

# HPTN 071-2

# Findings of the PopART phylogenetics study

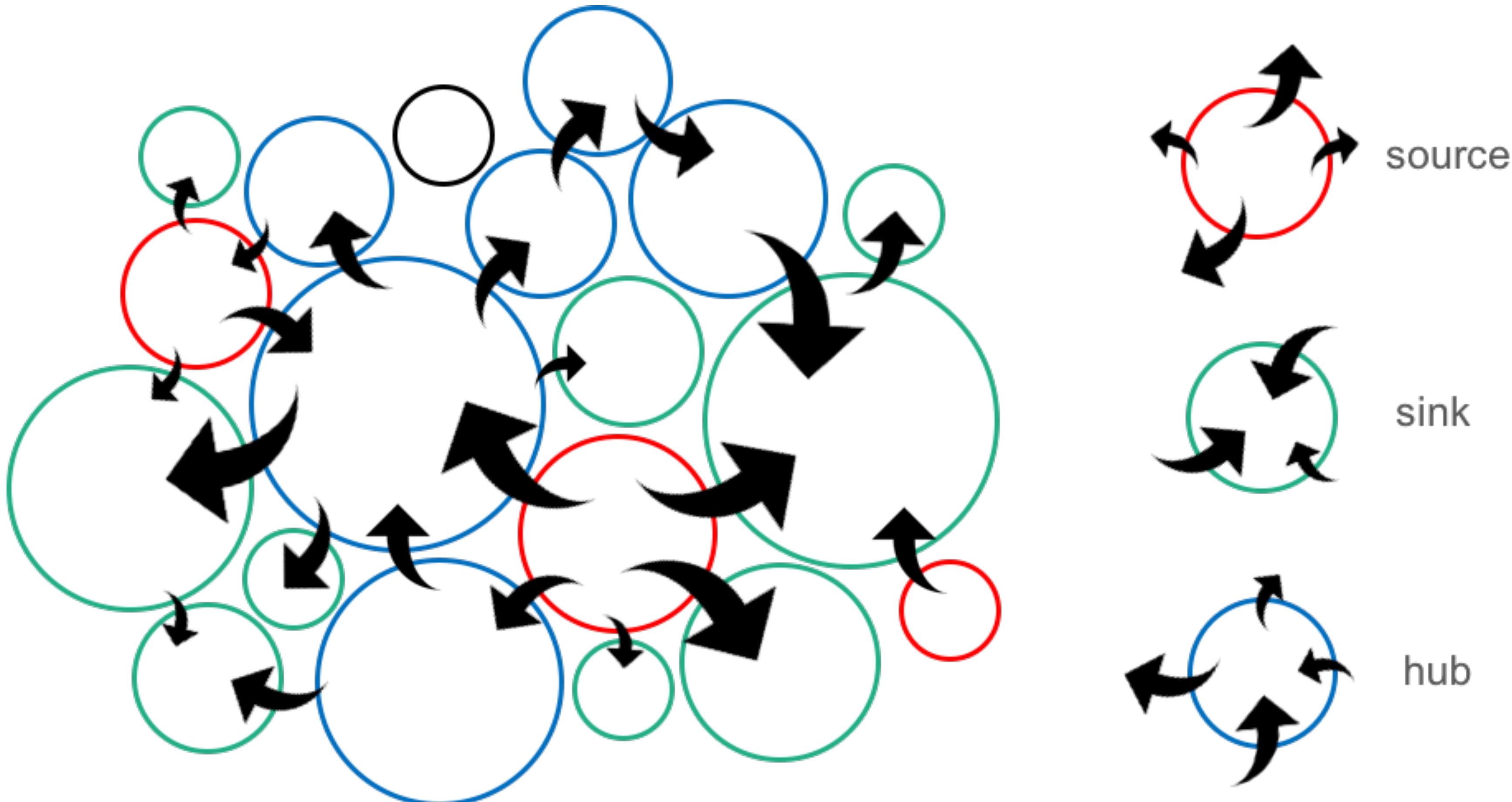
HPTN Executive Committee Webinar

1 July 2020



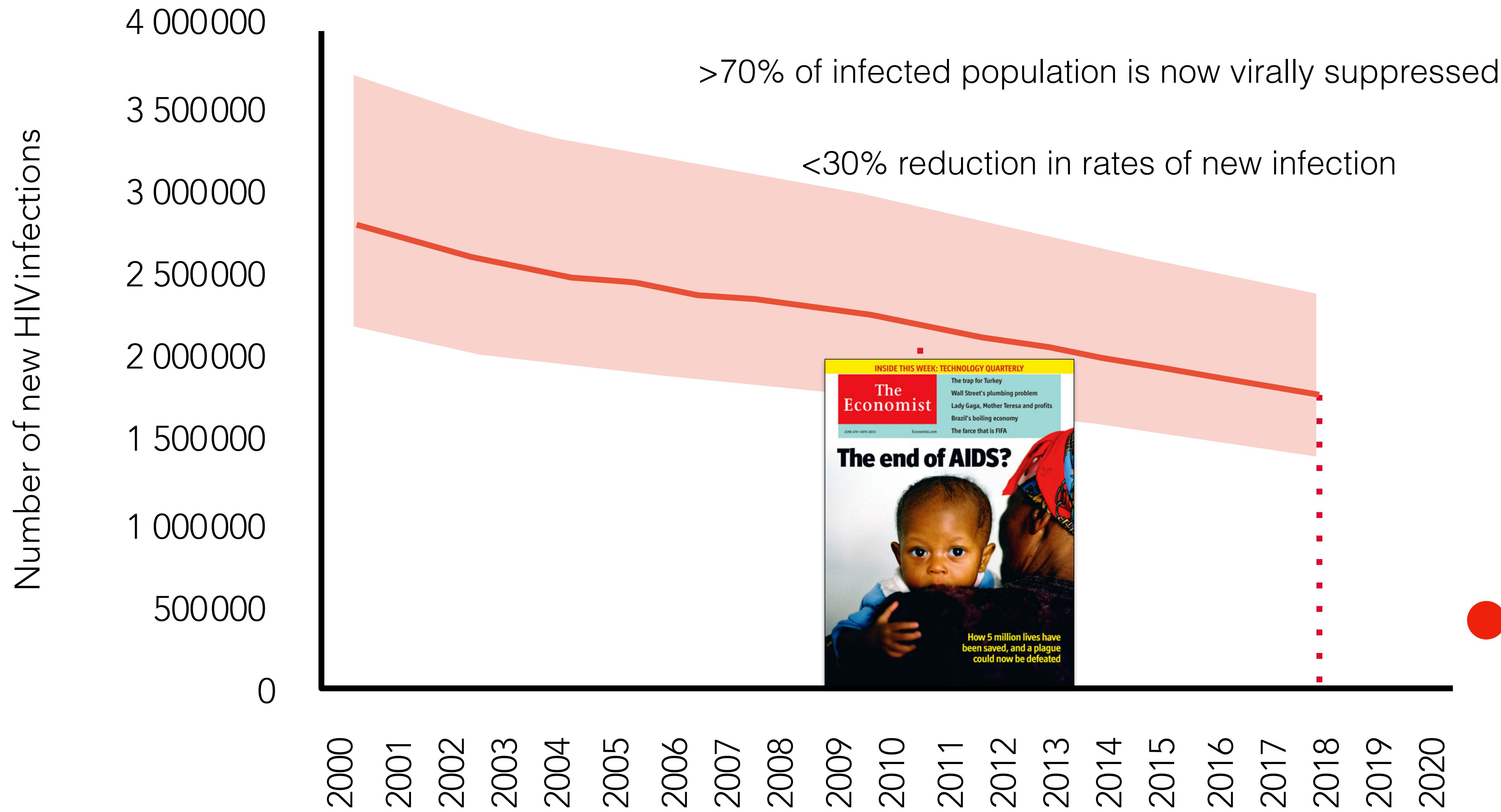
**U.S. NATIONAL INSTITUTES OF HEALTH:**  
*National Institute of Allergy and Infectious Diseases*  
*National Institute of Mental Health*  
*National Institute on Drug Abuse*

# Sources, sinks and hubs



Age groups, geographic locations, group identifiers, key populations,  
co-infected populations, self-identified high-risk

Which heterogeneities best explain the gap between expected and actual reductions in rate of new infections?



# What is phylogenetics?

As viruses spread, the genetic code (DNA) of each virus changes.

By comparing the genetic code of viruses from different people, phylogenetics is how we answer the question:

*“where did this virus come from, and when did it get here?”*

For PopART, we sequenced the genetic code of viruses from 8,152 participants, and performed phylogenetic analysis.



**Intervention communities receive household testing, prevention & linkage to care.**



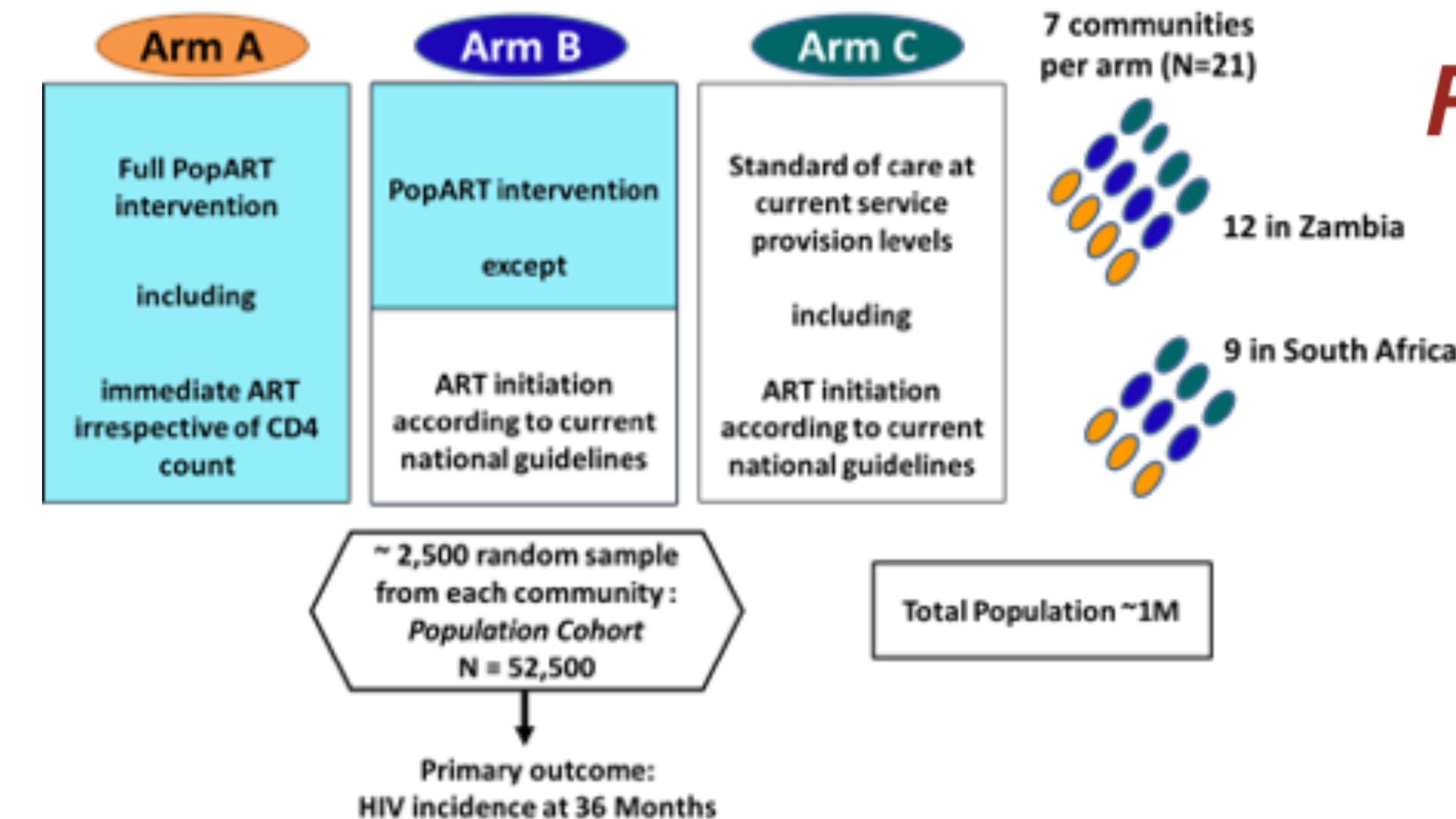
**PopART**  
HPTN 071

## Baseline adult prevalence

Arm A: 21%  
Arm B: 21%  
Arm C: 22%

## Viral suppression at 24 months

Arm A: 71.9%  
Arm B: 67.5%  
Arm C: 60.2%



## Incidence rate ratio

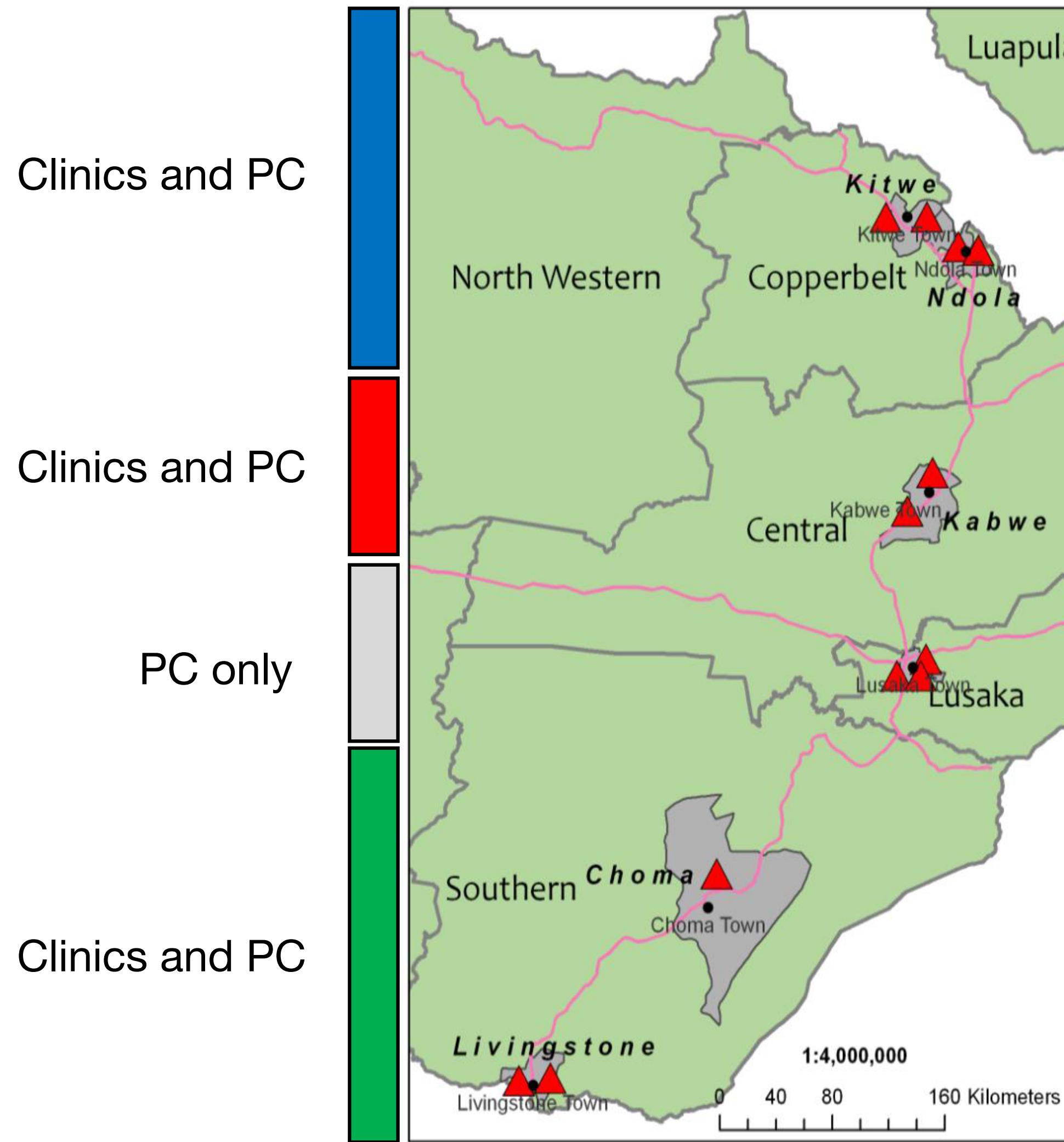
Arm A versus Arm C: 0.93 (0.74-1.18)  
Arm B versus Arm C: 0.70 (0.55-0.88)  
Arms A&B vs Arm C: 0.81 (0.66-0.99)

## Incidence time trend

12.3% reduction per year across all arms

Incidence (PC12-PC36)	Arm A	Arm B	Arm C	Incidence rate per 100 py
Age 18-24	1.75	1.50	1.72	
Age 25+	1.31	0.83	1.46	
Men, age 18-24	0.85	0.56	0.79	
Women, age 18-24	2.20	1.98	2.22	
Men, age 25+	0.74	0.39	0.94	
Women, age 25+	1.46	0.93	1.61	

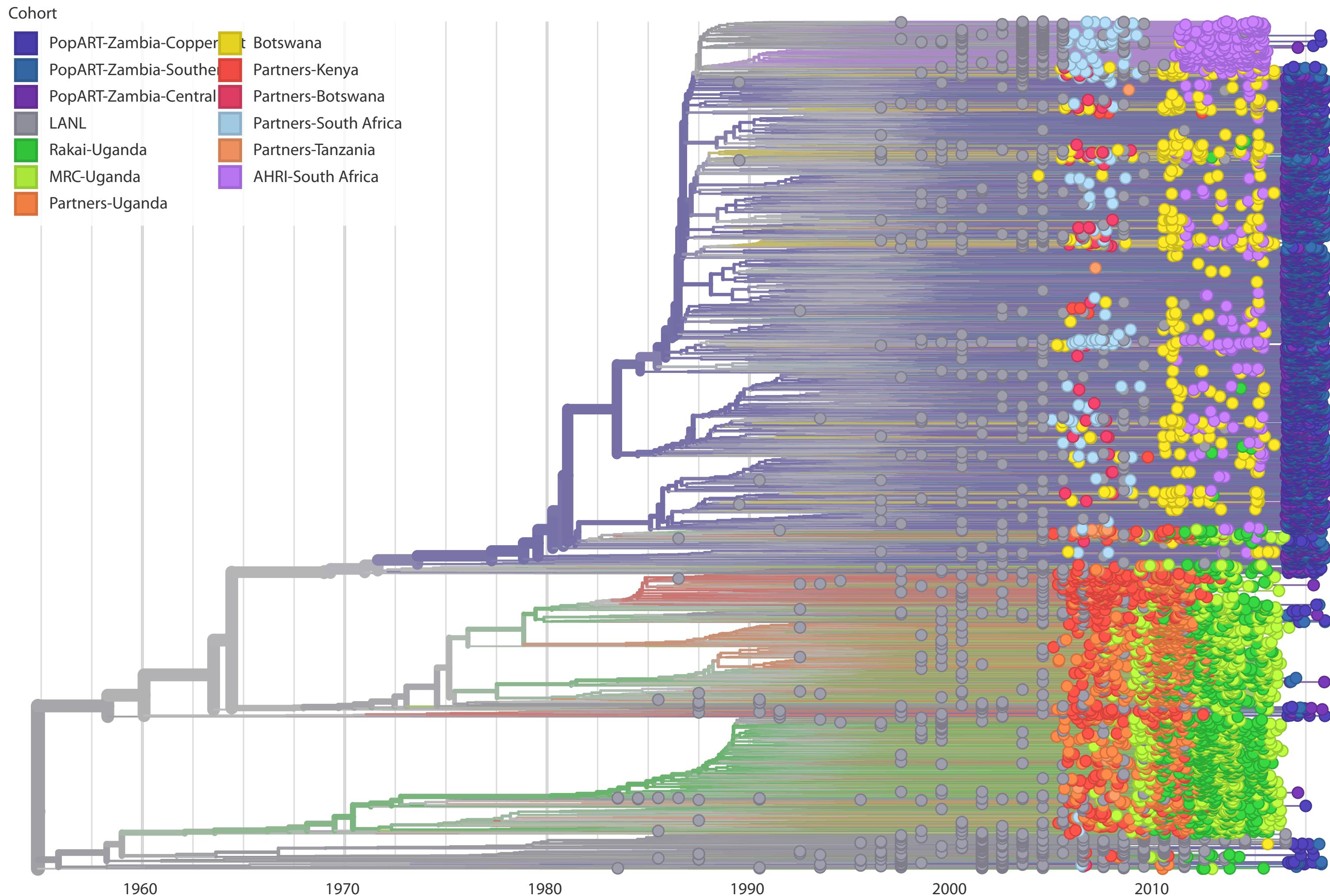
# The PopART Phylogenetics study took place in Zambia between 2016 and 2018



5648 participants recruited from clinics in 9 communities in Zambia.

