Population-level HIV risk and combination implementation of HIV services

Neena M. Philip

Submitted in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy
under the Executive Committee
of the Graduate School of Arts and Sciences

COLUMBIA UNIVERSITY

2020

© 2020

Neena M. Philip

All Rights Reserved

Abstract

Population-level HIV risk and combination implementation of HIV services

Neena M. Philip

Background: HIV transmission is greatly reduced when antiretroviral treatment (ART) suppresses an infected person's HIV viral load. It is unclear, however, whether the contextual risk of incident HIV is optimally reduced by widespread individual-level suppression of HIV viral load alone or in combination with other HIV prevention services. HIV service coverage and community norms can influence risk in small area geographies; and contextual factors, like gender inequality and stigma, may foster environments conducive to HIV transmission. Yet, the relationship between places with high HIV levels and the clustering of area risk factors is unknown.

The goal of this dissertation is to learn if and how a geographically focused combination implementation strategy could reduce population-level HIV risk. Analyses explored whether small area risk profiles explain area differences in HIV. The guiding hypothesis is that in high HIV prevalence settings, low HIV service uptake in a geographically defined area increases the prevalence of high HIV viremia, leading to greater HIV transmission and incident HIV.

Methods: A systematic review was conducted to examine the association between population-level measures of HIV viral load and incident HIV infection in generalized and concentrated epidemics. Publications were English, peer-reviewed articles published from January 1, 1995 through February 15, 2019 that explicitly defined HIV viral load and assessed outcomes of HIV recency, incidence, seroconversion, or new diagnosis. Studies sampled general or key populations through population-based surveillance registries, household-based enumeration, cluster sampling, or respondent driven sampling. Descriptive statistics summarized review findings.

The Swaziland HIV Incidence Measurement Survey (SHIMS) data were used for the remaining analyses. Using a two-stage cluster-based design, a nationally representative, household-based sample of adults, ages 18-49 years was enrolled from December 2010 to June 2011 in Eswatini. Consenting adults completed an interview and received home-based rapid HIV testing and counseling. All seropositive samples were tested for HIV viral load using the COBAS AmpliPrep/Taqman HIV-1 Test, v 2.0. Adults

testing HIV-seronegative were enrolled in a prospective cohort for the direct observation of HIV seroconversion, completing an interview and home-based rapid HIV testing six months later.

Multi-level latent class modeling was performed to identify statistically significant combinations of HIV risk factors and to classify the combinations into small area risk profiles. In the cross-sectional sample, linear regression with robust standard errors assessed the correlation between area profiles and places with high levels of uncontrolled HIV infection, or HIV core areas, measured by the area prevalence of detectable virus (≥20 copies/milliliter) among HIV-positive adults and among all adults, regardless of HIV status. In the prospective cohort, generalized linear regression of longitudinal data assessed the association between area profiles and places prone to new HIV infections (i.e., HIV susceptible areas), measured by area-level HIV seroconversions.

Results: The systematic review found an evidence base primarily of lower quality studies and inconsistent HIV viral exposure measurement. Overall findings supported a relationship between increasing levels of suppressed HIV in HIV-infected populations and fewer new infections over time. Better quality studies consistently showed higher population viremia (i.e. HIV viral quantity among all persons, regardless of HIV status) associated with HIV incidence in high prevalence populations; population viral load (i.e., HIV viral quantity among only HIV-positive persons) did not show an association with incident HIV in high prevalence, general populations and was inconsistent in key populations.

To determine whether area risk profiles can pinpoint HIV core areas, latent class modeling was used to categorize 18,172 adults into one of six HIV risk types. The risk typology, classified through unique combinations of HIV service uptake and sexual risk behaviors, conveyed an adult's propensity for HIV transmission and/or acquisition risk. The model next identified the area-level composite prevalences of HIV risk types; estimated the three most frequent, unique composite combinations; and categorized them into area risk profiles characterizing HIV risk: low-moderate acquisition risk, moderate acquisition/transmission risk, and high acquisition/transmission risk. The high acquisition/transmission areas comprised the largest proportions of highest risk transmission and acquisition types. The prevalence of detectable viremia progressively increased from low-moderate acquisition, moderate acquisition/transmission, and high acquisition/transmission profiles [17.7%, 25.4%, and 35.1%, respectively]. When compared with low-moderate acquisition areas, the prevalence of detectable viremia was 7.4% [p<.001] higher in moderate

acquisition/transmission areas and 17.1% [p<.001] higher in high acquisition/transmission areas. The prevalence of detectable viral load significantly decreased from low-moderate acquisition to moderate acquisition/transmission areas [76.6% versus 68.7%, p<.001], and was significantly higher in high acquisition/transmission areas by 7.3% [p<.001], when compared with low-moderate acquisition areas.

To determine whether area risk profiles can predict HIV susceptible areas, a total of 18,172 adults were surveyed of which 4396 [24%] had detectable viremia. 11,880 [96%; n=12,357] HIV-seronegative adults enrolled in the prospective cohort and 11,155 [94%] of them completed an endline visit. Four area profiles were identified, defined by unique patterns in prevalence of HIV viremia and of sexual risk behaviors. The proportion of HIV susceptible areas progressively increased from Profiles A, B, C, and D [14.3%, 21.8%, 24.6%, and 30.8%, respectively]. HIV susceptible areas were more than twice as likely to occur in Profile D than Profile A environments [RR 2.13, 95% confidence interval (CI) (1.13, 4.00); p=0.02]. Profile D areas had prevalences of unknown partner HIV status and detectable viremia at 28% and 24%, respectively. In contrast, Profile A areas had prevalences of only 8% with unknown HIV status and 31% with detectable viremia.

Conclusion: This dissertation shows that geographic risk profiles can explain differences in population-level HIV outcomes. Risk factors spatially cluster in predictable, meaningful combinations that can inform an area typology of HIV risk. The co-location of adults predisposed to greater HIV risk may heighten levels of uncontrolled HIV infection, thereby creating potential area sources of ongoing transmission; however, the concurrent levels of other risk factors may have more influence in reducing population-level incidence than previously considered. A composite indicator of contextual HIV risk may reveal places core to HIV transmission and susceptible to HIV acquisition. Such area profiles may help identify the combination of locally specific risk factors that readily promulgate HIV and better inform the design of place-based HIV intervention packages to enhance current strategies towards global HIV control.

Table of contents

List of tables, figures iii

Acknowledgments v

Dedication vi

Chapter One. Introduction 7

Combination HIV prevention 8

Treatment as prevention – the cornerstone of combination HIV prevention package 9

Gaps in our understanding of combination HIV prevention 10

References 13

Chapter Two. Measuring the population-level relationship between community viral load and new HIV infections: a systematic review 16

Summary 17

Introduction 18

Methods 19

Results 21

Discussion 26

Tables and figures 32

Supplements 47

References 51

Chapter Three. Small area HIV risk profiles and population-level HIV viral measures: a national-level, cross-sectional analysis in Eswatini 54

Summary 55

Introduction 57

Methods 58

Results 61

Discussion 65

Tables and figures 69

Supplements 76

References 87
Chapter Four. Area risk profiles and HIV susceptible areas: detecting the places most prone to incident HIV in a national hyperendemic setting 90
Summary 91
Introduction 93
Methods 94
Results 97
Discussion 100
Tables and figures 105
Supplements 111
References 118
Chapter Five. Conclusion 121
Conceptual model of environmental influences on population HIV outcomes 123
Program implications 125
Future directions 126
Tables and figures 128
References 129

List of tables, figures

Figure 2.1. Study selection 32

Table 2.1. Study characteristics 33

Figure 2.2. Number of studies by data source and population type 38

Table 2.2. Summary of study findings 39

Figure 2.3. Number of statistical analyses by population HIV viral exposure measure, disaggregated by new HIV infection measure 46

Supplemental Table S2.1. Search strategies used in PubMed and Embase 47

Supplemental Table S2.2. Newman-Ottawa risk bias scores 49

Table 3.1. Participant characteristics 69

Table 3.2. Characteristics of HIV risk types 71

Figure 3.1. Prevalence of HIV risk type by area profiles* 72

Table 3.3. Prevalence of detectable viral quantity by area profile 73

Figure 3.2. Risk factor distribution across area profiles 74

Supplemental Figure S3.1. Swaziland HIV incidence measurement survey (SHIMS): serial HIV test algorithm 76

