced population.

In summary, the prevalence of overall SDRMs in NSW is decreasing. The rate of TDR has remained stable, while the rate of acquired drug resistance has decreased. The rates of TDR in the seroconverter population is comparable to that reported in an antiretroviral-naïve population in the United Kingdom, with similar frequencies of common mutations reported [263]. The United States has also reported decreasing

trends in prevalence of resistance in all classes with the exception NNRTI class which increased [112]. Our study also did report an increased rate of NNRTI class resistance in treatment naïve population in 2013, but follow up data is needed to determine if this is a continuing trend [264]. The ongoing evaluation of drug resistance mutations in New South Wales will be further explored in Chapter 5, where this dataset is linked to HIV notifications data to provide a state-wide surveillance database of transmitted drug resistance mutations in newly diagnosed individuals.

Acknowledgements

The following people and institutions have contributed to this work and are gratefully acknowledged:

1 The Kirby Institute, UNSW Australia- Ansari Shaik

2 Royal Prince Alfred Laboratory, NSW health pathology- Hanan Salem, Frederick Lee

3 NSW State Reference Laboratory for HIV/AIDS & Molecular Diagnostics- Alex Carrera

4 ICMPR Westmead - NSW Health Pathology- Kiran Thapa, Dominic Dwyer

Funding

This study was partly funded by an NHMRC postgraduate research scholarship APP1074467.

CHAPTER 5

RATE OF DRUG RESISTANT MUTATIONS AND ASSOCIATED RISK FACTORS IN NEWLY DIAGNOSED HIV IN NEW SOUTH WALES

5.1 Abstract

Objective: Monitoring of pre-treatment HIV drug resistance (PDR) rates is an important component of any HIV strategy. This study defines the PDR rate in newly diagnosed HIV in New South Wales (NSW) from 2004 to 2015.

Methods: 2573 sequences from routinely performed genotypic antiretroviral resistance testing were successfully linked to HIV notifications. PDR analysis was restricted to 1475 newly diagnosed infections, using Stanford HIV db and 2009 WHO surveillance drug resistance mutations list. Logistic regression was performed to determine factors associated with PDR.

Results: The overall rate of PDR was 10.4% with a decreasing trend between 2004 and 2015 (OR 0.95, CI 0.90-1.01, p=0.08). Overall rates for nucleoside reverse transcriptase inhibitor (NRTI; mainly M41L, T215 and thymidine-associated mutations), non-nucleoside reverse transcriptase inhibitor (NNRTI) and protease inhibitor (PI) mutations were 5.2%, 3.5% and 2.8% respectively. Non-metropolitan residence was associated with increased PDR (OR 4.4, CI 1.84-10.51, p<0.01), while oldest age groups and non-B HIV subtype had decreased PDR (OR 0.36, CI 0.21-0.64, p<0.01; OR 0.28, CI 0.17-0.48, p<0.01).

Conclusion: PDR rates from 2004-2015 have been stable in NSW, Australia. People residing in non-metropolitan regions and those aged 19-29 years had the highest odds of PDR, highlighting the need for surveillance systems linking drug resistance testing to HIV notifications.

Keywords: antiretroviral, transmitted, resistance, HIV, mutations

5.2 Introduction

In the context of widespread rollout of antiretroviral drug therapy, better tolerated and more active regimens have seen declines in levels of acquired drug resistance [97]. In

contrast, several studies have shown increases in the rates of pre-treatment drug resistance (PDR) [97, 112, 121]. Other selected populations, usually clinical trial cohorts or those engaged in follow-up, have shown decreased in PDR rates over time [104, 116, 118, 263].

Australia has been reported to have one of the highest PDR rates in the world (15.5%), as reported in a recent systematic review [111] and in a clinical trial cohort (17.5%) [110]. Other studies report overall rates of 7.9% to 13.6%; however these are limited by sample size and lack of linkage to relevant demographic data [142, 143]. While declining rates of overall resistance mutations have been reported [143], the true PDR rate in recent years is unknown as there is no established surveillance that incorporates HIV notification data and linked sequence data.

NSW is the Australian state with the largest HIV burden, and has seen a decrease in new diagnoses in 2016 with the lowest standardised population rate of new infections on record for NSW [265]. Through a multifaceted public health intervention, including provision of subsidised PrEP to over 6000 people at risk, it is progressing towards the goal of “virtual elimination” of HIV transmission by 2020 [266]. In NSW there are approximately 8300 people with dispensed antiretroviral treatment, equivalent to around 92% coverage of people diagnosed and living with HIV [267].

The objective of this study was to report rates of PDR mutations for reverse transcriptase (RT) inhibitors and PI in NSW from 2004 to 2015 as well as acquired drug resistance (ADR) for integrase inhibitors from 2012 to 2016. This period represents different treatment eras, and occurred during the implementation of a new prevention strategy that includes a range of interventions aimed at increased testing, increased treatment uptake with earlier initiation of therapy as well as the beginning of PrEP rollout[252].

5.3 Methods

Study population and design

A state-wide database of sequences performed for the routine determination of genotypic anti-retroviral drug resistance (GART) was established, as previously published and outlined in Chapter 4[143]. The database was expanded to include all sequences sampled from 2004 to 2015 from each of the three reference laboratories

that perform all the GART for routine care at diagnosis, pre-treatment or for treatment failure (GART lab database).

Data linkage was performed as outlined in Chapter 2. Variables requested included baseline demographic data, HIV diagnostic testing information, CD4+ T-cell count, plasma HIV load at baseline and HIV risk exposure. HIV sequences included for PDR analysis were the first sequence obtained within 12 months of diagnosis (newly diagnosed). Poor quality sequences and vertically transmitted cases were excluded.

Ethical considerations

This study was approved by NSW Population & Health Services Research Ethics Committee [AU RED Reference: HREC/15/CIHS/38; Cancer Institute NSW reference number: 2015/08/605]. A waiver of informed consent was sought and granted.

Laboratory methods

Genotypic resistance testing was performed on the HIV *polymerase* gene using methods as previously described in Chapter 4.3 [143]. PDR was defined using the World Health Organization 2009 surveillance drug resistance mutation (SDRM) list [107], with mutations analysed using Stanford HIVdb version v8.3 (2017-03-02). Thymidine analogue-associated mutations (TAMS) include M41L, D67N, K70R, L210W, T215YF and K219QE. Early infection was defined by evidence of HIV infection acquired within 12 months of diagnosis. This included notification of a seroconversion-like illness or negative or indeterminate HIV test within 12 months of diagnosis, irrespective of CD4+ T-cell count or presentation with an AIDS defining illness at diagnosis [7]. Subtype was determined for reverse transcriptase gene according to the Stanford HIV Drug Resistance Database v7.0, and ambiguous subtypes were confirmed using REGA subtyping tool [268].

Statistical analysis

Univariate and multivariate logistic regression was performed to determine factors associated with PDR, and included: baseline CD4+ T-cell count, subtype, age at diagnosis, country of birth, likely place of acquisition, risk group, plasma HIV load at diagnosis, stage of HIV infection, year of test and geographical region of residence. Stage of infection at diagnosis was as defined by NSW Ministry of health: Early = Evidence of HIV infection acquired within 12 months of diagnosis, which was defined as notification of a seroconversion like illness or negative or indeterminate HIV test within 12 months of diagnosis, irrespective of CD4 or presentation with an AIDS defining illness at diagnosis. CD4 500+, CD4 350 to 499, CD4 200 to 349 each excludes early and advanced categories. Advanced = CD4 count less than 200 or

AIDS defining illness in absence of evidence of ‘Early’ diagnosis [237]. Analysis was performed using STATA statistical software version 13.0 (StataCorp LP, College Station, TX, USA).

5.4 Results

SDRM mutations

Between January 2004 and December 2015, there were 2573 HIV *polymerase* sequences that linked to the HIV notification database. Of these, 1493 (58%) were collected within 12 months of diagnosis (newly diagnosed) and 1475 were included for analysis. 18 poor quality sequences were excluded due to the presence of stop codons or frameshift mutations. The median sample year was 2012. The baseline demographics, stratified by year of resistance test are summarized in Table 5.1.

Table 5.1 Baseline demographics of newly diagnosed individuals

Year of resistance test	2004 to 2007	2008 to 2011	2012 to 2015	Total
Total N (%)	230(15.6)	461(31.3)	784(53.2)	1475
Age median (IQR)	37.5 (31-46)	38 (30-46)	37 (29-47)	38 (30-46)
Sex				
Male n (%)	208(90.4)	407(88.3)	707(90.2)	1322(89.6)
Female n (%)	21(9.1)	50(10.9)	35(4.5)	106(7.2)
Unknown	1(0.4)	4(0.9)	42(5.4)	47(3.2)
Risk Group				
MSM	159(69.1)	329(71.4)	634(80.9)	1122(76.1)
Heterosexual	40(17.4)	106(23)	117(14.9)	263(17.8)
PWID	8(3.5)	14(3)	12(1.5)	34(2.3)
Other or unknown	23(10)	12(2.6)	21(2.7)	56(3.8)
Country of Birth				
Australia	139(60.4)	266(57.7)	426(54.3)	831(56.3)
South East Asia	16(7)	44(9.5)	98(12.5)	158(10.7)
Sub-Saharan Africa	8(3.5)	32(6.9)	27(3.4)	67(4.5)
Other	67(29.1)	119(25.8)	233(29.7)	419(28.4)
Place of acquisition				
Australia	2(0.9)	271(58.8)	475(60.6)	748(50.7)
Overseas	1(0.4)	100(21.7)	204(26)	305(20.7)
Unknown	227(98.7)	90(19.5)	105(13.4)	422(28.6)
Region of residence				
Metro	188(81.7)	390(84.6)	661(84.3)	1239(84)
Rural and regional	17(7.4)	46(10)	88(11.2)	151(10.2)
NSW not specified	16(7)	4(0.9)	7(0.9)	27(1.8)
Outside NSW	9(3.9)	21(4.6)	28(3.6)	58(3.9)
Stage at diagnosis*				
Early	93(52.5)	162(36.3)	350(45.6)	605(43.5)
500+	12(6.8)	32(7.2)	56(7.3)	100(7.2)
350 to 499	7(4)	41(9.2)	96(12.5)	144(10.4)
200 to 349	16(9)	85(19.1)	108(14.1)	209(15)
<200	49(27.7)	126(28.3)	157(20.5)	332(23.9)
Plasma HIV load at baseline				
<=100 000	13(46.4)	16(32.7)	338(56.4)	367(54.3)
>100 000	15(53.6)	33(67.4)	261(43.6)	309(45.7)
HIV subtype				

A	3(1.3)	5(1.1)	27(3.4)	35(2.4)
B	187(81.3)	343(74.4)	543(69.3)	1073(72.8)
C	11(4.8)	36(7.8)	50(6.4)	97(6.6)
G	3(1.3)	1(0.2)	3(0.4)	7(0.5)
AE	24(10.4)	64(13.9)	119(15.2)	207(14)
AG	1(0.4)	8(1.7)	8(1)	17(1.2)
Other	1(0.4)	4(0.9)	34(4.3)	39(2.6)

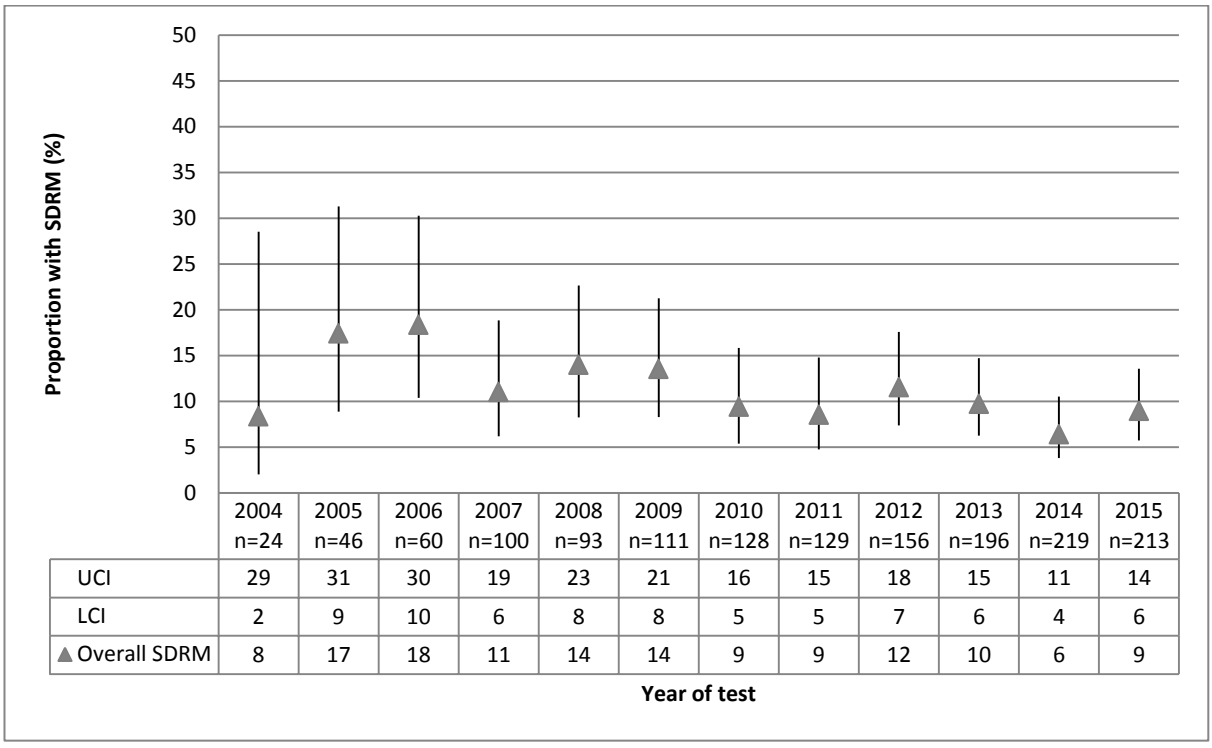
PWID people who inject drugs; MSM men who have sex with men; SSA sub-Saharan Africa; SEA South East Asia

*Stage: Early = Evidence of HIV infection acquired within 12 months of diagnosis, which was defined as notification of a seroconversion like illness or negative or indeterminate HIV test within 12 months of diagnosis, irrespective of CD4 or presentation with an AIDS defining illness at diagnosis. CD4 500+, CD4 350 to 499, CD4 200 to 349 each excludes early and advanced categories.

Advanced = CD4 count less than 200 or AIDS defining illness in absence of evidence of 'Early' diagnosis [251]

The overall rate of PDR was 10.4%, with a decreasing trend over time: 13.9% (CI 10.0-19.0) in 2004 to 2007; 11.0% (CI 8.5-14.3) in 2008 to 2011; 8.9% (CI 7.1-11.1). Rates for NRTI NNRTI and PI mutations were 5.2%, 3.5% and 2.8% respectively. Figure 5.1 shows the overall proportion of sequences with any SDRM.

Figure 5.1 Overall SDRM in newly diagnosed individuals



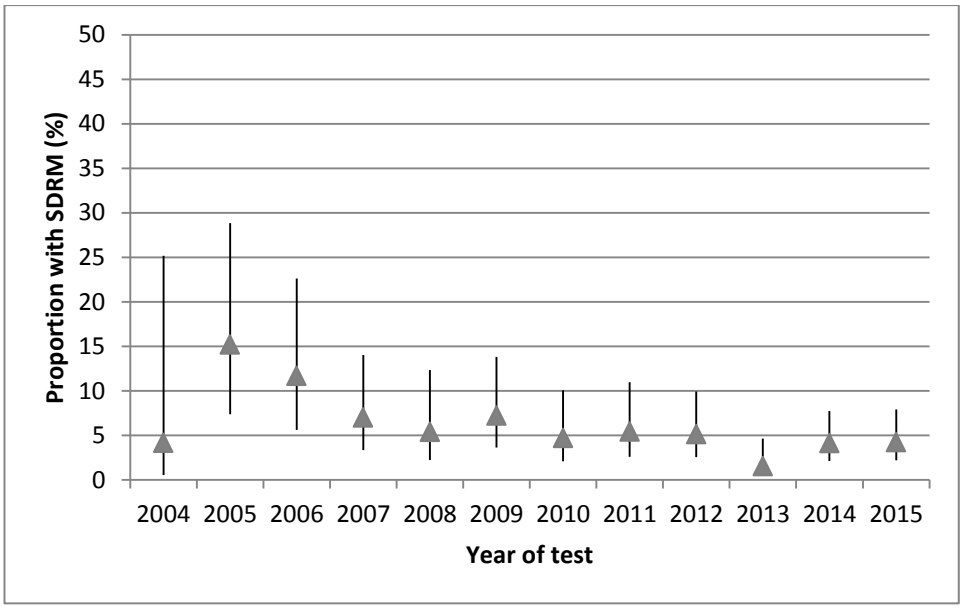
	2004 n=24	2005 n=46	2006 n=60	2007 n=100	2008 n=93	2009 n=111	2010 n=128	2011 n=129	2012 n=156	2013 n=196	2014 n=219	2015 n=213
UCI	29	31	30	19	23	21	16	15	18	15	11	14
LCI	2	9	10	6	8	8	5	5	7	6	4	6
▲ Overall SDRM	8	17	18	11	14	14	9	9	12	10	6	9

UCI= upper confidence interval, LCI= lower confidence interval, bars indicate confidence interval

The greatest reduction was in the NRTI class, which had the highest proportion of SDRM of 15.2% in 2005, decreasing to 4.2% in 2015 (Figure 5.2). This was driven by a

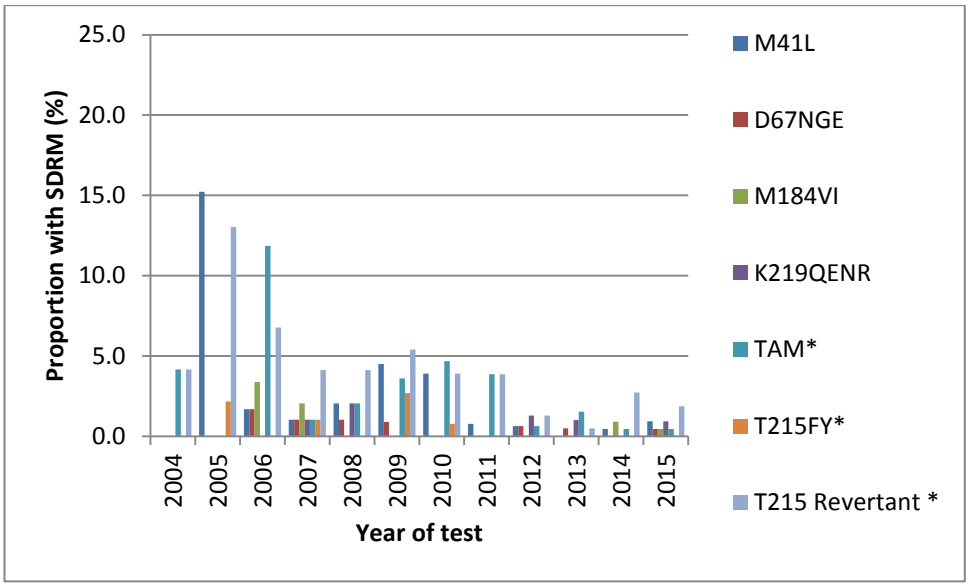
reduction in M41L and other thymidine associated mutations (TAMs), particularly T215 associated mutations (Figure 5.3). The greatest reduction was seen in number of T215 revertant mutations; however T215FY mutations that are associated with zidovudine treatment also had a significant decline ($p<0.05$). No tenofovir-specific primary mutations (K65R/E/N or K70E) were observed over this period. Seven M184V mutations (overall frequency 0.5%) were observed, three in the last two years.

Figure 5.2 Frequency of NRTI SDRM 2004 to 2015



Bars indicate confidence interval

Figure 5.3 Most common NRTI mutations



*P-trend chi-square test for significance p<0.05

Abbreviations: TAM= thymidine analogue mutations and includes M41L, D67N, K70R, L210W, T215FY K219QE

The proportion of NNRTI and PI mutations has remained low and stable as shown in Figures 5.4 and 5.5. The most common NNRTI mutation was E138A which occurred 39 times, with an overall frequency of 2.6%; however this is not included in the WHO SDRM list. This mutation is reported more commonly in subtype C, and was seen with

a frequency of 8.3% in subtype C. There were 36 K103N/S mutations, with a frequency of 2.4%. The most frequent protease mutation was L90M, which only was present for 27 sequences from 2008 with a frequency of 1.8%.