Residual bloods from CD4 testing were used for sequencing.

2504 samples from Population Cohort (PC) in 12 communities in Zambia.

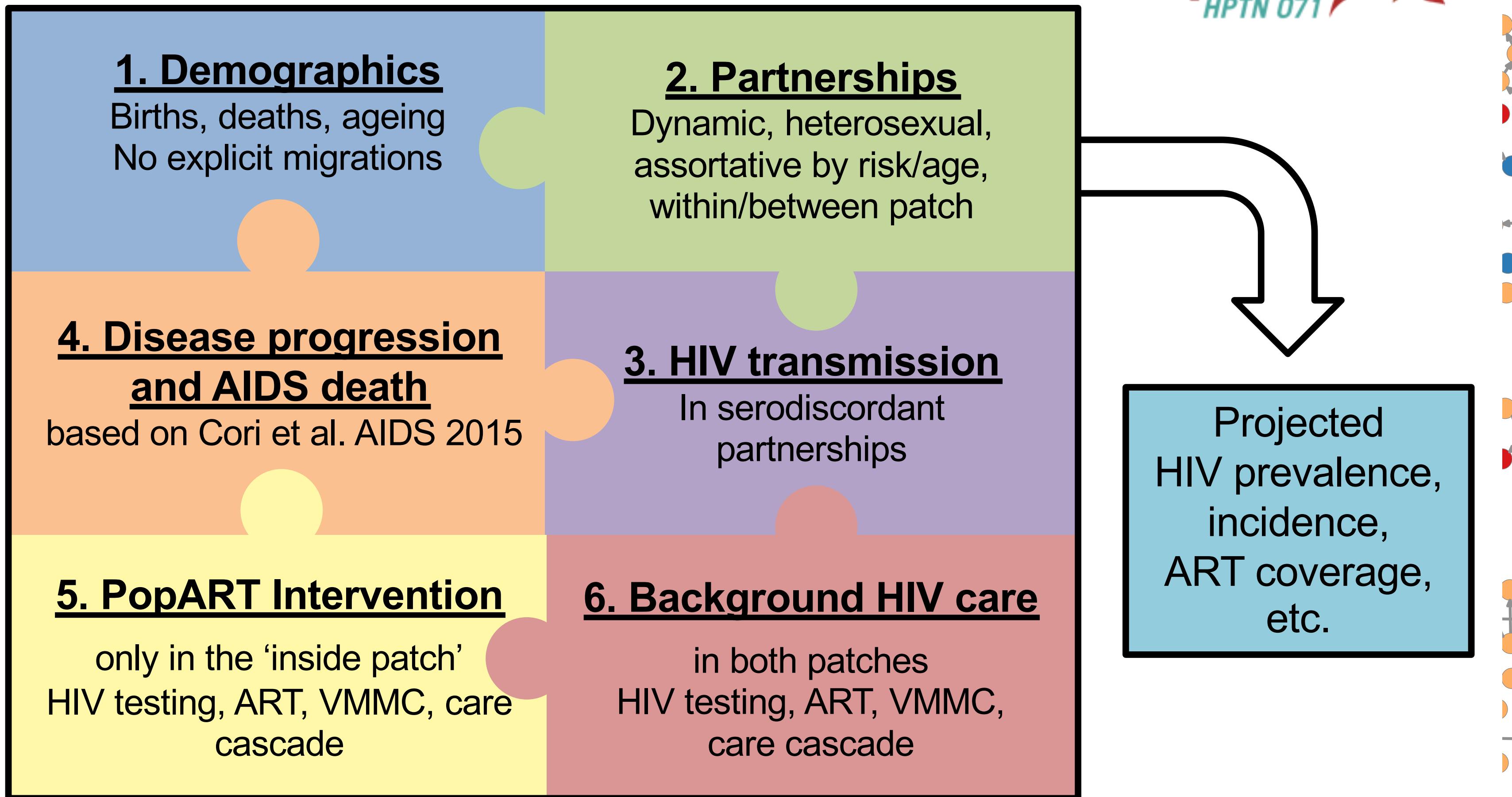


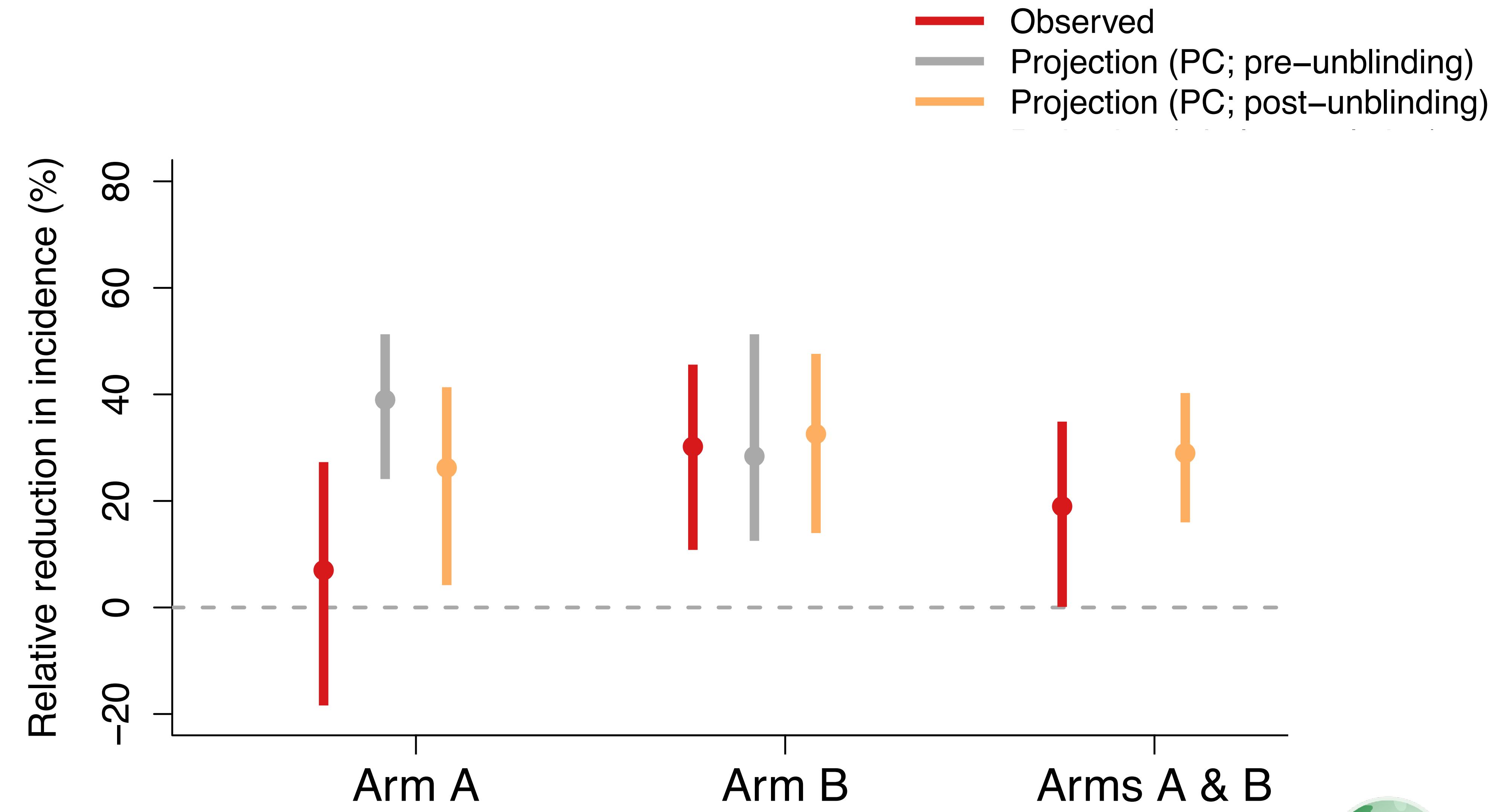
Often focus on clusters of closely related viruses.

*Rambaut et al - preliminary*

**Phylogenetics is the inference of the family tree of a sample.**

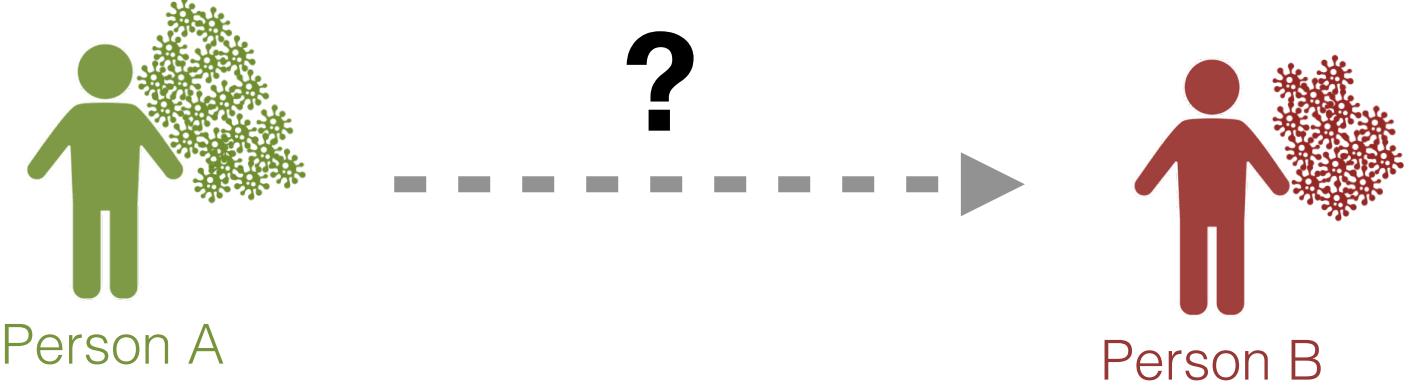
# Individual-based simulation model (IBM)





Probert et al, in preparation

# Phylogenetic output

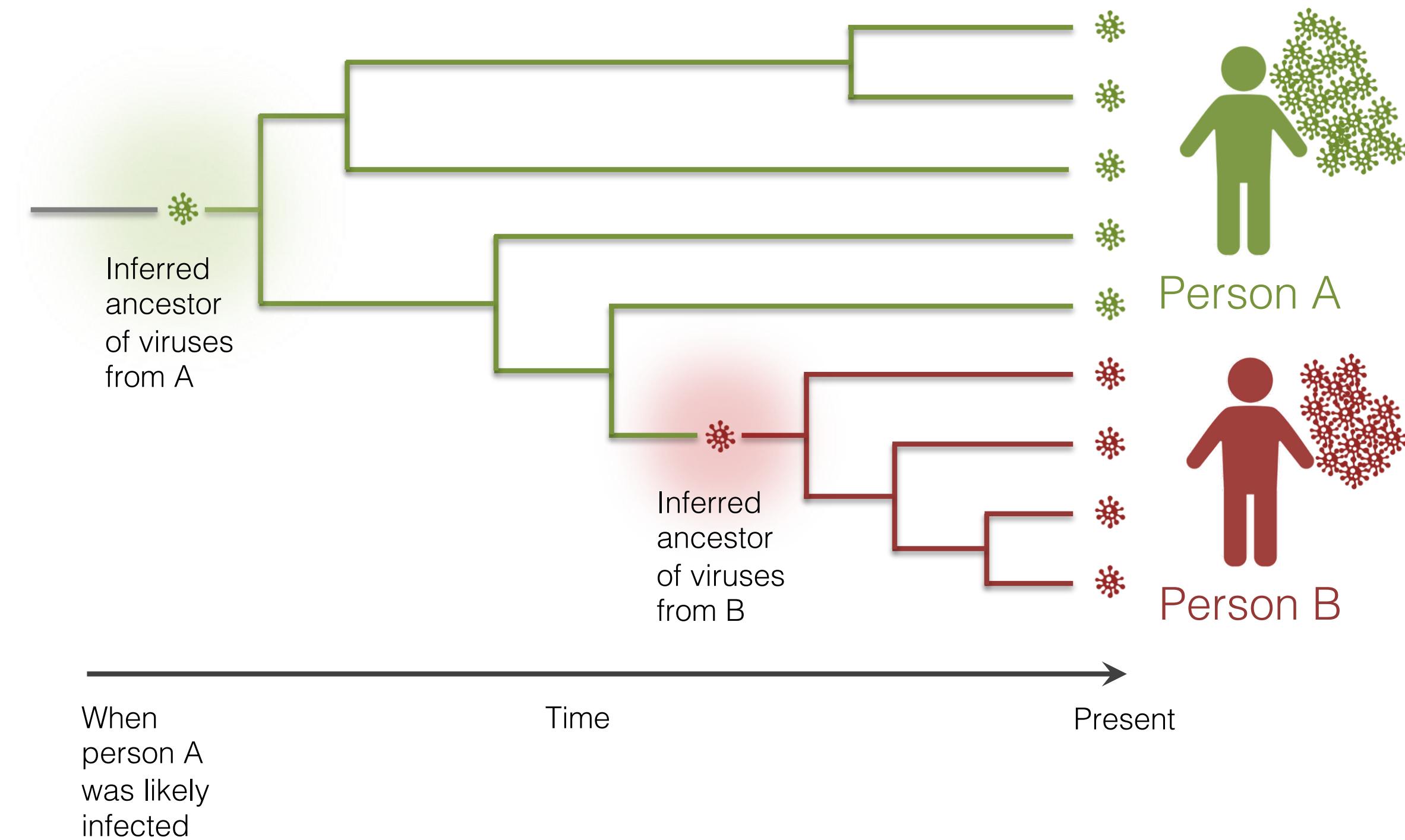


We infer that person A *may be* the source of infection for person B.

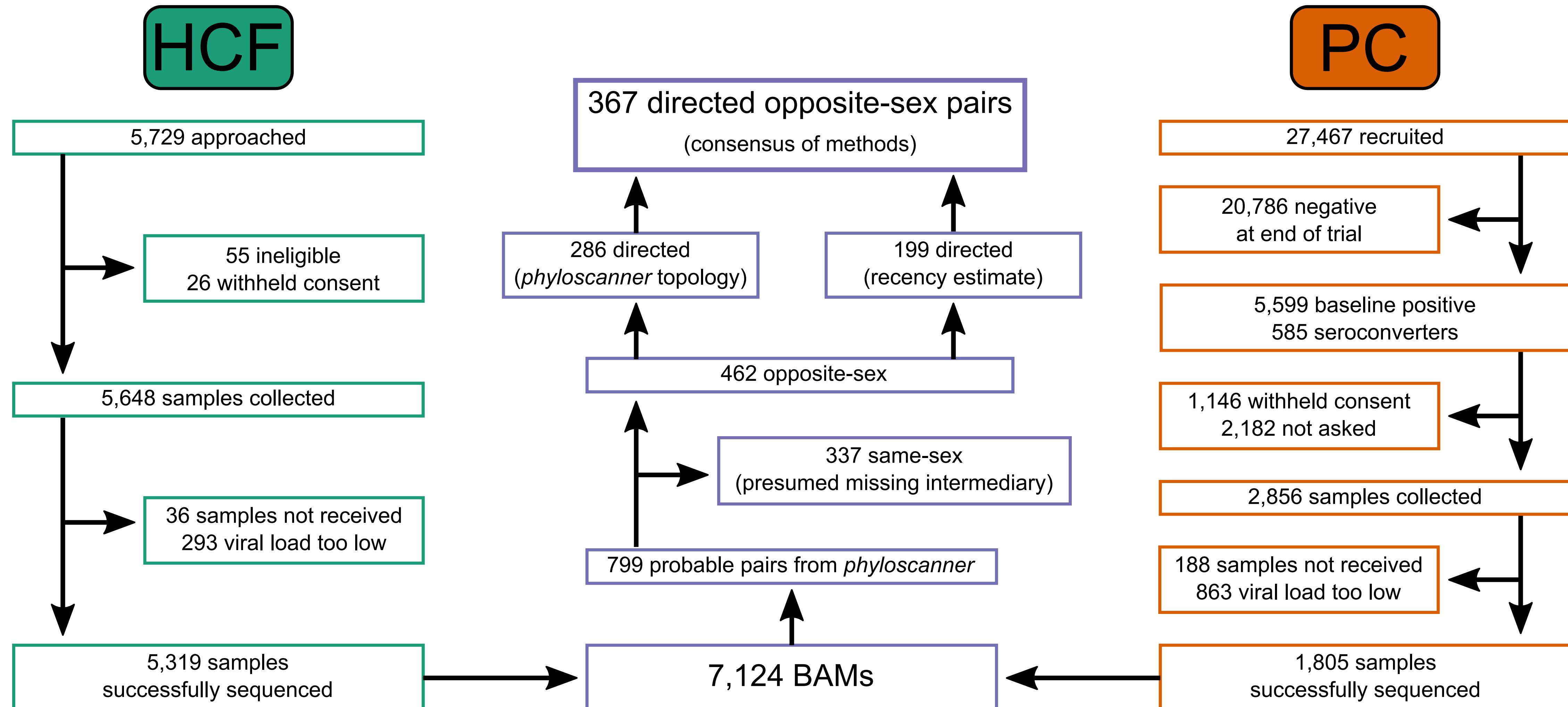
The analysis is anonymised.

We are only interested in the characteristics of transmission by comparing many such pairs.