Supplemental Table S3.1. Model fit statistics for latent class models 77

Supplemental Figure S3.2. Model fit statistics curves for latent class models 78

Supplemental Table S3.2. Model fit statistics for latent cluster models 79

Supplemental Figure S3.3. Model fit statistics curves for latent cluster models 79

Supplemental Table S3.3. Multinomial logistic regression of HIV risk factor indicators on HIV risk type. 80

Supplemental Table S3.4. Multinomial logistic regression of area profiles on HIV risk type 80

Supplemental Table S3.5. Pairwise comparison of risk types by risk factor indicator 82

Supplemental Table S3.6. Pairwise comparison of area profiles 83

Supplemental Table S3.7. Pairwise comparison of area profiles by HIV viral quantity 83

Supplemental Table S3.8. Cross tabulation of individual assignment rules for HIV risk types 84

Supplemental Table S3.9. Cross tabulation of EA assignment rules to area profiles 85

Supplemental Table S3.10. Prevalence of area HIV viral quantity by area profile 86

Table 4.1. Participant characteristics..... 105

Table 4.2. Characteristics of area profiles..... 107

Figure 4.1. Average prevalence of HIV, detectable viremia,* and antiretroviral treatment (ART) by area profile..... 108

Table 4.3. Generalized linear regression of HIV susceptible areas on area profile..... 109

Figure 4.2. Unadjusted and adjusted risk ratios of the association between observed area-level prevalences of risk factors and HIV susceptible areas..... 110

Supplemental Figure S4.1. Swaziland HIV incidence measurement survey (SHIMS): serial HIV test algorithm..... 111

Supplemental Table S4.1. Model fit statistics for latent cluster models..... 113

Supplemental Table S4.2. Multinomial logistic regression of HIV risk factor indicators on area profile..... 114

Supplemental Table S4.3. Pairwise comparisons of area profiles by risk factor indicator..... 115

Supplemental Table S4.4. Cross tabulation of enumeration area (EA) assignment rules to area profiles..... 117

Figure 5.1. Conceptual model of environmental influences on population HIV outcomes..... 128

Acknowledgments

Over the years of this endeavor, I feel very fortunate for the support that I have received. I am grateful to my committee members for their time and valuable insight. Dr. Barun Mathema was formative. He helped me to think like an epidemiologist, inspiring assurance and enthusiasm. His support has been thoughtful and beyond the pages of this dissertation. Dr. Gina Lovasi was unconditional and helped cultivate the methodological creativity and rigor of this work. Drs. Quarraisha Abdool Karim and Frank Tanser fostered new lines of thought through their critique and guidance. Their prior scientific contributions and dedication to improving the health of populations were key motivators of this research.

I also am deeply grateful to ICAP at Columbia University. For over 15 years, ICAP provided me unique opportunities to develop an expertise on HIV service delivery and a range of research paradigms. The underpinnings of this dissertation fully rest on those experiences and the day-to-day insights gathered from brilliant leaders and colleagues, Drs. Wafaa El-Sadr, Elaine Abrams, and Jessica Justman. Dr. Justman's influence was indispensable, without which this work could not have been done.

I am tremendously thankful to my brother George Philip, who cared for our family and created the space for me to complete this work.

And I feel indescribable appreciation for my husband, Chris Riggert. His infinite encouragement, patience, love, and support was as central as the air we breathe and simply irreplaceable.

Dedication

To my father, Dr. Thoni V. Philip, and mother, Mary Philip.

Without their love, ethic, and aspiration for their children,

this achievement would not have been possible.

Chapter One. Introduction

Three decades since its discovery, HIV remains a formidable global health threat. Nearly 38 million people were living with HIV (PLHIV) in 2018, and another 1.7 million became infected. An estimated 770,000 died from AIDS. Eastern and Southern Africa were disproportionately burdened by the pandemic with 20.6 million PLHIV – over half (54%) of PLHIV worldwide. Nearly two-thirds (64%) of new infections occurred in sub-Saharan Africa. Nonetheless, global commitment to stem the pandemic is yielding results. Uptake of life-saving ART more than tripled from 7.5 million globally in 2010 to 23.3 million in 2018. Growing treatment coverage since 2005 has coincided with decreasing AIDS-related deaths. These gains, coupled with advances in HIV prevention methods, stirred a shift in global policy from HIV control to a path towards its elimination.[1]

UNAIDS pledged in 2014 an ambitious three-part treatment target by end 2020: 90% of all PLHIV would know their HIV status, 90% of all people diagnosed with HIV would receive sustained ART, and 90% of all people receiving ART would be virally suppressed. Assuming that ART coverage and viral suppression would increase to 95% by 2030, UNAIDS estimated annual infections would reduce by nearly 90% by 2030, thereby ending the AIDS epidemic. The success of this strategy, however, rests on a combination HIV prevention approach – the scale-up of HIV prevention tools in tandem with expedited expansion of HIV treatment.[2] A single, stand-alone method to fully protect against HIV acquisition remains elusive. While recent gains in pre-exposure prophylaxis (PrEP) and vaccine development accelerate progress, it is doubtful that one method will single-handedly end the epidemic.[3,4] Rather, partially effective methods remain our best option and thus, the concurrent use of multiple partially effective methods has made traction in the past decade.[5]

Combination HIV prevention

Combination HIV prevention is the concurrent use of multiple partially effective methods across biomedical, behavioral, and structural risk domains to reduce population HIV incidence. ART serves as the cornerstone of the strategy, and UNAIDS has emphasized this approach as essential to ending the epidemic.[6-8] The toolbox includes biomedical interventions, like male circumcision, ART, and PrEP; behavioral interventions, like couples counseling, prevention messaging on condom use, HIV testing, and medication adherence; and structural interventions, like conditional cash transfers, and policies on universal

test and treat. Programming is characterized by targeting high transmission areas and major at-risk populations and the high coverage of an evidence-based intervention package.[3]

Contextual, or social and environmental, factors drive the spread of epidemics and, therefore, define the most salient features of combination prevention programming. While individual-level biologic mechanisms produce infection, HIV transmission-acquisition is socially constructed through causal pathways that shape how individuals engage with their social and economic circumstances and environments characterized by HIV service uptake. These pathways of proximal causes and epidemiologic drivers are specific to each epidemiologic setting, yet regularly undermine the ability of individuals to protect themselves and others.[9,10] Combination prevention addresses these factors by developing multi-component intervention packages effecting multiple levels of influence. A combination prevention strategy aims to optimize HIV prevention impact by having additive (complementary) or synergistic effects at the population-level, which may result in reducing transmission to below reproductive rate and to elimination.[3,11,12]

Treatment as prevention – the cornerstone of combination HIV prevention package

Numerous observational studies have provided proof of concept for treatment as prevention, showing reduced HIV acquisition risk in prospective cohorts and in populations with increased ART use.[13-18] Yet, the definitive evidence emerged in a 2011 randomized controlled clinical trial evaluating the effect of early ART (i.e., at $CD4^+$ count 350-550 cells/milliliter) versus delayed initiation (i.e., at $CD4^+$ count <250 cells/milliliter or illness) on sexual HIV acquisition risk. Early treatment showed a stunning 96% relative reduction in linked sexual HIV transmission in stable serodiscordant couples. The most likely biologic causal mechanism was suppressed genital viral load resulting from sustained ART use.[19,20]

Building on this scientific breakthrough, population-level analyses were conducted hypothesizing the possible effect of ART coverage on individual HIV acquisition risk in high prevalence settings. In Kwa-Zulu Natal, South Africa, a 1% increase in ART coverage was associated with a 1.4% decrease in HIV acquisition risk. Uncontrolled HIV viral load was common in HIV-positive persons, and thus, a main independent predictor of new infections was the community-level HIV prevalence.[21] Similarly, in a population-based open cohort in Kwa-Zulu Natal, ART uptake among HIV-positive partners in stable serodiscordant relationships was associated with a 77% decrease in HIV acquisition risk. This real-world evidence affirmed the population-level effects of ART on HIV acquisition risk in hyper-endemic, generalized

epidemics. Nonetheless, the effects were notably less than the clinical trial findings, suggesting that supplementary prevention interventions may be needed for substantial, sustained reductions in population-level incident infections.[22] Moreover, several randomized cluster trials subsequently failed to show a definitive impact of a test-and-treat approach on HIV incidence, calling into question whether universal treatment coverage is sufficient to achieving epidemic control.[23-27]

Gaps in our understanding of combination HIV prevention

This dissertation investigates the contextual dynamics of combination HIV prevention at the population-level. Certain individual-level risk factors promote HIV acquisition and transmission and targeting areas of high HIV prevalence and incidence with effective interventions can reduce transmission-acquisition risk.[3] The central premise of high risk geographies, however, is less defined than risk dynamics at the individual level. For example, individual risk scores for HIV acquisition are developed[28-30] yet a parallel metric for geographic locations is less evident. Potential spatial risk factors, such as prevalence of sexual risk behaviors and prevalence of HIV intervention uptake, may predict locations of high HIV viremia or incidence. Similarly, higher prevalence of individuals with high risk profiles may also predict geographic hotspots of HIV risk. Th intensity of spatial risk factors likely vary by epidemic typology (e.g., generalized versus concentrated epidemics), as well as within epidemic settings through small area micro-epidemics of incident infection.