Figure 5.4 Frequency of NNRTI SDRM 2004 to 2015

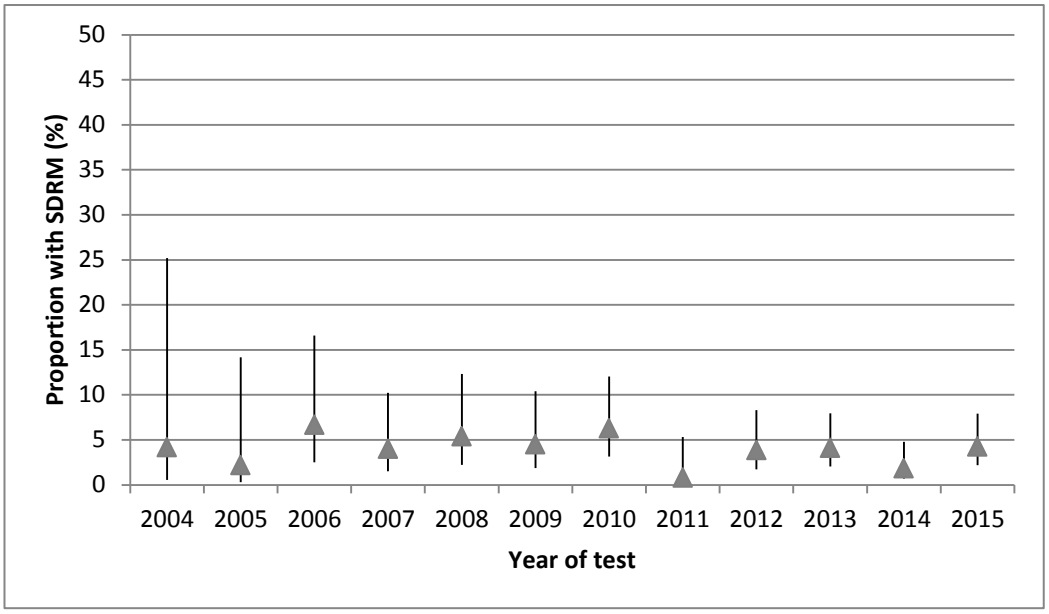
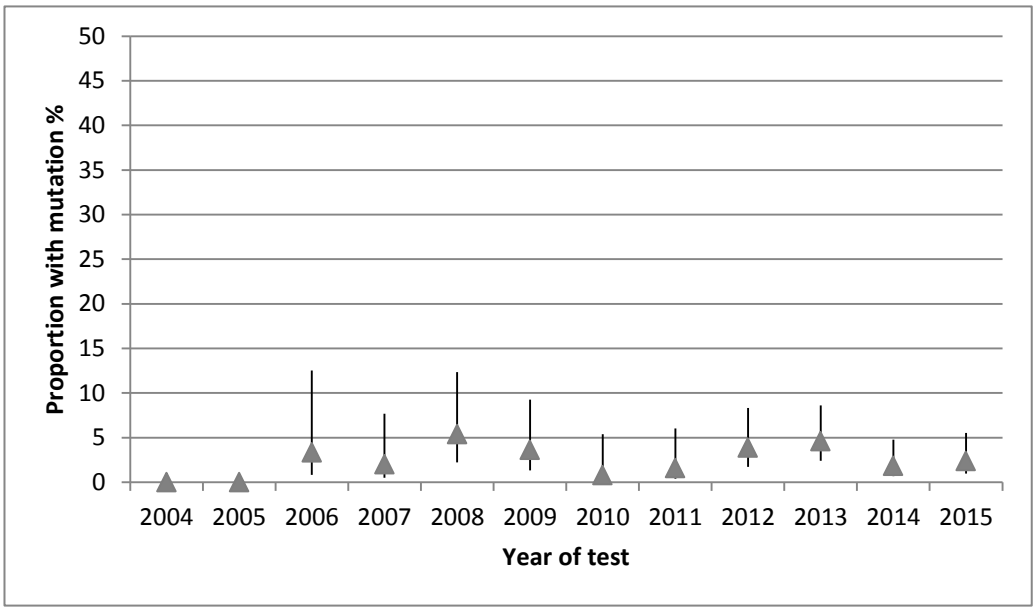


Figure 5.5 Frequency of PI SDRM 2004 to 2015



Factors associated with PDR

On univariate analysis, year of test, overseas place of acquisition and non-B subtype were associated with a reduced odds ratio of PDR, while younger age at diagnosis and non-metropolitan region of residence were associated with increased PDR on

multivariate analysis (Table 5.2). Residing in a non-metropolitan location was associated with between 1.19-4.40 increased odds of PDR. Although year of test did not remain significant on multivariate analysis (OR 0.95, CI 0.08-1.01, p=0.08), there was a trend towards decreased rate of PDR with calendar year.

To account for potential residual confounding due to collinearity of non-B subtype to other variables (overseas born and calendar year), subtype B sequences were also analysed separately. The variables age at diagnosis and region of residence still remained significant in the final model (Table S5.1).

Table 5.2 Univariate and multivariate logistic regression analysis of variables associated with SDRMs

		Univariate analysis					Multivariate analysis				
Variable name	n	Odds Ratio	95% CI		p value	Overall p value	Odds Ratio	95% CI		p value	Overall p value
Age (years) at diagnosis											
19 to 29	391	1.00				<0.01					**<0.01**
30 to 39	461	0.52	0.33	0.80	<0.01		0.48	0.31	0.75	0.00	
40 to 49	348	0.72	0.47	1.12	0.15		0.67	0.43	1.04	0.07	
50 and over	275	0.40	0.23	0.70	0.00		0.36	0.21	0.64	0.00	
Country of Birth											
Australia	831	1.00									
Overseas	643	0.72	0.51	1.02	0.07						
Place of acquisition											
Australia	748	1.00				0.01					0.13
Overseas	305	0.50	0.29	0.86	0.01		0.74	0.41	1.31	0.30	
Unknown	422	1.32	0.92	1.90	0.13		1.31	0.89	1.91	0.17	
Risk group											
MSM	1122	1.00				0.49					
Heterosexual male	133	0.44	0.20	0.97	0.04						
Heterosexual female	92	0.86	0.42	1.76	0.69						
PWID	34	0.50	0.12	2.11	0.34						

Other	5	1.00									
Unknown	89	1.01	0.51	2.00	0.98						
Year of test											
Year of test	1475	0.93	0.88	0.98	0.01		0.95	0.90	1.01	0.08	
Region of residence											
Metro	1239	1.00				<0.01					**<0.01**
Rural and regional	151	1.13	0.65	1.96	0.67		1.19	0.68	2.10	0.54	
NSW not specified	27	4.00	1.71	9.34	0.00		4.40	1.84	10.51	0.00	
Outside NSW	58	2.22	1.12	4.40	0.02		2.69	1.32	5.48	0.01	
Stage of Infection											
Early	605	1.00				0.88					
CD4+ T-cell count 500+	100	0.49	0.21	1.16	0.10						
350 to 499	144	0.82	0.45	1.51	0.53						
200 to 349	209	0.68	0.39	1.18	0.17						
CD4<200	332	0.79	0.50	1.23	0.29						
Missing	85	1.64	0.89	3.02	0.11						
Baseline CD4+ T-cell count(cells/µl)											
<200	356	1.00				0.67					
200-349	407	1.16	0.72	1.86	0.55						
350-499	344	1.32	0.82	2.15	0.26						
>=500	242	0.74	0.40	1.36	0.33						
Missing	126	1.74	0.95	3.18	0.07						

Plasma HIV load at diagnosis (c/mL)											
<100 000	367	1.00				0.52					
>100 000	309	0.83	0.47	1.47	0.52						
Missing	799	1.55	1.02	2.37	0.04						
HIV subtype											
B	1073	1.00									
Non-B	402	0.30	0.18	0.51	<0.01		**0.28**	**0.17**	**0.48**	**<0.01**	

Abbreviations**:** PWID people who inject drugs; MSM men who have sex with men

Bold typeface indicates variable that remained significant on multivariate analysis

5.5 Discussion

PDR rates from 2004-2015 have been stable in NSW, the state with the highest HIV burden in Australia. This is the largest study of PDR reported in Australia, and is strengthened by its linkage to the HIV notifications dataset, allowing analysis of associated risk factors. People residing in non-metropolitan regions and those aged 19-29 years had the highest odds of PDR.

The overall rate of 10.4% PDR is lower than that reported in previous systematic reviews and clinical trials cohorts, which used historical laboratory datasets and were more limited and biased in their sampling strategies [110, 111]. Our rates are comparable to those in Victoria another state jurisdiction on the east coast of Australia with the second highest burden of HIV infection, which showed an overall rate of 11.4% over a similar time period (2005 to 2010) [142].

Over a twelve year period rates of PDR have remained stable despite a period of increased treatment coverage. NSW has seen increases in antiretroviral treatment coverage from 70% in 2011 to 95% in 2017 in primary care clinics [267]. This treatment expansion is paralleled by the increased access to genotype testing, with numbers increasing over the study period such that over half the data was collected from the last three years. This sampling bias has been addressed by including year of test as a variable in the univariate analysis. This also accounts for the large confidence intervals seen in Figure 5.1, which decrease towards the more recent years of the study period. Although some other studies have shown decreases in PDR, these have occurred within the context of linkage to care or close clinical follow up such as research study enrolment [118], large clinical cohorts [104], people who use drugs within a clinical cohort [116], or only within a subgroup of men who have sex with men [113].

This study provides insights into vulnerable populations not captured by previous studies, which have included predominantly men who have sex with men (MSM) populations from metropolitan areas, enrolled into clinical trials and cohorts. Younger age at diagnosis (19-29 years) was found to be associated with higher PDR rates compared to older age groups, and could be associated with lower treatment rates (46% in 2011) and lower adherence which may lead to onward transmission of drug resistant virus [267, 269, 270]. This study found increased rates of PDR in non-metropolitan regions, including rural and regional areas, which may be a marker of

lower engagement in care and greater difficulties in accessing regional specialist HIV services. PDR was not found to be associated with any clusters or transmission networks, adding further weight to the role of lower engagement in care, in these regions. Other non-metropolitan groups included those outside NSW and “NSW not specified” which may represent transient or travelling populations. The role of travel and population movements was highlighted by a recent study showing growing HIV subtype diversity and a changing epidemic in Australia [254]. A national surveillance program for PDR is needed to better identify these populations. Globally, PDR is most prevalent in MSM risk groups [111], however this study did not find any direct association with transmission risk groups. Although there was lower PDR in overseas acquisition on univariate analysis, this was not significant on multivariate analysis.

This study is limited by the proportion of sequences that linked to the HIV notifications database. Assessment of demographic data characteristics (described in Chapter 2) found the linked data analysed to be representative of the population diagnosed with HIV in NSW. One of the strengths of this approach compared to previous reports is that it includes broader geographical sampling from rural and regional parts of the state as well as non-MSM risk groups that were not represented by published clinical trial cohorts. A number of reasons were identified to account for the low proportion of linked records including linkage to HIV notifications database that only dated back to 1998. A large number of individuals had invalid identifiers (incomplete two by two name code, date of birth or gender fields) that were available for linkage.

To improve data linkage for future analyses, the following modifications are proposed: i) request for data linkage to include all records of the HIV notifications database dating back to 1981 ii) inclusion of sexual health clinic codes which have been used in place of two by two name codes and iii) inclusion of additional variables to optimize linkage [eg: postcode (current, at diagnosis, AIDS. at death), AKA (otherwise known as) variable, date of death (to exclude incorrect linkages), duplicate match variable (flags duplicate records), event type variable (whether record is deleted or active)].

The definition of PDR was based on newly diagnosed infections which were defined by having had a GART performed within 12 months of diagnosis. This assumes that no antiretroviral therapy has been commenced within this time; however complete treatment history was not available in this dataset. A recent study reported overall PDR of 13.6% in the treatment-naïve population from 2009 to 2013 in a subset of the same

population, hence the likelihood that treatment experienced individuals were included in this analysis is extremely low, as it is also standard of care to have GART performed prior to commencing therapy [143].

In conclusion, this study found that rates of PDR have remained stable in the setting of increasing ART coverage over a twelve year period. Risk groups with higher rates of PDR include the youngest age group (19-29 years) and non-metropolitan regions, highlighting the need for ongoing surveillance to identify emerging mutations in high risk populations.

Acknowledgements

The following people and institutions have contributed to this work and are gratefully acknowledged:

1 The Kirby Institute, UNSW Australia- Ansari Shaik, Phillip Keen, Andrew Grulich

2 Royal Prince Alfred Laboratory, NSW health pathology- Hanan Salem, Frederick Lee

3 NSW State Reference Laboratory for HIV/AIDS & Molecular Diagnostics- Alex Carrera

4 ICMPR Westmead - NSW Health Pathology- Kiran Thapa, Dominic Dwyer

5 Health Protection NSW- Christine Selvey, Barbara Telfer

6 Positive Life, Sydney – Craig Cooper

7 ACON, Sydney- Karen Price

8 NSW Ministry of Health- Jo Holden

9 Steering group of the NSW HIV Prevention Partnership Project

10 Centre for Health Record Linkage- Elizabeth Moore

Funding

This study was partly funded by an NHMRC postgraduate research scholarship APP1074467, NHMRC Partnership Project grant (Grant #1092852), the NSW Ministry of Health, UNSW, Sydney.

5.6 Appendix: supplementary data

Table S5.1 Univariate and multivariate logistic regression subtype B only

		Univariate analysis					Multivariate analysis				
	n	Odds Ratio	95% CI		p value	Overall p value	Odds Ratio	95% CI		p value	Overall p value
CD4 at baseline											0.11
CD4<200 [ref]	238					0.18					
CD4 200-349	294	0.87	0.52	1.45	0.60		0.84	0.50	1.42	0.52	
CD4 350-499	262	1.16	0.70	1.92	0.56		1.09	0.65	1.82	0.75	
CD4>=500	186	0.48	0.25	0.95	0.04		0.46	0.23	0.91	0.03	
Missing	93	1.34	0.70	2.57	0.38		0.93	0.45	1.90	0.83	
Age at diagnosis						<0.01					<0.01
0 to 29 [ref]	273										
30 to 39	333	0.46	0.29	0.74	<0.01		0.44	0.27	0.71	0.00	
40 to 49	268	0.68	0.42	1.07	0.10		0.66	0.41	1.05	0.08	
50 and over	199	0.41	0.23	0.73	0.00		0.37	0.20	0.67	0.00	
Country of Birth						0.65					
Australia [ref]	707										
Overseas	365	0.92	0.62	1.34	0.66						
Place of acquisition						0.35					
Australia [ref]	620										
Overseas	123	0.74	0.38	1.43	0.37						
Unknown	330	1.40	0.95	2.06	0.09						
Risk group						0.34					
MSM[ref]	924										
Heterosexual male	55	0.69	0.27	1.76	0.44						
Heterosexual female	22	2.59	0.99	6.74	0.05						
PWID	24	0.63	0.15	2.70	0.53						
Other	4	1.00									
Unknown	44	1.09	0.45	2.63	0.85						
Viral load at diagnosis (c/mL)						0.50					
VL < 100 000 [ref]	259										
load >100 000	205	0.81	0.44	1.49	0.50						
Missing	609	1.34	0.86	2.10	0.20						
Stage of Infection						0.88					
Early	333										
not early	740	0.97	0.66	1.43	0.88						
Region						<0.01					<0.01
Metro [ref]	901										
Rural and regional	117	1.21	0.69	2.14	0.50		1.33	0.75	2.38	0.33	
NSW not specified	22	4.38	1.79	10.69	<0.01		5.32	2.14	13.23	0.00	
Outside NSW	33	2.45	1.08	5.58	0.03		2.46	1.07	5.65	0.03	
Year of test						0.09					0.20
Year of test	1073	0.95	0.90	1.01	0.09		0.96	0.91	1.02	0.20	

CHAPTER 6

THE ROLE OF PRIMARY HIV INFECTION IN TRANSMISSION IN NEW SOUTH WALES

6.1 Abstract

Background: Despite treatment as prevention rollout, the number of newly diagnosed HIV-1 infections in NSW was stable between 2004 and 2015. To gain a deeper understanding into the role of PHI in onward HIV transmission, a molecular epidemiological approach was taken to determine characteristics associated with transmission clusters.

Methods: The state-wide HIV sequence database was linked with the HIV notifications. Phylogenetic tree construction was performed for three periods (2004 to 2007, 2008 to 2011, and 2012 to 2015). Early infection was defined by evidence of HIV infection acquired within 12 months of diagnosis, while PHI was defined by presence of seroconversion symptoms with specific laboratory defined criteria. Clusters were identified using genetic distance threshold of 0.015 and support > 0.7. Univariate and multivariate analysis was undertaken to determine factors associated with cluster membership.

Results: There were 2552 linked sequences, 1904 subtype B. The proportion of sequences in a cluster increased from 16% (2004-2007) to 30% (2008 to 2011) to 38% (2012 to 2015), with increases in cluster size. For subtype B sequences 2012 to 2015 (n=760), 94% male, 60% < 40 years, 86% MSM, 29% seroconversion symptoms. The two largest clusters (n=10 &13) were predominantly metropolitan MSM. For non-B sequences 2004 to 2015 (n=580): 68% male, 65% <40 years, 43% MSM, 20% seroconversion symptoms. These were predominantly transmission pairs, with largest cluster size of 5. On multivariate analysis, factors independently associated with subtype B and non-B cluster membership were acquisition in Australia, diagnosis post-2012, early infection stage (but not laboratory defined PHI). Age <30 years was associated with B subtype clusters.

Conclusion: Transmission clusters in NSW are predominantly comprised of early stage of infection acquired in Australia in the previous five years, but not specifically PHI. Campaigns focused on increased HIV testing among MSM may have led to the identification of new cases in these clusters in recent years. Factors associated with cluster membership are the same in B and non-B subtypes, although cluster size is larger in B subtypes. Ongoing analysis of transmission dynamics can help inform future public health strategies.

6.2 Introduction

In response to the concerning rate of new HIV infections, with the highest number seen in over twenty years recorded in 2012, a new public health approach was urgently needed. The NSW Ministry of health (MoH) developed a new strategy aiming to reduce the incidence of HIV by 80% in NSW by 2020 [166]. This strategy includes a multi-pronged approach including a wide range of behavioural and biomedical interventions strategies including expanded testing programs, rapid testing, removing barriers to early treatment, strengthened data collection and an ambitious pre-exposure prophylaxis demonstration program[271]. This approach is highly resource intensive and required collaborative support from community, academic, clinical and government bodies, including funding support. It is imperative to understand where new infections arise from in the context of this public health strategy, in particular what role primary HIV infection plays in driving onward transmission.

The proportion of infections that occur from a person who is unaware of their infection is uncertain. Onward transmission during HIV seroconversion is thought to possibly account for one third of onward transmission [79] although previous studies in Sydney suggested lower levels [139, 272]. Studies in similar populations demonstrate higher infectivity during early infection [81, 273-275]. However recent analyses have challenged previously accepted estimates of the infectivity of acute infection [78, 218]. Further investigation of this hypothesis is needed to help guide and improve the efficacy and efficiency of public health interventions. The study of molecular epidemiology uses HIV sequence data linked to clinical and demographic data. This can be used to generate an inferred phylogeny and evolutionary history of an epidemic that can help determine the impact of early versus late presenters in propagating the epidemic [88, 276]. In order to accurately interpret phylogenetic data, reliable clinical and epidemiological data is essential to provide a complete molecular epidemiological picture [276].

As outlined in Chapter 4, a state-wide HIV genotypic database has been established to define trends in HIV resistance mutations (Chapter 4 and 5). The characterisation of transmission networks or clusters provides a better understanding of the contribution of PHI, among other factors toward onward transmission of HIV in NSW. This knowledge will help to focus the various prevention strategies aimed at reducing transmission, with the aim of virtual elimination of HIV by 2020.

6.3 Methods

6.3.1 Study population and design

HIV-1 sequences sampled from 2004 to 2015 were obtained from the state-wide resistance database (as described in Chapter 4) and linked to the HIV notifications database as described in Chapter 2. Sequences that linked to notifications data were included for analysis.

6.3.2 Ethical considerations

This study was approved by NSW Population & Health Services Research Ethics Committee [AU RED Reference: HREC/15/CIHS/38; Cancer Institute NSW reference number: 2015/08/605]. A waiver of informed consent was sought and granted.

6.3.3 Community engagement

Due to concerns about the ethical and legal implications of potentially sensitive molecular epidemiological data, a study steering group was established. Included were representatives from HIV community organisations Positive Life and the AIDS Council of New South Wales (ACON) who were an integral part of the protocol development, design, analysis and publication. As a part of this process, a risk benefit analysis document was prepared (see Appendix 6.6.1) and extensive community engagement was undertaken. This included several multidisciplinary and multijurisdictional workshops, a face to face presentation at a “Treatments and Research” evening and distribution of a community information pamphlet (see Appendix 6.6.2).

6.3.4 Laboratory methods

HIV RNA testing and *pol* sequencing was performed as outlined in Chapter 4.

6.3.5 Phylogenetic methods

Subtyping Subtype was assigned using REGA HIV-1 automated subtype tool [277].

Editing Subtype B *pol* sequences were mapped to HIV-1 HXB2 HIV1/HTLV-III/LAV reference genome accession number K03455.1 and manually edited using BioEdit Sequence Alignment Editor ©1997-2013 and [278, 279]. Drug resistance mutation sites were identified and removed using Geneious® v 9.1.4(© 2005-2016 Biomatters Ltd, Auckland, New Zealand) [108, 163, 181, 280, 281]. These were removed to avoid potential bias because similarities in drug resistance mutations may arise separately from antiretroviral treatment and not from common shared ancestors.

Alignment Multiple pairwise alignment was performed on polymerase subtype B sequences using CIPRES gateway [282, 283]. The final alignment was 797 ungapped bases in length from positions 2400 to 3285 to included coverage of both *reverse transcriptase* and *protease* genes.