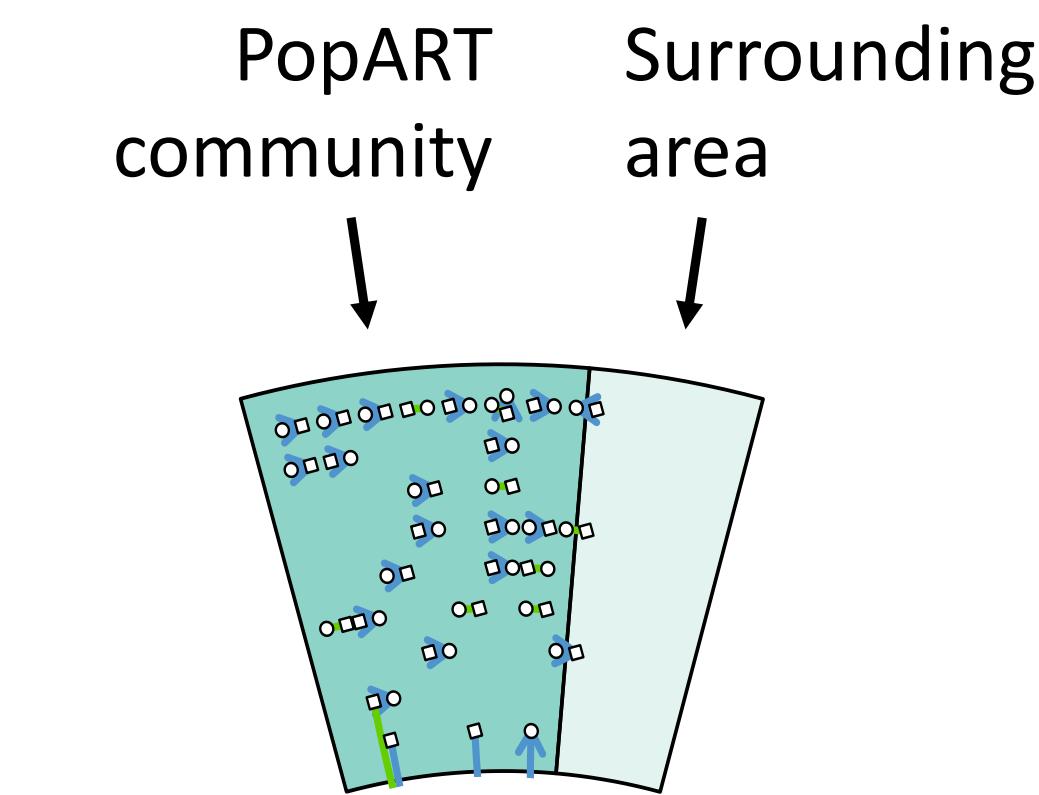
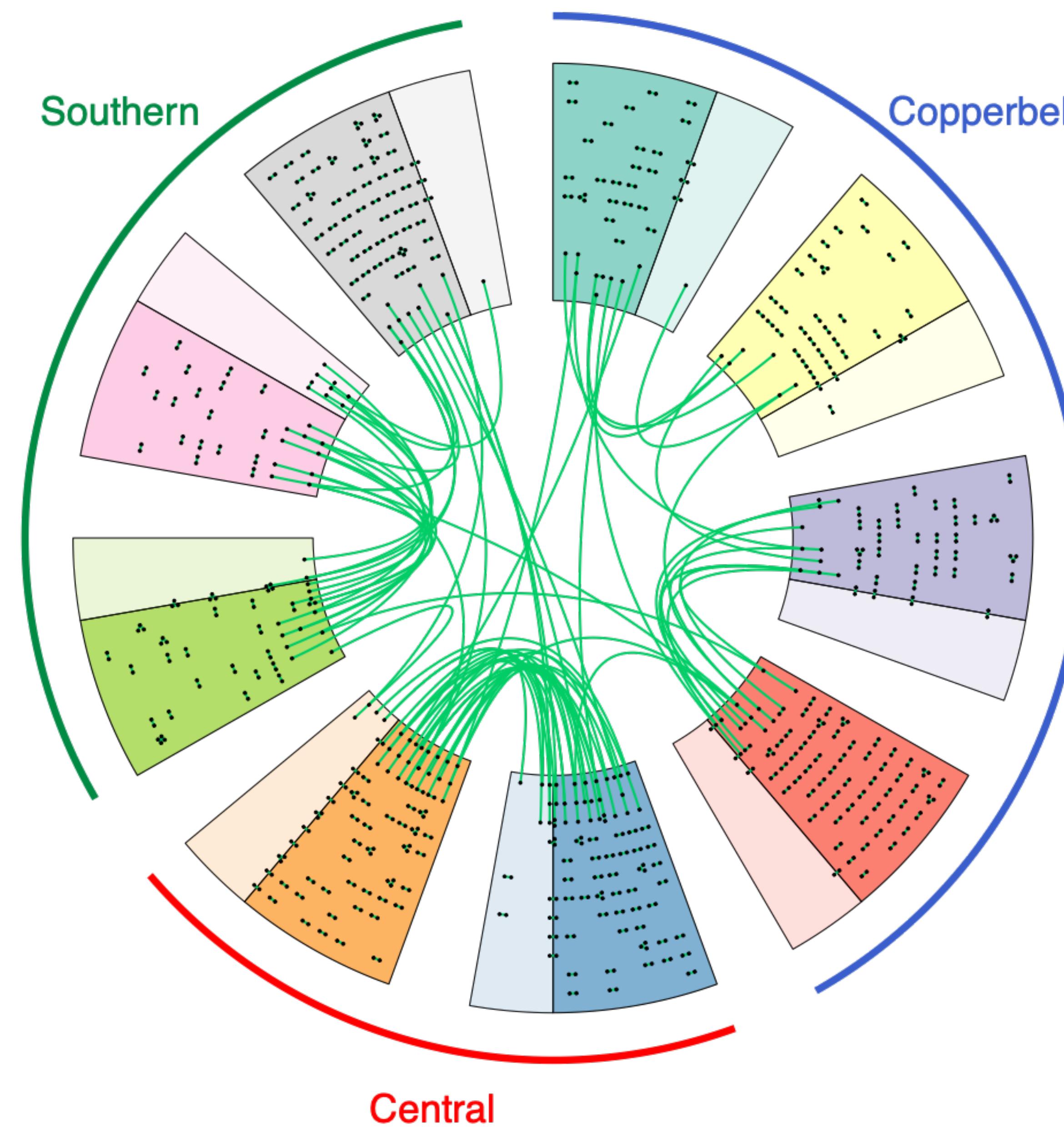
# Phylogenetic analysis



# 367 probable transmission pairs were identified from 7124 samples



# Transmission within and between PopART communities



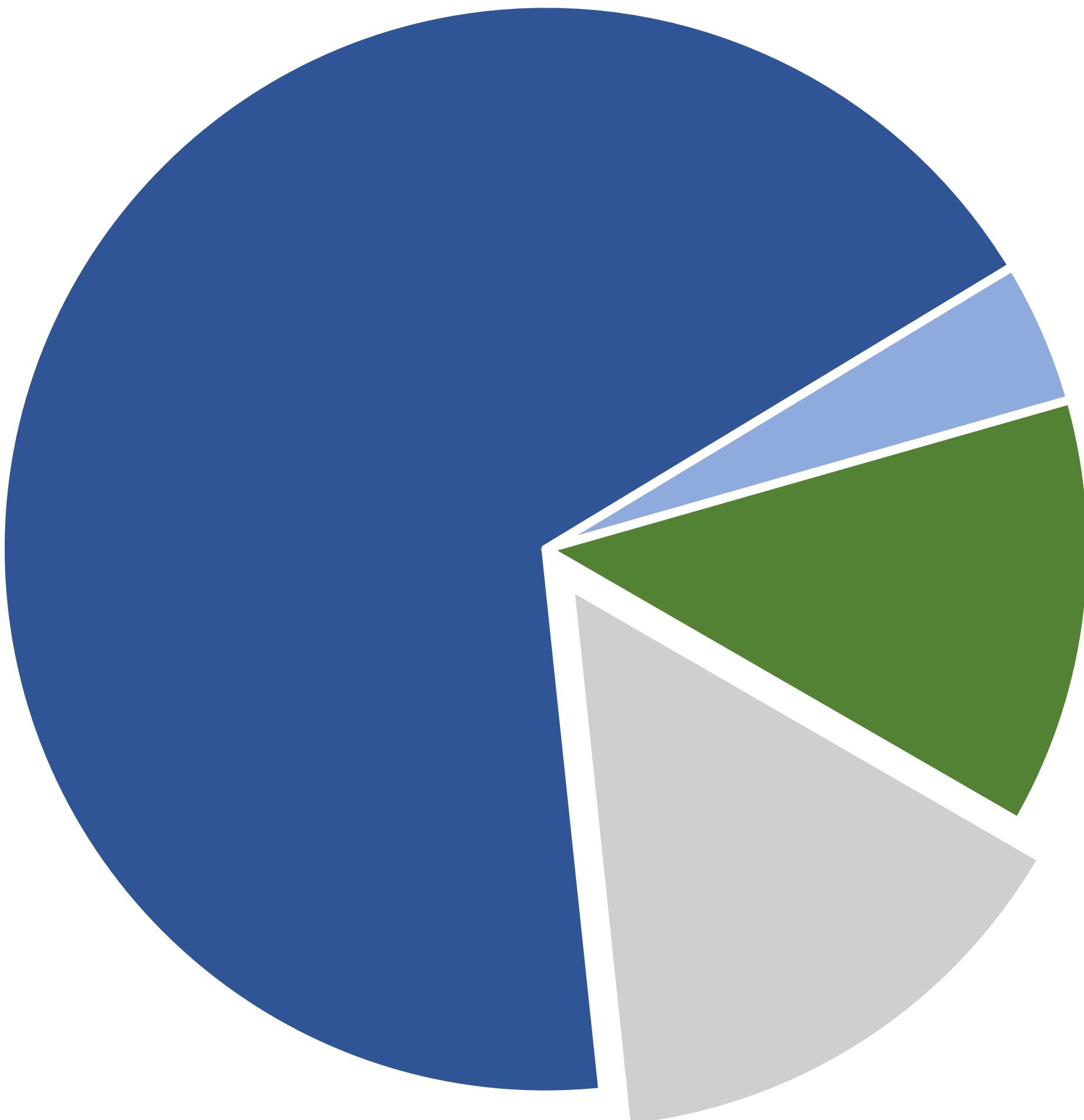
Community

01
02
04
05
03
06
10
11
12

462 probable transmission pairs

# Where do people get infected?

- Within community
- From surroundings
- From other communities
- Unsampled



## **Transmission within and between PopART communities**

Of transmissions we can link:

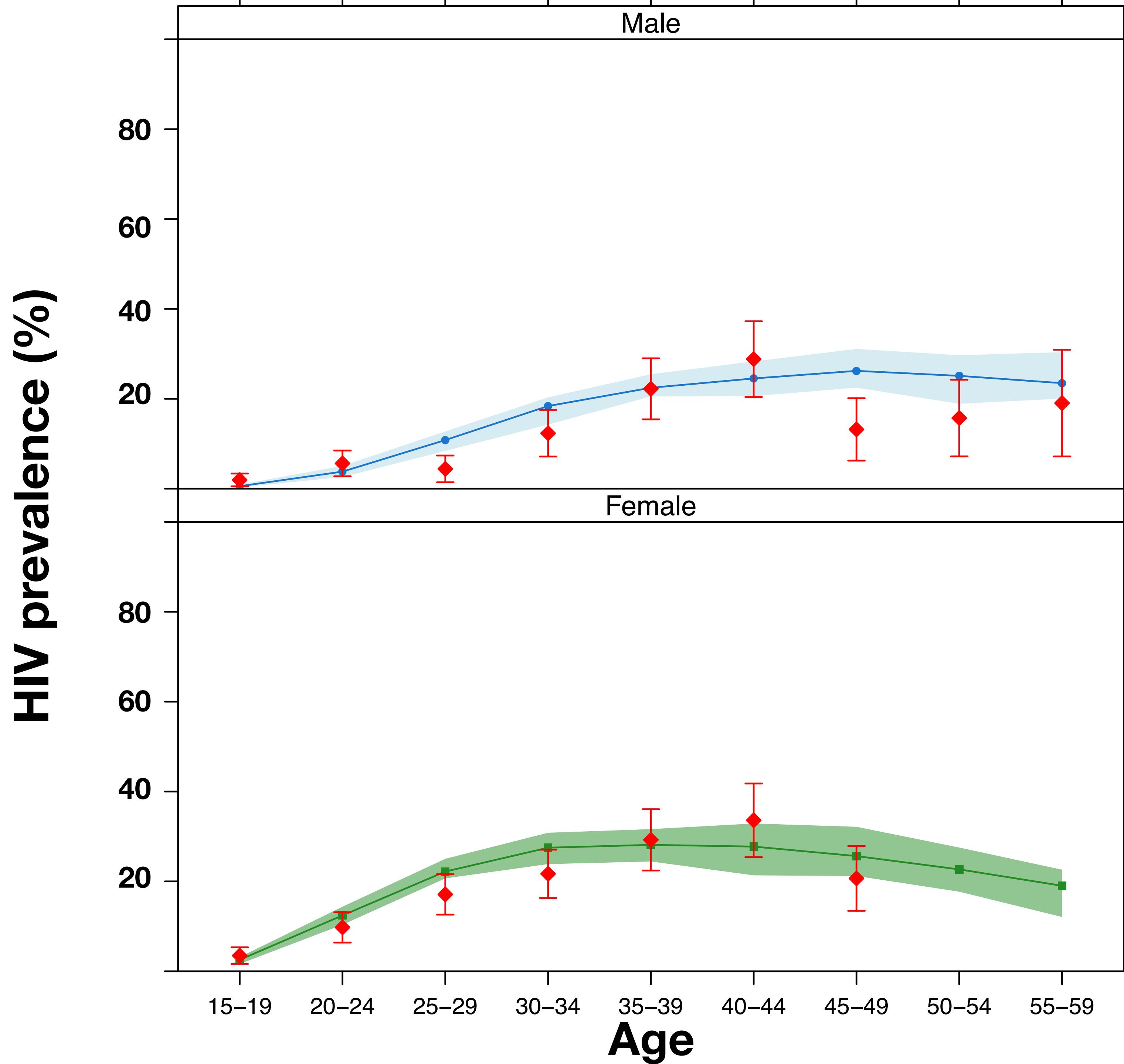
Most transmissions (80%) are within community. Participants should not assume infections are from elsewhere.

5% of transmissions from outskirts.

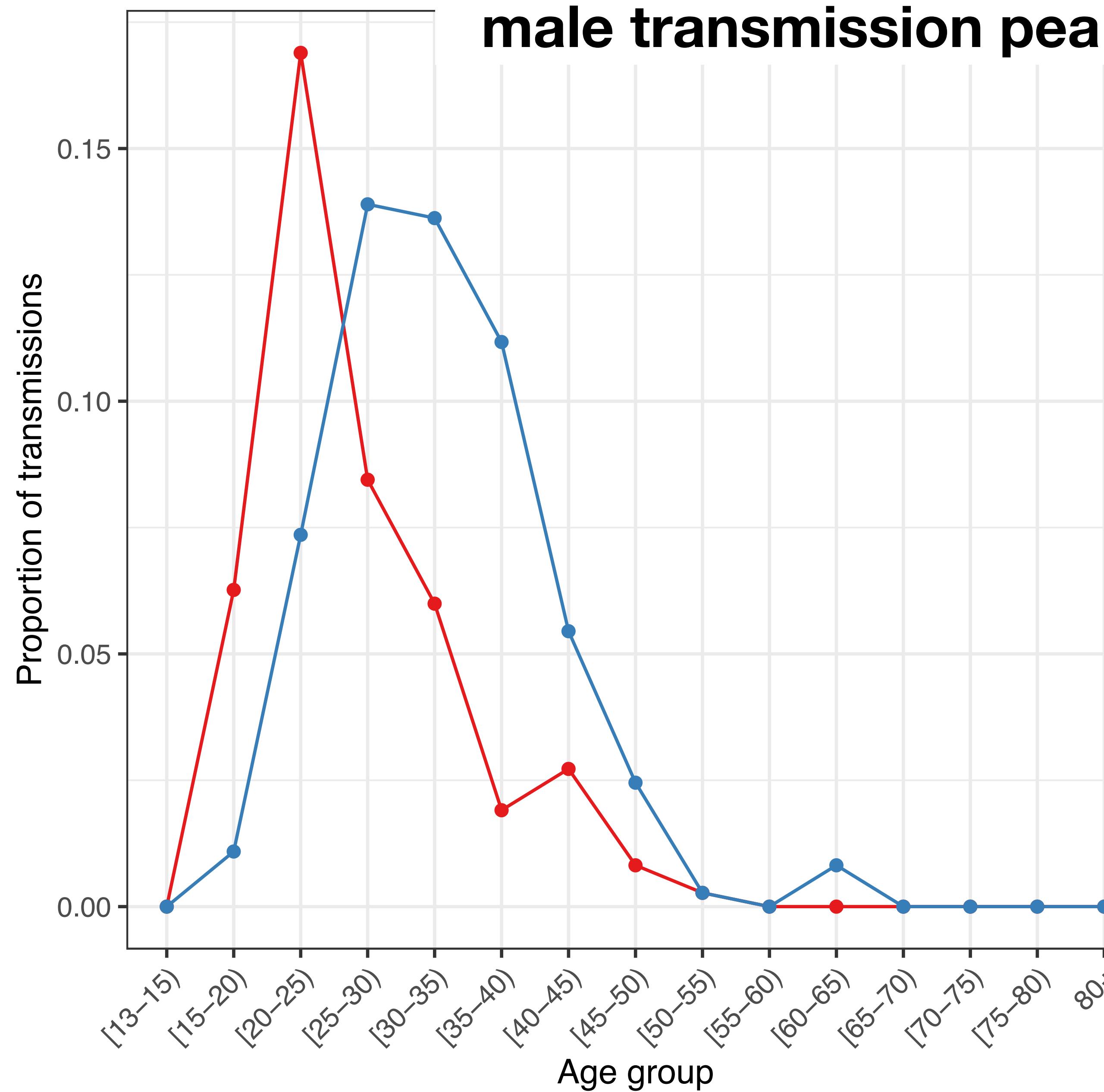
15% of transmissions from other communities.