The dissertation goal is to determine whether a spatial profile of HIV risk may inform the prediction of geographic hotspots of HIV transmission and of HIV acquisition. Identifying geographic profiles of spatial risk factors could increase the effectiveness of combination HIV prevention through refined geographic targeting and customization of intervention packages to the spatial context. Analyses are described across three chapters, including a systematic literature review and two secondary analyses of the Swaziland HIV Incidence Measurement Survey (SHIMS). SHIMS comprises data from a household-based, nationally-representative cross-sectional sample of adults, ages 18-49 years, from Eswatini (formerly known as Swaziland) in 2011. Over 12,500 households were surveyed and 18,172 adults completed a face-to-face interview and home-based HIV rapid testing. Among those who tested HIV-seronegative, 11,880 enrolled in a six-month cohort designed to directly observe the rate of HIV seroconversions.

Chapter 2 describes a systematic literature review on the association between population-level measures of HIV viral load and incident HIV infection in generalized and concentrated epidemics. The review details the different viral load measures used in the scientific literature, the strengths and limitations of the measures, and the interpretations of each measure to summarize how they may relate to incidence prediction.

Chapter 3 investigates factors that may shape geographically defined, macro-level dynamics of population-level HIV transmission potential. HIV core areas, or settings with comparatively higher levels of HIV viral quantities, are hypothesized to be driven by people who inconsistently use condoms, do not know the HIV status of a recent sexual partner, and have never tested for HIV. Individual-level HIV testing facilitates two causal pathways, one which may increase the likelihood of diagnosed HIV-infected persons achieving viral suppression through ART [31-35] and a second which may increase the likelihood of HIV-uninfected persons remaining uninfected through partner disclosure of HIV status and condom use. These two pathways may affect the population proportion of uncontrolled viremia over time. To investigate this, multi-level latent class modeling is used to determine whether area risk profiles, classified by population subgroups with distinct combinations of HIV service uptake and sexual risk behaviors, can explain area differences in HIV viral quantities. The analysis first identifies population subgroups of individual-level risk profiles, defined by sexual risk behaviors, HIV disclosure, and intervention uptake; and then area risk profiles, comprising unique prevalence patterns of the population subgroups. The area profiles are then assessed for variation in area-based HIV viral quantity.

Chapter 4 explores factors that may shape geographically defined, macro-level dynamics of incident HIV seroconversion at the population-level. HIV susceptible areas, or places with incident HIV infections, are hypothesized to have higher HIV viremia, lower knowledge of a partner's HIV status, and greater frequency of multiple partnerships. Individual-level HIV acquisition risk is associated with having partners with unknown HIV status, lower ART coverage and higher prevalence of HIV viremia among all persons regardless of HIV status.[21,36] Similarly, increasing prevalence of viral suppression among HIV-infected persons is associated with lowered HIV incidence over time.[37] To investigate this further, multi-level latent class modeling is used to determine whether area risk profiles, defined by distinct patterns in

prevalence of HIV viremia and of sexual risk behaviors, predict area-based variation in incident HIV infection risk.

Chapter 5 summarizes the overall findings of the dissertation and integrates the key results from prior chapters. A conceptual model of environmental influences on population HIV outcomes is proposed, including a discussion of potential population mechanisms of HIV disease in Eswatini in 2011. The chapter closes with program implications of the dissertation findings and recommendations for future research.

References

1. UNAIDS. AIDS by the Numbers: AIDS Is Not Over, but It Can Be. Geneva: Joint United Nations Programme on HIV/AIDS; 2016.

2. UNAIDS. 90-90-90 an Ambitious Treatment Target to Help End the AIDS Epidemic. Geneva: Joint United Nations Programme on HIV/AIDS; 2014.

3. Kurth AE, Celum C, Baeten JM, Vermund SH, Wasserheit JN. Combination HIV prevention: significance, challenges, and opportunities. Curr HIV/AIDS Rep 2011;8:62-72.

4. Vermund SH. Vaccine efficacy trials for human immunodeficiency virus/acquired immunodeficiency syndrome are feasible in the United States: a commentary on the HIVNET Vaccine Preparedness Study. Am J Epidemiol 2001;153:628-31.

5. Vermund SH, Hayes RJ. Combination prevention: new hope for stopping the epidemic. Curr HIV/AIDS Rep 2013;10:169-86.

6. UNAIDS. Combination HIV Prevention: Tailoring and Coordinating Biomedical, Behavioral and Structural Strategies to Reduce New HIV Infections. A UNAIDS Discussion Paper. Geneva: Joint United Nations Programme on HIV/AIDS; 2010.

7. UNAIDS. Prevention Gap Report. Geneva: Joint United Nations Programme on HIV/AIDS; 2016.

8. UNAIDS. On the Fast-track to End AIDS: UNAIDS 2016-2021 Strategy. Geneva: Joint United Nations Programme on HIV/AIDS; 2015.

9. Hankins CA, de Zalduondo BO. Combination prevention: a deeper understanding of effective HIV prevention. AIDS 2010;24 Suppl 4:S70-80.

10. UNAIDS. UNAIDS Practical Guidelines for Intensifying HIV Prevention : Towards Universal Access. Geneva: Joint United Nations Programme on HIV/AIDS; 2007.

11. Piot P, Bartos M, Larson H, Zewdie D, Mane P. Coming to terms with complexity: a call to action for HIV prevention. Lancet 2008;372:845-59.

12. Holtgrave DR. Is the elimination of HIV infection within reach in the United States? Lessons from an epidemiologic transmission model. Public Health Rep 2010;125:372-6.

13. Castel AD, Befus M, Willis S, et al. Use of the community viral load as a population-based biomarker of HIV burden. AIDS 2012;26:345-53.

14. Das M, Chu PL, Santos GM, et al. Decreases in community viral load are accompanied by reductions in new HIV infections in San Francisco. PLoS One 2010;5:e11068.

15. Wood E, Kerr T, Marshall BD, et al. Longitudinal community plasma HIV-1 RNA concentrations and incidence of HIV-1 among injecting drug users: prospective cohort study. BMJ 2009;338:b1649.

16. Law MG, Woolley I, Templeton DJ, et al. Trends in detectable viral load by calendar year in the Australian HIV observational database. J Int AIDS Soc 2011;14:10.

17. Montaner JS, Lima VD, Barrios R, et al. Association of highly active antiretroviral therapy coverage, population viral load, and yearly new HIV diagnoses in British Columbia, Canada: a population-based study. Lancet 2010;376:532-9.

18. Fang CT, Hsu HM, Twu SJ, et al. Decreased HIV transmission after a policy of providing free access to highly active antiretroviral therapy in Taiwan. J Infect Dis 2004;190:879-85.

19. Graham SM, Holte SE, Peshu NM, et al. Initiation of antiretroviral therapy leads to a rapid decline in cervical and vaginal HIV-1 shedding. AIDS 2007;21:501-7.

20. Vernazza PL, Troiani L, Flepp MJ, et al. Potent antiretroviral treatment of HIV-infection results in suppression of the seminal shedding of HIV. The Swiss HIV cohort study. AIDS 2000;14:117-21.

21. Tanser F, Barnighausen T, Grapsa E, Zaidi J, Newell ML. High coverage of ART associated with decline in risk of HIV acquisition in rural KwaZulu-Natal, South Africa. Science 2013;339:966-71.

22. Oldenburg CE, Barnighausen T, Tanser F, et al. Antiretroviral therapy to prevent HIV acquisition in serodiscordant couples in a hyperendemic community in rural South Africa. Clin Infect Dis 2016;63:548-54.