Phylogenetics Phylogenetic trees were inferred using maximum-likelihood analysis in RAxML via the CIPRES science gateway [282]. A General Time Reversible model of nucleotide substitution was used. Branch support was obtained using ultrafast bootstrapping with 10 000 replicates in IQ-TREE [185]. Clusters were identified using ClusterPickerv1.2.5. A sensitivity analysis was performed to test thresholds for cluster definition, including boot strap support values of up to 0.9 and large cluster threshold of 10. The following thresholds that are widely used in the literature were chosen as they had lower cut-offs to allow for increased cluster detection [186, 187]: initial support of 0.7, genetic distance of 0.015 and large cluster threshold of 2. Trees were visualised with Tree Figure Drawing Tool Figtree v1.4.3.

6.3.6 Statistical analysis

Analysis was stratified to three periods of sequence sampling: 2004 to 2007, 2008 to 2011 and 2012 to 2015 and descriptive statistics were used to describe the population and identify the presence of clusters in these groups. Further statistical analysis was restricted to period 2012 to 2015 as this correlated with the current NSW Ministry of Health strategy [166].

Univariate and multivariate logistic regression was performed to determine specifically whether PHI was one of the factors associated with cluster membership, however for completeness, the analysis also included age at diagnosis, risk group, region of residence, stage at diagnosis, seroconversion illness, negative test within 12 months,

year of onset and newly acquired infection. Analysis was performed using STATA statistical software v13.0 (StataCorp LP, College Station, TX, USA).

PHI was defined using the following criteria: Either i) signs and symptoms of acute retroviral symptoms with the presence of positive p24 antigen/HIV-1 proviral DNA/HIV-1 RNA or ii) indeterminate or evolving Western blot with positive p24 antigen/HIV-1 proviral DNA/HIV-1 RNA [4, 284].

Stage at diagnosis was defined as per NSW Ministry of Health: Early infection was defined by evidence of HIV infection acquired within 12 months of diagnosis as described in Chapter 5. This included notification of a seroconversion-like illness or negative or indeterminate HIV test within 12 months of diagnosis, irrespective of CD4+ T-cell count or presentation with an AIDS defining illness at diagnosis [7].

6.4 Results

6.4.1 Baseline characteristics

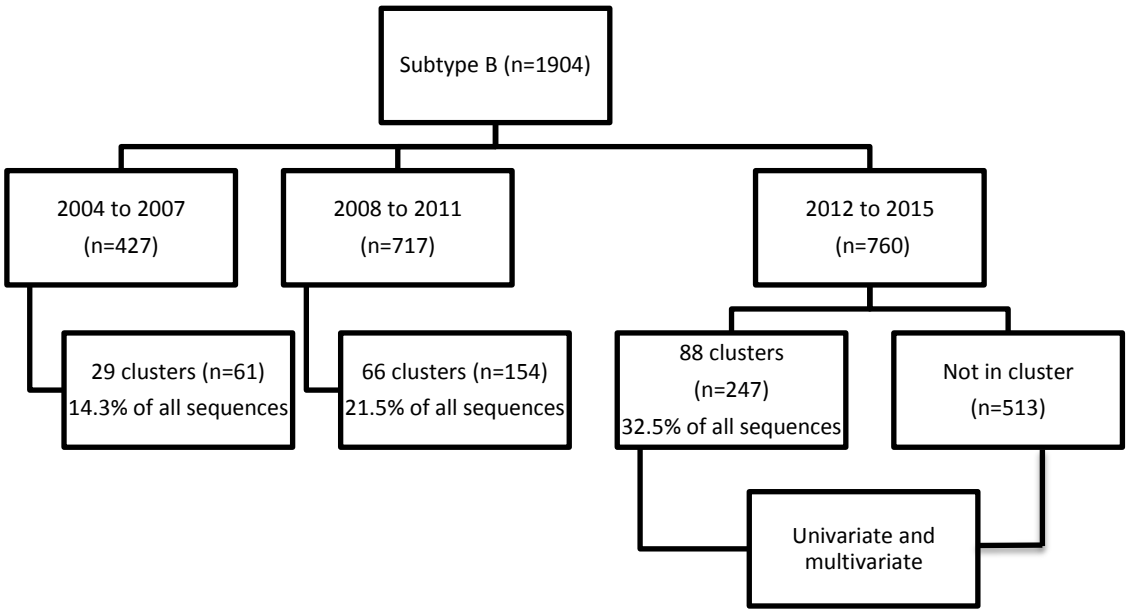
There were 1904 unique subtype B sequences sampled from 2004 to 2015 included for analysis as outlined in Figure 6.1. Baseline demographics of this population were predominantly male, MSM, with almost one third reporting symptoms of seroconversion within the previous 12 months as summarised in Table 6.1.

Table 6.1 Baseline characteristics of individuals with subtype B sequence

	N	%
Male sex	1718	90.2
Age At Diagnosis		
15 to 29 years	481	25.3
30 to 39 years	694	36.5
40 to 49 years	454	23.8
50 and over	275	14.4
Reported risk		
MSM	1599	84.0
Heterosexual sex	148	7.8
PWID	48	2.5
Other or unknown	109	5.7
Symptoms of seroconversion	567	29.8
Australian born	1,212	63.7
Place of Acquisition		

Australia	790	41.5
Overseas	157	8.3
Unknown	957	50.3

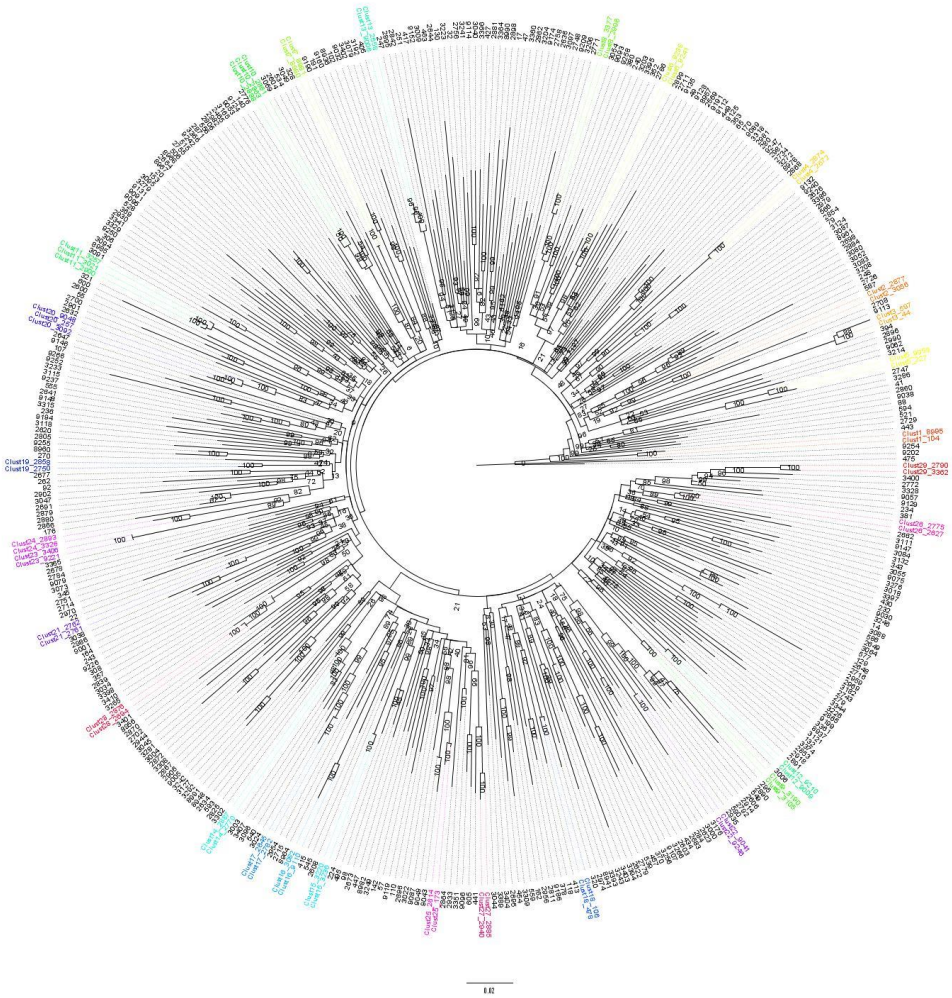
Figure 6.1 Sequences included for phylogenetic analysis



6.4.2 Cluster analysis

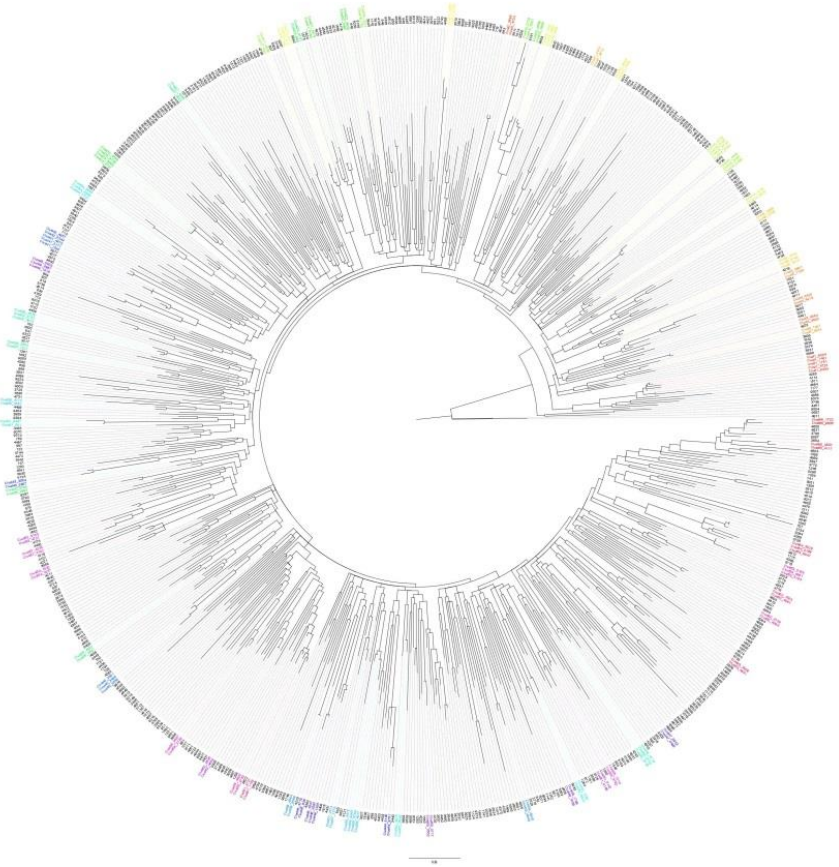
Figures 6.2-6.4 shows phylogenetic trees of HIV sequences sampled over three periods: 2004 to 2008, 2008 to 2011 and 2012 to 2015. The majority of sequences were not in a cluster, while the proportion of sequences belonging to a cluster (shown in colour) increased over time from 14.3% to 32.5%. Concurrently, the size of clusters increased over time as shown in Figure 6.5.

Figure 6.2 Phylogenetic tree of subtype B HIV sequences sampled between 2004 to 2007



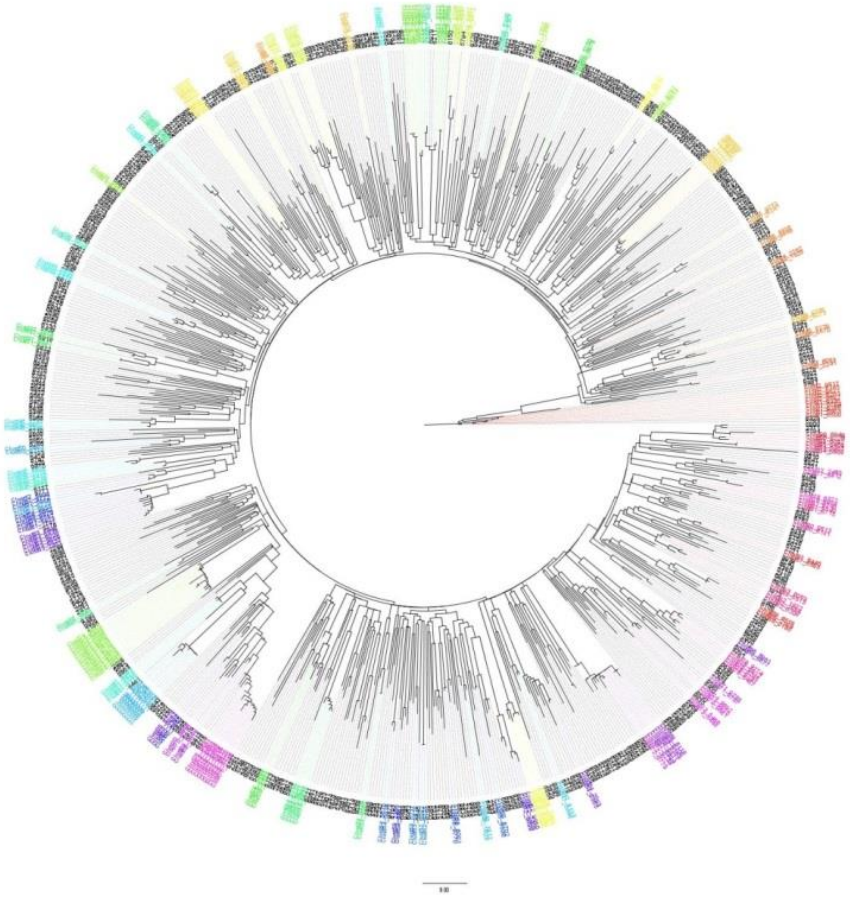
Maximum likelihood tree showing transmission clusters identified using ClusterPicker are labelled in colour.

Figure 6.3 Phylogenetic tree of subtype B HIV sequences sampled between 2008 to 2011



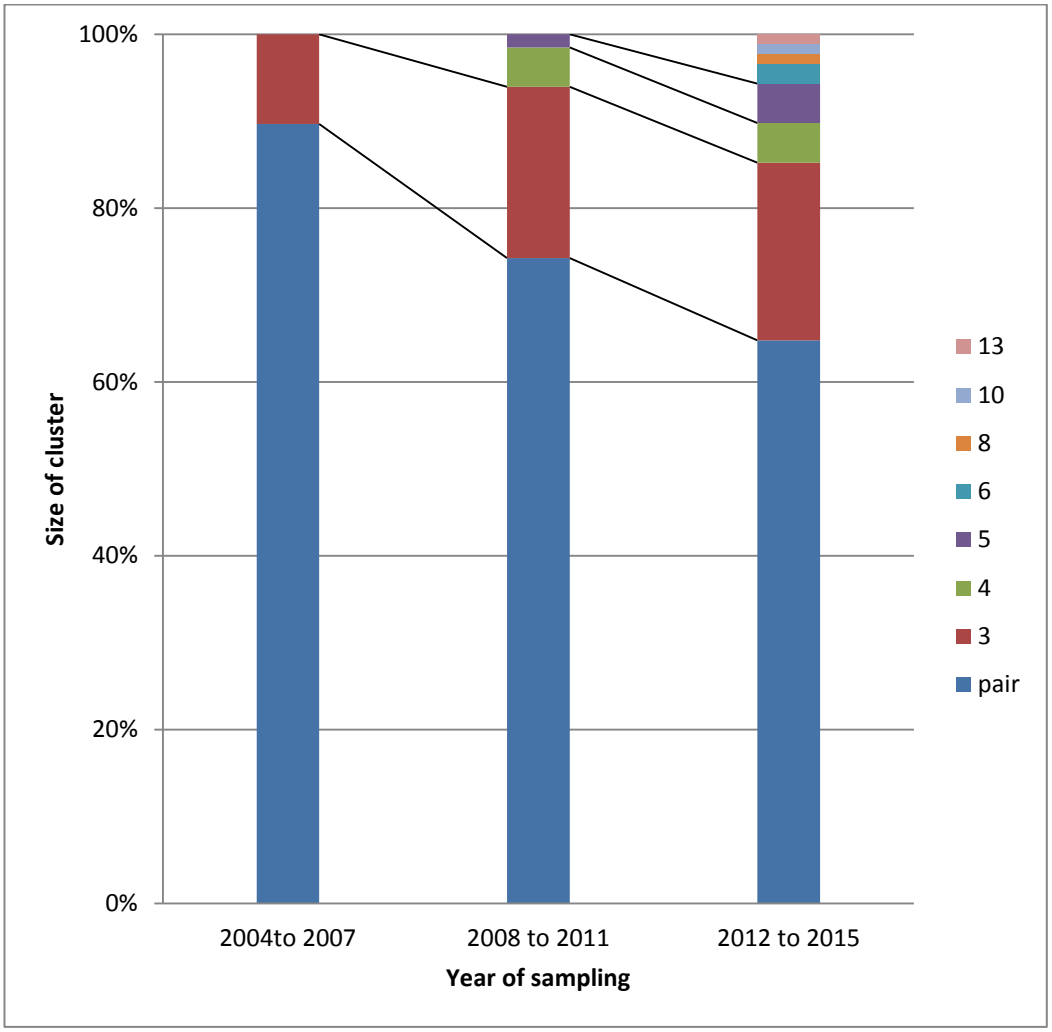
Maximum likelihood tree showing transmission clusters identified using ClusterPicker are labelled in colour.

Figure 6.4 Phylogenetic tree of subtype B HIV sequences sampled between 2012 to 2015



Maximum likelihood tree showing transmission clusters identified using ClusterPicker are labelled in colour.

Figure 6.5 Proportion of sequences in a cluster based on cluster size



The two largest clusters contained 10 and 13 members respectively. A summary is shown below in Table 6.2.

Table 6.2 Description of two largest clusters

	Cluster #1 n=13	Cluster #2 n=10
Male sex %	92	80
Age at diagnosis %		
15 to 29 years	23	0
30 to 39 years	38	0
40 to 49 years	38	10
> 50 years	0	90
Risk Group %		
MSM	100	90
Other or unknown	0	10
Symptoms of seroconversion%		
yes	23	40
no	77	60
Region of birth %		
European Region	8	10
South-East Asia Region	15	10
Western Pacific Region	77	80
Region of residence		
Metro	92	80
Rural and regional	8	20
Place of acquisition		
Unknown	8	20
Australia	92	80

6.4.3 Factors associated with cluster membership for sequences sampled between 2012 and 2015

For the period of interest between 2012 and 2015, baseline demographics are summarized below in Table 6.3. Univariate analysis was performed on these sequences to determine factors associated with cluster membership during the period of the new HIV strategy implementation. On univariate analysis, younger age, Australian place of acquisition, early stage of diagnosis, seroconversion illness, laboratory evidence of primary HIV infection, negative test within 12 months were more likely to be associated with membership of a cluster ($p<0.001$). On multivariate

analysis, being young, recently diagnosed with an earlier stage of infection acquired in Australia was more likely to be associated with cluster membership (Table 6.4).

Table 6.3 Baseline demographics for individuals sampled 2012 to 2015

	n	%
Sex		
Male	717	94.3
Age At Diagnosis		
15 to 29 years	207	27.2
30 to 39 years	255	33.6
40 to 49 years	161	21.2
50 and over	137	18.0
Reported risk		
MSM	654	86.1
Heterosexual sex	58	7.6
PWID	21	2.8
Other or unknown	27	3.6
Symptoms of seroconversion	222	29.2
WHO Region of Birth		
African Region	8	1.1
Eastern Mediterranean Region	10	1.3
European Region	61	8.0
Region of the Americas	46	6.1
South-East Asia Region	40	5.3
Western Pacific Region	562	74.0
Unknown	33	4.3
Region of residence		
Metro	640	84.2
Rural and regional	84	11.1
NSW not specified	17	2.2
Outside NSW	19	2.5
Place of Acquisition		
Unknown	201	26.5
Australia	454	59.7
Overseas	105	13.8

Table 6.4 Univariate and multivariate analysis of factors associated with cluster membership for sequences sampled 2012 to 2015

Variable	Freq. (n)	Percent (%)	Odds Ratio	p value	LCI	UCI	Overall p value	Odds Ratio	p value	LCI	UCI	Overall p value
Age at diagnosis							0.024					0.029
15 to 29 years	207	27.2	1.00					1.00				
30 to 39 years	255	33.6	0.74	0.13	0.50	1.09		0.72	0.13	0.47	1.10	
40 to 49 years	161	21.2	0.80	0.30	0.52	1.22		0.77	0.27	0.47	1.23	
50 and over	137	18.0	0.56	0.02	0.35	0.89		0.53	0.02	0.32	0.90	
Risk group							0.036					
MSM	654	86.1	1.00									
Other	106	14.0	0.60	0.04	0.37	0.97						
Region residence							0.034					
Metro	640	84.2	1.00									
Rural and regional	84	11.1	0.51	0.02	0.30	0.89						
NSW not specified	17	2.2	0.40	0.16	0.11	1.42						
Outside NSW	19	2.5	0.50	0.23	0.16	1.53						
Place of acquisition							<0.001					<0.001
Australia	454	59.7	1.00					1.00				
Overseas	105	13.8	0.28	0.00	0.16	0.48		0.29	0.00	0.17	0.51	
Unknown	201	26.5	0.30	0.00	0.20	0.44		0.59	0.03	0.37	0.94	
Stage at Diagnosis												<0.001
Early	347	45.7	1.00									
CD4 500+	63	8.3	0.57	0.06	0.32	1.01		0.59	0.10	0.32	1.11	

CD4 350 to 499	94	12.4	0.65	0.07	0.40	1.04		0.60	0.05	0.36	1.00	
CD4 200 to 349	85	11.2	0.46	0.00	0.27	0.78		0.39	0.00	0.22	0.67	
CD4<200	171	22.5	0.22	0.00	0.14	0.36		0.27	0.00	0.16	0.44	
Seroconversion illness							<0.001					
yes	222	29.2	1.00									
no	538	70.8	0.60	0.00	0.44	0.84						
Lab evidence of PHI							<0.001					
yes	71	9.3	1.00									
No	689	90.7	0.46	0.00	0.28	0.75						
Tested within 12 months?							<0.001					
yes	228	30.0	1.00									
no	265	34.9	0.60	0.01	0.42	0.86						
unknown	267	35.1	0.25	0.00	0.17	0.37						
Year of onset of HIV							<0.001					
before 2012	176	23.2	1.00									<0.001
2012 to 2013	300	39.5	6.50	0.00	3.75	11.27		5.99	0.00	3.28	10.92	
2014 and 2015	284	37.4	5.65	0.00	3.25	9.85		5.24	0.00	2.82	9.73	

Abbreviations: PHI = Primary HIV infection LCI= lower confidence interval, UCI= upper confidence interval

6.5 Discussion

This is the first state-wide phylogenetic analysis systematically linked to HIV notifications in NSW. Transmission clusters accounted for up to one third of sequences sampled. Factors associated with cluster membership were age younger than 30, acquisition in Australia, diagnosis post- 2012 and early infection stage. Laboratory defined PHI was not independently associated with cluster membership.