**This movement may have caused us to under-estimate the impact of the PopART intervention.**

**Consider HIV prevalence by age and sex:**



# Female transmission peaks between ages 20 and 25; male transmission peaks between ages 25 and 35.



Source  
gender

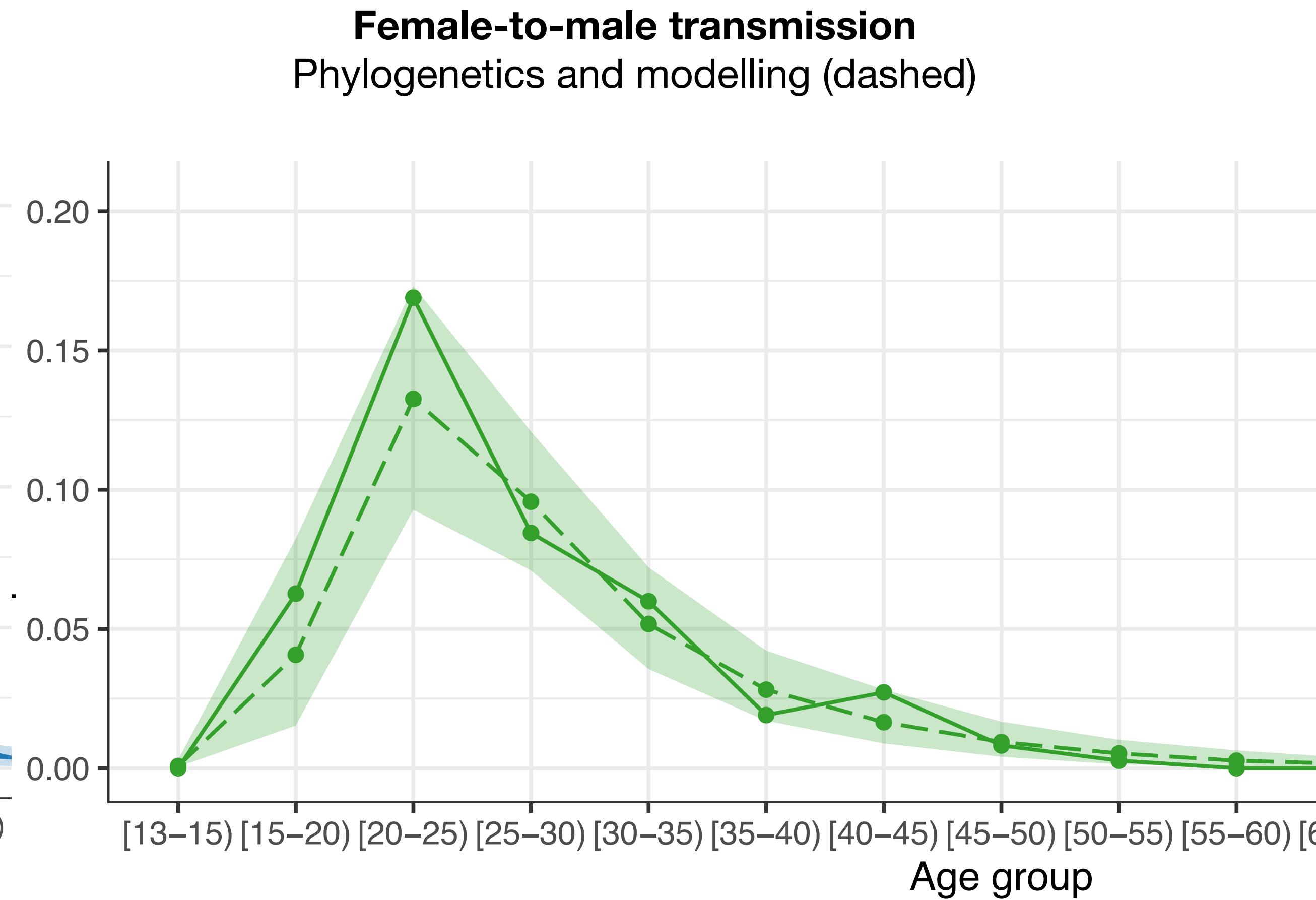
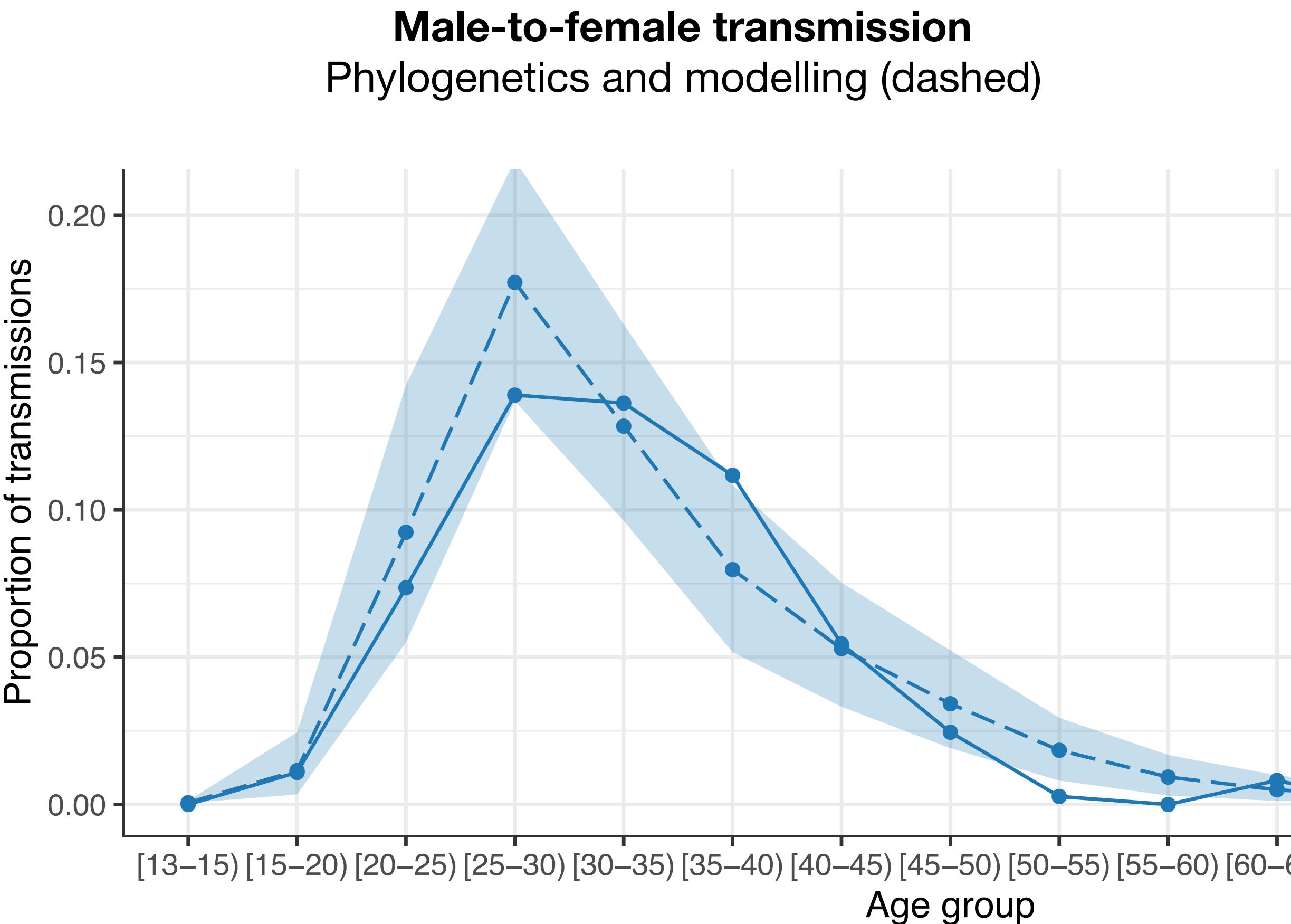
—●— F  
—●— M

56.7% of pairs had a male source

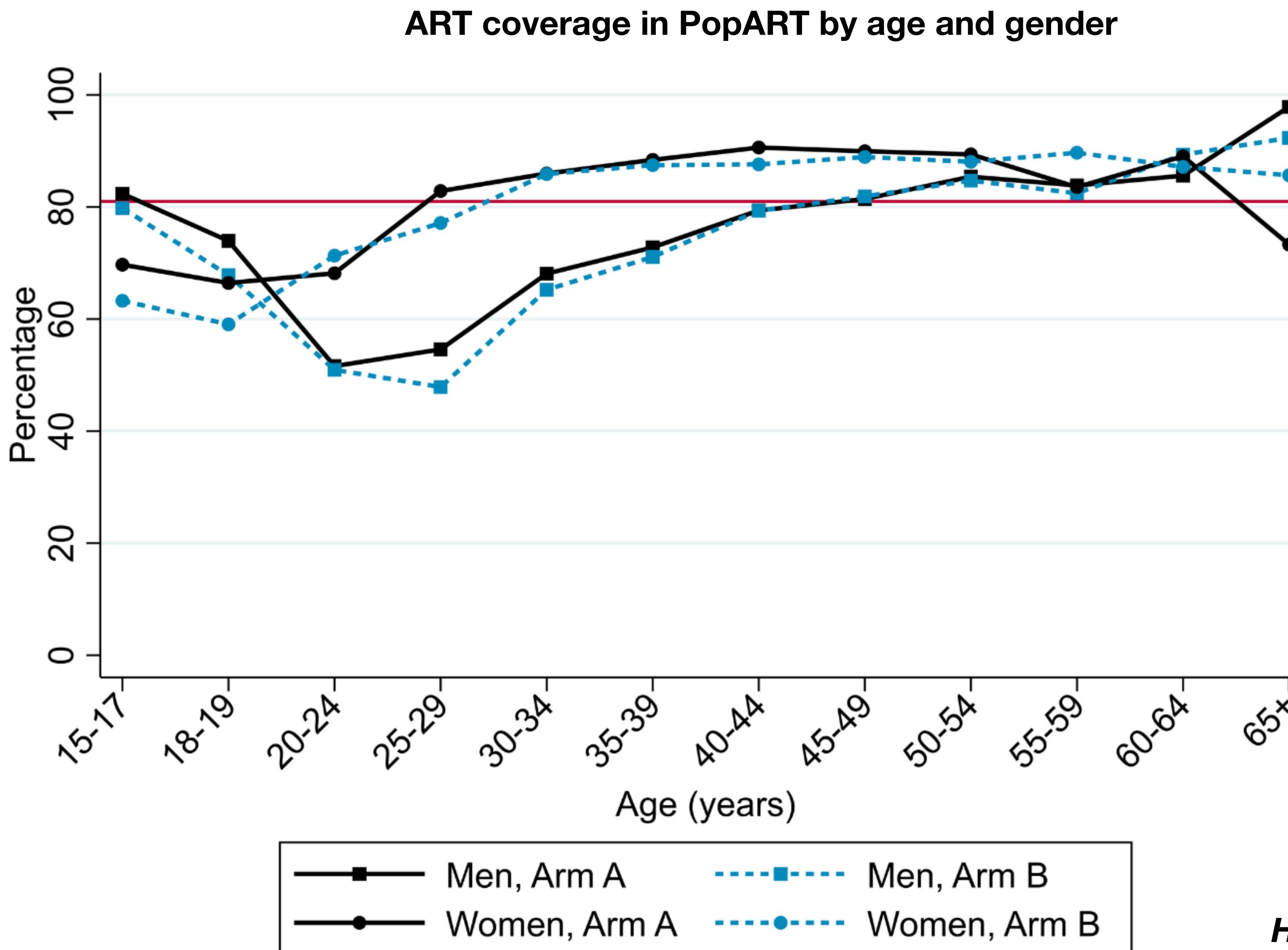
38.7% of pairs had a male source  
of age between 25 and 40

Probert, Hall et al, in preparation

# Independently, our simulation arrives at similar conclusions



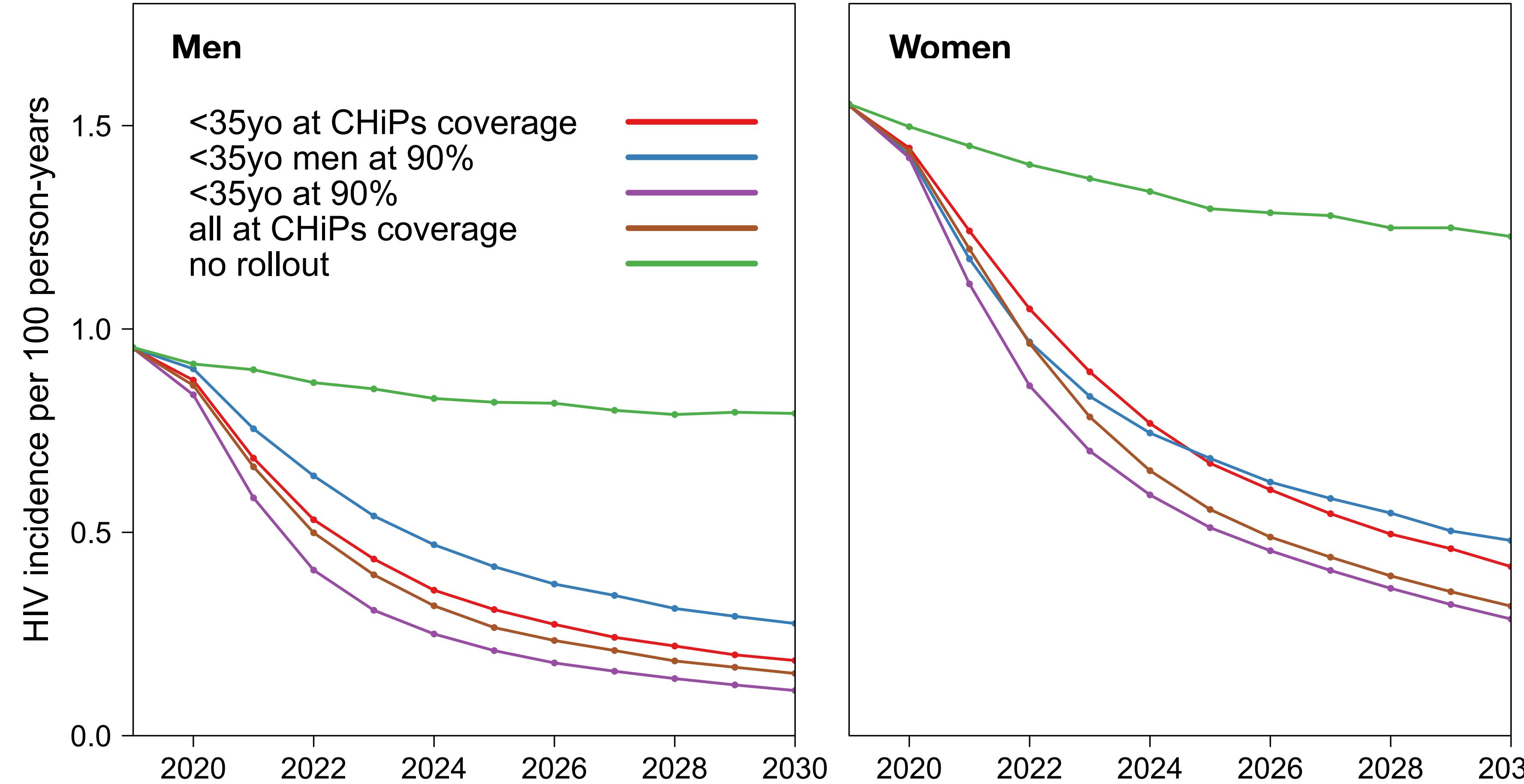
# The peak of transmission is explained by the dip in ART coverage, and by HIV prevalence: “Viral Burden”



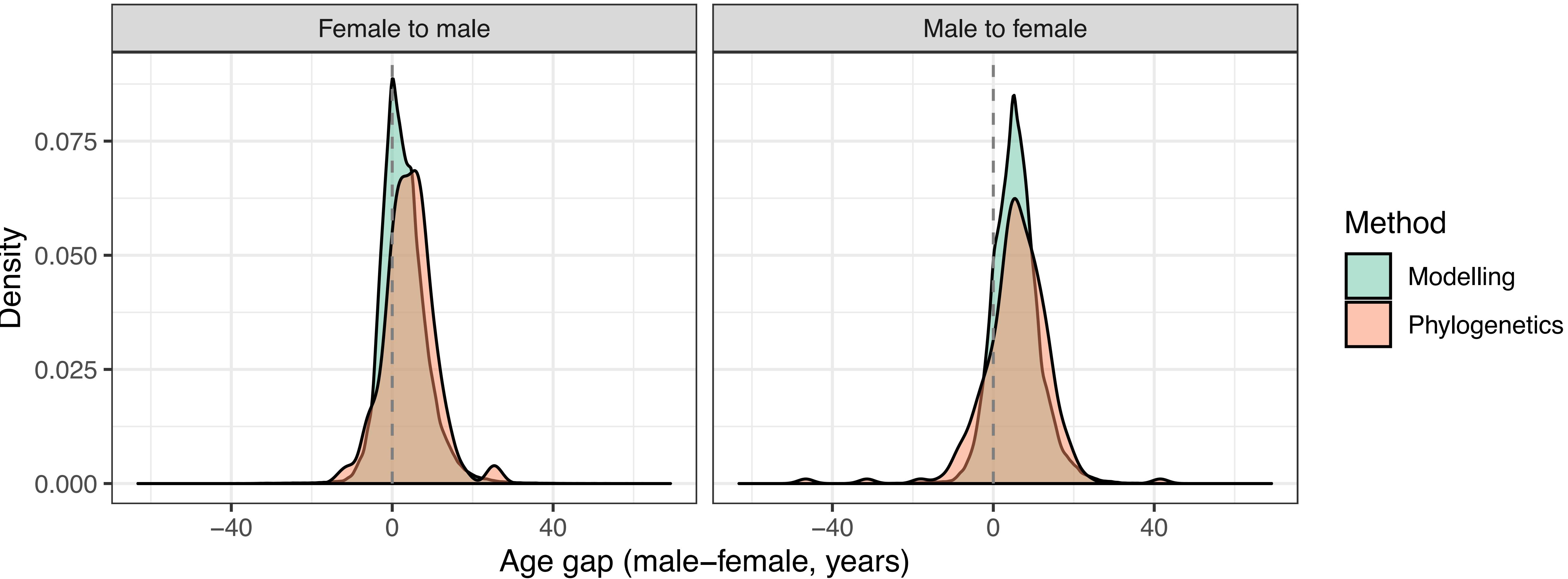
- The 90:90:90 targets were not achieved in men under 35 and women under 25.

Hayes et al - NEJM 2019

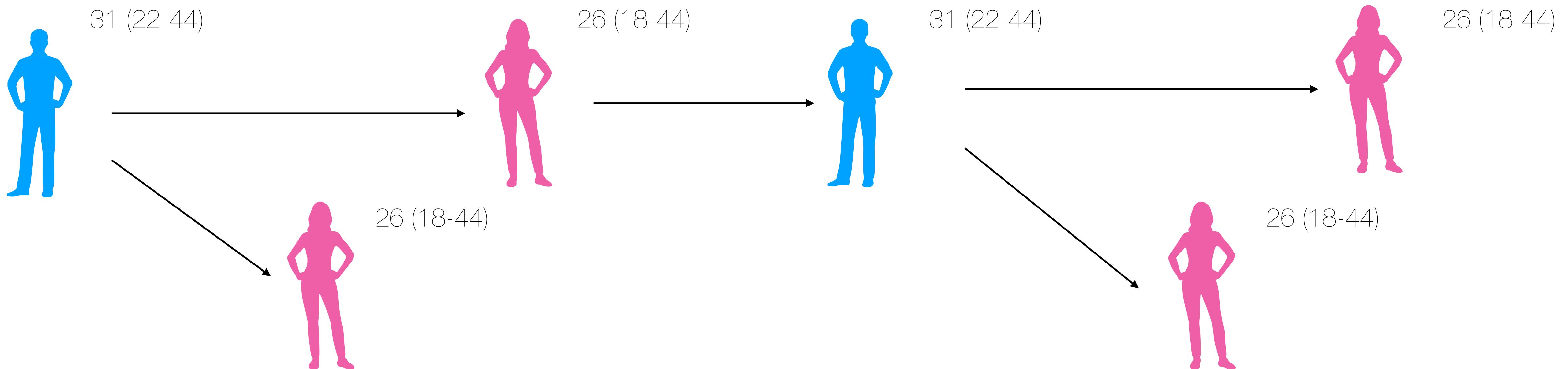
# Modelling in PopART suggests that reaching under 35 year-olds is almost as effective as reaching everybody



# Age difference between couples is 3-6 years



# This study suggests a ‘basic motif’ of transmission



**The asymmetric cycle of transmission**

# What is the generation time?



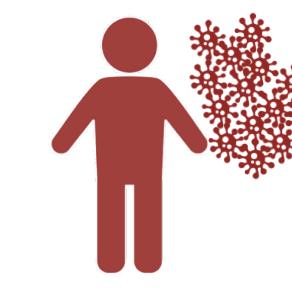
The generation time is how long it takes, in years, between A becoming infected and transmitting to B.

Transmissions occur in the first year are difficult to prevent with UTT.