23. Iwuji CC, Orne-Gliemann J, Larmarange J, et al. Universal test and treat and the HIV epidemic in rural South Africa: a phase 4, open-label, community cluster randomised trial. Lancet HIV 2018;5:e116-e25.

24. Hayes RJ, Donnell D, Floyd S, et al. Effect of universal testing and treatment on HIV incidence - HPTN 071 (PopART). N Engl J Med 2019;381:207-18.

25. Havlir DV, Balzer LB, Charlebois ED, et al. HIV testing and treatment with the use of a community health approach in rural Africa. N Engl J Med 2019;381:219-29.

26. Makhema J, Wirth KE, Pretorius Holme M, et al. Universal testing, expanded treatment, and incidence of HIV infection in Botswana. N Engl J Med 2019;381:230-42.

27. Abdool Karim SS. HIV-1 epidemic control - insights from test-and-treat trials. N Engl J Med 2019;381:286-8.

28. Balkus JE, Brown E, Palanee T, et al. An empiric HIV risk scoring tool to predict HIV-1 acquisition in African women. J Acquir Immune Defic Syndr 2016;72:333-43.

29. Pintye J, Drake AL, Kinuthia J, et al. A risk assessment tool for identifying pregnant and postpartum women who may benefit from preexposure prophylaxis. Clin Infect Dis 2017;64:751-8.

30. Wahome E, Thiong'o AN, Mwashigadi G, et al. An empiric risk score to guide PrEP targeting among MSM in coastal Kenya. AIDS Behav 2018; 22(1):35-44.

31. ICAP at Columbia University. Zimbabwe Population-based HIV Impact Assessment: ZIMPHIA 2015-2016. New York: ICAP at Columbia University; 2016.

32. ICAP at Columbia University. Malawi Population-based HIV Impact Assessment: MPHIA 2015-2016. New York: ICAP at Columbia University; 2016.

33. ICAP at Columbia University. Zambia Population-based HIV Impact Assessment: ZAMPHIA 2015-2016. New York: ICAP at Columbia University; 2016.

34. ICAP at Columbia University. Namibia Population-based HIV Impact Assessment: NAMPHIA 2017. New York: ICAP at Columbia University; 2018.

35. ICAP at Columbia University. Swaziland HIV Incidence Measurement Survey 2: a Population-based HIV Impact Assessment SHIMS2 2016-2017. New York: ICAP at Columbia University; 2017.

36. Tanser F, Vandormael A, Cuadros D, et al. Effect of population viral load on prospective HIV incidence in a hyperendemic rural African community. Sci Transl Med 2017;9:eaam8012.

37. Grabowski MK, Serwadda DM, Gray RH, et al. HIV prevention efforts and incidence of HIV in Uganda. N Engl J Med 2017;377:2154-66.

Chapter Two. Measuring the population-level relationship between community viral load and new HIV infections: a systematic review

Summary

Background: HIV community viral load has been investigated to explain population-level differences in HIV incidence. Guidance from 2011 on measuring community viral load does not reflect recent findings. A literature review is needed on whether the cumulative evidence supports a reliable measure of population-level HIV viral exposure for predicting new HIV infections.

Methods: A systematic review was conducted using PubMed and Embase citations culled as of February 19, 2019. Publications included were English, peer-reviewed articles published from January 1, 1995 through February 15, 2019 that used explicit numerator and denominator definitions of a viral load metric and an outcome of HIV recency, incidence, seroconversion, or new diagnosis. All studies sampled general or key populations through population-based surveillance registries, household-based enumeration, cluster sampling, or respondent driven sampling. A primary and secondary reviewer independently screened citations for eligibility. A primary reviewer manually extracted data using a standardized form and assessed study quality using a modified Newcastle-Ottawa Scale.

Results: Twenty eligible studies identified from 4207 screened references averaged a quality score of 4.4 [range: 2-9] of 10. Ten studies statistically analyzed the relationship between population-level HIV viral exposure and new HIV infections. The majority [62%, 21/34] of analyses were statistically significant. Population exposure was measured using 14 numerators among five groups. Analyses inclusive of HIV-negative persons [n=5] showed significant relationships, while general population analyses among only HIV-positive persons [n=4] were not significant.

Conclusion: The evidence base mainly comprises lower quality studies and is inconsistent in exposure measurement. The literature, however, supports a relationship between population-level HIV viral exposure and new HIV infections.

Introduction

"Community viral load" was first coined in 2008 to explain disparities in HIV prevalence among men who have sex with men (MSM) in the United States. High proportions of untreated infections and subsequent high viremia in the African American MSM community were putatively leading to greater new infections.[1-3] The novel concept motivated the field to investigate whether decreases in an aggregate community measure of HIV viral load (i.e., the blood level concentration of HIV ribonucleic acid in persons living with HIV (PLHIV)) could reduce incident infections. Shortly thereafter, population-based observational data from Vancouver and surveillance data from San Francisco corroborated the intriguing hypothesis.[4,5] Additional evidence from British Columbia, Canada showed an ecological relationship between HIV antiretroviral treatment (ART) coverage, declines in community viral load, and decreased new HIV diagnoses.[6] Combined with the landmark findings from HPTN 052 – a randomized clinical trial showing a 96% reduction in HIV transmission among HIV-positive persons on sustained ART – the strategy of HIV treatment as prevention was embraced.[7] Given these developments, health surveillance systems in the United States were directed to measure community viral load for epidemic monitoring.

In 2011, the Centers for Disease Control and Prevention (CDC) published guidance on the surveillance of community viral load to codify terms on population-level measures of HIV viral load.[1] The guidance described five population subgroups of PLHIV: 1) undiagnosed, 2) diagnosed but not in HIV care, 3) in care but with no viral load test, 4) in care with detectable viral load, and 5) in care with undetectable viral load. Among these five groups, four viral load measurements were proposed – population viral load, community viral load, in care viral load, and monitored viral load.

Population viral load was defined as the aggregate viral load of all PLHIV in a population (i.e., groups 1 to 5). Although characterized as an unmeasurable construct, it was deemed the most comprehensive measure of a population's HIV transmission potential. Community viral load, the aggregate viral load measure of all diagnosed PLHIV (i.e., groups 2 to 5), was the subset of population viral load who were aware of their HIV status. In care viral load reflected persons who had accessed HIV care but who may or may not have had a reported viral load test (i.e., groups 3 to 5), while monitored viral load reflected those who had accessed both HIV care and viral load monitoring (i.e., groups 4 and 5). In care and monitored viral load were proposed as quality of care indicators on viral load monitoring and on access to

care and ART adherence, respectively. Categorical measures of the proportions with viral suppression or with undetectable virus were also suggested as indicators on uptake of national treatment guidelines.

The CDC framework was mainly informed by viral load analyses of case surveillance data from high-resource, low HIV prevalence settings, in the context of ART eligibility criteria of <500 copies per milliliter (c/mL) or less. Today, universal ART is standard of care, and more robust data sources are available from high prevalence settings and key populations like MSM and persons who inject drugs (PWID). The guidance on population-level viral load measurement has yet to consider these advances. To address this gap, this systematic review will assess the literature examining the relationship between population-level measures of HIV viral load and incident HIV infection in generalized and concentrated epidemics across low, middle, and high-income countries. The review will build on the CDC guidance to clarify whether the literature substantiates a reliable measure of population-level HIV viral exposure that predicts an association, if one exists, with HIV incidence.

Methods

Search strategy and selection criteria

A systematic review of the peer-reviewed literature was conducted using PubMed (including Medline) and Embase reference databases. Citations were culled, as of February 19, 2019, using keyword terms and medical subject headings (MeSH) to identify the exposure (e.g., population viral load, community viral load, population viral suppression, detectable viremia) and outcome (e.g., HIV infection, incidence) of interest from publications on population-level analyses (e.g., epidemiological monitoring, health impact assessment, population surveillance) (Supplemental Table S2.1).

Eligibility criteria were designed to select population-level assessments using varied exposure and outcome definitions across different types of HIV epidemics. Inclusion criteria for selection were: 1) peer-reviewed articles in English; 2) publication date from January 1, 1995 through February 15, 2019; 3) general or key populations sampled through population-based surveillance registries, household-based enumeration, cluster sampling (e.g. time-location, venue-based), or respondent driven sampling; 4) explicit definitions of the numerator and denominator for viral load metric; and 5) outcome measure of new HIV diagnosis, HIV seroconversion, HIV incidence, or HIV recency. Publications with empirical outcomes among fewer than 100 people were excluded.