The size of clusters has increased in more recent sampling periods, which may be due to a number of factors. Firstly, there is a greater sampling of the population living with HIV in recent years, secondly there is more comprehensive diagnostic testing at a population level and thirdly, resistance testing is now standard of care immediately at diagnosis [217, 285]. As testing and treatment components of the NSW HIV strategy have been rolled out, there are a greater proportion of the population presenting for HIV testing engagement in care and commencing treatment [286]. A recent analysis performed in 2016, immediately after this study period, demonstrated that New South Wales has achieved the UNAIDS 90-90-90- targets [217]. As a result this population was better represented in our analysis in later years (e.g. 2012 to 2015). Nevertheless the size of transmission clusters has increased in recent years, with the largest clusters containing up to 13 members.

Although primary HIV infection, identified by the presence of symptoms of seroconversion with strict laboratory based criteria was associated with cluster membership on univariate analysis, these factors did not remain significant on multivariate analysis. A reason for this may be collinearity of the “early infection” variable that would have incorporated those with primary HIV infection. However this is a much broader categorisation that includes individuals infected up to 12 months from seroconversion. The final multivariate analysis did not support the hypothesis that primary HIV infection (as defined by either seroconversion symptoms or laboratory criteria) was associated with cluster membership, as has been found in other similar populations in the state of Victoria, Australia [79]. This could be as a result of increased awareness around primary HIV infection leading to earlier diagnosis, earlier initiation of ART and effective contact tracing – this has only occurred in recent years, and may not have been as effective during an earlier time period between 2005 to 2010 when the seroconverter analysis was performed in Victoria [79].

The findings presented here are consistent with other studies conducted in settings of mature epidemics that showed only a small proportion of new infections are attributable to PHI [287]. Over a half of new infections were found to be from people who were unaware of their status [288] Other studies have suggested that the diagnosis and treatment of PHI can lead to decreases in onward transmission of up to 89%, due to the high degree of viraemia contributing to infectivity [220, 289, 290]. Unfortunately, our study did not include viral load data, which has been shown to be strongly correlated between transmission pairs [291]. As our study period of 2012 to 2015 was at a time of rollout of increased testing and treatment as prevention strategies, it is possible that levels of high viraemia that lead to onward transmission were significantly reduced by these strategies. For example, there was an increase in the proportions of newly diagnosed on ART within six weeks of diagnosis from 48% to 65% from 2013 to 2015, with up to 93% having a suppressed viral load by 6 months [261].

The main characteristics associated with cluster membership were in younger age groups, Australian acquisition, earlier stage of infection at diagnosis and more recent sampling year. These factors are similar to those identified in a national study of HIV molecular epidemiology, which did not include all sequences from NSW [86]. This may be accounted for by a number of factors: firstly, scale up of public health testing campaigns may have resulted in more diagnoses at earlier stage, in more recent years. Secondly, as a result of new diagnoses, enhanced contact tracing efforts may have led to the identification of other cases that clustered together to be recognised and diagnosed around that same period. Understanding factors around cluster membership can provide greater insights into reasons for ongoing transmission and determine whether past interventions have been effective.

The limitations of this study are the proportion of sequences linked via data linkage without use of full name identifiers, as described in Chapter 2. This was addressed by demonstrating that the linked sequences were representative of the population diagnosed with HIV in NSW (see Chapter 2), and that analyses performed here can be applied to this population. Despite this, no phylogenetic analysis will ever be “complete” as there always will be members of the population that have not been sampled because they have not been diagnosed or not sequenced. To address this for future analyses, the linkage process has been modified to enhance the proportion of sequences that link to the HIV notifications database.

The methods used to define cluster membership were based on similar studies that use a combination of genetic distance and maximum likelihoods. A recent review paper suggested that the most common method for defining transmission clusters was the maximum likelihood method, using general time reversible substitution method with bootstrap support of >90% [158]. This study used a lower threshold of 70% to define clusters, due to the limited sampling of this dataset, but a more stringent genetic distance threshold of 0.015 was used in this definition. As the main research question was the contribution of primary HIV infection in transmission clusters over a relatively short time period, a shorter genetic distance threshold is preferred and this combination is accepted as a reasonable approach for most datasets [158, 292].

This study has provided baseline data on transmission cluster formation among HIV-1 infected individuals in NSW, and will form the background for comparison for future studies. Ongoing linkage on an annual basis will provided updated datasets to allow the characterisation of transmission clusters over time, and evaluate the impact of contemporaneous public health strategies aimed at preventing HIV transmission.

In summary, clusters of HIV-1 were found in less than one third of sequences sampled for resistance testing in 2004 to 2015. Of those that belonged to a cluster, younger age, Australian acquisition, earlier stage of infection and more recent sampling year were associated with cluster membership, suggesting that testing campaigns have correctly identified at risk populations, and their contacts to be diagnosed and engaged in care. In contrast to other states, those with primary HIV were not found to be significantly associated with cluster membership in this population.

Acknowledgements

The following people and institutions have contributed to this work and are gratefully acknowledged:

1 The Kirby Institute - Francesca Di Giallonardo, Ansari Shaik, Phillip Keen, Andrew Grulich,

2 Royal Prince Alfred Laboratory, NSW health pathology- Hanan Salem, Frederick Lee

3 NSW State Reference Laboratory for HIV/AIDS & Molecular Diagnostics- Alex Carrera

4 ICMPR Westmead - NSW Health Pathology- Kiran Thapa, Dominic Dwyer

5 Health Protection NSW- Christine Selvey, Barbara Telfer

6 Positive Life, Sydney – Craig Cooper

7 ACON, Sydney- Karen Price

8 NSW Ministry of Health- Jo Holden

9 Steering group of the NSW HIV Prevention Partnership Project

10 Centre for Health Record Linkage- Elizabeth Moore

Funding

This study was partly funded by an NHMRC postgraduate research scholarship APP1074467.

This analysis was supported by funding from an NHMRC Partnership Project grant (Grant #1092852), the NSW Ministry of Health, and UNSW Sydney. The Kirby Institute receives funding from the Australian Government Department of Health and is affiliated with the Faculty of Medicine, UNSW Sydney. No pharmaceutical grants were received in the development of this study.

6.6 Appendix

6.6.1 Risk benefit analysis

This document was prepared with input from representatives of key community organizations Positive Life (Craig Cooper, Chief executive officer) and the AIDS Council of NSW (Karen Price, Deputy Chief executive officer). It was tabled at steering committee working group meetings and distributed at a community "Treatments and Research" evening May 3rd 2018. This event was held to educate people living with HIV about research and clinical findings. This was followed by a facilitated group discussion so people have an opportunity to seek further detail and ask clarifying questions.

The HIV prevention revolution: Using the molecular epidemiology of HIV transmission in NSW to inform the Public Health Response to HIV prevention

NSW has the largest burden of HIV in Australia. While annual HIV diagnoses in NSW have been stable over the last decade, an opportunity now exists to use new scientific knowledge about biomedical HIV prevention strategies to significantly reduce new HIV transmissions.

The *NSW HIV Strategy 2012-2015* has set a target of an 80% reduction of HIV incidence by 2020[166]. To achieve this target, the HIV prevention response needs to be focused and flexibly respond to changes in HIV transmission dynamics. Molecular epidemiology provides a tool to identify transmission networks, but needs to be paired with reliable clinical and epidemiological data for full benefit. This study uses data linkage of routinely collected health data (HIV epidemiology and laboratory data) to inform the HIV prevention response in NSW. This study will identify characteristics of transmission, allowing prioritised public health responses. It will focus elements of the

NSW strategy to advise whether to “test more” or “treat more” by understanding what the drivers of continued HIV transmission are.

What is being proposed?

HIV surveillance data has been collected via the HIV administrative database since 1998. In this database, all new infections are notified using a 2X2 name code, DOB, sex, country of birth and postcode (no full names are contained in this dataset).

Simultaneously, all patients diagnosed with HIV now receive baseline HIV drug resistance testing, and this test involves the sequencing of the HIV virus at one of three reference laboratories in NSW.

We propose linking these two already available datasets, then removing any personal identifying information to create a larger de-identified dataset that can be analysed from a population health perspective.

What are the benefits of this study?

This study can provide important and timely information to help evaluate the NSW HIV Strategy using data that is already available. By providing data on the characteristics of new HIV transmissions, it will allow assessment of the success of specific elements of the current HIV strategy. It has the potential benefit to public health policy making by guiding the strategic allocation of resources towards areas that will have the highest impact: for example, whether more action is needed in testing programs, treatment rollout or pre exposure prophylaxis.

Other data from this study could improve clinical decision making for people living with HIV. Through improved knowledge of drug resistance patterns, the most efficacious antiretroviral regimens could be recommended up front. Ultimately the goal is to eliminate HIV transmission in NSW, but this can only be achieved by understanding transmission patterns first.

What are the risks of this study?

This study uses data that is already being collected within reference laboratories and NSW Ministry of Health. While this study may identify clusters of HIV transmission, the final dataset will not contain any name codes or DOB to allow re-identification of individuals.

What are the expected outcomes of this study?

The types of data that will be obtained from this study will include:

- the proportion (%) of new infections that arise from recently infected people
- the proportion (%) of new infections from those not on treatment
- the proportion of new infections acquired overseas
- the proportion of drug resistant strains among new infections

Risks: to people living with HIV (PLH) and high risk populations

1. Potential implications for privacy through re-identification of data
   The design of this study is such that the final dataset will not contain any personal identifiers to allow tracing of an individual. In the preparatory steps of this final dataset, breaches in several layers of de-identification and security would be needed to potentially allow identifiers (name code, DOB, postcode, sex, country of birth) used in data linkage to be re-identified.

Risk assessment: The data being used are already routinely collected and stored by reference laboratories and Ministry of Health, and hence the risk would be the same as currently exists for data collection of HIV notifications, and routine laboratory testing which is already in existence. This risk of data re-identification is very low as it would require access and breach to several separately housed datasets including:
i) Ministry of health HIV administrative dataset
ii) Laboratory identifier dataset housed in three separate reference laboratories
iii) Laboratory HIV sequence data obtained from reference laboratories and stored separately in a deidentified database at the Kirby Institute
iv) UNSW Australia network
v) Centre for Health Record Linkage

There has never been a breach of these datasets and the integrity of these data collection facilities remains intact.
Risk mitigation: Major safeguards will be taken to endure that the risk of any breach of privacy is minimized. For example:

- The final linked datasets *without* identifiers such as names, name codes or DOB will be received by Kirby Institute, and transferred to a secure firewall protected network.

- Individual files will be password protected.
- All results will be published in a form that will not allow individuals to be identified, that is, in tabular or graphical aggregate form only; no individual results will be disclosed. Only aggregate data by region/local health district (LHD) at site of testing will be reported.
- Once the de-identified linked data file is received by the researchers at the Kirby Institute, all study data will be transferred to a secure network protected by a firewall and individual files will also be password protected.
- Any potential identifiers such as postcode will be modified or presented in aggregate (eg age brackets 20-25, 26-30, 31-34, or local health district) so that potential for reidentification is eliminated

2. Legal subpoena of sequence data for a linked transmission case
   Risk assessment : This request could potentially be made *right now* directly to any reference laboratory, where sequence data is currently tested under *full patient name*, DOB, sex and address for Medicare rebate.

It is considered that the final dataset for this study will in fact be at lower risk of legal subpoena than HIV sequence data currently held by laboratories, as the final dataset will have identifiers (Name code, DOB) removed. Any conclusions drawn from this phylogenetic analysis *cannot* in isolation be used to infer direct linked transmissions at the standard required for robust forensic interrogation or even the direction of those transmissions. Given the dataset will contain over 10 000 sequences, identification to the individual level will not be possible. Hence this study is designed to provide aggregate population level data and do not provide any data on the likelihood that a particular transmission was linked to a specific individual nor will it provide information regarding the direction of the transmission. Similarly this study is not designed to be a complete transmission network of HIV in NSW, but only reflective of PLHIV engaged in care or on treatment that have had GART testing (54-80% of PLHIV) [166]. Hence it is not possible to make forensic assumptions about individual cases of linked transmission from this data, as the dataset will be incomplete.
Previous studies such as Opposites Attract were thought by some to create risks of potential prosecution in relation to HIV transmissions by PLHIV. Legal advice was

sought prior to commencing these studies, and similar legal arguments will be applied here.

Risk mitigation: To completely mitigate the risk of legal subpoena, the data linkage process is designed to remove any link between laboratory identifiers and HIV notification identifiers (Namecode, DOB) within the realms of NSW Health. Components of the data that include patient identifiers will then be irrevocably severed through the data linkage process. This will make it inaccessible to subpoena in the event of future legal proceedings.

3. Data linkage unsuccessful

   Risk assessment: As limited identifiers will be used for data linkage, the sensitivity of the study can be compromised. Based on previous similar data linkage studies, sensitivity of linkage using name code, DOB and sex is between 77-85% (email communication, Kathy Petoumenos, Kirby Institute). We expect this linkage to be higher, as it is linking directly to HIV notifications database, not larger health datasets like NSW Emergency Department Data collection (EDDC) or NSW Admitted patient data collection (APDC). Additionally, these previous studies were based on enrolled cohort studies where false names may have been used. Given that routine testing is performed using Medicare card (for rebate), it is more likely that accurate identifiers have been recorded.

   There will be further limitations with a subset of sequences tested at Sexual health clinics (~17%) where coded identifiers are used. Hence only DOB, Sex, and postcode can be used to perform linkage, limiting the sensitivity of results.

Risk mitigation: The investigators have sought the advice of those experienced in HIV data linkage studies using 2X2 name codes and consult with CHeReL in the feasibility, design and conduct of this study.

Benefits: outcomes and how they relate to objectives of NSW HIV strategy

Public health benefits

1. Enhancement of routinely collected data

   The *NSW HIV Strategy 2012-2015* prioritises continued investment in surveillance and research to inform the HIV response [166]. The data required for this study are already being collected routinely; hence it would be a lost opportunity not to enhance this dataset with epidemiological data to draw

meaningful conclusions that inform public health responses in the implementation of the NSW HIV Strategy.

2. Assess the impact of NSW HIV strategy and guide policy making
Several other regions have already employed data linkage to HIV notifications to track epidemic and inform public health response. This proposal is similar to numerous recently published in similar populations interstate and overseas, including developing countries [174, 175, 254, 293-295]. NSW however has the greatest burden of HIV and has shown great leadership in HIV policy over the years. This is a unique and timely opportunity to continue this leadership and research the impact of the NSW HIV Strategy (2012-2015), in particular on policies focusing on PrEP access, increased testing, earlier treatment and increasing treatment uptake in NSW. Research outcomes from this study should be able to provide answers about what is driving the epidemic in NSW- early undiagnosed infections or untreated chronic infections- and inform the use of the different public health strategies that are required to address each.

3. Informing treatment guidelines
Resistance surveys as extensive as this have not been performed before in NSW. Retrospective resistance data has been collected however is limited in the lack of paired clinical and epidemiological data. Monitoring of drug resistance is important and extremely timely in an era of rollout of TasP and PreP. PLHIV will also directly benefit from this knowledge of population based resistance, as it may influence future antiretroviral choices.

4. Describing clusters and eliminating transmission
Molecular epidemiology provides a way to understand key features of the epidemic at a population level, which can then inform the public health response.
These analyses can be used as part of the evaluation of the effects of the new interventions by NSW government and how these might be strengthened or modified to give greater efficacy. To date, no studies have systematically examined the patterns of transmission in NSW, due to lack of a centralised sequence database paired with epidemiological data. This may be a reflection of lower than estimated testing levels or that new transmission are acquired from outside Sydney. Knowledge of these types of transmission patterns is

essential to guide the public health response.

5. Greater knowledge of subtype diversity
   While historically NSW has been a predominantly subtype B epidemic, recent data has shown that up to 50% of new infections in Western Sydney are non-B subtype [254]. Knowledge of transmission characteristics of these subtypes can help guide public health campaigns identifying and prioritising non-traditional risk groups, in particular understanding heterosexual transmission, and the proportion that may have occurred overseas.

Individual benefits

1. Improved clinical decision making for people with HIV
   Through real time knowledge of HIV drug resistance rates, clinicians will be able to choose the most efficacious regimen for an individual with the least likelihood of resistance, and preserving other drug classes for future use.

2. Better understanding the impact of PrEP on population level resistance
   The timing of this study coincides with the rollout of and implementation of PrEP, and will be able to provide data on impact of this strategy on transmission, as well population level drug resistance. This will provide benefits to both PLHIV as well as at risk groups.

3. Resource allocation
   This study will provide evidence to support enhanced testing and/or treatment programs as a tool to decrease HIV transmissions.

4. Understanding the demographic and clinical profile of transmission clusters
   Through knowledge of cluster involvement, a better understanding of risks for transmissions will be gained. At risk groups will benefit from knowledge of whether new transmission is a result of new undiagnosed infections, or untreated infections.

6.6.2 Community information pamphlet

New research to help reduce HIV transmission

NSW at the forefront of HIV research and prevention

NSW has a record of HIV research and prevention approaches that have engaged the HIV community as active partners. The NSW HIV Strategy aims to reduce HIV transmission among gay and other homosexually active men by 80% by 2020. To achieve this we need to understand finer aspects of the epidemic such as how new HIV infections are acquired. Are they coming from people living with HIV that are yet to start treatment, or are they coming from those who are infected but yet to be tested and diagnosed? Each of these situations requires a different approach: either to treat people earlier or to test people with undiagnosed HIV. New research methodologies and technology now allows further detailed understanding of the epidemic.

Using available data to understand more about HIV transmission

HIV is a notifiable condition, which means that whenever someone is diagnosed, some data is collected from the doctor who made the diagnosis by the NSW Public Health Unit. All HIV notifications are strictly confidential - no full names or addresses are contained in this dataset.

Similarly, whenever someone living with HIV first sees a doctor, blood is collected to find out more about the type of virus and whether it contains any resistance mutations. This information is important to decide on the best antiretroviral treatment for that person. But combined together, these genetic sequences of HIV could be used to create a "family tree" of viruses.

This is called phylogenetics, which is a way to assess how closely some viruses are related. This information can help us understand how HIV is being transmitted in clusters and sub-populations to help tailor future prevention and treatment strategies.

Greater understanding through data linkage

Through linkage of these two already available datasets, we can understand key features about HIV in NSW. As these two data sources are combined, all personal information is removed. We hope to answer: 1)What proportion of HIV transmission are from newly acquired versus untreated long standing infections? 2) What is the rate of HIV treatment drug resistance in NSW? 3) What are the common subtypes of HIV and which geographic regions do they originate from?
This information can benefit both people living with HIV through better knowledge of HIV treatments, as well as those at risk of HIV, by better understanding transmission. This study may also help inform public health education and prevention efforts.

CHAPTER 7

SUMMARY AND CONCLUSIONS

This research was done to understand various aspects of PHI in the NSW context, encompassing individual clinical outcomes to population level public health surveillance of transmitted drug resistance and molecular epidemiology.

The broad objectives were to understand aspects of PHI to help inform clinical and public health policies in NSW, with research findings having implications for clinical management paradigms, surveillance programs, informing public health interventions and empiric treatment guidelines. The overall research findings have been outlined in relevant chapters. This chapter will discuss broader implications of this research for both NSW and in the global context.

7.1 Mortality in PHI

Chapter 3 evaluated the long term health outcomes of a cohort of PHI. Long term outcomes in HIV are approaching that of the general population in the post HAART era. Comparable data in PHI cohorts are limited as cohorts of PHI are challenging to recruit. Therefore existing cohorts are a valuable resource to evaluate long term health outcomes.