# What is the distribution of generation times.



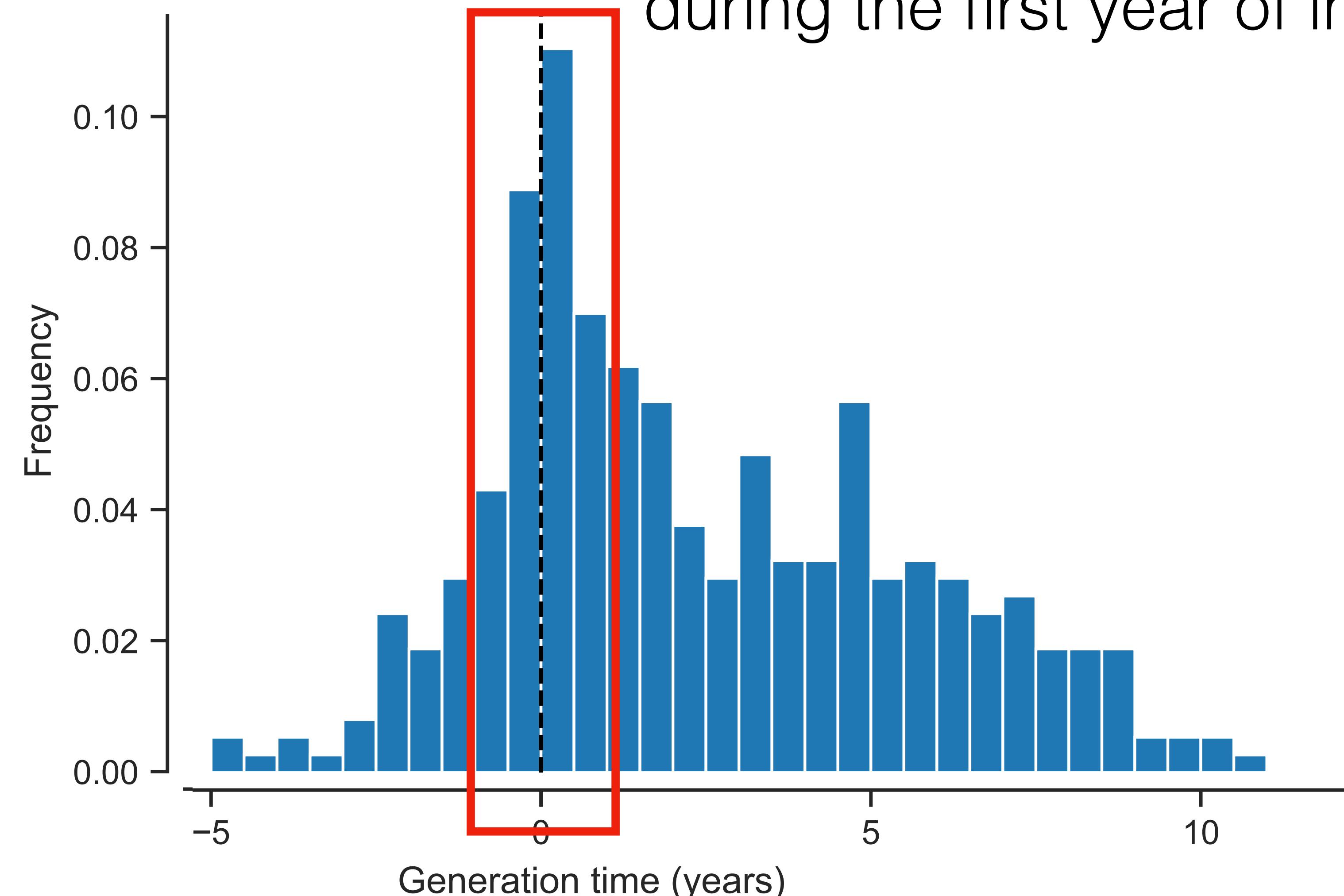
Person A



Person B

How long has person A been infected when they transmit?

30-40% of transmissions occur during the first year of infection.



## **veSEQ-HIV: a new protocol for drug resistance prediction using low-cost sequencing of HIV**

veSEQ-HIV: High-throughput, low-cost, full-genome HIV sequencing for quantification of viral load, characterization of transmission, recency of infection, and drug resistance.

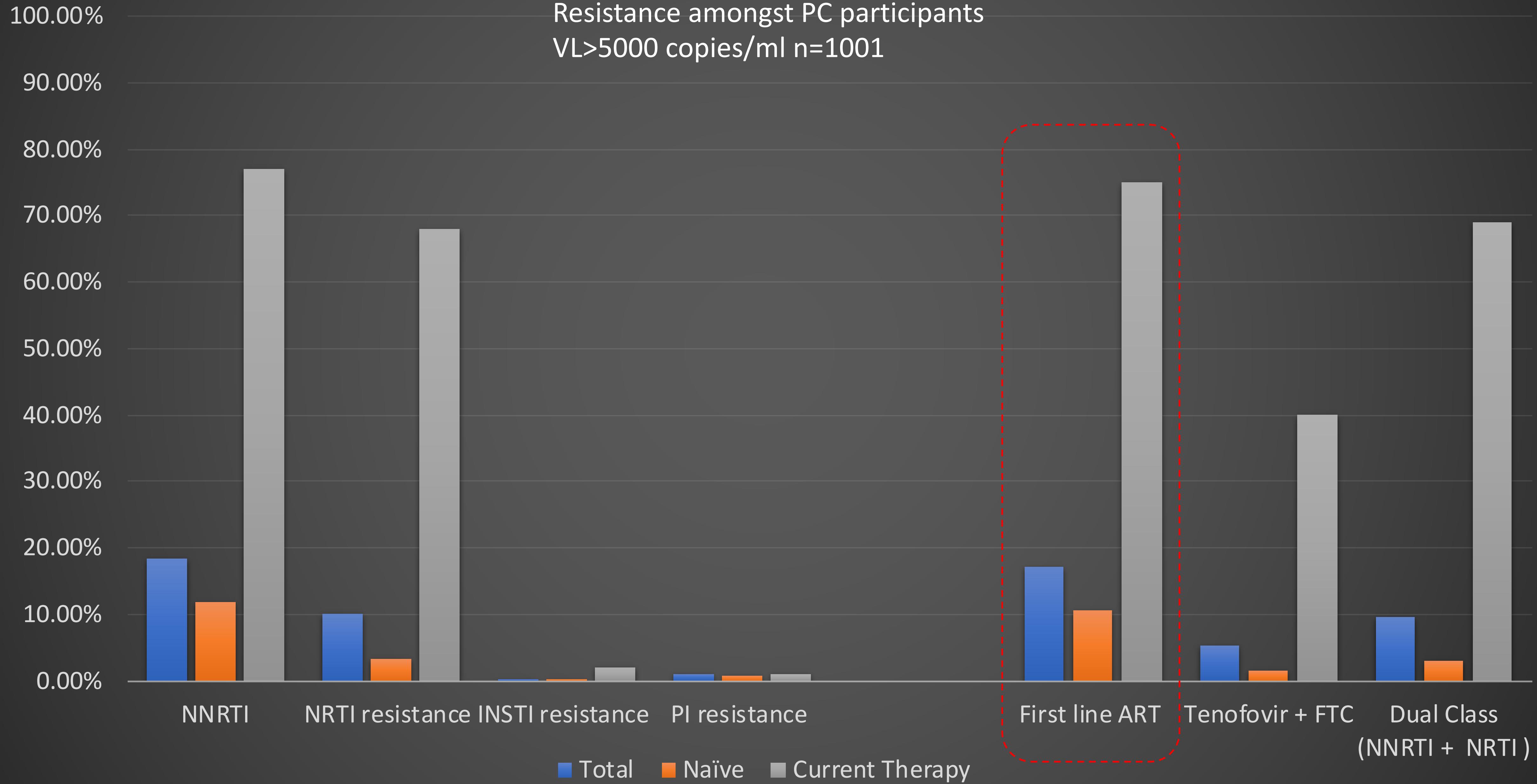
drmSEQ: Integrated bioinformatics for drug resistance calling.

Validation: The full pipeline (veSEQ-HIV + drmSEQ) was validated against an FDA-accredited clinical drug resistance test (ViroSeq) (HPTN-078 study - submitted for publication)

# We used sequencing protocol and drug resistance calling on the PopART dataset

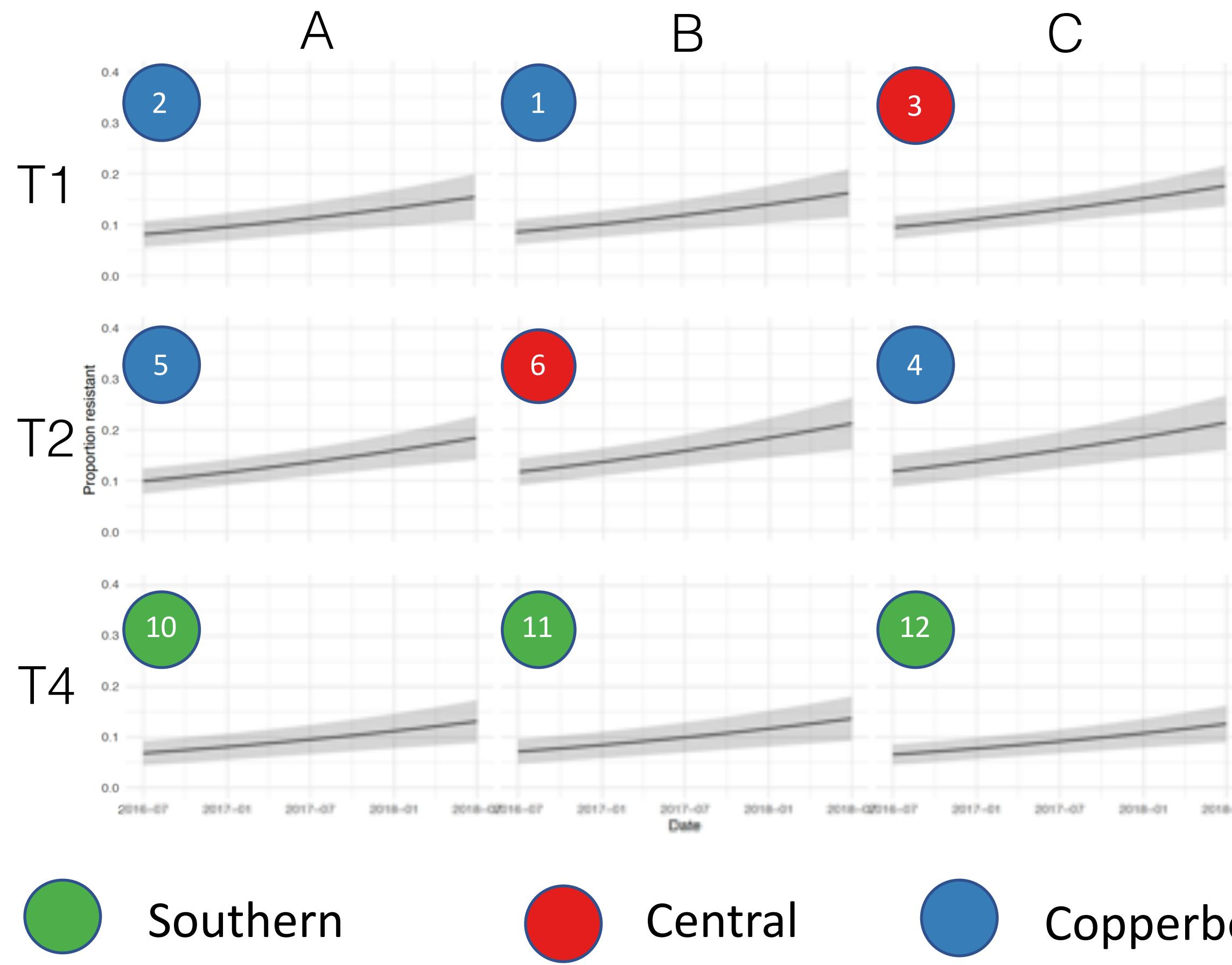
		Health Care Facility		Population Cohort	
		N	%	N	%
	Total sequenced	5,610		2,574	
Viral Load	VL<5000 (excluded)	115	2.36	1,573	61.11
	VL>=5000 (included)	4,756	97.64	1,001	38.89
Gender	Male	2,116	44.49	162	16.18
	Female	2,640	55.51	839	83.82
ART Status (self-reported)	ART naive	4,575	96.19	850	84.92
	Previous ART	81	1.70	1	0.10
	ART current	99	2.08	96	9.59
	Missing Info	1	0.02	54	5.39
Study Arm	Arm A	1,491	31.35	351	35.06
	Arm B	1,518	31.92	297	29.67
	Arm C	1,747	36.73	353	35.26
Year at sampling	2014	NA	NA	456	45.55
	2015	NA	NA	160	15.98
	2016	1,601	33.66	156	15.58
	2017	2,319	48.76	197	19.68
	2018	835	17.56	31	3.10

# Drug resistance observed in the viraemic individuals in the Population Cohort

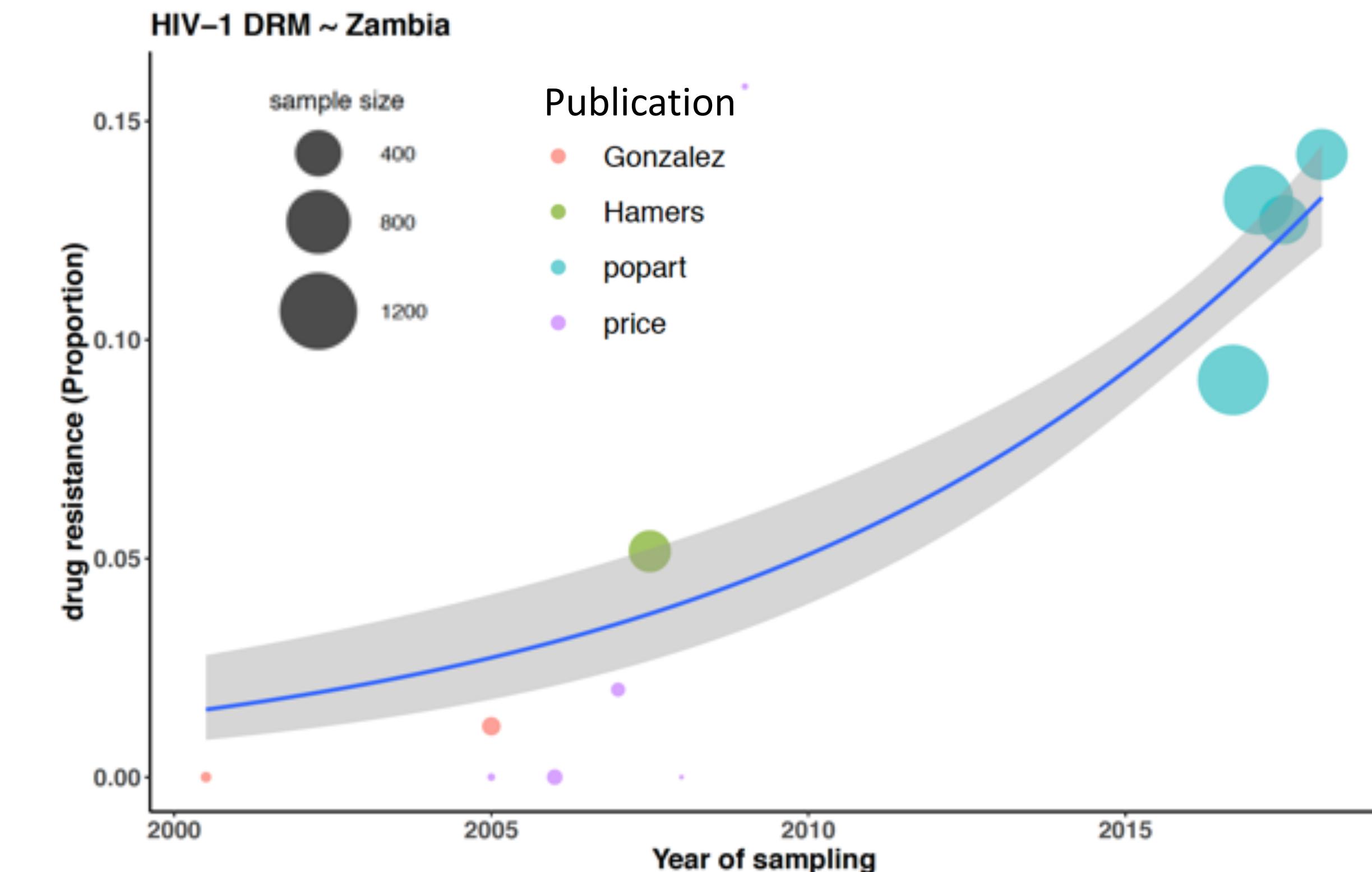


# In PopART, we observe a rise across all 9 communities in Zambia

Increasing drug resistance observed in all communities throughout the trial (HCF data only)



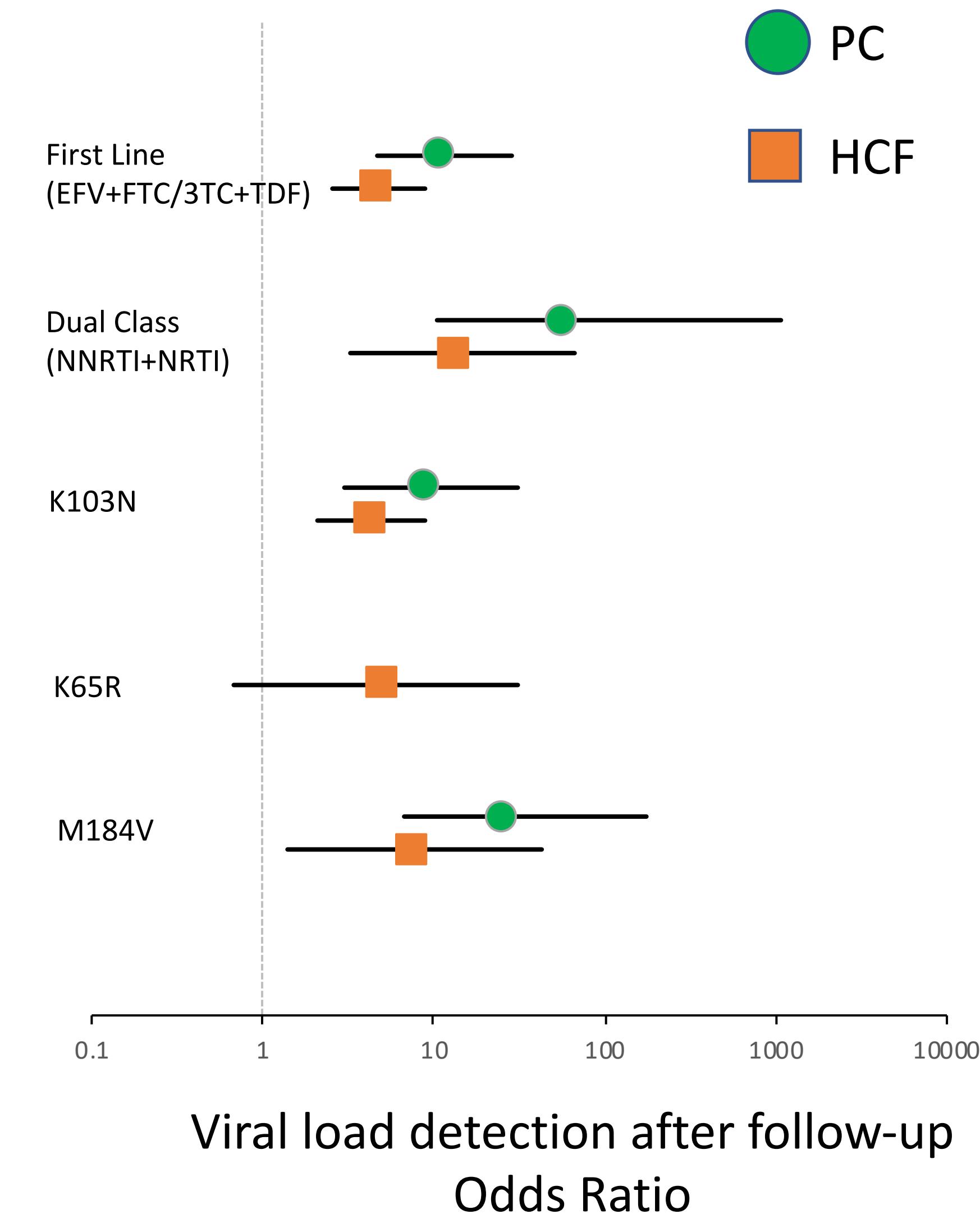
Significant rises compared with the historic studies