Data analysis

PRISMA reporting guidelines were followed.[8] Covidence, an online systematic review tool, was used for citation screening and review. Search results from databases were merged, and duplicates were removed. Citation titles were screened by a primary reviewer to remove irrelevant references. A primary and secondary reviewer independently screened the remaining citations for eligibility. Citation review was an iterative process, where reference lists of included studies were searched for additional articles. Discrepancies were jointly reviewed and resolved through consensus. Citations that met inclusion criteria were retrieved, and a primary reviewer manually abstracted data using a standardized form detailing citation components; study design; study group, if applicable (arms or comparison populations); interventions, if applicable; sampling strategies; setting; participant eligibility criteria; study population characteristics; viral load metric, including definition of numerator and denominator; outcome measures; and outcome results. A primary reviewer assessed the risk of bias using the Newcastle-Ottawa Scale for nonrandomized studies, which was adapted to reflect exposure group comparability related to HIV risk factors and use of biomarker measures and population registries.[9,10]

Study characteristics were assessed first by grouping studies according to study population and setting, to facilitate interpretations by risk populations, and then by detailing key characteristics: catchment area, study period, sample characteristics, study design, data source, exposures, and outcomes. For exposures, the numerator was defined according to the number or summary value of HIV viral copies per milliliter (c/mL). The denominator was defined as the population among which the numerator was measured. A sub-analysis of study characteristics summarized studies by delineating the sample source by the sample population.

Study findings were evaluated by grouping studies by analysis strategy to facilitate inferences by analysis method. Studies were then described by factors that would help gauge the scientific strength of the findings: study population, sampling strategy, quality (i.e., risk of bias), and relevant study findings. A sub-analysis summarized the statistical significance across studies by: 1) categorizing published analyses according to the exposure population, 2) sub-setting the exposure numerator within categories, and 3) tabulating within subsets the outcome measure and statistical significance of the exposure-outcome association/correlation.

Results

Keyword search criteria and reference lists of included studies identified 4660 potentially eligible references (Figure 2.1). A total of 453 duplicates were detected and removed, resulting in 4207 non-duplicate references for screening. Among them, 4117 citations were determined irrelevant after title and abstract screening, and the remaining 90 references were assessed for eligibility based on abstract and full-text reviews. A total of 70 references were excluded per the eligibility criteria. Studies were most commonly excluded due to an exposure, outcome, or study design that was inconsistent with study inclusion criteria. A total of 20 studies were identified for review and synthesis.

Study characteristics

The 20 studies were published from 2000-2018 and included populations across Africa, Asia, Europe and North America (Table 2.1). One study each was from China,[11] India,[12] the Netherlands,[13] and South Africa.[14] Two studies were from Denmark,[15,16] four from Canada,[4,6,17,18] four from Uganda,[19-22] and six from the United States.[5,23-27] Three studies were at the national level, six were at the state/provincial level, and the remaining were at sub-administrative levels (e.g. county, city, or lower). The study durations ranged from 1.25 to 18 years, with half [50%] spanning ten or more years. The earliest study data were collected in 1987, and the most recent data were collected in 2016. Sample sizes ranged from 256 to 33,937 persons.

Across all studies, four types of populations were assessed – low HIV prevalence, general populations;[5,6,13,15,17,18,23-25] high HIV prevalence, general populations;[14,22] key populations;[4,12,16,26,27] and HIV serodiscordant couples[11,19-21] – among which four outcome measures were identified: 1) new HIV diagnosis; 2) recent infection calculated according to a HIV recency assay testing algorithm; 3) observed seroconversion; and 4) incident infection calculated by adjusting the number of new diagnoses by correlated factors (Table 2.1). Exposures were measured among five groups across the study populations: 1) reported HIV cases with viral load measures; 2) HIV-positive persons in HIV care; 3) HIV-positive persons who were diagnosed; 4) HIV-positive persons, regardless of prior diagnosis; and 5) all persons, regardless of HIV status.

Of the nine studies conducted in low HIV prevalence, general populations, six were designed as serial cross-sectional cohorts,[5,6,17,23-25] and three were population-based open cohorts (Table 2.1).[13,15,18] All

studies used case surveillance registries as the data source and included new HIV diagnosis as an outcome. When measuring exposure, 18 unique numerators were identified. The exposure denominator was the number of reported HIV-positive cases with viral load measures in six studies and the number of HIV-positive persons in care in the three open cohorts.

The two studies in the high HIV prevalence, general populations were prospective, population-based open cohorts collecting household-based, surveillance site data with observed HIV seroconversions as the outcome (Table 2.1).[14,22] Four unique exposure numerators were measured. The exposure denominator in both studies included the number of HIV-positive persons, regardless of prior diagnosis and, in one study, the number of persons, regardless of HIV status.

Five studies were conducted among key populations, specifically MSM and PWID (Table 2.1). Two studies were cross-sectional cohorts,[12,27] and one study each was a population-based open cohort,[16] a serial cross-sectional cohort,[26] and a two parallel arm prospective observational cohort.[4] Two studies used case surveillance registries and three relied on target population sampling to measure number of new diagnoses [n=2], observed seroconversions [n=2], and recent infections [n=1]. In each study, a unique exposure numerator was used, while the exposure denominator varied across studies, including number of persons regardless of HIV status [n=2]; number of HIV-positive persons regardless of prior diagnosis [n=3]; number of HIV-positive persons who were diagnosed [n=1] or who were in care [n=2]; and number of reported HIV cases with viral load measures [n=1].

All four serodiscordant couple studies were population-based cohorts of which one was sampled from a case surveillance registry[11] and three were nested, either prospectively[19] or retrospectively,[20,21] in an open cohort sampled from a household-based surveillance site (Table 2.1). All study outcomes were observed seroconversions. Six unique numerators were used in the exposure measure. The exposure denominator was reported HIV cases and HIV-positive persons who were diagnosed in one and three studies, respectively.

Frequency of sample sources

National surveillance case registries were the most common data source, accounting for 60% [n=12] of all reviewed studies (Figure 2.2). All [100%, n=9] studies in low HIV prevalence populations were based on case surveillance data,[5,6,13,15,17,18,23-25] as well as two-fifths [40%, n=2] of those conducted in key

populations.[16,27] All [100%, n=2] studies from high HIV prevalence populations were sampled from household-based surveillance sites.[14,22] Direct sampling of target populations was conducted only among key populations and included both non-probability (i.e., convenience sampling)[4] and probability (i.e., respondent driven sampling, targeted sampling)[12,26] methods.

Study findings

Three analytic strategies – assessing ecological trends, evaluating ecological correlations, and estimating population-level exposure-outcome associations – were identified across the 20 studies, resulting in 22 sets of findings evenly distributed across strategies (8, 7, and 7 sets, respectively) (Table 2.2). The first strategy, assessing ecological trends, comprised of studies documenting temporal changes or venue-based differences in HIV viral exposure and new HIV infections but did not statistically test for an exposure-outcome association. In contrast, studies evaluating ecological correlations did test for statistically significant correlations between group-level trends in HIV viral exposure and new HIV infections. Studies using individual-level longitudinal data and statistically testing the relationship between HIV viral exposure and new HIV infections within well-defined populations were categorized as estimating population-level exposure-outcome associations. Studies in low HIV prevalence populations were exclusively analyzed for ecological trends or correlations, and serodiscordant couple studies were only used to estimate exposure-outcome associations. The mean quality (i.e., risk of bias) score across the 20 studies was 4.4 [range: 2–9] out of 10 maximum score (Supplemental Table S2.2).

Of the eight studies that assessed for ecological trends, half [50%, n=4] were conducted in low HIV prevalence populations[15,18,23,25] and the remaining in high HIV prevalence [n=2][14,22] and key [n=2][16,27] populations (Table 2.2). Studies were mainly [63%, n=5] population-based open cohorts or cross-sectional cohorts (either serial [n=2] or single [n=1]), and the majority [75%, n=6] were sampled from case surveillance registries. These six studies each scored 2 on quality, consequently reducing the mean score across all eight studies to 3.6 [range: 2–9].

The majority [71%, n=5] of the seven studies evaluating ecological correlations were conducted in low HIV prevalence populations,[5,6,13,17,24] with one study each in high HIV prevalence[22] and key populations (Table 2.2).[12] Over half [57%, n=4] were serial cross-sectional cohorts sampled from case registries. The mean quality score was 3.7 [range: 2–9], where five of the seven studies scored at 2.