This research showed that AIDS as a cause of death in PHI cohort is decreasing, while non-AIDS events like cancer and cardiovascular disease are predominant causes of death. Similar to studies in other large international cohorts published during the course of this thesis, AIDS still remains the predominant cause of death [77, 296]. However this study was the first to report these findings in a cohort of PHI.

Treatment of HIV has been shown to have public health benefits due to decreased risk of onward transmission, and was advocated as part of the treatment as prevention strategy [173]. This benefit has also been demonstrated in a PHI cohort, with reduction in onward transmission shown with immediate ART [289]. Whether there is a benefit for the individual with PHI has not been clearly demonstrated. Early treatment has been shown to have definite mortality benefits for people living with HIV, with results of the START study published, and adopted into guidelines during the course of this thesis [83, 297]. The limitations of the START study are that the median time from HIV

diagnosis was one year, with an unknown timing of acquisition of HIV; hence there is limited applicability of these findings to the setting of PHI.

The results of Chapter 3 have shown that there is a mortality benefit with early treatment (within one year from PHI), adding further weight and narrowing the time point to starting therapy to as close as possible to acquisition of the virus. This research supports the current test and treat paradigm for individual as well as for public health benefit.

Although these research findings will not impact upon current practice of immediate treatment upon diagnosis, it does add support to the weight of evidence of benefit for the individual with PHI. A national study in Australia showed that there still remained opportunities for earlier diagnosis in the Australian healthcare setting, which could lead to earlier time to treatment and improved clinical outcomes[298]. If PHI is correctly identified and treated, there may be substantial mortality benefits in this group. The START study also reported favourable quality of life outcomes in those who received immediate ART, but this is not yet confirmed in the PHI setting. This is a potential area for further research in PHI cohorts, particularly given the longer duration of treatment over the course of a lifetime in this group [299].

Several studies have shown short term immunological benefits of treatment during PHI [42, 44, 49, 230]. This research raises further questions around the immunopathological basis for the longer term mortality benefits seen with early treatment. During the course of this thesis, further research has been published evaluating the impact of antiflammatory approaches in the setting of HIV. The ADVICE study evaluated the role of proteinase activated receptor-1 (PAR-1) inhibition on D dimer concentrations in those with stable HIV, but found no effect on this biomarker or other markers of activation [300]. Other research has evaluated the role of early ART and reservoir size. Although reduction in reservoir size has been shown with early ART, no change in T cell activation was achieved in those treated with PHI as compared to those treated during chronic infection [301, 302]. An improved understanding around the immunological events in PHI that impact upon long term health outcomes such as cardiovascular events, inflammation and cancer is needed.

7.2 Transmitted and acquired drug resistance

The aim of chapters 4 and 5 was to evaluate the rate of drug resistance in NSW. As there is a move towards earlier initiation of treatment, it is imperative to know population rates of transmitted drug resistance (TDR). In NSW there is no surveillance

of drug resistance mutations, so to evaluate TDR, a state-wide database needed to be established. Chapter 4 described the setup of a state-wide drug resistance database containing over 7000 sequences, and showed the declining rate of overall prevalent drug resistance mutations. This decline was most marked in those that were treatment experienced, reflecting availability of better tolerated antiretroviral therapy with less toxicity. Rates in the treatment naïve population have remained stable.

The specific research question around drug resistance mutations in PHI was addressed by evaluating a subpopulation of seroconverters, where the most prevalent mutations were identified as M41L and T215CS, V106I and V90I. These mutations such as thymidine associated mutations (TAMs) are associated with treatment with older NRTI regimens, for example, prior AZT use. T215 revertant mutations (eg T215C/D/E/S) while not conferring any increase in drug resistance, may confer increased risk of failure to these regimens in the future due to rapid conversion to a resistance mutation [148]. The NNRTI mutation V106I may occur in between 1-2% of treatment naïve individuals and mainly has implications for treatment failure if associated with other specific mutations [148]. The TAM mutations seen were similar to that reported in a seroconverter study in Germany[303].

These findings suggested a legacy effect of circulating mutations within the populations from historic regimens that are still being transmitted to newly infected individuals with PHI. Australia has a track record of an early and effective response to the HIV epidemic by providing universal healthcare and free access to antiretrovirals. This may have resulted in early antiretroviral exposure and development of drug resistance mutations in the population. During the course of this thesis, data from the START study were published showing that Australia had the highest rates of resistance in the world[110] which are in stark contrast to the levels found in the studies resented in this thesis. The patients enrolled in START however were sampled during the earlier years of the study, and from a single centre so may have represented a selection bias [112]. Nevertheless, these published studies illustrate the importance of collating contemporaneous and representative surveillance data for drug resistance mutations. This study is the first of its kind in NSW to include all three reference laboratories of all individuals in the state sampled for drug resistance testing.

Chapter 5 builds on this dataset through linkage to the HIV notification database to include demographic and epidemiological data. This allowed analysis of factors associated with transmitted drug resistance to be identified. People residing in non-metropolitan regions and those aged 19-29 years had the highest odds of pre-

treatment drug resistance. Results from this thesis provide more comprehensive data at a state-wide level on contemporaneous rates of drug resistance from geographically disparate parts of the state. These findings have implications for better linkage to care and services for those in non-metropolitan regions of NSW. As other national studies have shown a changing epidemic in Australia, with increasing non–B subtypes and heterosexual transmission, it is imperative that services and prevention strategies are targeted at geographically diverse regions within NSW, rather than just in the inner city MSM neighbourhoods where the epidemic was previously concentrated [86, 304, 305].

7.3 Role of PHI in transmission in NSW

PHI has been reported to be a significant factor in onward transmission of HIV in other states in Australia and globally [79, 80, 306]. The objective of this research was to determine whether PHI was associated with transmission clusters in NSW. Research in this area has been challenging due to legal and ethical implications associated with the phylogenetic analysis of HIV. Chapter 6 describes the first comprehensive molecular epidemiological analysis of its kind in NSW and complements work done in other states in Australia [86, 197].

Although laboratory defined PHI was associated with cluster membership on univariate analysis, this was not significant on multivariate analysis and could not be considered a major driver of onward transmission. Acquisition in Australia, diagnosis post- 2012, early infection stage (but not laboratory defined PHI) were found to be factors associated with cluster membership in NSW.

The methodology used to define transmission clusters varies throughout the literature, but there is consensus for the bootstrap thresholds used in this research. A similar study has recently been published in a multinational PHI cohort, which also used maximum likelihood methodology and similar thresholds to define transmission clusters. As in this research, transmission clusters were more likely to be seen in recent diagnoses, rather than older more established ones [307]. This could also reflect the results of effective contact tracing, as recently diagnosed individuals allow the identification of other contacts within their transmission network to also be diagnosed and treated. Research such as this is imperative in informing and evaluating public health strategies.

Phylogenetics has been increasing used in public health to complement epidemiological surveillance and socio-behavioural factors in gaining a deeper understanding of an epidemic [308]. Phylogenetic analysis can assist in managing a

response to an epidemic such as HIV by providing early recognition of a cluster or outbreak of transmission within a population. At a multijurisdictional level, phylogenetic analysis can provide insights into populations with high degrees of migration and mobility that can benefit from targeted interventions, particularly where resources are scarce [309].

Poon *et al* demonstrated the implementation of real time phylogenetic monitoring which was automated, cost effective and triggered an enhanced public health response [162]. A growing cluster with transmitted drug resistance was identified, cases followed up with immediate linkage to care and treatment, and a reduction in onward transmission was shown. Balanced against this enhanced tracking of outbreaks that provides greater resolution of characteristics of at risk populations, is potential threats to privacy for the individual that may ensue [180, 310]. At a programmatic level, large scale phylogenetic analysis can be used for monitoring and evaluation of a public health strategy and the effect of particular interventions such as PrEP in halting cluster growth over time. For this to be most effectively utilised, there needs to be dense sampling of the population in question and short timelines from infection to diagnosis, to allow interpretation of transmission dynamics and implementation of preventative interventions [311, 312].

During the course of preparation of this thesis there has been a significant decline in new diagnosis seen in MSM in NSW. Reasons for this are attributed to rollout of PReP as well as a combination of other strategies including increased testing availability and treatment uptake [217, 313]. With declining rates of new diagnoses, a deeper understanding of transmission dynamics is needed through phylogenetic reconstructions, with the hope of identifying unrecognised risk factors for subpopulations that drive transmission, and potential groups that may benefit from targeted prevention strategies such as PrEP [314].

Implications

Cohorts of PHI such as described in our thesis are unlikely to be replicated again. In the era of test and treat, individuals diagnosed with HIV are offered immediate ART in line with current evidence and guidelines. Although cohorts may continue to be recruited and followed long term, the ability to recruit comparitor groups with delayed antiretroviral treatment is not ethically possible in the current era of HIV management. This research has implications for clinical management of PHI, with the demonstration of individual mortality benefit from early ART. It also has broader public health implications strengthening the argument for treatment as prevention.

As immediate treatment is now the paradigm, there is a need to ensure the initial treatment regimen remains effective through the identification of prevalent and transmitted drug resistance mutations. This research also supports the test and treat paradigm by ensuring local empiric treatment guidelines remain effective with the least chance of virological failure from known circulating drug resistance mutations. Although PrEP failure is rare, vigilance is still needed to monitor TDR at a population level [315-318]. In the era of widespread PrEP use, surveillance of transmitted drug resistance is essential to recognise the transmission of multidrug-resistant strains and to inform optimal PrEP regimens.

It appears that the public health messaging in NSW has been effective in terms of early testing, and contract tracing to identify undiagnosed infections, as the factors identified with clusters were more recent diagnoses and earlier infection stages. Research such as this helps to evaluate preventative public health interventions and inform whether resource intensive strategies such as enhanced contact tracing are effective. Early identification and treatment of individuals with PHI are beneficial from both individual and public health perspective.

Limitations

A more detailed discussion of the limitations is presented in each chapter. The main limitations relate to the data linkage methodology required for Chapters 2, 5 and 6.

The research presented in Chapter 2 relied upon data linkage to death notifications. While the rate of linkage was suboptimal, it was comparable to other HIV linkage studies that did not use full name and date of birth, and was considered acceptable given the long follow up period. It should also be noted that data linkage was only used to supplement existing cohort data collection due to the long period of follow up and was not the sole basis for determining mortality outcome. As our cohort was comprised mainly of gay and bisexual men who have been shown to have higher rates of age related comorbidities, the inclusion of a control groups may have strengthened our findings [236, 319, 320].

For data presented in Chapters 5 & 6, while rate of linkage was not optimal, we determined that the sampling for this study was representative of the population of NSW based on demographic and epidemiological criteria (shown in chapter 2). This is one of the largest sized population based phylogenetic studies that contained epidemiological metadata. A recent review reported that in 2015, there was an increase in the average number of sequences/ article from 41 to 5389 [158]. This research study

exceeds this amount, and improves on coverage and sampling of the population by including sequences from routine diagnostic testing rather than from a dedicated clinical trial.

Future research directions

PHI cohorts continue to provide valuable insights into the natural history of HIV with implications for vaccine studies, cure research and immunological interventions that may alter disease progression. Further research in the field is evaluating the area of post treatment control where early treatment during PHI may result in the ability to stop ART and maintain viral suppression [19, 49, 50, 57, 66, 238]. Post treatment controllers may also provide insights for cure research, as further knowledge is gained into the basis of functional versus virological cure. While there is much interest in serious non AIDS events (SNAEs) comorbidities in HIV in general, there is a paucity of data about long term morbidity in cohorts of PHI. Particular research questions would be around identification of biomarkers or other immune parameters during PHI that can predict SNAEs, so that monitoring and interventions can be made to disease progression [321].

In Australia, there is still work to be done to establish a comprehensive national HIV drug resistance surveillance program. While there is now a national approach to reporting critical antibiotic resistance, no such program exists for antiretrovirals [322]. Currently, each state independently collates its own data, but with increasing travel and movement between jurisdictions, a national approach is needed. This is an urgently required, especially given the rapid uptake and widespread rollout of subsidized PrEP throughout Australia [323, 324].

From a methodological perspective, newer sequencing methods are being developed for the detection of drug resistance mutations. Next generation sequencing (NGS) has been shown to have a higher sensitivity compared to Sanger sequencing, particularly for the detection of minority variants. Studies have shown a lower limit of detection for mutations with low frequencies [325, 326]. Although there may be improved detection of minority variants, the clinical relevance of these mutations remains to be seen [327-331]. Work into whole genome sequencing (WGS) of HIV for drug resistance is the next frontier, with WGS being evaluated to identify minority variants in non-polymerase genes [332].

In the area of molecular epidemiology, some jurisdictions have implemented real time phylogenetic surveillance to identify and respond to transmission clusters [160, 162].

Whether this approach can be applied to other regions is an area for further evaluation. Robust data collection systems, ethical processes and the appropriate legal environment are needed for this research to be further developed in Australia. While a national study has been published, more contemporaneous research is needed with participation of all reference laboratories and linkage to national HIV epidemiological datasets [86].

HIV notification reports that the decline in new diagnoses was not seen in MSM born overseas, making this group a focus for future public health strategies [32]. Further ongoing studies are needed to evaluate where transmission is stemming from, in particular, developing a greater understanding of phylogeography. Additionally, as non B subtypes are increasing, the impact of regional strains such as CRF_01AE needs to be evaluated, with international molecular epidemiological collaborative research efforts.

PHI has provided many insights into the natural history of HIV, and remains a key time to study clinical, virological and public health aspects of the virus. Ongoing research into understanding clinical and immunological events, transmitted drug resistance mutations and impact on transmission dynamics can inform this field from both an individual as well as public health perspective.

References

1. Cooper, D.A.*, et al.*, *Acute AIDS retrovirus infection. Definition of a clinical illness associated with seroconversion.* Lancet, 1985. **1**(8428): p. 537-40.
2. Lavreys, L.*, et al.*, *Virus load during primary Human Immunodeficiency Virus (HIV) type 1 infection is related to the severity of acute HIV illness in Kenyan women.* Clin Infect Dis, 2002. **35**(1): p. 77-81.
3. Henn, A., C. Flateau, and S. Gallien, *Primary HIV Infection: Clinical Presentation, Testing, and Treatment.* Curr Infect Dis Rep, 2017. **19**(10): p. 1-37.
4. Fiebig, E.W.*, et al.*, *Dynamics of HIV viremia and antibody seroconversion in plasma donors: implications for diagnosis and staging of primary HIV infection.* Aids, 2003. **17**(13): p. 1871-9.
5. Joint United Nations Programme on HIV/AIDS, W.H.O., *When and how to use assays for recent infection to estimate HIV incidence at a population level*, Joint United Nations Programme on HIV/AIDS (UNAIDS), Editor. 2011.
6. World health Organization, *Technical Update on HIV Incidence Assays For Surveillance and Monitoring Purposes* in *UNAIDS/WHO Working Groups on Global HIV/AIDS and STI Surveillance* 2015: Glion, Switzerland. p. 1-36.
7. NSW Ministry of Health, *NSW HIV Strategy 2016 – 2020 Quarter 1 Jan-Mar Data Report*. 2017: Sydney, NSW. p. 1-52.
8. Morgan, D.*, et al.*, *HIV-1 infection in rural Africa: is there a difference in median time to AIDS and survival compared with that in industrialized countries?* Aids, 2002. **16**(4): p. 597-603.
9. Nsanzimana, S.*, et al.*, *Life expectancy among HIV-positive patients in Rwanda: a retrospective observational cohort study.* Lancet Glob Health, 2015. **3**(3): p. e169-77.
10. Mocroft, A.J.*, et al.*, *Survival of AIDS patients according to type of AIDS-defining event. The AIDS in Europe Study Group.* Int J Epidemiol, 1997. **26**(2): p. 400-7.
11. Smolen-Dzirba, J. and T.J. Wasik, *Current and future assays for identifying recent HIV infections at the population level.* Med Sci Monit, 2011. **17**(5): p. Ra124-33.
12. Brown, A.E.*, et al.*, *Phylogenetic reconstruction of transmission events from individuals with acute HIV infection: toward more-rigorous epidemiological definitions.* J Infect Dis, 2009. **199**(3): p. 427-31.
13. Brew, B.J. and J.Y. Garber, *Neurologic sequelae of primary HIV infection.* Handb Clin Neurol, 2018. **152**: p. 65-74.
14. Crowell, T.A.*, et al.*, *Acute Retroviral Syndrome Is Associated With High Viral Burden, CD4 Depletion, and Immune Activation in Systemic and Tissue Compartments.* Clin Infect Dis, 2018. **66**(10): p. 1540-1549.
15. Vanhems, P.*, et al.*, *Severity and prognosis of acute human immunodeficiency virus type 1 illness: a dose-response relationship.* Clin Infect Dis, 1998. **26**(2): p. 323-9.
16. Vanhems, P., C. Lecomte, and J. Fabry, *Primary HIV-1 infection: diagnosis and prognostic impact.* AIDS Patient Care STDS, 1998. **12**(10): p. 751-8.
17. Stekler, J.D.*, et al.*, *No Time to Delay! Fiebig Stages and Referral in Acute HIV infection: Seattle Primary Infection Program Experience.* AIDS Res Hum Retroviruses, 2018. **34**(8): p. 657-666.
18. de Jong, M.D., H.J. Hulsebosch, and J.M. Lange, *Clinical, virological and immunological features of primary HIV-1 infection.* Genitourin Med, 1991. **67**(5): p. 367-73.
19. Chen, J.*, et al.*, *Immunological and virological benefits resulted from short-course treatment during primary HIV infection: a meta-analysis.* PLoS One, 2013. **8**(12): p. e82461.
20. Kaufmann, G.R.*, et al.*, *Primary HIV-1 infection: a review of clinical manifestations, immunologic and virologic changes.* AIDS Patient Care STDS, 1998. **12**(10): p. 759-67.

21. Krastinova, E.*, et al.*, *Does transient cART started during primary HIV infection undermine the long-term immunologic and virologic response on cART resumption?* BMC Infect Dis, 2015. **15**(178): p. 1-10.
22. Spits, H.B.*, et al.*, *A lower viral set point but little immunological impact after early treatment during primary HIV infection.* Viral Immunol, 2015. **28**(3): p. 134-44.
23. Fauci, A.S.*, et al.*, *Immunopathogenic mechanisms of HIV infection.* Ann Intern Med, 1996. **124**(7): p. 654-63.
24. Smith, M.Z., F. Wightman, and S.R. Lewin, *HIV reservoirs and strategies for eradication.* Curr HIV/AIDS Rep, 2012. **9**(1): p. 5-15.
25. Chun, T.W., *Tracking replication-competent HIV reservoirs in infected individuals.* Curr Opin HIV AIDS, 2013. **8**(2): p. 111-6.
26. Chun, T.W.*, et al.*, *Early establishment of a pool of latently infected, resting CD4(+) T cells during primary HIV-1 infection.* Proc Natl Acad Sci U S A, 1998. **95**(15): p. 8869-73.
27. Chun, T.W. and A.S. Fauci, *HIV reservoirs: pathogenesis and obstacles to viral eradication and cure.* AIDS, 2012. **26**(10): p. 1261-8.
28. Tyrer, F.*, et al.*, *The relationship between HIV seroconversion illness, HIV test interval and time to AIDS in a seroconverter cohort.* Epidemiol Infect, 2003. **131**(3): p. 1117-23.
29. Arnott, A.*, et al.*, *High viral fitness during acute HIV-1 infection.* PLoS One, 2010. **5**(9): p. 1-12.
30. Emery, S.*, et al.*, *Randomized, Placebo-Controlled, Phase I/IIa Evaluation of the Safety and Immunogenicity of Fowlpox Virus Expressing HIV gag-pol and Interferon-gamma in HIV-1 Infected Subjects.* Human Vaccines, 2005. **1**(6): p. 232-238.
31. Fidler, S.*, et al.*, *Short-course antiretroviral therapy in primary HIV infection.* N Engl J Med, 2013. **368**(3): p. 207-17.
32. Bloch, M.T.*, et al.*, *The role of hydroxyurea in enhancing the virologic control achieved through structured treatment interruption in primary HIV infection: final results from a randomized clinical trial (Pulse).* J Acquir Immune Defic Syndr, 2006. **42**(2): p. 192-202.
33. Munier, C.M.L.*, et al.*, *The primary immune response to Vaccinia virus vaccination includes cells with a distinct cytotoxic effector CD4 T-cell phenotype.* Vaccine, 2016. **34**(44): p. 5251-5261.
34. Saag, M.S.*, et al.*, *Randomized, double-blind comparison of two nelfinavir doses plus nucleosides in HIV-infected patients (Agouron study 511).* Aids, 2001. **15**(15): p. 1971-8.
35. Goh, L.E.*, et al.*, *Study protocol for the evaluation of the potential for durable viral suppression after quadruple HAART with or without HIV vaccination: the QUEST study.* HIV Clin Trials, 2001. **2**(5): p. 438-44.
36. Krastinova, E.*, et al.*, *Is clinical practice concordant with the changes in guidelines for antiretroviral therapy initiation during primary and chronic HIV-1 infection? The ANRS PRIMO and COPANA cohorts.* PLoS One, 2013. **8**(8): p. e71473.
37. El-Sadr, W.M.*, et al.*, *CD4+ count-guided interruption of antiretroviral treatment.* N Engl J Med, 2006. **355**(22): p. 2283-96.
38. Girard, P.M.*, et al.*, *Treatment interruption after one year of triple nucleoside analogue therapy for primary HIV infection.* AIDS, 2001. **15**(2): p. 275-7.
39. Taiwo, B.O. and C.B. Hicks, *Primary human immunodeficiency virus.* South Med J, 2002. **95**(11): p. 1312-7.
40. Bodsworth, N. and D.A. Cooper, *Update on HIV infection.* Aust Fam Physician, 1991. **20**(6): p. 747, 750-1, 753-9.
41. Strain, M.C.*, et al.*, *Effect of treatment, during primary infection, on establishment and clearance of cellular reservoirs of HIV-1.* J Infect Dis, 2005. **191**(9): p. 1410-8.
42. Hocqueloux, L.*, et al.*, *Long-term immunovirologic control following antiretroviral therapy interruption in patients treated at the time of primary HIV-1 infection.* Aids, 2010. **24**(10): p. 1598-601.