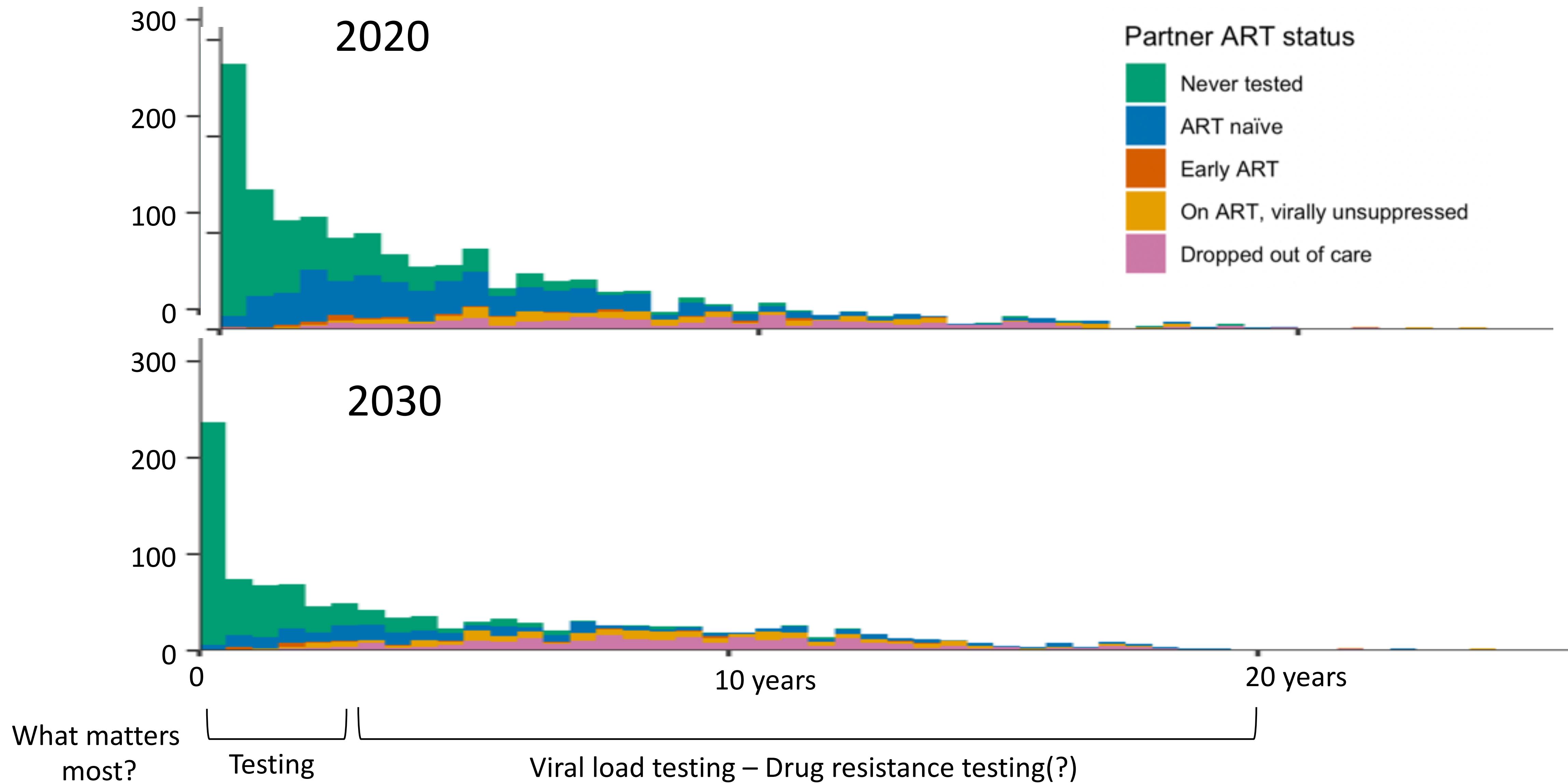
# A substantial number of people is failing treatment as a result of undiagnosed drug resistance

- Viral was measure in a proportion of individuals after baseline sampling
- Viraemia was detected more frequently in individuals with drug resistance mutations detected at baseline.
- Odds ratios:

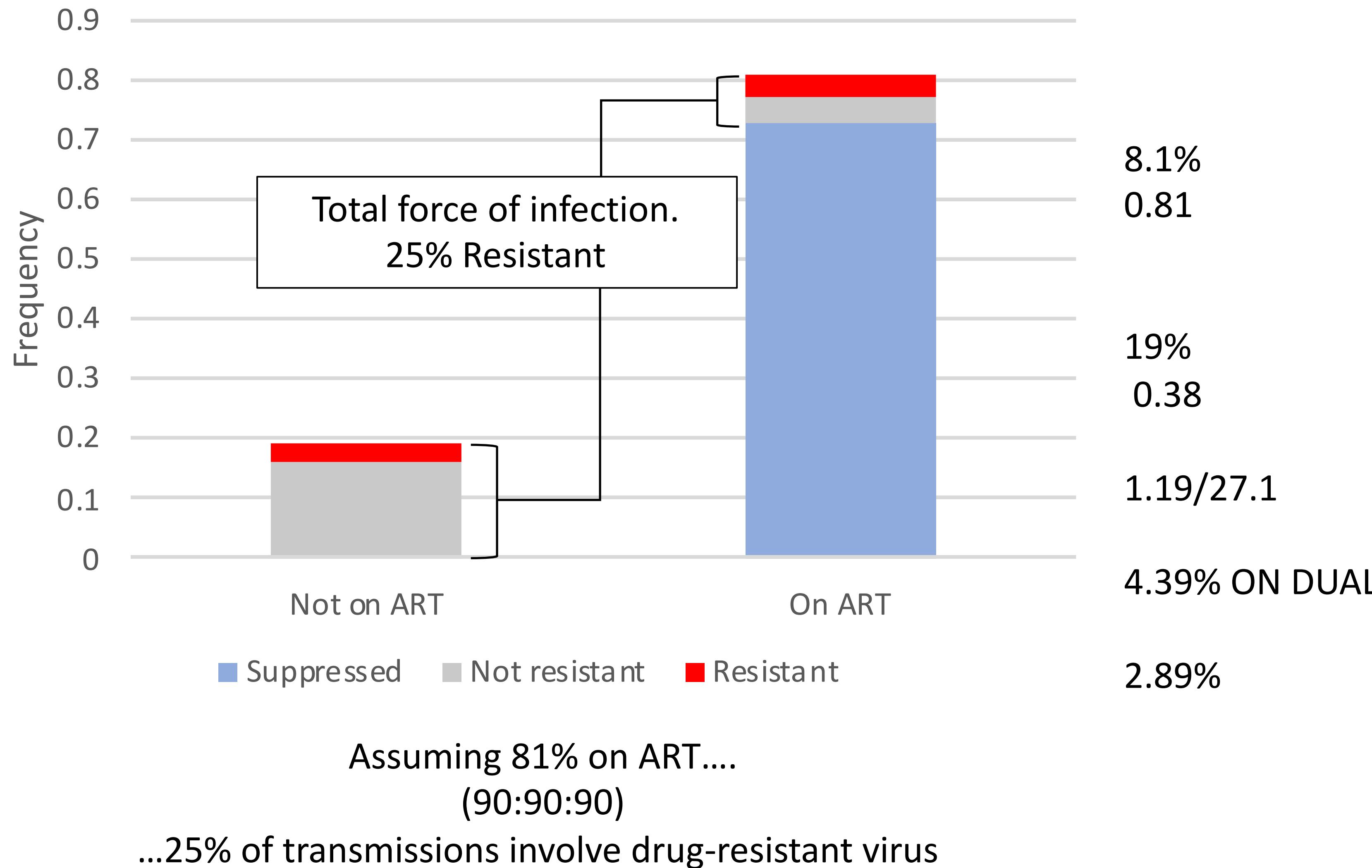
Baseline resistance detected vs. not detected



# Modelling HIV transmission events: Time to transmission and characteristics of source



# In presence of high ART coverage a significant proportion of infections arise from individuals on failed ART regimens

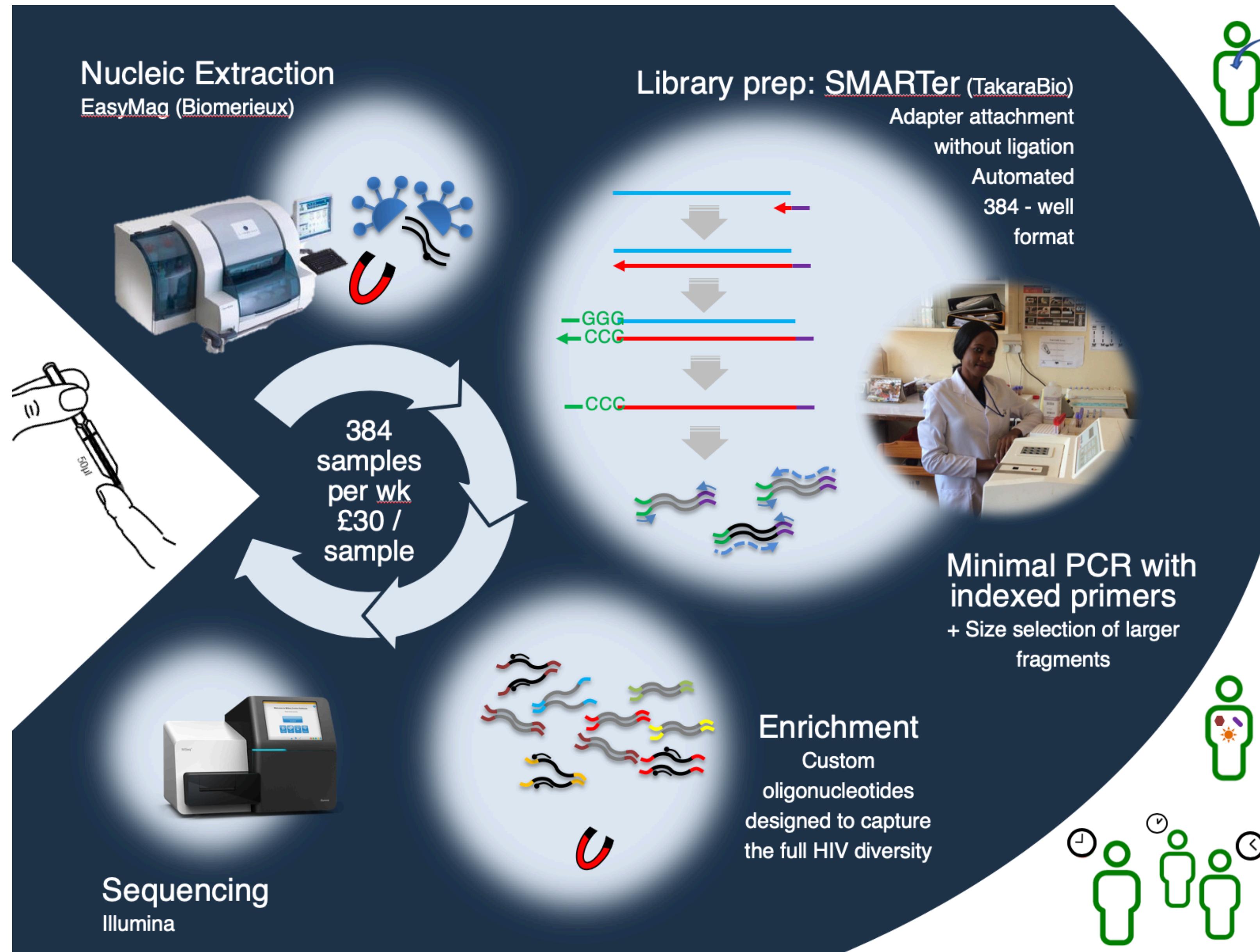


## Summary drug resistance analysis

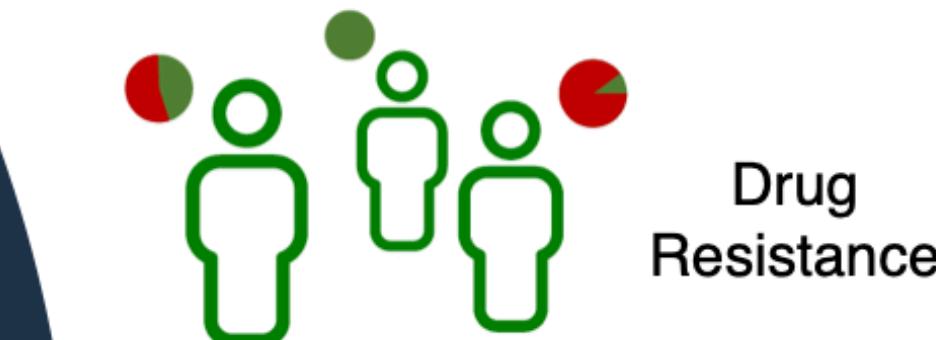
- Resistance is high and increasing.
- An increasing proportion of infections will come from ART experienced and treated individuals in whom resistance levels are especially high
- Observed resistance NRTI mutations which are the backbone to dolutegravir regimens.
- Viral load monitoring and drug resistance surveillance should continue *after* therapies are changed to DTG
- We have developed a cost-effective combined method for testing viral load and drug resistance – utility and feasibility of clinical drug resistance testing in Zambia could be re-evaluated

# A new laboratory method

Throughput = 384 samples/week, one technician, assay cost = \$45



Transmission



Drug Resistance



Genotype



Viral Load

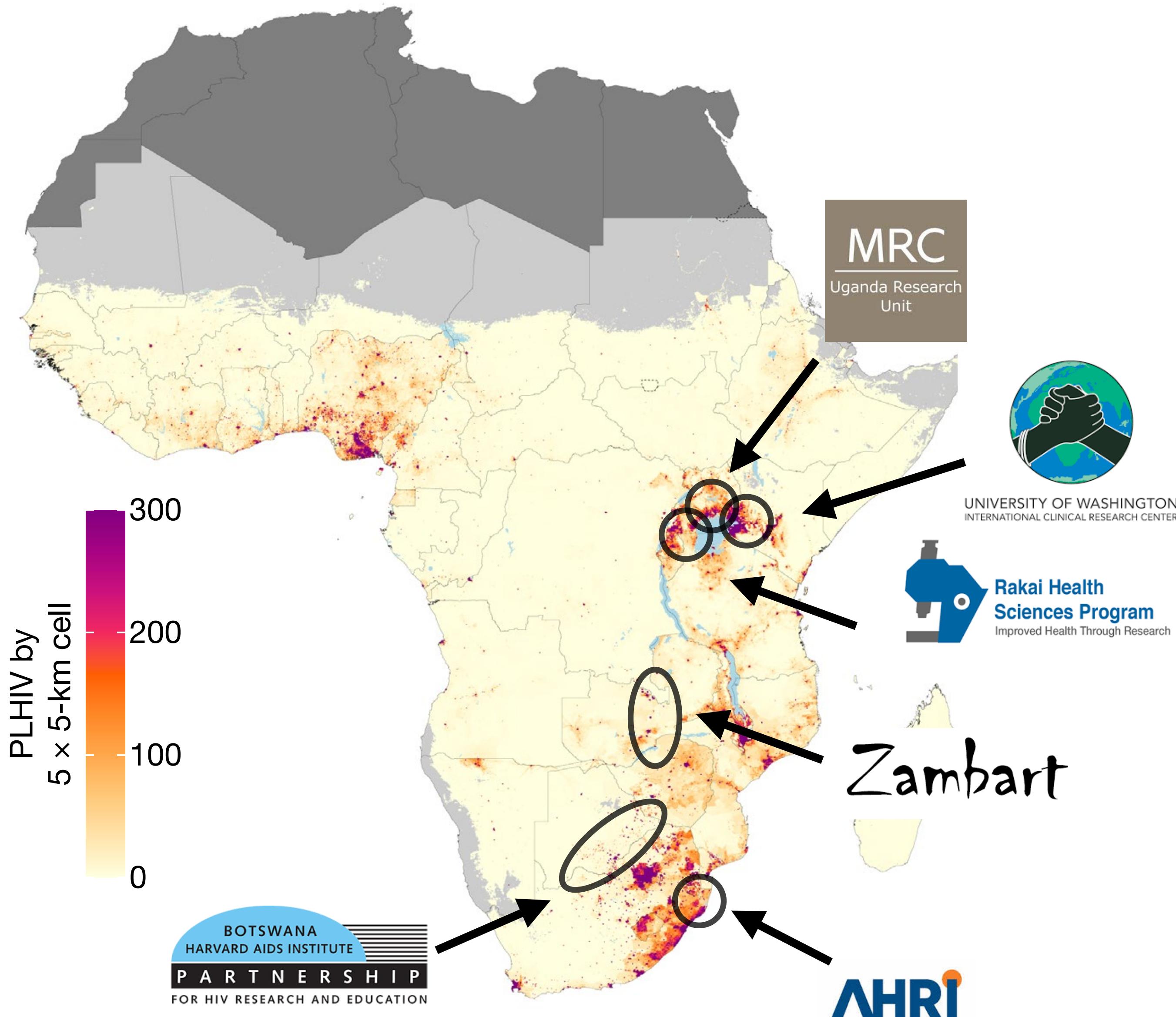


Co-Infection



Timing of Infection

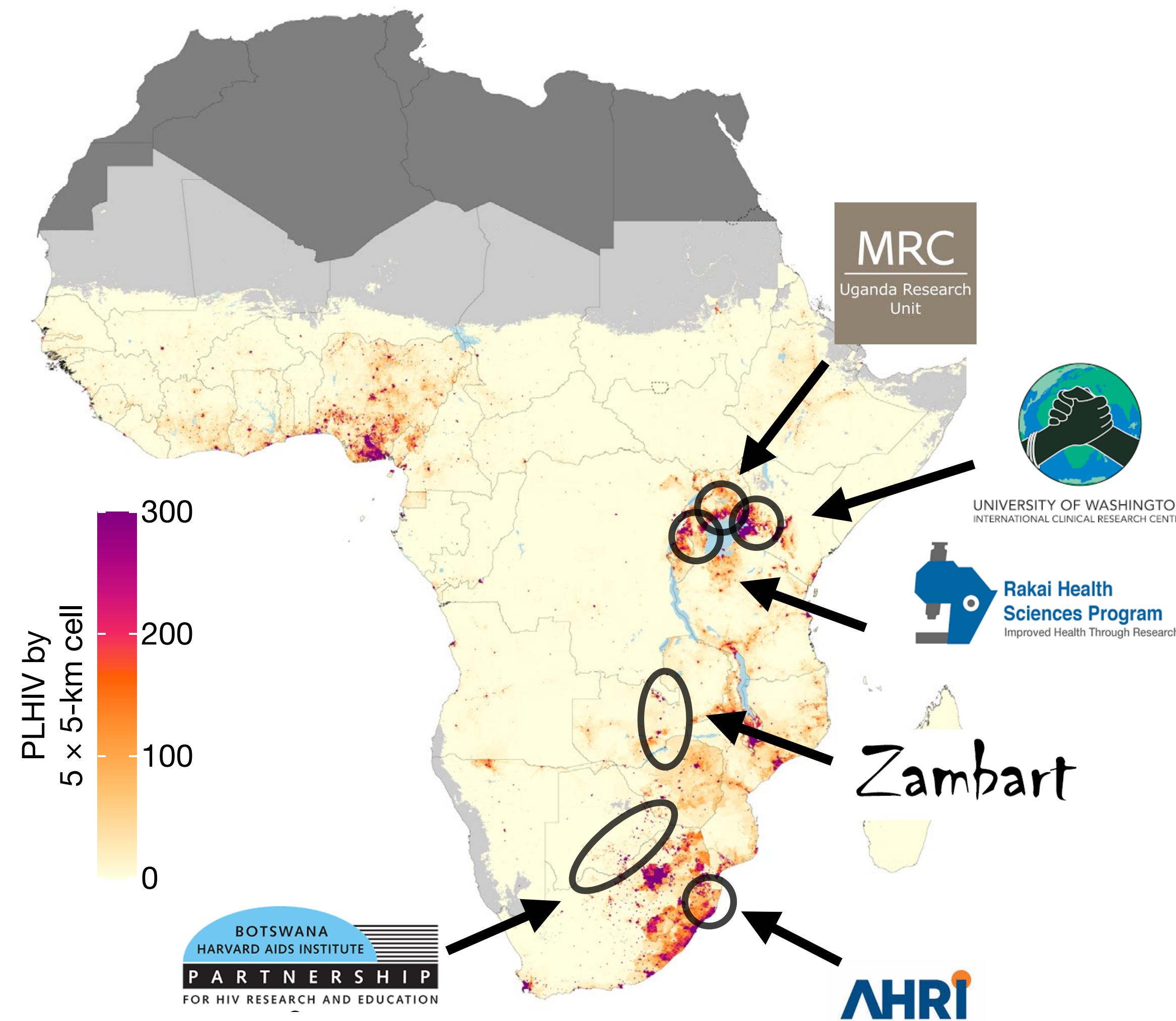
# PANGEA-HIV consortium



1. Molecular epidemiology & modelling (Fraser)
2. Clinical science, drug resistance & ethics (Pillay)
3. Mobility and migration (Grabowski)
4. Phylodynamics (Rambaut)