The seven studies that estimated population-level exposure-outcome associations were conducted in serodiscordant couples [n=4],[11,19-21] as well as high HIV prevalence [n=1][14] and key [n=2] populations (Table 2.2).[4,26] The sampling strategy ranged across studies, including population-based cohorts (retrospective [n=1] and prospective [n=2]) sampled from household surveillance sites, and one each of serial cross-sectional cohorts using targeted sampling, parallel arm observational cohorts using snowball sampling, population-based prospective cohort sampled from a case registry, and open cohort sampled from household surveillance sites. The mean quality score was 7.0 [range: 5–9].

Study findings overall generally support ecological trends and correlations between population-level HIV viral exposure and new HIV infections. Serial cross-sectional studies[23,25] and household-based open cohorts[14,22] showed trends of significantly increased HIV viral suppression over time, complemented by decreases in new HIV diagnoses. Although one household-based open cohort did not find a statistical correlation between community prevalence of unsuppressed viral load and observed seroconversions over time,[22] most serial cross-sectional studies[5,6,17] and an open cohort sampled from case registries[13] indicate the observed trends are statistically correlated. Similarly, nearly all the studies in serodiscordant couples,[11,19,21] as well as studies in high prevalence[14] and key populations,[4,26] demonstrate statistically significant associations between unsuppressed viral load and HIV seroconversions among serodiscordant dyads and at the population level.

Across the three analytic strategies, three studies contrast the measurement of population-level HIV viral exposure among all HIV-positive persons versus all persons, regardless of HIV status. In one study, a dose-response pattern was demonstrated among all persons, regardless of HIV status – but not among all HIV-positive persons – when plotting the geometric mean viral load and prevalence of viremia (≥1550 c/mL) against the rate of seroconversion.[14] Another study found the strongest correlation between population-level viral exposure and recent infection when measuring prevalence of viremia (>150 c/mL) among all MSM and PWID, regardless of HIV status, than when compared to measurements of mean log10 viral load among all or only diagnosed HIV-positive MSM and PWID.[12] Moreover, two studies in high prevalence[14] and key populations[26] demonstrated a statistical association between observed seroconversion and population-level viral exposure (geometric mean viral load, viremia at ≥1550 c/mL, and mean viral load) among all persons regardless of HIV status, but not among all HIV-positive persons.

Frequency of analyses by exposure and outcome measures

Ten of the 20 reviewed studies analyzed the statistical correlation or association between population-level HIV viral exposure and new HIV infections.[4-6,12-14,17,22,24,26] Figure 2.3 shows seven categories of population exposure measures identified across the ten studies: 1) reported HIV cases whose viral load was monitored (monitored viral load); 2) HIV-positive persons who were currently in care, regardless of viral load monitoring (in-care viral load); 3) HIV-positive persons who were diagnosed, regardless of HIV care status (diagnosed viral load, in key populations); 4 and 5) HIV-positive persons, regardless of prior diagnosis (population viral load, in general and key populations); and 6 and 7) all persons, regardless of HIV status (population viremia, in general and key populations). Across the seven categories, 14 unique exposure numerators were measured; and thirty-four analyses were conducted of which the majority [62%, n=21] were statistically significant. The most [21%, n=7] common numerator was mean viral load, where under half [43%, n=3] showed a significant finding. Half [50%, n=17] of the assessed study outcomes was new diagnosis, and nearly a third [32%, n=11] was observed seroconversion.

Monitored and in-care viral load exposure measures comprised over half [59%, n=20] of analyses (Figure 2.3).[5,6,12,13,17,24] Most [85%, n=17] of the study outcomes was new diagnosis, and three-quarters [76%, n=13] of them demonstrated a significant exposure-outcome relationship. The three remaining outcomes (i.e. recent infection) did not show statistical significance.

Fourteen analyses were conducted in the three other exposure groups – diagnosed viral load, population viral load, and population viremia (Figure 2.3).[4,12,14,22,26] Most [79%, n=11] of the study outcomes was observed seroconversion, with nearly half [45%, n=5] showing significant exposure-outcome relationships. The three other outcomes of recent infection were measured in key population studies (both MSM and PWID) across the three exposure groups and were statistically significant. Population viral load consisted the majority [57%, n=8] of the 14 analyses and was evenly divided among general population [n=4] and key population [n=4] studies. Population viremia analyses comprised over a third of analyses [36%, n=5], also with a relatively even divide across general [n=2] and key [n=3] population studies.

None of the population viral load analyses in the general population demonstrated statistical significance when using a summary measure (geometric mean), categorical measure (viral load prevalence ≥1550 c/mL), or contextual measure (median community prevalence ≥1000 c/mL) (Figure 2.3).[14,22] All

analyses in the population viremia group – both general and key populations – showed significant exposure-outcome relationships.[14]

Discussion

In this systematic review of the literature on population-level HIV viral exposure and new HIV infections, the evidence largely indicates that increasing suppression of HIV viral load in a population is positively associated with a decrease in new HIV infections. The evidence base, however, is tempered by lower quality studies derived from convenience sampling of case registries that preclude outcome comparison among different exposure groups and collect numbers of new HIV diagnoses, a measure that overrepresents prevalent cases. The literature also shows marked inconsistency across exposure definitions and the populations in which exposure is measured. The relative paucity of higher quality studies points toward the challenge of conducting robust measurement in population-level HIV biomarker studies.

The review highlights that measurement of population-level HIV viral exposure is commonly framed by the available data source and, therefore, subject to the validity of the data source's exposure measure. Over half of the findings from this review was based on case registry data on HIV-positive persons who received HIV testing and viral load monitoring services – a selected sub-population more likely to access treatment and have lower HIV transmission risk. Household-based samples and direct sampling of target populations can address this bias and, among reviewed studies, comprised all exposure measurements in HIV-positive persons who were undiagnosed or diagnosed but not accessing HIV care (i.e., diagnosed and population viral loads). These studies supplemented the quality of the evidence base with generally more representative population measures of HIV viral quantity among all PLHIV.

The range of data sources in the reviewed studies also shows an evolving definition of population-level HIV exposure. The case registry studies typically reflected an older period where population-level HIV viral exposure was defined as a summary measure of HIV viral quantity in an infected population. Yet with recent household-based samples, transmission risk was also measured as a population-level exposure, defined as a summary measure of HIV viral quantity in a population inclusive of persons infected with HIV and at-risk of HIV acquisition. The household-based and directly sampled studies reflected this evolution through their analyses of all persons, regardless of HIV status (i.e., population viremia).

Review findings also demonstrate that the type of study population influences the feasibility of discerning population-level exposure-outcome relationships. For example, studies in high HIV prevalence, general populations were sampled through household-based surveillance sites. This sampling strategy facilitated not only stronger measurement and construct validity of exposure, as described above, but also routinely employed the gold standard of incidence measurement, directly observed HIV seroconversion. The larger case samples in these settings increased statistical power to detect differences in outcomes between comparison groups and more reliably determined population-level exposure-outcome associations.

Key population studies provided a reasonable strength of causal inference, given the: 1) high HIV prevalence and greater statistical power; 2) outcomes in three studies of observed HIV seroconversion or recent HIV infection (via validated HIV recency assay); and 3) exposure measurement in those studies among all persons regardless of HIV status or of all HIV-positive persons. The scientific rigor, however, varied by direct sampling strategy where only one of the three studies used probability sampling. In contrast, all studies in low HIV prevalence, general populations were sampled from case registries with study outcomes of new HIV diagnosis, an imprecise proxy of incident infection. Of note, serodiscordant couple studies were generally higher quality and afforded stronger causal inference at the dyad level.

Considerations when measuring population-level HIV exposure

While this systematic review did not identify a reliable population-level HIV exposure measure, the range of analyzed variables may imply that well-defined exposures are contextually driven. When new HIV infection is the outcome of interest, transmission risk is a more appropriate exposure construct than viral exposure, since a population's transmission potential is of greater relevance to incident HIV. Given that incidence rates are proportionally determined by prevalence levels,[28] an exposure measure accounting for viremia distribution and the relative size of infected and uninfected populations (i.e. prevalence) may carry greater predictive value.[29,30] Population viremia measures in high prevalence populations in this review consistently showed an association with HIV incidence, whereas population viral load did not show an association in high prevalence, general populations and was inconsistent in key populations.[12,14,26] Being a composite measure, population viremia may capture more than relative size of an infected population. When HIV prevalence was included in models assessing population viremia as a predictor of HIV incidence,

the significance of population viremia persisted.[14] Evaluations of transmission risk may only be feasible in high prevalence populations, since case registries will likely remain a primary data source in low prevalence, general populations.