43. Lori, F.*, et al.*, *Treatment of human immunodeficiency virus infection with hydroxyurea, didanosine, and a protease inhibitor before seroconversion is associated with normalized immune parameters and limited viral reservoir.* J Infect Dis, 1999. **180**(6): p. 1827-32.
44. Jain, V.*, et al.*, *Antiretroviral therapy initiated within 6 months of HIV infection is associated with lower T-cell activation and smaller HIV reservoir size.* J Infect Dis, 2013. **208**(8): p. 1202-11.
45. Blankson, J.N.*, et al.*, *Biphasic decay of latently infected CD4+ T cells in acute human immunodeficiency virus type 1 infection.* J Infect Dis, 2000. **182**(6): p. 1636-42.
46. Grijsen, M.L.*, et al.*, *No Treatment versus 24 or 60 Weeks of Antiretroviral Treatment during Primary HIV Infection: The Randomized Primo-SHM Trial.* Plos Medicine, 2012. **9**(3): p. e1001196.
47. Hecht, F.M.*, et al.*, *A multicenter observational study of the potential benefits of initiating combination antiretroviral therapy during acute HIV infection.* J Infect Dis, 2006. **194**(6): p. 725-33.
48. Hamlyn, E.*, et al.*, *Plasma HIV viral rebound following protocol-indicated cessation of ART commenced in primary and chronic HIV infection.* PLoS One, 2012. **7**(8): p. e43754.
49. Stohr, W.*, et al.*, *Duration of HIV-1 viral suppression on cessation of antiretroviral therapy in primary infection correlates with time on therapy.* PLoS One, 2013. **8**(10): p. e78287.
50. Martin, G.E.*, et al.*, *Post-treatment control or treated controllers? Viral remission in treated and untreated primary HIV infection.* Aids, 2017. **31**(4): p. 477-484.
51. Concerted Action on SeroConversion to AIDS and Death in Europe (CASCADE), *Time from HIV-1 seroconversion to AIDS and death before widespread use of highly-active antiretroviral therapy: a collaborative re-analysis. Collaborative Group on AIDS Incubation and HIV Survival including the CASCADE EU Concerted Action. Concerted Action on SeroConversion to AIDS and Death in Europe.* Lancet, 2000. **355**(9210): p. 1131-7.
52. Jarrin, I.*, et al.*, *Uptake of combination antiretroviral therapy and HIV disease progression according to geographical origin in seroconverters in Europe, Canada, and Australia.* Clin Infect Dis, 2012. **54**(1): p. 111-8.
53. Weber, J., K. Porter, and A. Babiker, *Short-course antiretroviral therapy in primary HIV infection.* N Engl J Med, 2013. **368**(21): p. 2036-7.
54. Lehmann, C., C. Wyen, and G. Fatkenheuer, *Short-course antiretroviral therapy in primary HIV infection.* N Engl J Med, 2013. **368**(21): p. 2036-2037.
55. Writing Committee for the, C.C., *Timing of HAART initiation and clinical outcomes in human immunodeficiency virus type 1 seroconverters.* Arch Intern Med, 2011. **171**(17): p. 1560-9.
56. Hogan, C.M.*, et al.*, *The Setpoint Study (ACTG A5217): Effect of Immediate Versus Deferred Antiretroviral Therapy on Virologic Set Point in Recently HIV-1-Infected Individuals.* Journal of Infectious Diseases, 2012. **205**(1): p. 87-96.
57. Goujard, C.*, et al.*, *HIV-1 control after transient antiretroviral treatment initiated in primary infection: role of patient characteristics and effect of therapy.* Antivir Ther, 2012. **17**(6): p. 1001-9.
58. Zhang, L.*, et al.*, *Quantifying residual HIV-1 replication in patients receiving combination antiretroviral therapy.* N Engl J Med, 1999. **340**(21): p. 1605-13.
59. Markowitz, M.*, et al.*, *A Randomized Open-Label Study of Three- versus Five-Drug Combination Antiretroviral Therapy in Newly HIV-1 Infected Individuals.* J Acquir Immune Defic Syndr, 2014. **66**(2): p. 140-147.
60. Gurdasani, D.*, et al.*, *A systematic review of definitions of extreme phenotypes of HIV control and progression.* AIDS, 2014. **28**(2): p. 149-62.
61. Rosenberg, E.S.*, et al.*, *Immune control of HIV-1 after early treatment of acute infection.* Nature, 2000. **407**(6803): p. 523-6.

62. Shiau, S. and L. Kuhn, *Antiretroviral treatment in HIV-infected infants and young children: novel issues raised by the Mississippi baby.* Expert Rev Anti Infect Ther, 2014. **12**(3): p. 307-18.
63. Persaud, D.*, et al.*, *Absence of detectable HIV-1 viremia after treatment cessation in an infant.* N Engl J Med, 2013. **369**(19): p. 1828-35.
64. Lisziewicz, J.*, et al.*, *Control of HIV despite the discontinuation of antiretroviral therapy.* N Engl J Med, 1999. **340**(21): p. 1683-4.
65. Lambotte, O.*, et al.*, *HIV controllers: a homogeneous group of HIV-1-infected patients with spontaneous control of viral replication.* Clin Infect Dis, 2005. **41**(7): p. 1053-6.
66. Vanham, G.*, et al.*, *What is the significance of posttreatment control of HIV infection vis-a-vis functional cure?* AIDS, 2014. **28**(4): p. 603-5.
67. Madec, Y.*, et al.*, *Natural history of HIV-control since seroconversion.* AIDS, 2013. **27**(15): p. 2451-60.
68. Schacker, T.W.*, et al.*, *Biological and virologic characteristics of primary HIV infection.* Ann Intern Med, 1998. **128**(8): p. 613-20.
69. Porter, K.*, et al.*, *Determinants of survival following HIV-1 seroconversion after the introduction of HAART.* Lancet, 2003. **362**(9392): p. 1267-74.
70. Bhaskaran, K.*, et al.*, *Changes in the risk of death after HIV seroconversion compared with mortality in the general population.* Jama, 2008. **300**(1): p. 51-9.
71. Smith, C.J.*, et al.*, *Trends in underlying causes of death in people with HIV from 1999 to 2011 (D:A:D): a multicohort collaboration.* Lancet, 2014. **384**(9939): p. 241-8.
72. Krentz, H.B., G. Kliewer, and M.J. Gill, *Changing mortality rates and causes of death for HIV-infected individuals living in Southern Alberta, Canada from 1984 to 2003.* HIV Med, 2005. **6**(2): p. 99-106.
73. Palella, F.J., Jr.*, et al.*, *Mortality in the highly active antiretroviral therapy era: changing causes of death and disease in the HIV outpatient study.* J Acquir Immune Defic Syndr, 2006. **43**(1): p. 27-34.
74. Yang, C.H.*, et al.*, *Trends of mortality and causes of death among HIV-infected patients in Taiwan, 1984-2005.* HIV Med, 2008. **9**(7): p. 535-43.
75. Aldaz, P.*, et al.*, *[Trends in mortality and causes of death among persons with HIV infection, 1985-2004].* Enferm Infecc Microbiol Clin, 2007. **25**(1): p. 5-10.
76. Petoumenos, K.*, et al.*, *Cancer, immunodeficiency and antiretroviral treatment: results from the Australian HIV Observational Database (AHOD).* HIV Med, 2013. **14**(2): p. 77-84.
77. Croxford, S.*, et al.*, *Mortality and causes of death in people diagnosed with HIV in the era of highly active antiretroviral therapy compared with the general population: an analysis of a national observational cohort.* Lancet Public Health, 2017. **2**(1): p. e35-e46.
78. Bellan, S.E.*, et al.*, *Reassessment of HIV-1 Acute Phase Infectivity: Accounting for Heterogeneity and Study Design with Simulated Cohorts.* PLoS Med, 2015. **12**(3): p. e1001801.
79. Chibo, D., M. Kaye, and C. Birch, *HIV Transmissions During Seroconversion Contribute Significantly to New Infections in Men Who Have Sex with Men in Australia.* Aids Research and Human Retroviruses, 2012. **28**(5): p. 460-464.
80. Hamlyn, E.*, et al.*, *Antiretroviral treatment of primary HIV infection to reduce onward transmission.* Curr Opin HIV AIDS, 2010. **5**(4): p. 283-90.
81. Brenner, B.G.*, et al.*, *High Rates of Forward Transmission Events after Acute/Early HIV-1 Infection.* Journal of Infectious Diseases, 2007. **195**(7): p. 951-959.
82. Wilson, D.P.*, et al.*, *Importance of promoting HIV testing for preventing secondary transmissions: modelling the Australian HIV epidemic among men who have sex with men.* Sex Health, 2009. **6**(1): p. 19-33.
83. Lundgren, J.D.*, et al.*, *Initiation of Antiretroviral Therapy in Early Asymptomatic HIV Infection.* N Engl J Med, 2015. **373**(9): p. 795-807.

84. World Health Organization, *Prevent HIV, Test and Treat All*, in *WHO Support for Country Impact*. 2016, World Health Organization: 20 Avenue Appia, 1211 Geneva 27, Switzerland. p. 1-64.
85. DHHS, *Guidelines for the Use of Antiretroviral Agents in HIV-1-Infected Adults and Adolescents*, D.o.H.a.H. Services, Editor. 2015. p. 1-297.
86. Castley, A.*, et al.*, *A national study of the molecular epidemiology of HIV-1 in Australia 2005-2012.* PLoS One, 2017. **12**(5): p. e0170601.
87. Blaser, N.*, et al.*, *Impact of viral load and the duration of primary infection on HIV transmission: systematic review and meta-analysis.* AIDS, 2014. **28**(7): p. 1021-9.
88. Robineau, O.*, et al.*, *Combining the Estimated Date of HIV Infection with a Phylogenetic Cluster Study to Better Understand HIV Spread: Application in a Paris Neighbourhood.* PLoS One, 2015. **10**(8): p. e0135367.
89. Puhr, R.*, et al.*, *The impact of changes in HIV management guidelines on time to treatment initiation in Australia.* HIV Med, 2017. **18**(9): p. 701-703.
90. ASHM. *What to Start: Initial Combination Regimens for the Antiretroviral-Naïve Patient*. Antiretroviral Guidelines with Australian commentary 2018 30/5/2018 [cited 2016 9/11/2016]; 30/5/2018:[Guidelines]. Available from: http://arv.ashm.org.au/arv-guidelines/what-to-start.
91. O'Brien, M. and M. Markowitz, *Should we treat acute HIV infection?* Curr HIV/AIDS Rep, 2012. **9**(2): p. 101-10.
92. Pinto, A.N.*, et al.*, *HIV-1 subtype diversity, transmitted drug resistance and phylogenetics in Australia.* Future Virology, 2018. **13**(8): p. 1-10.
93. Nijhuis, M., S. Deeks, and C. Boucher, *Implications of antiretroviral resistance on viral fitness.* Curr Opin Infect Dis, 2001. **14**(1): p. 23-8.
94. World Health Organization, *Global action plan on HIV drug resistance 2017–2021*. 2017: Geneva, Switzerland. p. 1-40.
95. Wertheim, J.O.*, et al.*, *Transmission fitness of drug-resistant HIV revealed in a surveillance system transmission network.* Virus Evol, 2017. **3**(1): p. 1-12.
96. Lee, F.J., J. Amin, and A. Carr, *Efficacy of initial antiretroviral therapy for HIV-1 infection in adults: a systematic review and meta-analysis of 114 studies with up to 144 weeks' follow-up.* PLoS One, 2014. **9**(5): p. e97482.
97. Rocheleau, G.*, et al.*, *Longitudinal trends of HIV drug resistance in a large Canadian cohort, 1996-2016.* Clin Microbiol Infect, 2018. **24**(2): p. 185-191.
98. Scherrer, A.U.*, et al.*, *Emergence of Acquired HIV-1 Drug Resistance Almost Stopped in Switzerland: A 15-Year Prospective Cohort Analysis.* Clin Infect Dis, 2016. **62**(10): p. 1310-7.
99. Lodi, S.*, et al.*, *Effect of immediate initiation of antiretroviral treatment on the risk of acquired HIV drug resistance.* AIDS, 2018. **32**(3): p. 327-335.
100. Aghokeng, A.F.*, et al.*, *Extraordinary heterogeneity of virological outcomes in patients receiving highly antiretroviral therapy and monitored with the World Health Organization public health approach in sub-saharan Africa and southeast Asia.* Clin Infect Dis, 2014. **58**(1): p. 99-109.
101. Stanford University, *Surveillance Drug Resistance Mutation (SDRM) Worksheet*, S. University, Editor. 2012: USA.
102. Gregson, J.*, et al.*, *Occult HIV-1 drug resistance to thymidine analogues following failure of first-line tenofovir combined with a cytosine analogue and nevirapine or efavirenz in sub Saharan Africa: a retrospective multi-centre cohort study.* Lancet Infect Dis, 2017. **17**(3): p. 296-304.
103. Rafiei, N.*, et al.*, *Integrase-resistant HIV in an antiretroviral-naive patient in Australia.* Aids, 2017. **31**(4): p. 598-600.

104. Aldous, A.M., A.D. Castel, and D.M. Parenti, *Prevalence and trends in transmitted and acquired antiretroviral drug resistance, Washington, DC, 1999-2014.* BMC Res Notes, 2017. **10**(1): p. 1-10.
105. Scherrer, A.U*., et al.*, *Successful Prevention of Transmission of Integrase Resistance in the Swiss HIV Cohort Study.* J Infect Dis, 2016. **214**(3): p. 399-402.
106. Cheung, P.K., B. Wynhoven, and P.R. Harrigan, *2004: which HIV-1 drug resistance mutations are common in clinical practice?* AIDS Rev, 2004. **6**(2): p. 107-116.
107. Bennett, D.E*., et al.*, *Drug resistance mutations for surveillance of transmitted HIV-1 drug-resistance: 2009 update.* PLoS One, 2009. **4**(3): p. e4724.
108. Wensing, A.M*., et al.*, *2017 Update of the Drug Resistance Mutations in HIV-1.* Top Antivir Med, 2017. **24**(4): p. 132-133.
109. Liu, T.F. and R.W. Shafer, *Web resources for HIV type 1 genotypic-resistance test interpretation.* Clin Infect Dis, 2006. **42**(11): p. 1608-18.
110. Baxter, J.D*., et al.*, *Global HIV-1 transmitted drug resistance in the INSIGHT Strategic Timing of AntiRetroviral Treatment (START) trial.* HIV Med, 2015. **16 Suppl 1**: p. 77-87.
111. Pham, Q.D*., et al.*, *Global burden of transmitted HIV drug resistance and HIV-exposure categories: a systematic review and meta-analysis.* Aids, 2014. **28**(18): p. 2751-62.
112. Rhee, S.Y*., et al.*, *Geographic and temporal trends in the molecular epidemiology and genetic mechanisms of transmitted HIV-1 drug resistance: an individual-patient- and sequence-level meta-analysis.* PLoS Med, 2015. **12**(4): p. e1001810.
113. Tostevin, A*., et al.*, *Recent trends and patterns in HIV-1 transmitted drug resistance in the United Kingdom.* HIV Med, 2017. **18**(3): p. 204-213.
114. Avila-Rios, S*., et al.*, *Surveillance of HIV Transmitted Drug Resistance in Latin America and the Caribbean: A Systematic Review and Meta-Analysis.* PLoS One, 2016. **11**(6): p. e0158560.
115. Li, J.F*., et al.*, *Sensitive sentinel mutation screening reveals differential underestimation of transmitted HIV drug resistance among demographic groups.* Aids, 2016. **30**(9): p. 1439-45.
116. Milloy, M.J*., et al.*, *Increased Prevalence of Controlled Viremia and Decreased Rates of HIV Drug Resistance Among HIV-Positive People Who Use Illicit Drugs During a Community-wide Treatment-as-Prevention Initiative.* Clin Infect Dis, 2016. **62**(5): p. 640-7.
117. Socias, M.E*., et al.*, *Patterns of transmitted drug resistance and virological response to first-line antiretroviral treatment among HIV-positive people who use illicit drugs in a Canadian setting.* Clin Infect Dis, 2017. **65**(5): p. 796-802.
118. Kassaye, S.G*., et al.*, *Transmitted HIV Drug Resistance Is High and Longstanding in Metropolitan Washington, DC.* Clin Infect Dis, 2016. **63**(6): p. 836-843.
119. Ambrosioni, J*., et al.*, *Trends in Transmission of Drug Resistance and Prevalence of Non-B Subtypes in Patients with Acute or Recent HIV-1 Infection in Barcelona in the Last 16 Years (1997-2012).* PLoS One, 2015. **10**(6): p. e0125837.
120. Onywera, H*., et al.*, *Surveillance of HIV-1 pol transmitted drug resistance in acutely and recently infected antiretroviral drug-naive persons in rural western Kenya.* PLoS One, 2017. **12**(2): p. e0171124.
121. Margot, N.A*., et al.*, *Commonly Transmitted HIV-1 Drug Resistance Mutations in Reverse-Transcriptase and Protease in Antiretroviral Treatment-Naive Patients and Response to Regimens Containing Tenofovir Disoproxil Fumarate or Tenofovir Alafenamide.* J Infect Dis, 2017. **215**(6): p. 920-927.
122. Park, M*., et al.*, *The trend of transmitted drug resistance in newly diagnosed antiretroviral-naive HIV/AIDS patients during 1999-2012 in South Korea.* J Clin Virol, 2016. **81**: p. 53-7.
123. Mendoza, Y*., et al.*, *HIV-1 Antiretroviral Drug Resistance Mutations in Treatment Naive and Experienced Panamanian Subjects: Impact on National Use of EFV-Based Schemes.* PLoS One, 2016. **11**(4): p. e0154317.

124. WHO. *Consolidated guidelines on the use of antiretroviral drugs for treating and preventing HIV infection*. 2013; Available from: http://www.who.int/hiv/pub/guidelines/arv2013/download/en/.
125. Cambiano, V.*, et al.*, *Predicted levels of HIV drug resistance: potential impact of expanding diagnosis, retention, and eligibility criteria for antiretroviral therapy initiation.* AIDS, 2014. **28 Suppl 1**: p. S15-23.
126. Stanojevic, M.*, et al.*, *Ten years survey of primary HIV-1 resistance in Serbia: the occurrence of multiclass resistance.* AIDS Res Hum Retroviruses, 2014. **30**(7): p. 634-41.
127. Pennings, P.S., *HIV Drug Resistance: Problems and Perspectives.* Infect Dis Rep, 2013. **5**(Suppl 1): p. e5.
128. Manasa, J.*, et al.*, *Primary drug resistance in South Africa: data from 10 years of surveys.* AIDS Res Hum Retroviruses, 2012. **28**(6): p. 558-65.
129. Myers, J.E.*, et al.*, *Transmitted drug resistance among antiretroviral-naive patients with established HIV type 1 infection in Santo Domingo, Dominican Republic and review of the Latin American and Caribbean literature.* AIDS Res Hum Retroviruses, 2012. **28**(7): p. 667-74.
130. Hamers, R.L.*, et al.*, *HIV-1 drug resistance in antiretroviral-naive individuals in sub-Saharan Africa after rollout of antiretroviral therapy: a multicentre observational study.* Lancet Infect Dis, 2011. **11**(10): p. 750-9.
131. Frentz, D., C.A. Boucher, and D.A. van de Vijver, *Temporal changes in the epidemiology of transmission of drug-resistant HIV-1 across the world.* AIDS Rev, 2012. **14**(1): p. 17-27.
132. The Kirby Institute, *HIV, viral hepatitis and sexually transmissible infections in Australia*, in *Annual surveillance report 2018*. 2018: Sydney. p. 1-218.
133. The Kirby Institute, *Australian HIV observational Database Annual report*. 2013, The University of New South Wales: Sydney, NSW. p. 1-19.
134. ASHM, *Antiretroviral Guidelines US DHHS guidelines with Australian commentary*, ASHM, Editor. 2016: Australia.
135. Hales, G.*, et al.*, *A randomised trial comparing genotypic and virtual phenotypic interpretation of HIV drug resistance: the CREST study.* PLoS Clin Trials, 2006. **1**(3): p. e18.
136. Liu, T.F. and R.W. Shafer, *Web resources for HIV type 1 genotypic-resistance test interpretation.* Clinical Infectious Diseases, 2006. **42**(11): p. 1608-1618.
137. Little, S.J.*, et al.*, *Antiretroviral-drug resistance among patients recently infected with HIV.* N Engl J Med, 2002. **347**(6): p. 385-94.
138. Frentz, D.*, et al.*, *Increase in transmitted resistance to non-nucleoside reverse transcriptase inhibitors among newly diagnosed HIV-1 infections in Europe.* BMC Infect Dis, 2014. **14**: p. 1-12.
139. Ammaranond, P.*, et al.*, *No increase in protease resistance and a decrease in reverse transcriptase resistance mutations in primary HIV-1 infection: 1992-2001.* AIDS, 2003. **17**(2): p. 264-7.
140. Ammaranond, P.*, et al.*, *Rates of transmission of antiretroviral drug resistant strains of HIV-1.* J Clin Virol, 2003. **26**(2): p. 153-61.
141. Russell, J.S.*, et al.*, *Prevalence of transmitted HIV drug resistance since the availability of highly active antiretroviral therapy.* Commun Dis Intell Q Rep, 2009. **33**(2): p. 216-20.
142. D'Costa, J.*, et al.*, *Analysis of transmitted HIV drug resistance from 2005 to 2015 in Victoria, Australia: a comparison of the old and the new.* Sex Health, 2017. **14**(6): p. 558-565.
143. Pinto, A.*, et al.*, *Evolution of HIV-1 surveillance drug resistance mutations over ten years in New South Wales, Australia.* AIDS Res Hum Retroviruses, 2017.
144. Tong, W.W.*, et al.*, *Transmission of triple-class, drug-resistant HIV-1 in Australia.* Intern Med J, 2010. **40**(9): p. 657-61.
145. Australasian Society for HIV Viral Hepatitis and Sexual Health Medicine *Post-Exposure Prophylaxis after Non-Occupational and Occupational Exposure to HIV: Australian Guidelines.*, in *Australian National Guidelines (Second edition)*. 2016. p. 1-24.