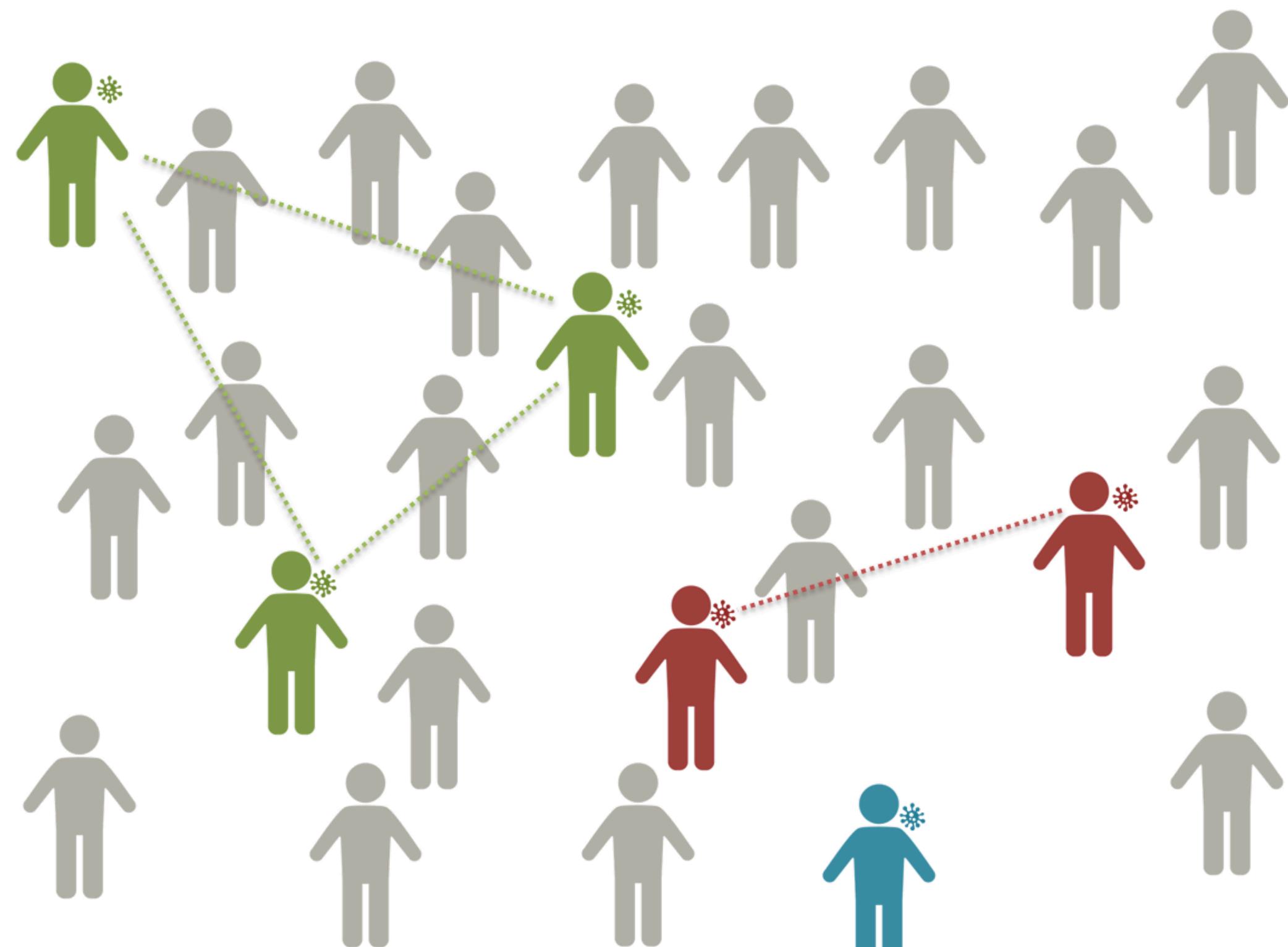
*What are the sources, sinks & hubs of transmission?  
How do we better target prevention?*

# PopART is in the PANGEA-HIV consortium



Active sequencing - training - collaboration - analysis - data sharing - ethics - social science

# Compare clustering and source attribution

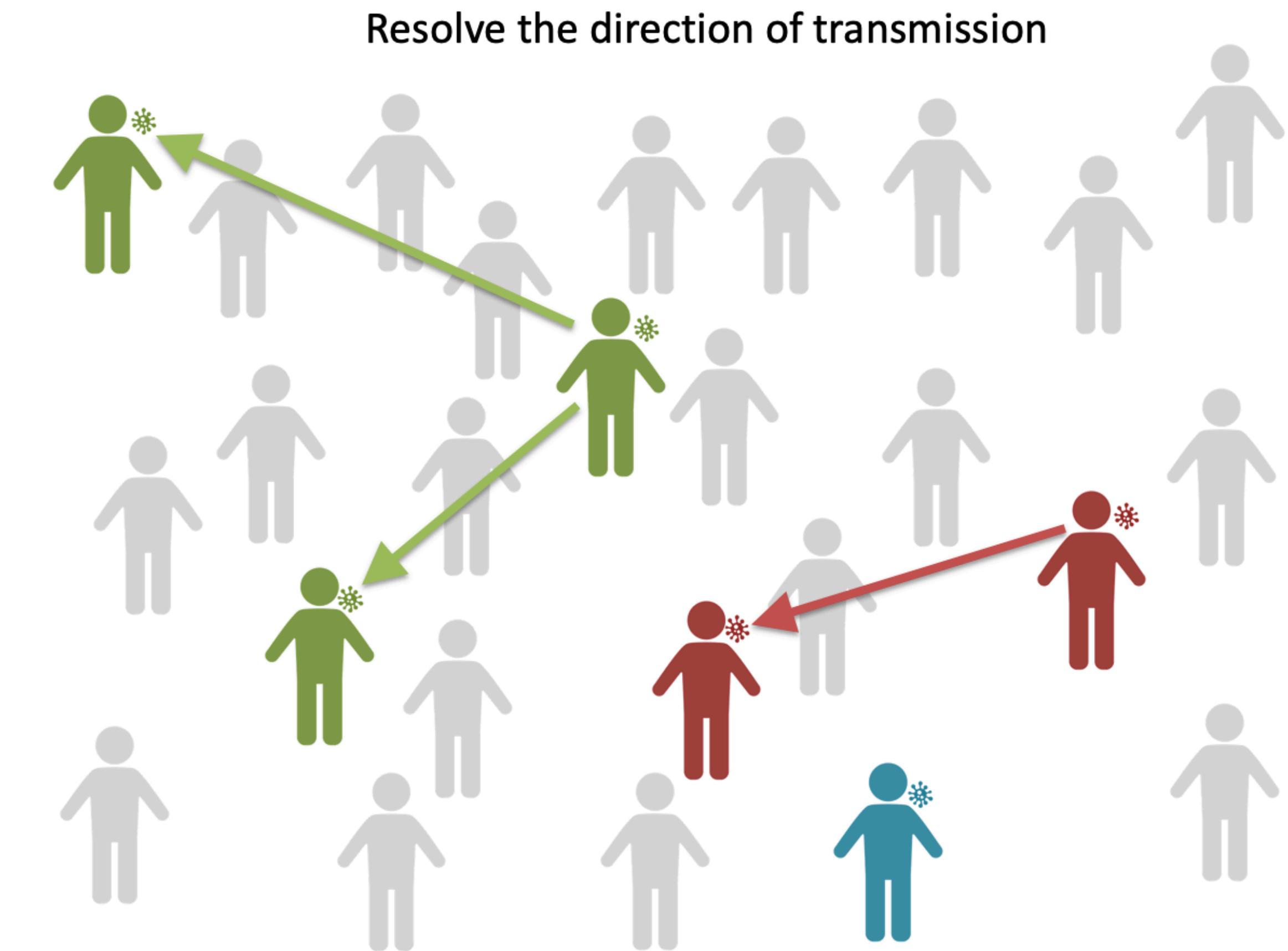


Transmission clusters (missing intermediates):

What characteristics are shared amongst individuals in clusters?

What are the sub-epidemics?

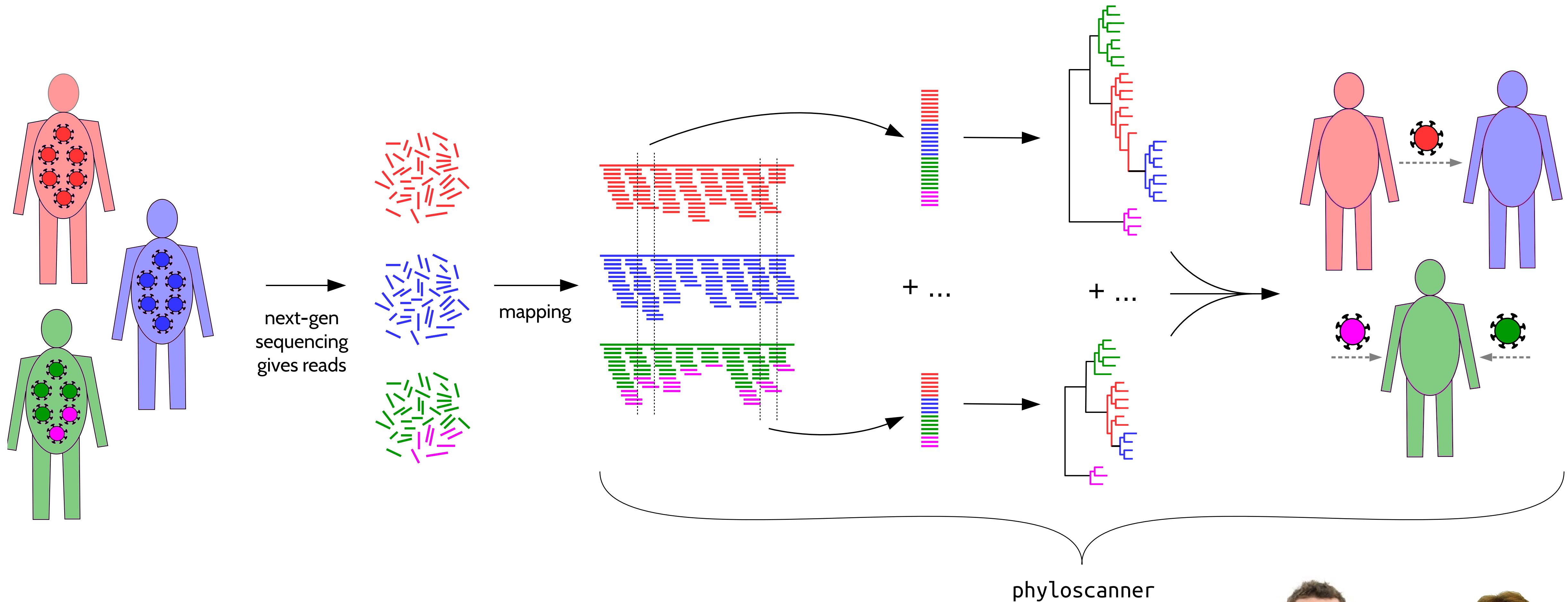
**Clustering**



What are the characteristics of a source case?  
Can we better target prevention?

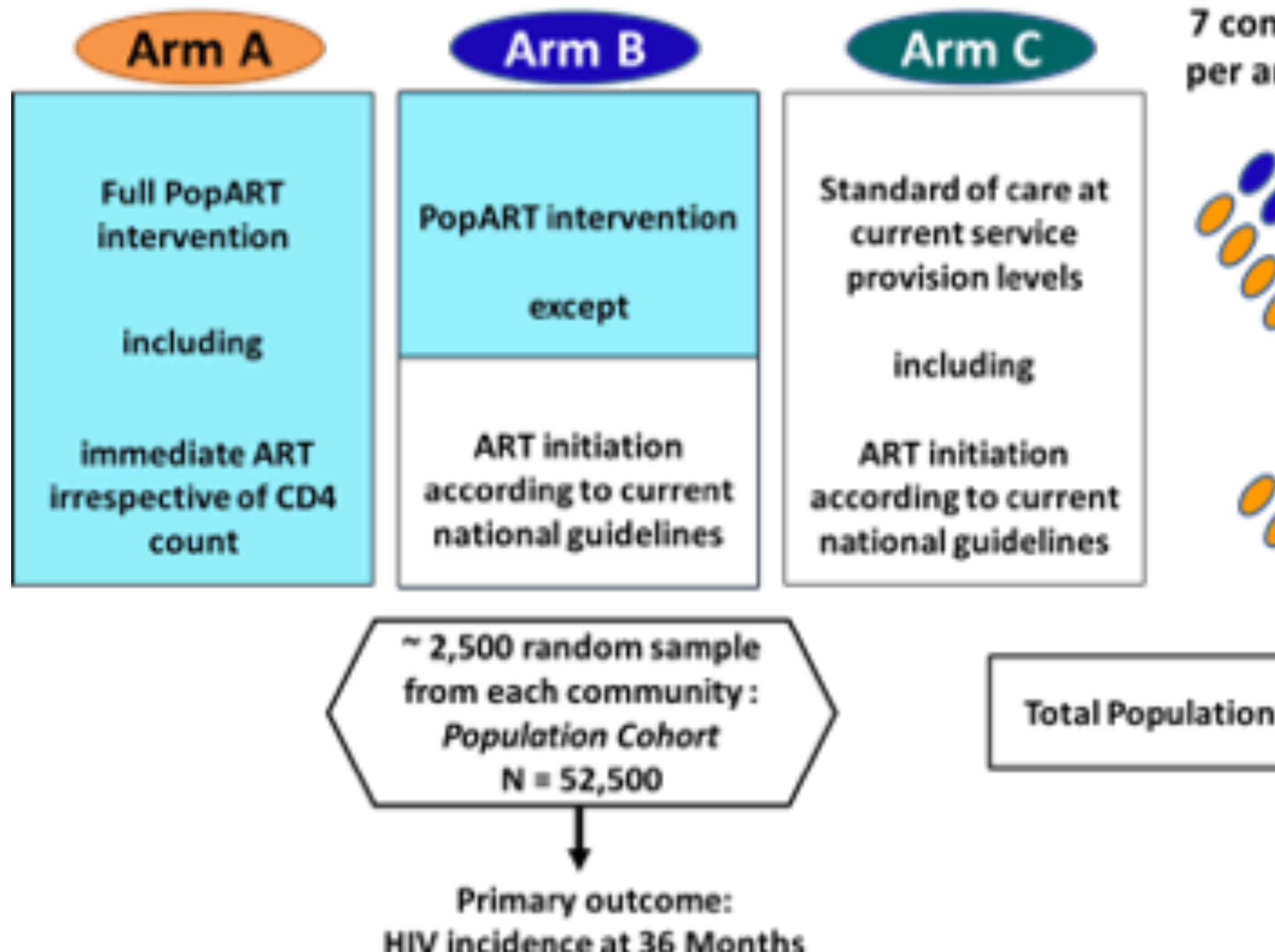
**Source attribution**

Ancestral state reconstruction with within-host diversity can be done in *phyloscanner*, our software for analysing NGS data:



**Chris Wymant** **Matthew Hall**

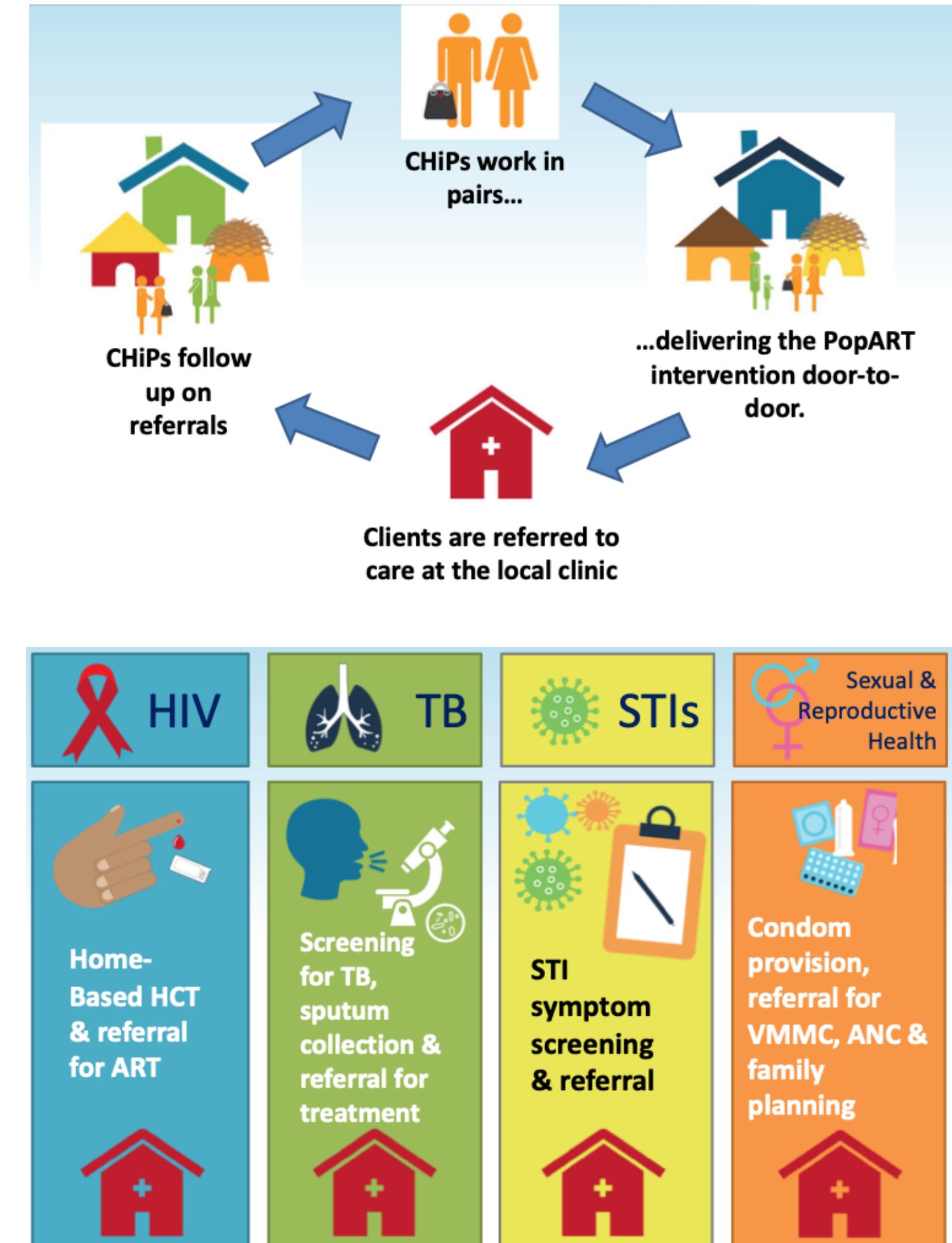
# HPTN 071 (PopART)