Similarly, the extent of HIV service uptake in a population can influence how population-level HIV exposure is measured. ART coverage impacts viral load distribution which is often skewed by variations in treatment access and adherence. Summarizing viral load data by an arithmetic mean may be an infrequent option since arithmetic means best represent symmetrical data distributions. Measures of location for non-normal distributions, like the median, is appropriate when skewness is towards the right or when working with polymodal distributions. It may, however, inadequately reflect reductions in population viral quantity as the prevalence of undetectable virus increases with expanded ART coverage. Variation in median viral load would diminish over time, as it would remain below the limit of detection.[4,5,13]

Log transformation of viral load data may often be warranted since it reduces skewness, thereby increasing data interpretability, while also meeting assumptions of inferential statistical analyses. Estimates of mean log viral load consequently decrease the influence of outliers and provide more stable estimates.[1] Yet when the prevalence of viral suppression and undetectable virus is more common, the geometric mean, the anti-log of the arithmetic mean of the log values, may be a more robust measure of central tendency. This measure of location is less affected by very high values, being appropriate when viral load distribution is heavily skewed to the right. Such distributions are more likely in populations with higher ART coverage and when analyzing population viremia (i.e., viral measures among all persons, regardless of HIV status).

Transmission risk may potentially be better predicted by viral quantity measures than by categorical measures.[14] Prevalence of undetectable virus, however, can simultaneously provide insight into transmission risk as well as program quality in a population. Evidence now supports that an undetectable viral load confers zero transmission risk,[31] and in settings where universal ART is standard care, undetectable viremia is an indicator of ART access and adherence. Thus, where prior exposure measures used cut-off values demarking transmission risk (i.e., 400 c/mL), prevalence of undetectable virus is an appropriate replacement. A limitation with this measure is the different minimum levels of detection across assays.

Other cut-off values noted in this review (i.e., 200 and 500 c/mL) demarked changes in viral suppression definitions. Of note, the threshold of 1000 c/mL is commonly used in viral load analyses and was recommended by the World Health Organization to demarcate population levels of virological failure (i.e., treatment failure when distinguished through viral blips), while also affording insight into ongoing HIV transmission.[32,33] Although one study in this review did not ascertain any difference when using a cut point of 150, 1000, or 5000,[12] future analyses using a threshold of 1000 c/mL may have value by providing a snapshot of both transmission risk and program quality. Lastly, CDC guidance suggests prevalence of high viral load (≥100,000 c/mL) as a program indicator.[1] Although not used in the reviewed analyses, it could provide insight into geographic "hotspots" for new HIV infections and serve as an early warning signal for new or re-emerging micro-epidemics.

Limitations and strengths of review

The findings of this systematic review are limited to the English peer-reviewed literature and therefore exclude analyses from conference proceedings and the grey literature. This constraint was intended to better assure higher quality studies given the challenges of population-level infectious disease research. Findings may be resourceful because of the broad applicability to population-based HIV research. The conclusions pertain to all epidemic types (hyperendemic, endemic, concentrated, and mixed) and accommodate multiple, common outcome measurements for new HIV infection. The 24-year duration of study inclusion also illustrates the historical progression of population-level HIV exposure measurement with possible markers for future research directions.

Future directions

Three well-conducted cluster randomized trials published after the study inclusion period for this review evaluated the population-level effect of a test-and-treat approach on HIV incidence in high prevalence, general populations.[34-36] Only one study demonstrated reduced HIV incidence after implementing a combination prevention intervention with ART initiation based on local guidelines (i.e., use of $CD4^+$ eligibility criteria prior to 2016), when compared to standard care. The finding was subdued, however, by a second study arm which failed to show an impact of a similar combination intervention with universal ART. Surprisingly, with nearly indistinguishable ART coverage between intervention communities, viral suppression (<400 c/mL) prevalence among all HIV-positive persons in the universal ART communities

was higher than in the communities with an ART eligibility threshold (72% versus 68%).[35] The discrepant – and perplexing – findings across trials may suggest that universal coverage is necessary but insufficient to meaningfully curb HIV incidence, and a possible deterrent is successful engagement of hard-to-reach sub-populations at higher risk of HIV acquisition and transmission.[37]

This review's findings lend support for this hypothesis. For example, one study in this review suggested that the overall decrease in new HIV infections in British Columbia, Canada was due to the marked increase of PWID accessing treatment and suppressing viral load, compared to other risk groups like MSM.[17] Also, the initial 2-year period after migration into Rakai, Uganda was associated with higher HIV acquisition risk among in-migrants.[38] The same household-based study population, meanwhile, did not show in this review an ecological correlation between community prevalence of unsuppressed viral load and observed seroconversions.[22] Similarly, the "churn" of mobile HIV-positive persons in and out of care may have obscured an ecological relationship between viral exposure and HIV transmission in another reviewed study.[18]

Given the evidence that undetectable means untransmissible at the individual-level, the divergent population-level findings may illustrate that population-level viral load reductions obscure heterogeneity of acquisition and transmission risks in sub-groups of high-risk populations.[39] Core-group theory postulates that a disproportionate number of infections result – either directly or indirectly – from a small sub-group of people and that intervening with this small sub-group can result in greater decrease in community-level infections.[40,41] Since a robust viral load measure is not only reliable over time but also reflects sexual behaviors or other HIV transmission risks in the community,[1] aggregate viral load measures in core groups may more precisely pinpoint population mechanisms of ongoing HIV transmission. If so, such information could better inform public health strategies for identifying sub-populations with emergent or sustained micro-epidemics and could more judiciously direct HIV prevention resources.

Although core groups (or the sub-groups) of high-risk populations are difficult to find, novel approaches to measuring viral exposure and transmission risk in these groups may be especially useful. One consideration is network viral load, or the aggregate measure of viral load among members of a risk environment. Defined as the "composition of risk factors external to the individual, such as community-level norms and practices",[42] a risk environment encompasses individuals networked through shared experience

that elevates HIV-related risk.[43] Network viral load could be collected through respondent-driven sampling in key population studies. A second consideration for future research is the idea of risk spaces, or physical spaces that elevate HIV-related risk,[40,41] and the measurement of viral load characterizing those spaces. Venue viral load, a novel concept, showed support of a relationship with new HIV diagnoses in this review, albeit by a scientifically limited study design.[27] Nonetheless, evidence more broadly shows that incident HIV tends to cluster both geographically and within sexual networks.[44-46] Thus, venue-based sampling, often conducted at social venues in key population studies, may be an option. Lastly, considerations of viral measures among migrant populations in relation to migration patterns may be an important area of exploration. A possible measure for consideration in these populations is a community transmission index, a composite measure which incorporates a community's viral load, estimated sexual contact rates, and transmission probability. This may be a more sensitive measure than viral quantity for evaluating transmission risk in a population.[14]

In conclusion, the reviewed evidence base, while limited by lower quality studies, supports a relationship between population-level HIV viral exposure and new HIV infections. Decreasing levels of unsuppressed HIV in HIV-infected populations has been associated with fewer new infections over time. Differences in transmission risk across sub-populations is not well understood and may be critical to continuing reductions in incident HIV. Future research assessing population-level HIV viral exposure and measurement methods in subpopulations defined by locations, such as venues, or risk networks may help unlock the key population-level mechanism for controlling epidemics. When coupled with greater use of HIV recency assays, for greater feasibility in valid measurement of new infections, these advances may identify new opportunities for population-level program interventions and to eventual control of the global epidemic.