146. van Hal, S.J.*, et al.*, *HIV-1 integrase polymorphisms are associated with prior antiretroviral drug exposure.* Retrovirology, 2009. **6**(12): p. 1-3.
147. Hawke, K.G.*, et al.*, *HIV non-B subtype distribution: emerging trends and risk factors for imported and local infections newly diagnosed in South Australia.* AIDS Res Hum Retroviruses, 2013. **29**(2): p. 311-7.
148. Rhee, S.Y.*, et al.*, *Human immunodeficiency virus reverse transcriptase and protease sequence database.* Nucleic Acids Res, 2003. **31**(1): p. 298-303.
149. Chibo, D. and C. Birch, *Increasing diversity of Human Immunodeficiency Virus type 1 subtypes circulating in Australia.* AIDS Res Hum Retroviruses, 2012. **28**(6): p. 578-83.
150. Barnes, P.F. and M.D. Cave, *Molecular epidemiology of tuberculosis.* N Engl J Med, 2003. **349**(12): p. 1149-56.
151. Chin, C.S.*, et al.*, *The origin of the Haitian cholera outbreak strain.* N Engl J Med, 2011. **364**(1): p. 33-42.
152. Sabat, A.J.*, et al.*, *Overview of molecular typing methods for outbreak detection and epidemiological surveillance.* Euro Surveill, 2013. **18**(4): p. 20380.
153. Pinto, A.N.*, et al.*, *Emergence and control of an outbreak of infections due to Panton-Valentine leukocidin positive, ST22 methicillin-resistant Staphylococcus aureus in a neonatal intensive care unit.* Clinical Microbiology and Infection, 2013. **19**(7): p. 620-7.
154. Mellmann, A.*, et al.*, *Prospective genomic characterization of the German enterohemorrhagic Escherichia coli O104:H4 outbreak by rapid next generation sequencing technology.* PLoS One, 2011. **6**(7): p. e22751.
155. Borgdorff, M.W. and D. van Soolingen, *The re-emergence of tuberculosis: what have we learnt from molecular epidemiology?* Clin Microbiol Infect, 2013. **19**(10): p. 889-901.
156. Neo, F.J.X.*, et al.*, *Outbreak of caliciviruses in the Singapore military, 2015.* BMC Infect Dis, 2017. **17**(1): p. 1-8.
157. Australian Commission on Safety and Quality in Health Care. *National Alert System for Critical Antimicrobial Resistances: CARAlert*. 2017 [cited 2018 12/2/2018]; Report]. Available from: https://www.safetyandquality.gov.au/antimicrobial-use-and-resistance-in-australia/what-is-aura/national-alert-system-for-critical-antimicrobial-resistances-caralert/.
158. Hassan, A.S.*, et al.*, *Defining HIV-1 transmission clusters based on sequence data.* Aids, 2017. **31**(9): p. 1211-1222.
159. Little, S.J.*, et al.*, *Using HIV networks to inform real time prevention interventions.* PLoS One, 2014. **9**(6): p. e98443.
160. Grabowski, M.K., J.T. Herbeck, and A.F.Y. Poon, *Genetic Cluster Analysis for HIV Prevention.* Curr HIV/AIDS Rep, 2018. **15**(2): p. 182-189.
161. Grabowski, M.K. and A.D. Redd, *Molecular tools for studying HIV transmission in sexual networks.* Curr Opin HIV AIDS, 2014. **9**(2): p. 126-33.
162. Poon, A.F.*, et al.*, *Near real-time monitoring of HIV transmission hotspots from routine HIV genotyping: an implementation case study.* Lancet HIV, 2016. **3**(5): p. e231-8.
163. Hue, S.*, et al.*, *HIV-1 pol gene variation is sufficient for reconstruction of transmissions in the era of antiretroviral therapy.* Aids, 2004. **18**(5): p. 719-28.
164. The Kirby Institute, *Annual Surveillance Report of HIV, viral hepatitis, STIs 2017*. 2017, UNSW: Sydney.
165. The Kirby Institute, *HIV, viral hepatitis and sexually transmissible infections in Australia Annual Surveillance Report 2013.*, L.Z. Prepared by Ann McDonald, Megan Tapia, David Wilson, Editor. 2014, The Kirby Institute for infection and immunity in society: The University of New South Wales, Sydney NSW 2052.
166. NSW Ministry of Health, *NSW HIV Strategy 2012–2015 A NEW ERA*, D.o. Health, Editor. 2012, NSW Ministry of Health: North Sydney.
167. Pinto Angie N and Cooper David A, *Antiretroviral therapy: research, rollout and resistance.* Microbiology Australia, 2014. **35**(2): p. 79-82.

168. Health, T.D.o., *Seventh National HIV Strategy 2014-2017*. 2014, Publications approval number: 10716: Online ISBN: 978-1-74186-169-3. p. 52.
169. Lemoh, C.*, et al.*, *Acquisition of HIV by African-born residents of Victoria, Australia: insights from molecular epidemiology.* PLoS One, 2013. **8**(12): p. e84008.
170. AIDS council of NSW. *EPIC NSW Study*. 2015 4/4/2016]; Available from: http://endinghiv.org.au/nsw/stay-safe/epic/.
171. Knox, D.*, et al. HIV-1 Infection with multi-class Resistance despite Pre-Exposure Prophylaxis (PrEP)*. in *CROI*. 2016. Boston, Massachusetts.
172. Zablotska, I.B.*, et al.*, *The Informal Use of Antiretrovirals for Preexposure Prophylaxis of HIV Infection Among Gay Men in Australia.* Jaids-Journal of Acquired Immune Deficiency Syndromes, 2013. **62**(3): p. 334-338.
173. Cohen, M.S.*, et al.*, *Prevention of HIV-1 Infection with Early Antiretroviral Therapy.* New England Journal of Medicine, 2011. **365**(6): p. 493-505.
174. Poon, A.F.Y.*, et al.*, *The Impact of Clinical, Demographic and Risk Factors on Rates of HIV Transmission: A Population-based Phylogenetic Analysis in British Columbia, Canada.* Journal of Infectious Diseases, 2015. **211**(6): p. 926-935.
175. Drescher, S.M.*, et al.*, *Treatment-naive individuals are the major source of transmitted HIV-1 drug resistance in men who have sex with men in the Swiss HIV Cohort Study.* Clin Infect Dis, 2014. **58**(2): p. 285-94.
176. Bezemer, D.*, et al.*, *Dispersion of the HIV-1 Epidemic in Men Who Have Sex with Men in the Netherlands: A Combined Mathematical Model and Phylogenetic Analysis.* PLoS Med, 2015. **12**(11): p. e1001898.
177. Keener, A.B., *Connectivity webs from HIV sequences could inform counseling.* Nat Med, 2014. **20**(10): p. 1083-1083.
178. Smith, D.M.*, et al.*, *A public health model for the molecular surveillance of HIV transmission in San Diego, California.* AIDS, 2009. **23**(2): p. 225-32.
179. Frost, S.D.W. and D. Pillay, *Understanding Drivers of Phylogenetic Clustering in Molecular Epidemiological Studies of HIV.* Journal of Infectious Diseases, 2015. **211**(6): p. 856-858.
180. Brooks, J.I. and P.A. Sandstrom, *The power and pitfalls of HIV phylogenetics in public health.* Can J Public Health, 2013. **104**(4): p. e348-50.
181. Kaye, M., D. Chibo, and C. Birch, *Phylogenetic investigation of transmission pathways of drug-resistant HIV-1 utilizing pol sequences derived from resistance genotyping.* J Acquir Immune Defic Syndr, 2008. **49**(1): p. 9-16.
182. Tree of Life design and icons, *What is Phylogeny? http://tolweb.org/tree/learn/concepts/whatisphylogeny.html*. 2004.
183. Patino-Galindo, J.A. and F. Gonzalez-Candelas, *Molecular evolution methods to study HIV-1 epidemics.* Future Virol, 2018. **13**(6): p. 399-404.
184. Lemey P, S.M., Vandamme Anne-Mieke, , *The Phylogenetic Handbook: A Practical Approach to Phylogenetic Analysis and Hypothesis Testing*. 2 ed. 2009, Cambridge: Cambridge University Press.
185. Hoang, D.T.*, et al.*, *UFBoot2: Improving the Ultrafast Bootstrap Approximation.* Mol Biol Evol, 2018. **35**(2): p. 518-522.
186. Hassan, A.S.*, et al.*, *Defining HIV-1 transmission clusters based on sequence data.* AIDS (London, England), 2017. **31**(9): p. 1211-1222.
187. Rose, R.*, et al.*, *Identifying Transmission Clusters with Cluster Picker and HIV-TRACE.* AIDS Res Hum Retroviruses, 2017. **33**(3): p. 211-218.
188. Dennis, A.M.*, et al.*, *Phylogenetic studies of transmission dynamics in generalized HIV epidemics: an essential tool where the burden is greatest?* J Acquir Immune Defic Syndr, 2014. **67**(2): p. 181-95.

189. Bernard, E.J.*, et al.*, *HIV forensics: pitfalls and acceptable standards in the use of phylogenetic analysis as evidence in criminal investigations of HIV transmission.* HIV Med, 2007. **8**(6): p. 382-7.
190. Campbell, M.S.*, et al.*, *Viral linkage in HIV-1 seroconverters and their partners in an HIV-1 prevention clinical trial.* PLoS One, 2011. **6**(3): p. e16986.
191. Gunthard, H.F. and R. Kouyos, *Can directionality of HIV transmission be predicted by next generation sequencing data?* J Infect Dis, 2018. **Epub ahead of print Dec 24 2018**.
192. Rose, R.*, et al.*, *Phylogenetic Methods Inconsistently Predict the Direction of HIV Transmission Among Heterosexual Pairs in the HPTN 052 Cohort.* The Journal of Infectious Diseases, 2018. **Epub ahead of print Dec 24 2018**.
193. Pasquale, D.K.*, et al.*, *Leveraging Phylogenetics to Understand HIV Transmission and Partner Notification Networks.* J Acquir Immune Defic Syndr, 2018. **78**(4): p. 367-375.
194. Easterbrook, P.J.*, et al.*, *Impact of HIV-1 viral subtype on disease progression and response to antiretroviral therapy.* J Int AIDS Soc, 2010. **13**(4): p. 1-9.
195. Guy, R.*, et al.*, *A new surveillance system for monitoring HIV infection in Victoria, Australia.* Sex Health, 2007. **4**(3): p. 195-9.
196. Butler, I.F.*, et al.*, *HIV genetic diversity: biological and public health consequences.* Curr HIV Res, 2007. **5**(1): p. 23-45.
197. Hawke, K., *Phylogenetic analysis and molecular epidemiological characterization of HIV infections diagnosed in South Australia*, in *Adelaide*. 2016, Flinders University of South Australia.
198. Oelrichs, R., *The subtypes of human immunodeficiency virus in Australia and Asia.* Sex Health, 2004. **1**(1): p. 1-11.
199. Siemieniuk, R.A., B. Beckthold, and M.J. Gill, *Increasing HIV subtype diversity and its clinical implications in a sentinel North American population.* Can J Infect Dis Med Microbiol, 2013. **24**(2): p. 69-73.
200. Frentz, D.*, et al.*, *Limited cross-border infections in patients newly diagnosed with HIV in Europe.* Retrovirology, 2013. **10**(36): p. 1-10.
201. Herring, B.L.*, et al.*, *Segregation of human immunodeficiency virus type 1 subtypes by risk factor in Australia.* J Clin Microbiol, 2003. **41**(10): p. 4600-4.
202. Castley, A.S.*, et al.*, *Longitudinal Trends in Western Australian HIV-1 Sequence Diversity and Viral Transmission Networks and Their Influence on Clinical Parameters: 2000-2014.* AIDS Res Hum Retroviruses, 2015. **32**(3): p. 211-9.
203. Chow, W.Z.*, et al.*, *A newly emerging HIV-1 recombinant lineage (CRF58_01B) disseminating among people who inject drugs in Malaysia.* PLoS One, 2014. **9**(1): p. e85250.
204. Adungo, F.O.*, et al.*, *Diversity of Human Immunodeficiency Virus Type-1 Subtypes in Western Kenya.* World Journal of AIDS, 2014. **4**(4): p. 365-372.
205. Russell, K.L.*, et al.*, *Emerging genetic diversity of HIV-1 in South America.* AIDS, 2000. **14**(12): p. 1785-91.
206. Thomson, M.M., L. Perez-Alvarez, and R. Najera, *Molecular epidemiology of HIV-1 genetic forms and its significance for vaccine development and therapy.* Lancet Infect Dis, 2002. **2**(8): p. 461-71.
207. Baeten, J.M.*, et al.*, *HIV-1 subtype D infection is associated with faster disease progression than subtype A in spite of similar plasma HIV-1 loads.* J Infect Dis, 2007. **195**(8): p. 1177-80.
208. Santoro, M.M. and C.F. Perno, *HIV-1 Genetic Variability and Clinical Implications.* ISRN Microbiol, 2013. **2013**: p. 481314.
209. Ng, K.Y.*, et al.*, *High prevalence of CXCR4 usage among treatment-naive CRF01_AE and CRF51_01B-infected HIV-1 subjects in Singapore.* BMC Infect Dis, 2013. **13**(90): p. 1-6.
210. Phuphuakrat, A.*, et al.*, *Coreceptor tropism determined by genotypic assay in HIV-1 circulating in Thailand, where CRF01_AE predominates.* HIV Med, 2014. **15**(5): p. 269-75.

211. Kanki, P.J.*, et al.*, *Human immunodeficiency virus type 1 subtypes differ in disease progression.* J Infect Dis, 1999. **179**(1): p. 68-73.
212. Palm, A.A.*, et al.*, *Faster progression to AIDS and AIDS-related death among seroincident individuals infected with recombinant HIV-1 A3/CRF02_AG compared with sub-subtype A3.* J Infect Dis, 2014. **209**(5): p. 721-8.
213. Horyniak, D.*, et al.*, *The impact of immigration on the burden of HIV infection in Victoria, Australia.* Sex Health, 2009. **6**(2): p. 123-8.
214. Petoumenos, K., *The Australian HIV Observational Database Temporary Residents Access Study (ATRAS)* in *One year follow-up*. 2013, The Kirby Institute University of New South Wales: Sydney, Australia.
215. McPherson, M.E.*, et al.*, *Using HIV notification data to identify priority migrant groups for HIV prevention, New South Wales, 2000-2008.* Commun Dis Intell Q Rep, 2011. **35**(2): p. 185-91.
216. Salemi, M., *Toward a robust monitoring of HIV subtypes distribution worldwide.* AIDS, 2011. **25**(5): p. 713-4.
217. Keen, P.*, et al.*, *The 2016 HIV diagnosis and care cascade in New South Wales, Australia: meeting the UNAIDS 90-90-90 targets.* J Int AIDS Soc, 2018. **21**(4): p. e25109.
218. Abu-Raddad, L.J., *Role of Acute HIV Infection in Driving HIV Transmission: Implications for HIV Treatment as Prevention.* PLoS Med, 2015. **12**(3): p. e1001803.
219. Yerly, S.*, et al.*, *Acute HIV infection: impact on the spread of HIV and transmission of drug resistance.* AIDS, 2001. **15**(17): p. 2287-92.
220. Wawer, M.J.*, et al.*, *Rates of HIV-1 transmission per coital act, by stage of HIV-1 infection, in Rakai, Uganda.* J Infect Dis, 2005. **191**(9): p. 1403-9.
221. Brenner, B.G.*, et al.*, *High rates of forward transmission events after acute/early HIV-1 infection.* J Infect Dis, 2007. **195**(7): p. 951-9.
222. Mehta, S.R.*, et al.*, *Using HIV Sequence and Epidemiologic Data to Assess the Effect of Self-referral Testing for Acute HIV Infection on Incident Diagnoses in San Diego, California.* Clin Infect Dis, 2016. **63**(1): p. 101-107.
223. Coltart, C.E.M.*, et al.*, *Ethical considerations in global HIV phylogenetic research.* Lancet HIV, 2018. **5**(11): p. e656-e666.
224. Kidd, P., *Phylogenetic analysis as expert evidence in HIV transmission prosecutions*, in *HIV Australia*. 2016, Australian Federation of AIDS Organisations.
225. Kaye, M., D. Chibo, and C. Birch, *Comparison of Bayesian and maximum-likelihood phylogenetic approaches in two legal cases involving accusations of transmission of HIV.* AIDS Res Hum Retroviruses, 2009. **25**(8): p. 741-8.
226. NSW government, *Public Health Act 2010 No 127*. 2010: Sydney, New South Wales p. 1-98.
227. Boyd, M.*, et al.*, *Sexual transmission of HIV and the law: an Australian medical consensus statement.* Med J Aust, 2016. **205**(9): p. 409-412.
228. Tindall, B., C.E. Swanson, and D.A. Cooper, *Development of AIDS in a cohort of HIV-seropositive homosexual men in Australia.* Med J Aust, 1990. **153**(5): p. 260-5.
229. Emery, S.*, et al.*, *Influence of IFNgamma co-expression on the safety and antiviral efficacy of recombinant fowlpox virus HIV therapeutic vaccines following interruption of antiretroviral therapy.* Human Vaccine, 2007. **3**(6): p. 260-7.
230. Hecht, F.M.*, et al.*, *A multicenter observational study of the potential benefits of initiating combination antiretroviral therapy during acute HIV infection.* J Infect Dis, 2006. **194**(6): p. 725-33.
231. Linkage., C.f.H.R. *How record linkage works*. 2011 2011 [cited 2015 20/11/2015]; The CHeReL was established in 2006 as a dedicated data linkage unit and is managed by the NSW Ministry of Health. Other organisations may become members of the CHeReL and are

represented on the Data Linkage Advisory Committee.]. Available from: http://www.cherel.org.au/how-record-linkage-works.
232. Jaro, M.A., *Probabilistic linkage of large public health data files.* Stat Med, 1995. **14**(5-7): p. 491-8.
233. Irvine, K.A. and E.A. Moore, *Linkage of routinely collected data in practice: the Centre for Health Record Linkage.* Public Health Res Pract, 2015. **25**(4): p. e2541548.
234. Borthwick, A., Buechi, M., and Goldberg, A. *Key Concepts in the ChoiceMaker 2 Record Matching System*. in *Proceedings of the KDD-2003 Workshop on Data Cleaning, Record Linkage, and Object Consolidation*. 2003. Washington, DC.
235. Boyd, J.H.*, et al.*, *Understanding the origins of record linkage errors and how they affect research outcomes.* Aust N Z J Public Health, 2016: p. 1753-6405.
236. Moore, C.L.*, et al.*, *Hospitalisation rates and associated factors in community-based cohorts of HIV-infected and -uninfected gay and bisexual men.* HIV Med, 2015. **17**(5): p. 327-39.
237. Ministry of Health, *NSW HIV Strategy 2016 – 2020 Quarter 3 2016 Data Report*. 2016, Ministry of Health: Sydney. p. 1-53.
238. Saez-Cirion, A.*, et al.*, *Post-treatment HIV-1 controllers with a long-term virological remission after the interruption of early initiated antiretroviral therapy ANRS VISCONTI Study.* PLoS Pathog, 2013. **9**(3): p. e1003211.
239. Lopez-Rios, J.*, et al.*, *Limited Knowledge and Lack of Screening for Acute HIV Infection at Primary Care Clinics in High-Prevalence Communities of New York City.* AIDS Behav, 2019: p. Epub ahead of print on May 3.
240. Rutstein, S.E.*, et al.*, *Incorporating Acute HIV Screening into Routine HIV Testing at Sexually Transmitted Infection Clinics, and HIV Testing and Counseling Centers in Lilongwe, Malawi.* J Acquir Immune Defic Syndr, 2016. **71**(3): p. 272-80.
241. Weber, R.*, et al.*, *Decreasing mortality and changing patterns of causes of death in the Swiss HIV Cohort Study.* HIV Med, 2013. **14**(4): p. 195-207.
242. Smith, C.J.*, et al.*, *Trends in underlying causes of death in people with HIV from 1999 to 2011 (D:A:D): a multicohort collaboration.* The Lancet. **384**(9939): p. 241-248.
243. Petoumenos, K., M.G. Law, and H.I.V.O.D. Australian, *Risk factors and causes of death in the Australian HIV Observational Database.* Sex Health, 2006. **3**(2): p. 103-12.
244. McManus, H.*, et al.*, *Long-term survival in HIV positive patients with up to 15 Years of antiretroviral therapy.* PLoS One, 2012. **7**(11): p. e48839.
245. Kowalska, J.D.*, et al.*, *The Coding Causes of Death in HIV (CoDe) Project: initial results and evaluation of methodology.* Epidemiology, 2011. **22**(4): p. 516-23.
246. Selik, R.M., Motkotoff E, Branson B, Owen S, Whitmaore S, Hall HI., *Revised Surveillance Case Definition for HIV Infection — United States, 2014*, in *Recommendations and Reports*. 2014, Centers for Disease Control and Prevention: 1600 Clifton Road Atlanta, GA 30329-4027, USA. p. 1-10.
247. Joint United Nations Programme on HIV/AIDS (UNAIDS), *90-90-90 An ambitious treatment target to help end the AIDS epidemic*, U.I.P. Unit., Editor. 2014: Geneva , Switzerland. p. 1-40.
248. Farahani, M.*, et al.*, *Prevalence and distribution of non-AIDS causes of death among HIV-infected individuals receiving antiretroviral therapy: a systematic review and meta-analysis.* Int J STD AIDS, 2017. **28**(7): p. 636-650.
249. The Kirby Institute, *HIV, viral hepatitis and sexually transmissible infections in Australia*, in *Annual Surveillance Report 2015*. 2015, The Kirby Institute for infection and immunity in society, UNSW Australia: Sydney NSW 2052. p. 1-152.
250. The Kirby Institute, *HIV, viral hepatitis and sexually transmissible infections in Australia. Annual Surveillance Report 2016*. 2016: Sydney NSW 2052. p. 1-180.
251. NSW Ministry of Health, *NSW HIV Strategy 2016 – 2020 Quarter 4 & Annual 2016 Data Report*. 2016: Sydney, NSW. p. 1-52.