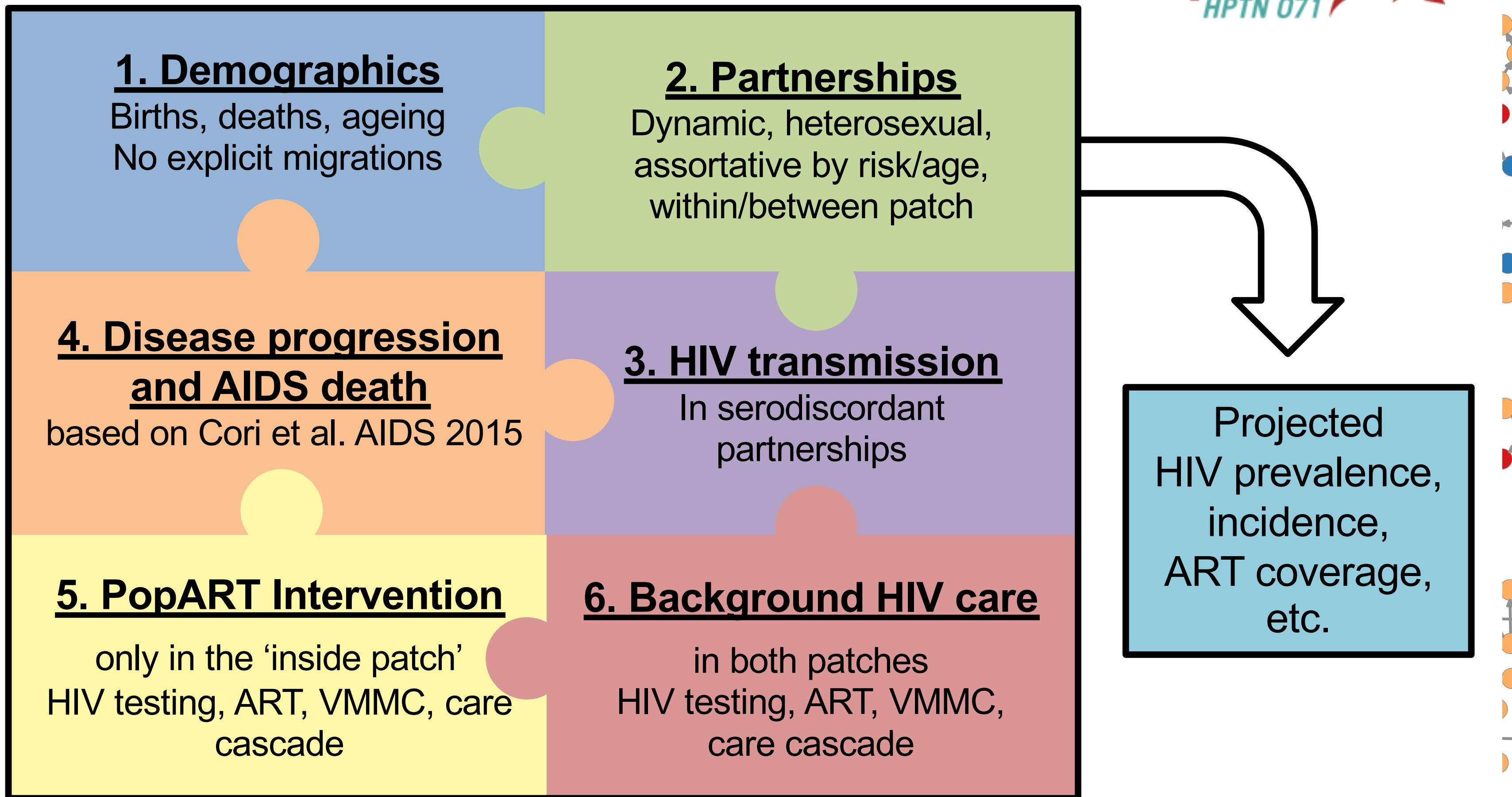
7 communities per arm (N=21)

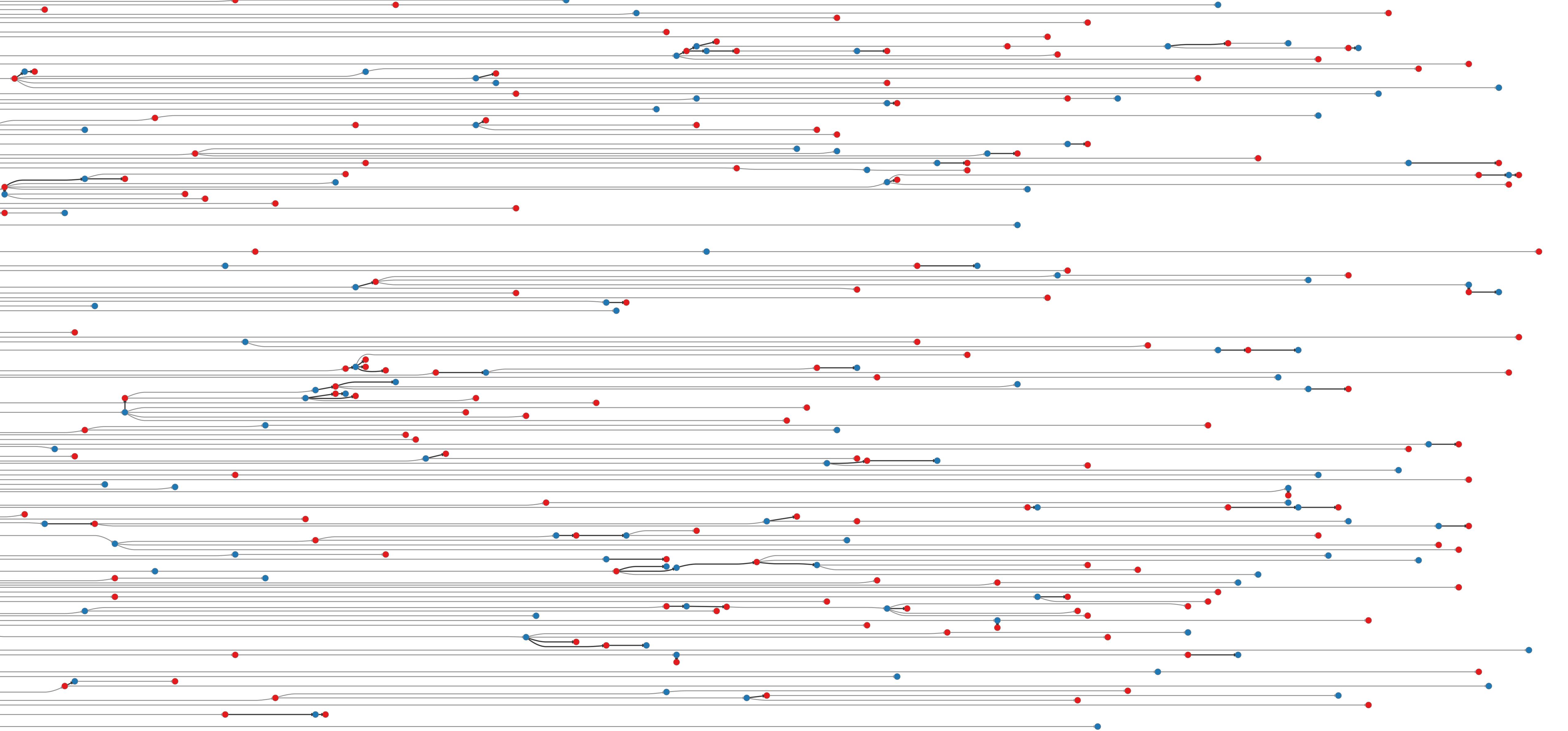
12 in Zambia

9 in South Africa



# Individual-based simulation model (IBM)





1975

Pickles et al, submitted 1980

# Policy questions arising

Prevention gap: viral suppression over 70%, yet incidence is above 1%/year.

1. Are we missing key groups? Young women? Young men?
2. Are infections coming from outside of the community?
3. Are infections arising during early and acute HIV (does treatment start too late)?
4. Is drug resistance impeding success?

# Policy questions arising

Prevention gap: viral suppression over 70%, yet incidence is above 1%/year.

## **1. Are we missing key groups? Young women? Young men?**

2. Are infections coming from outside of the community?

3. Are infections arising during early and acute HIV (does treatment start too late)?

4. Is drug resistance impeding success?

# Summary of PopART phylogenetic analysis

90:90:90 targets were reached, leading to 20% reduction in incidence

**Missing 25-40 men is a big gap, accounting for 40% of transmissions.**

- *Finding missing men is effective even if yield is much lower*

**30-40% of transmissions occurred in first year of infection. 60-70% of transmissions occur later, some much later.**

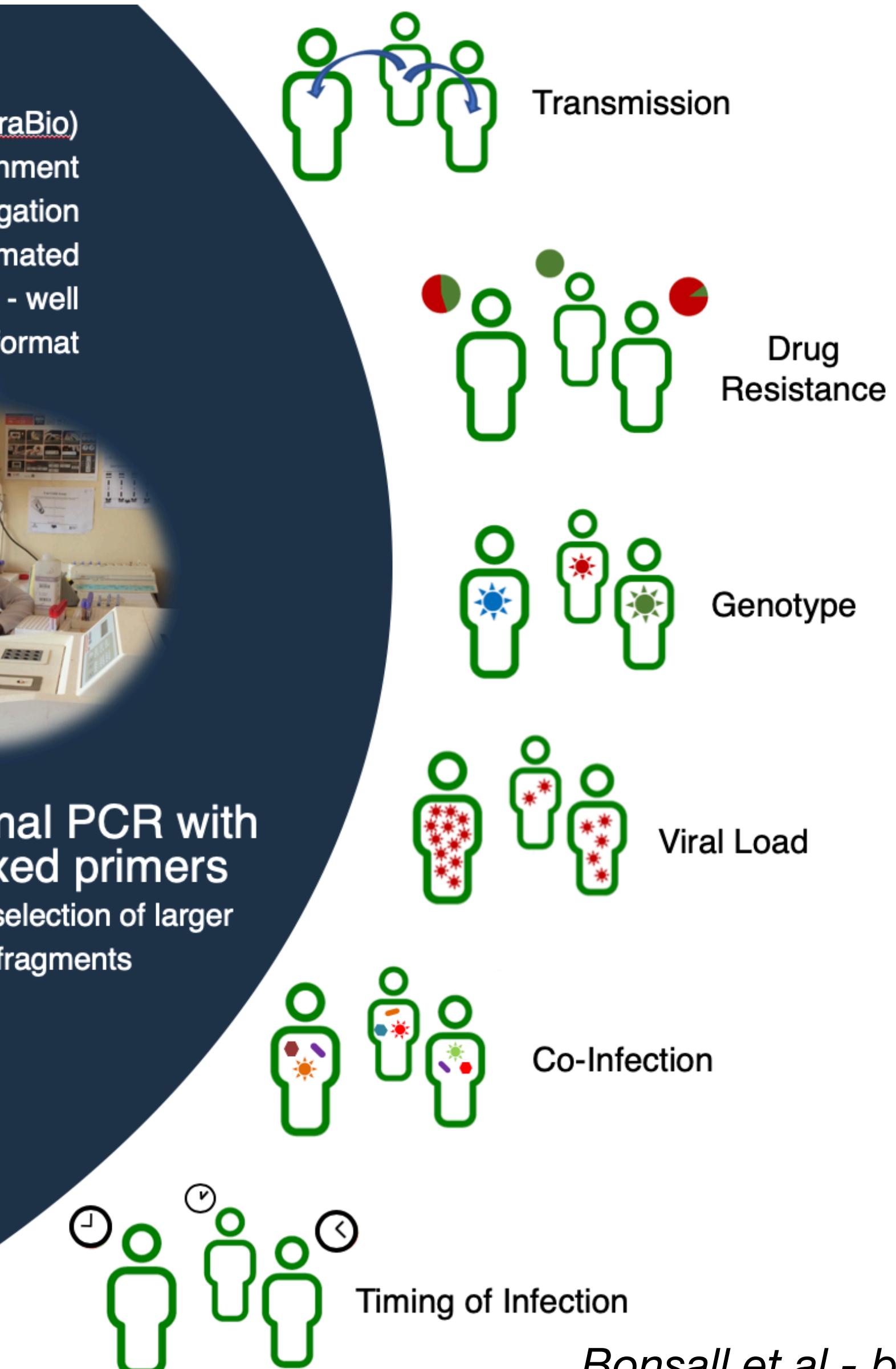
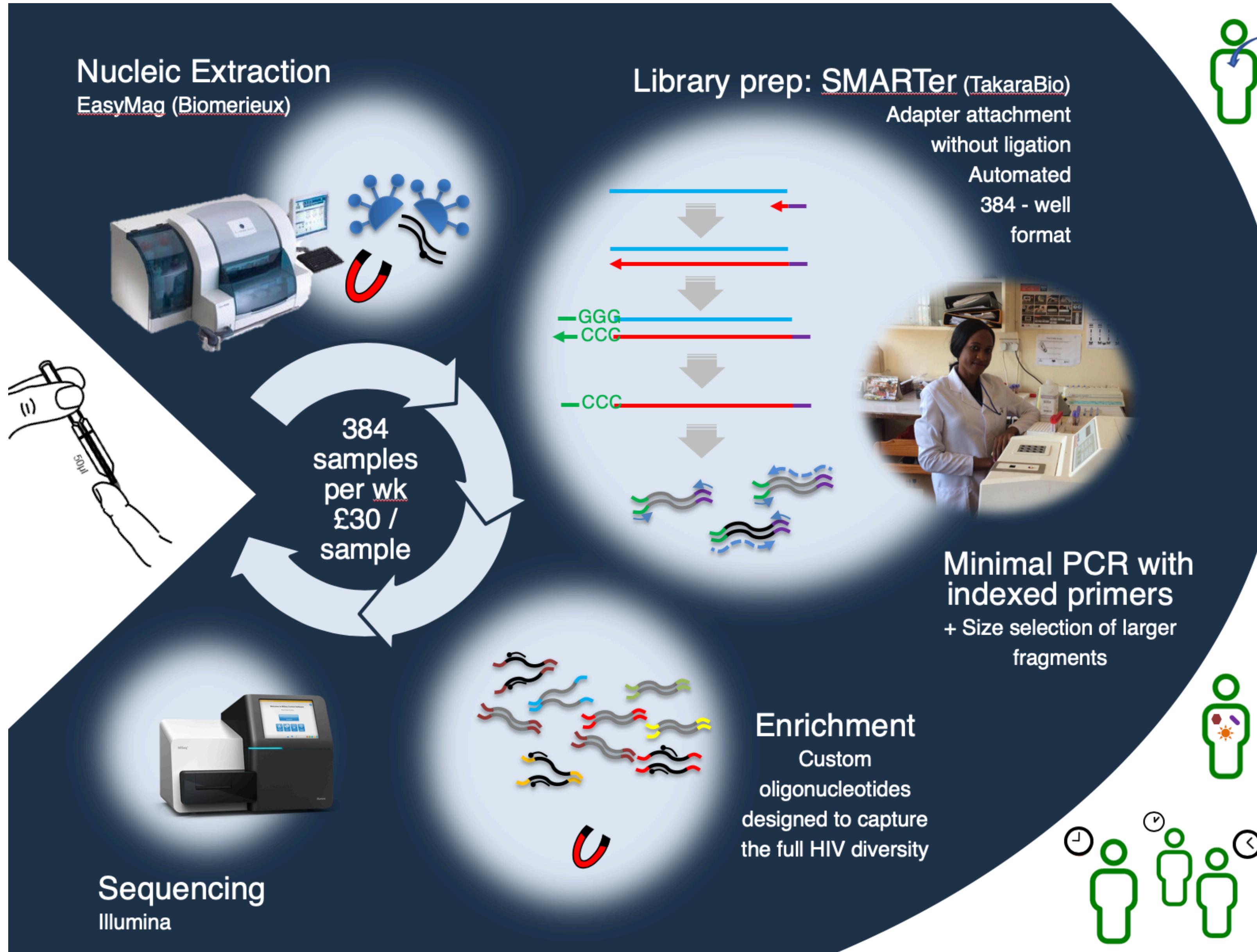
- *Annual testing is needed*
- *Sustained viral suppression is also needed*

**15-20% of transmissions linked different PopART sites.**

- *We may have under-estimated impact of the PopART intervention*
- *Continuity of care as people move is vital*

# For PANGEA, we developed a new sequencing and bioinformatics assay that generates high resolution clinical and epidemiological data. Now being deployed for AMPHEUS.

Throughput = 384 samples/week, one technician, assay cost = \$45



## With thanks to:

- All research participants and their families
- The 21 research communities and their religious, traditional, secular and civil leadership structures
- Volunteers in the community advisory board structures





BIG DATA  
INSTITUTE



UNIVERSITY OF  
OXFORD



Anel Nurtay

Anne Cori

Carol Mulligan-John

Chris Wymant

David Bonsall

George Macintyre

Laura Thompson

Lele Zhao

Lucie Abeler-Dörner

Luca Ferretti

Mariateresa de Cesare

Matthew Hall

Michelle Kendall

Mike Pickles

Newton Otecko

Rafael Sauter

Tanya Golubchik

Will Probert



**Oxford**  
Chris Wymant  
Lucie Abeler-Dörner  
Matthew Hall  
Will Probert  
Rafael Sauter  
David Bonsall  
Tanya Golubchik  
Rory Bowden

**Imperial College**  
Oliver Ratmann

**PANGEA co-PIs**  
M. Kate Grabowski  
Andrew Rambaut  
Deenan Pillay

**POPART**  
Richard Hayes  
Sarah Fidler  
Helen Ayles  
Nuala Beyer  
Peter Bock  
Sian Floyd  
Mike Pickles  
Mohammed Limbada  
Barry Kosloff  
Ab Schaap

**MRC Uganda**  
Pontiano Kaleebu  
Deogratius Ssemwanga

**AHRI**  
Frank Tanser  
Vincent Calvez  
Deenan Pillay

**Rakai**  
M. Kate Grabowski  
Joseph Kagaayi  
Ron Gray  
Maria Wawer  
Tom Quinn

**Botswana/Harvard**  
Vlad Novitsky  
Max Essex

**UW/Partners**  
Jai Lingappa  
Josh Herbeck

**Thank you!**