252. NSW Ministry of Health, *NSW HIV Strategy 2016-2020*, Health, Editor. 2015, © NSW Ministry of Health 2015: Sydney, Australia. p. 20.
253. HIVdb, S. *HIV-1 Drug Resistance in ARV-naive Populations*. 2014; Available from: http://hivdb.stanford.edu/surveillance/map/.
254. Castley A, S.S., Varma R, Herring B, Thapa K, Chibo D, Nguyen N, Hawke K, Ratcliff R, Dwyer D, Nolan D. , *An epidemic in transition: impacts of migration and local networks on HIV sequence diversity and infection transmission in Australia 2005-2012*, in *AIDS2014*. 2014: Melbourne, Australia.
255. Brook, M.G., R. Gilson, and E. Wilkins, *BHIVA guidelines on HIV and chronic hepatitis: coinfection with HIV and hepatitis B virus infection (2005).* HIV Med, 2005. **6 Suppl 2**: p. 84-95.
256. Australian Government, *Pathology services*, D.o. Health, Editor. 2011: MBS online.
257. Gifford, R.J.*, et al.*, *The calibrated population resistance tool: standardized genotypic estimation of transmitted HIV-1 drug resistance.* Bioinformatics, 2009. **25**(9): p. 1197-8.
258. David A.M.C. van de Vijver , A.M.J.W., Charles A.B. Boucher, *The Epidemiology of Transmission of Drug Resistant HIV-1*. 2006.
259. The Kirby Institute, *Australian HIV Observational Database Annual Report*. 2015, UNSW, Australia: Sydney, Australia. p. 1-20.
260. Huang, R., R. Puhr, and K. Petoumenos, *Australian HIV Observational Database*, A. Pinto, Editor. 2017, The Kirby Institute: Sydney.
261. NSW Ministry of Health, *NSW HIV Strategy 2016 – 2020 Third Quarter Data Report.* 2018: p. 1-52.
262. Gregson J, T.M., Ndembi N, Hamers RL, Rhee SY, Marconi VC, Diero L, Brooks KA, Theys K, Rinke de Wit T, Arruda M, Garcia F, Monge S, Günthard HF, Hoffmann CJ, Kanki PJ, Kumarasamy N, Kerschberger B, Mor O, Charpentier C, Todesco E, Rokx C, Gras L, Helvas EK, Sunpath H, Di Carlo D, Antinori A, Andreoni M, Latini A, Mussini C, Aghokeng A, Sonnerborg A, Neogi U, Fessel WJ, Agolory S, Yang C, Blanco JL, Juma JM, Smit E, Schmidt D, Watera C, Asio J, Kurungi W, Tostevin A, El-Hay T, Clumeck N, Goedhals D, Van Vuuren C, Bester PA, Sabin C, Mukui I, Santoro MM, Perno CF, Hunt G, Morris L, Camacho R, De Oliveira T, Pillay D, Schulter E, Murakami-Ogasawara A, Reyes-Terán G, Romero K, Avila-Rios S, Sirivichayakul S, Ruxrungtham K, Mekprasan S, Dunn D, Kaleebu P, Raizes E, Kantor R, Shafer RW, Gupta RK, *Global epidemiology of drug resistance after failure of WHO recommended first-line regimens for adult HIV-1 infection: a multicentre retrospective cohort study.* Lancet Infect Dis, 2016. **16**(5): p. 565-575.
263. Tostevin, A.*, et al.*, *Recent trends and patterns in HIV-1 transmitted drug resistance in the United Kingdom.* HIV Med, 2016. **18**(3): p. 204-213.
264. Paquet, A.C.*, et al.*, *A decade of HIV-1 drug resistance in the United States: trends and characteristics in a large protease/reverse transcriptase and co-receptor tropism database from 2003 to 2012.* Antivir Ther, 2014. **19**(4): p. 435-41.
265. NSW Ministry of Health. *NSW HIV STRATEGY 2016 ANNUAL SNAPSHOT*. 2016 20/10/2017]; Available from: http://www.health.nsw.gov.au/endinghiv/Documents/2016-hiv-snapshot.pdf.
266. NSW Ministry of Health, *NSW HIV Strategy 2016-2020, Annual snapshot*. 2017. p. 4.
267. NSW Ministry of Health, *NSW HIV Strategy 2016 – 2020 April - June 2017 Data Report*. 2017.
268. Pineda-Pena, A.C.*, 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, 2013. **19**: p. 337-48.
269. Bolsewicz, K.*, et al.*, *Factors associated with antiretroviral treatment uptake and adherence: a review. Perspectives from Australia, Canada, and the United Kingdom.* AIDS Care, 2015. **27**(12): p. 1429-38.

270. Mao, L.*, et al.*, *Younger age, recent HIV diagnosis, no welfare support and no annual sexually transmissible infection screening are associated with nonuse of antiretroviral therapy among HIV-positive gay men in Australia.* HIV Med, 2015. **16**(1): p. 32-7.
271. Vaccher, S.*, et al.*, *Protocol for an open-label, single-arm trial of HIV pre-exposure prophylaxis (PrEP) among people at high risk of HIV infection: the NSW Demonstration Project PRELUDE.* BMJ Open, 2016. **6**(6): p. e012179.
272. Wilson DP, H.A., Regan DG, Wand HC, Law MG, *Mathematical models to investigate recent trends in HIV notifications among men who have had sex with men in Australia*, in *Mathematical models to investigate recent trends in HIV notifications among men who have had sex with men in Australia*. 2008, National Centre in HIV Epidemiology and Clinical Research: Australia.
273. Hollingsworth, T.D.*, et al.*, *High Transmissibility During Early HIV Infection Among Men Who Have Sex With Men-San Francisco, California.* J Infect Dis, 2014. **211**(11): p. 1757-1760.
274. White, P.J.*, et al.*, *How Many HIV Infections May Be Averted by Targeting Primary Infection in Men Who Have Sex With Men? Quantification of Changes in Transmission-Risk Behavior, Using an Individual-Based Model.* Journal of Infectious Diseases, 2014. **210**(suppl 2): p. S594-S599.
275. Hollingsworth, T.D., R.M. Anderson, and C. Fraser, *HIV-1 Transmission, by Stage of Infection.* Journal of Infectious Diseases, 2008. **198**(5): p. 687-693.
276. Audelin, A.M.*, et al.*, *Phylogenetics of the Danish HIV epidemic: the role of very late presenters in sustaining the epidemic.* J Acquir Immune Defic Syndr, 2013. **62**(1): p. 102-8.
277. de Oliveira, T.*, et al.*, *An automated genotyping system for analysis of HIV-1 and other microbial sequences.* Bioinformatics, 2005. **21**(19): p. 3797-800.
278. Ratner, L.*, et al.*, *Complete nucleotide sequence of the AIDS virus, HTLV-III.* Nature, 1985. **313**(6000): p. 277-84.
279. Hall, T.A., *BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT.* . Nucleic Acids Symposium, 1999. **41**: p. 95-98.
280. Kuhnert, D.*, et al.*, *Quantifying the fitness cost of HIV-1 drug resistance mutations through phylodynamics.* PLoS Pathog, 2018. **14**(2): p. e1006895.
281. Castor, D.*, et al.*, *Transmitted drug resistance and phylogenetic relationships among acute and early HIV-1-infected individuals in New York City.* J Acquir Immune Defic Syndr, 2012. **61**(1): p. 1-8.
282. Miller, M.A., Pfeiffer, W., and Schwartz, T. . *Creating the CIPRES Science Gateway for inference of large phylogenetic trees.* in *Proceedings of the Gateway Computing Environments Workshop (GCE),* . 2010. New Orleans, LA: IEEE.
283. Edgar, R.C., *MUSCLE: multiple sequence alignment with high accuracy and high throughput.* Nucleic Acids Research, 2004. **32**(5): p. 1792-1797.
284. Suthar, A.B.*, et al.*, *Programmatic Implications of Acute and Early HIV Infection.* J Infect Dis, 2015. **212**(9): p. 1351-60.
285. Australian society of HIV medicine. *2013 Antiretroviral guidelines with Australian Commentary*. [Guidelines] 2013 [cited 2013 28/10/2013]; Available from: http://arv.ashm.org.au/.
286. NSW Ministry of Health, *2nd Annual Report of Progress in Implementing the NSW HIV Strategy 2012-2015: A New Era*. 2014: Sydney NSW. p. 4.
287. Escudero, D.J.*, et al.*, *Acute HIV infection transmission among people who inject drugs in a mature epidemic setting.* Aids, 2016. **30**(16): p. 2537-2544.
288. Escudero, D.J.*, et al.*, *The risk of HIV transmission at each step of the HIV care continuum among people who inject drugs: a modeling study.* BMC Public Health, 2017. **17**(1): p. 1-9.

289. Kroon, E.*, et al.*, *Acute HIV infection detection and immediate treatment estimated to reduce transmission by 89% among men who have sex with men in Bangkok.* J Int AIDS Soc, 2017. **20**(1): p. 21708.
290. Miller, W.C.*, et al.*, *Role of acute and early HIV infection in the sexual transmission of HIV.* Curr Opin HIV AIDS, 2010. **5**(4): p. 277-82.
291. Hecht, F.M.*, et al.*, *HIV RNA level in early infection is predicted by viral load in the transmission source.* Aids, 2010. **24**(7): p. 941-5.
292. Hillis, D.M. and J.J. Bull, *An Empirical Test of Bootstrapping as a Method for Assessing Confidence in Phylogenetic Analysis.* Systematic Biology, 1993. **42**(2): p. 182-192.
293. Lubelchek, R.J.*, et al.*, *Transmission clustering among newly diagnosed HIV patients in Chicago, 2008 to 2011: using phylogenetics to expand knowledge of regional HIV transmission patterns.* J Acquir Immune Defic Syndr, 2014. **68**(1): p. 46-54.
294. Bartolo, I.*, et al.*, *HIV-1 diversity, transmission dynamics and primary drug resistance in Angola.* PLoS One, 2014. **9**(12): p. e113626.
295. Pineda-Pena, A.C.*, et al.*, *Trends and predictors of transmitted drug resistance (TDR) and clusters with TDR in a local Belgian HIV-1 epidemic.* PLoS One, 2014. **9**(7): p. e101738.
296. Poorolajal, J.*, et al.*, *Survival rate of AIDS disease and mortality in HIV-infected patients: a meta-analysis.* Public Health, 2016. **139**: p. 3-12.
297. World Health Organization, *WHO Guidelines Approved by the Guidelines Review Committee*, in *Guidelines for the Prevention, Care and Treatment of Persons with Chronic Hepatitis B Infection*. 2015, World Health Organization,: Geneva.
298. Mallitt, K.A.*, et al.*, *Identifying missed clinical opportunities for the earlier diagnosis of HIV in Australia, a retrospective cohort data linkage study.* PLoS One, 2018. **13**(12): p. e0208323.
299. Lifson, A.R.*, et al.*, *Improved quality of life with immediate versus deferred initiation of antiretroviral therapy in early asymptomatic HIV infection.* Aids, 2017. **31**(7): p. 953-963.
300. Advice study group, *Vorapaxar for HIV-associated inflammation and coagulopathy (ADVICE): a randomised, double-blind, placebo-controlled trial.* Lancet HIV, 2018. **5**(10): p. e553-e559.
301. Hey-Cunningham, W.J.*, et al.*, *Early antiretroviral therapy with raltegravir generates sustained reductions in HIV reservoirs but not lower T-cell activation levels.* Aids, 2015. **29**(8): p. 911-9.
302. Novelli, S.*, et al.*, *Long-term Therapeutic Impact of the Timing of Antiretroviral Therapy in Patients Diagnosed With Primary Human Immunodeficiency Virus Type 1 Infection.* Clin Infect Dis, 2018. **66**(10): p. 1519-1527.
303. Bartmeyer, B.*, et al.*, *Prevalence of transmitted drug resistance and impact of transmitted resistance on treatment success in the German HIV-1 Seroconverter Cohort.* PLoS One, 2010. **5**(10): p. e12718.
304. Bourne, C.*, et al.*, *Changing pattern of sexually transmissible infections and HIV diagnosed in public sexual health services compared with other locations in New South Wales, 2010-14.* Sex Health, 2018. **15**(4): p. 366-369.
305. Gunaratnam, P.*, et al.*, *People Born in Non-Main English Speaking Countries Are Less Likely to Start HIV Treatment Early in Australia: A National Cohort Analysis, 2014-15.* J Acquir Immune Defic Syndr, 2018. **77**(3): p. e31-e34.
306. Prins, H.A.B.*, et al.*, *Ending the epidemic: Critical role of primary HIV infection.* Neth J Med, 2017. **75**(8): p. 321-327.
307. Paraskevis, D.*, et al.*, *HIV-1 molecular transmission clusters in nine European countries and Canada: association with demographic and clinical factors.* BMC Med, 2019. **17**(1): p. 4.
308. German, D., M.K. Grabowski, and C. Beyrer, *Enhanced use of phylogenetic data to inform public health approaches to HIV among men who have sex with men.* Sex Health, 2017. **14**(1): p. 89-96.

309. Abeler-Dorner, L.*, et al.*, *PANGEA-HIV 2: Phylogenetics And Networks for Generalised Epidemics in Africa.* Curr Opin HIV AIDS, 2019. **14**(3): p. 173-180.
310. Gilbert, M.*, et al.*, *Need for robust and inclusive public health ethics review of the monitoring of HIV phylogenetic clusters for HIV prevention.* Lancet HIV, 2016. **3**(10): p. e461.
311. van de Vijver, D. and C.A.B. Boucher, *Insights on transmission of HIV from phylogenetic analysis to locally optimize HIV prevention strategies.* Curr Opin HIV AIDS, 2018. **13**(2): p. 95-101.
312. Paraskevis, D.*, et al.*, *The application of HIV molecular epidemiology to public health.* Infect Genet Evol, 2016. **46**: p. 159-168.
313. Grulich, A.E.*, et al.*, *Population-level effectiveness of rapid, targeted, high-coverage roll-out of HIV pre-exposure prophylaxis in men who have sex with men: the EPIC-NSW prospective cohort study.* Lancet HIV, 2018. **5**(11): p. e629-e637.
314. Volz, E.M.*, et al.*, *Molecular Epidemiology of HIV-1 Subtype B Reveals Heterogeneous Transmission Risk: Implications for Intervention and Control.* J Infect Dis, 2018. **217**(10): p. 1522-1529.
315. Taramasso, L.*, et al.*, *Short Communication: Tenofovir Disoproxil Fumarate/Emtricitabine Fits for All as Appropriate HIV-1 Pre-Exposure Prophylaxis?* AIDS Res Hum Retroviruses, 2018. **34**(2): p. 168-170.
316. Parikh, U.M. and J.W. Mellors, *Should we fear resistance from tenofovir/emtricitabine preexposure prophylaxis?* Curr Opin HIV AIDS, 2016. **11**(1): p. 49-55.
317. van de Vijver, D.A.*, et al.*, *Preexposure prophylaxis will have a limited impact on HIV-1 drug resistance in sub-Saharan Africa: a comparison of mathematical models.* Aids, 2013. **27**(18): p. 2943-51.
318. Knox, D.C.*, et al.*, *Multidrug-Resistant HIV-1 Infection despite Preexposure Prophylaxis.* N Engl J Med, 2017. **376**(5): p. 501-502.
319. Petoumenos, K.*, et al.*, *Prevalence of self-reported comorbidities in HIV positive and HIV negative men who have sex with men over 55 years-The Australian Positive & Peers Longevity Evaluation Study (APPLES).* PLoS One, 2017. **12**(9): p. e0184583.
320. Puhr, R.*, et al.*, *Cardiovascular disease and diabetes in HIV-positive and HIV-negative gay and bisexual men over the age of 55 years in Australia: insights from the Australian Positive & Peers Longevity Evaluation Study.* HIV Med, 2019. **20**(2): p. 121-130.
321. Ronit, A.*, et al.*, *Associations between serum albumin and serious non-AIDS events among people living with HIV.* Aids, 2018. **32**(13): p. 1837-1848.
322. Turnidge, J.D. and K.T. Meleady, *Antimicrobial Use and Resistance in Australia (AURA) surveillance system: coordinating national data on antimicrobial use and resistance for Australia.* Aust Health Rev, 2018. **42**(3): p. 272-276.
323. Holt, M.*, et al.*, *Trends in Attitudes to and the Use of HIV Pre-exposure Prophylaxis by Australian Gay and Bisexual Men, 2011-2017: Implications for Further Implementation from a Diffusion of Innovations Perspective.* AIDS Behav, 2018. **Dec 11**(1): p. 1-12.
324. Holt, M.*, et al.*, *Community-level changes in condom use and uptake of HIV pre-exposure prophylaxis by gay and bisexual men in Melbourne and Sydney, Australia: results of repeated behavioural surveillance in 2013-17.* Lancet HIV, 2018. **5**(8): p. e448-e456.
325. Dessilly, G.*, et al.*, *First evaluation of the Next-Generation Sequencing platform for the detection of HIV-1 drug resistance mutations in Belgium.* PLoS One, 2018. **13**(12): p. e0209561.
326. Ji, H.*, et al.*, *Bioinformatic data processing pipelines in support of next-generation sequencing-based HIV drug resistance testing: the Winnipeg Consensus.* J Int AIDS Soc, 2018. **21**(10): p. e25193.
327. Silver, N.*, et al.*, *Characterization of minority HIV-1 drug resistant variants in the United Kingdom following the verification of a deep sequencing-based HIV-1 genotyping and tropism assay.* AIDS Res Ther, 2018. **15**(1): p. 1-18.

328. Chimukangara, B.*, et al.*, *Primary HIV-1 Drug Resistant Minority Variants.* AIDS Rev, 2017. **19**(2): p. 89-96.
329. Stella-Ascariz, N.*, et al.*, *The Role of HIV-1 Drug-Resistant Minority Variants in Treatment Failure.* J Infect Dis, 2017. **216**(suppl_9): p. S847-s850.
330. Dalmat, R.R.*, et al.*, *Limited Marginal Utility of Deep Sequencing for HIV Drug Resistance Testing in the Age of Integrase Inhibitors.* J Clin Microbiol, 2018. **56**(12): p. e01443-18.
331. Tzou, P.L.*, et al.*, *Comparison of an In Vitro Diagnostic Next-Generation Sequencing Assay with Sanger Sequencing for HIV-1 Genotypic Resistance Testing.* J Clin Microbiol, 2018. **56**(6): p. 1-13.
332. Power, R.A.*, et al.*, *Genome-Wide Association Study of HIV Whole Genome Sequences Validated using Drug Resistance.* PLoS One, 2016. **11**(9): p. e0163746.