Tables and figures

***Figure 2.1**. Study selection*

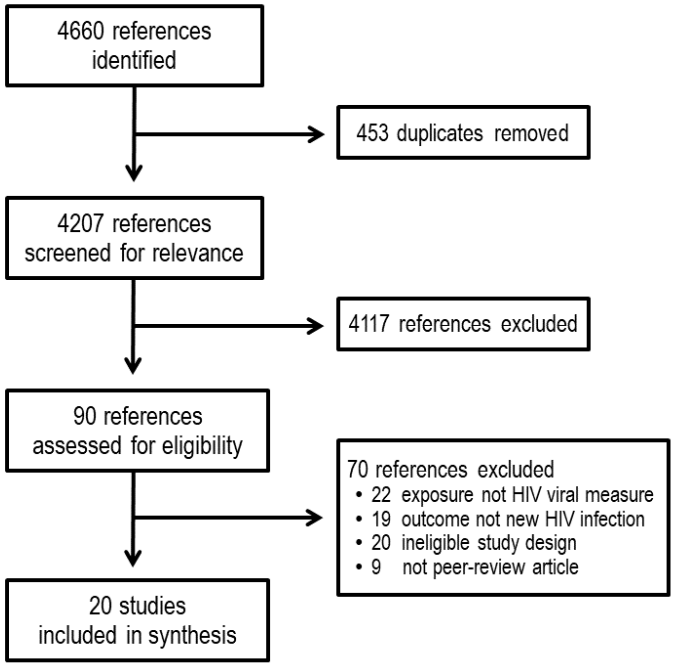


***Table 2.1**. Study characteristics*

Study	Catchment area	Years	Sample characteristics	Study design	Data source	Outcome	Exposure
Low prevalence setting							
Montaner 2010 [6]	British Columbia, Canada	1996-2009	Not specified	Ecological; serial cross-sectional cohorts	Case surveillance, population-based registries	Number new diagnoses	**Case count**: Number of reported HIV cases, with viral load measured as: log10 c/mL,[1] <50 c/mL, <500 c/mL.
Montaner 2014 [17]	British Columbia, Canada	1996-2012	Not specified	Ecological; serial cross-sectional cohorts	Case surveillance, population-based registries	Number new diagnoses, modelled incident cases[2]	**Case count**: Number of reported HIV cases, with viral load measured as: <50 c/mL, <500 c/mL.
Castel 2012 [24]	District of Columbia, United States	2004-2008	15,467 HIV-positive persons, >12 years; 38% MSM,[3] 20% PWID[4]	Ecological; serial cross-sectional cohorts	Case surveillance	Number new diagnoses	**Numerator**: mean c/mL; **Denominator**: reported HIV cases with viral load measures
Golden 2016 [23]	King County, United States	2004-2013	9539 HIV-positive persons, all ages; 67% MSM, 6% PWID	Ecological; serial cross-sectional cohorts	Case surveillance	Number new diagnoses	**Numerator**: < 200 c/mL; **Denominator**: reported HIV cases with viral load measures
Chakraborty 2017 [25]	South Carolina, United States	2004-2013	11,121 HIV-positive persons, >12 years; 32% MSM	Ecological; serial cross-sectional cohorts	Case surveillance	Number new diagnoses	**Numerator**: mean log10 c/mL **Denominator**: reported HIV cases with viral load measures

[1] c/mL: copies per milliliter
[2] Involved use of a CD4 cell backcalculation method
[3] MSM: men who have sex with men
[4] PWID: person who injects drugs

Study	Catchment area	Years	Sample characteristics	Study design	Data source	Outcome	Exposure
Das 2010 [5]	San Francisco, United States	2004-2008	12,512 HIV-positive persons, all ages; 72% MSM, 8% PWID	Ecological; serial cross-sectional cohorts	Case surveillance	Number new diagnoses, recent infection[5]	**Numerator**: mean c/mL, total c/mL **Denominator**: reported HIV cases with viral load measures
Lohse 2005 [15]	Denmark	1995-2003	3941 HIV-positive persons, >15 years; 44% MSM, 10% PWID	Ecological; prospective, nationwide, population-based open cohort	Danish HIV Cohort[6]	Number new diagnoses	**Numerator:** >10,000 c/mL **Denominator:** HIV-positive persons, in HIV care
Krentz 2013 [18]	Southern Alberta, Canada	2001-2010	1423 HIV-positive persons, >15 years	Ecological; prospective, population-based open cohort	Southern Alberta Clinic Cohort[7]	Number & percentage new diagnoses	**Numerator**: ≤50 c/mL, ≤200 c/mL, >200 c/mL, >100,000 c/mL, total c/mL **Denominator**: HIV-positive persons, in HIV care
Bolijn 2017 [13]	Netherlands	2002-2013	20,740 HIV-positive adults; 55% MSM	Ecological; prospective, nationwide, population-based open cohort	ATHENA Cohort[8]	Number new diagnoses at 1, 2, 3, and 4 years of follow-up	**Numerator**: log10 at diagnosis, mean log10 during calendar year >400 c/mL, mean log10 during calendar year ≤200 c/mL; group mean of mean log10 during calendar year, group mean of first/last/highest log10 measure during calendar year, group median of median log10 during calendar year, first log10 of newly diagnosed with mean log10 of

[5] Recent infection determined using STARHS recent infection testing algorithm.
[6] Open cohort sampled through Danish national health registry and includes all HIV-positive persons receiving care across the eight treatment centers in the country.
[7] Open cohort sampled through Alberta case registry and includes all HIV-positive person receiving HIV care in southern Alberta.
[8] Open cohort sampled through Netherlands national case registry and includes all diagnosed HIV-positive persons receiving HIV care across all 26 HIV treatment centers nationwide.

Study	Catchment area	Years	Sample characteristics	Study design	Data source	Outcome	Exposure
							prevalent cases during calendar year **Denominator**: HIV-positive persons, in HIV care
High prevalence setting							
Grabowski 2017 [22]	Rakai district, Uganda	1999-2016	33,937 persons (17,870 HIV-negative), 15-49 years	Prospective, population-based open cohort	Rakai Community Cohort Study[9]	Observed seroconversions	**Numerator**: <1000 c/mL **Denominator**: HIV-positive persons, regardless of prior diagnosis
Tanser 2017 [14]	Umkhanyakude District, Kwazulu Natal, South Africa	2011-2015	11,152 persons (2420 HIV-positive, 8732 HIV-negative), >14 years	Prospective, population-based open cohort	Africa Centre Demographic Information System[10]	Observed seroconversions	**Numerator**: geometric mean, ≥1550 c/mL **Denominator**: HIV-positive persons, regardless of prior diagnosis; all persons, regardless of HIV status
Key populations							
Cowan 2012 [16]	Denmark	1995-2010	HIV-positive MSM, >15 years	Ecological; prospective, nationwide, population-based open cohort [subsample]	Danish HIV Cohort	Number new diagnoses	**Case count**: Number of reported HIV cases, in HIV care, with viral load measured as: >400 c/mL
Wood 2009 [4]	Downtown Eastside Vancouver, Canada	May 1996-Jun 2007	2051 PWID persons (622 HIV-positive, 1429 HIV-negative)	Ecological; prospective observational parallel cohorts	Vancouver Injection Drug Users Study (VIDUS) and Barriers to Accessing Antiretroviral Therapy (BART)	Observed seroconversions	**Numerator**: median log10 c/mL **Denominator**: HIV-positive persons, regardless of prior diagnosis

[9] Surveillance site of 518,000 residents enumerated through household census.
[10] Demographic Surveillance Site of 87,000 residents enumerated through household census.

Study	Catchment area	Years	Sample characteristics	Study design	Data source	Outcome	Exposure
Glasheen 2018 [26]	San Francisco Bay Area, United States	1987-1999	2241 PWID,[11] >18 years	Ecological; serial cross-sectional cohorts	Urban Health Study	Observed seroconversions	**Numerator**: mean c/mL **Denominator**: HIV-positive persons, regardless of prior diagnosis; all persons, regardless of HIV status
Solomon 2016 [12]	22 cities in India	Oct 2012-Dec 2013	26,503 participants, >17 years; 45% MSM, 55% PWID	Ecological; cross-sectional cohort	Indian National Collaboration on AIDS Study (cluster-randomized trial), baseline measure	Recent infections[12]	**Numerator**: mean log10 **Denominator**: HIV-positive persons, in HIV care; HIV-positive persons, diagnosed; HIV-positive persons, regardless of prior diagnosis **Numerator**: >150 c/mL **Denominator**: all persons, regardless of HIV status
Brantley 2017 [27]	Baltimore, United States	Oct 2012-Jun 2014	256 newly diagnosed MSM, >15 years	Ecological;[13] population-based cross-sectional cohort	Case surveillance	Number new diagnoses	**Numerator**: geometric mean c/mL **Denominator**: reported HIV cases, with viral load measures who attended select venue
Serodiscordant couples							
He 2013 [11]	Dehong prefecture, China	Jun 2009-Mar2011	1101 HIV serodiscordant couples	Prospective, population-based, cohort	Case surveillance	Observed seroconversions	**Numerator**: number of c/mL **Denominator**: reported HIV cases

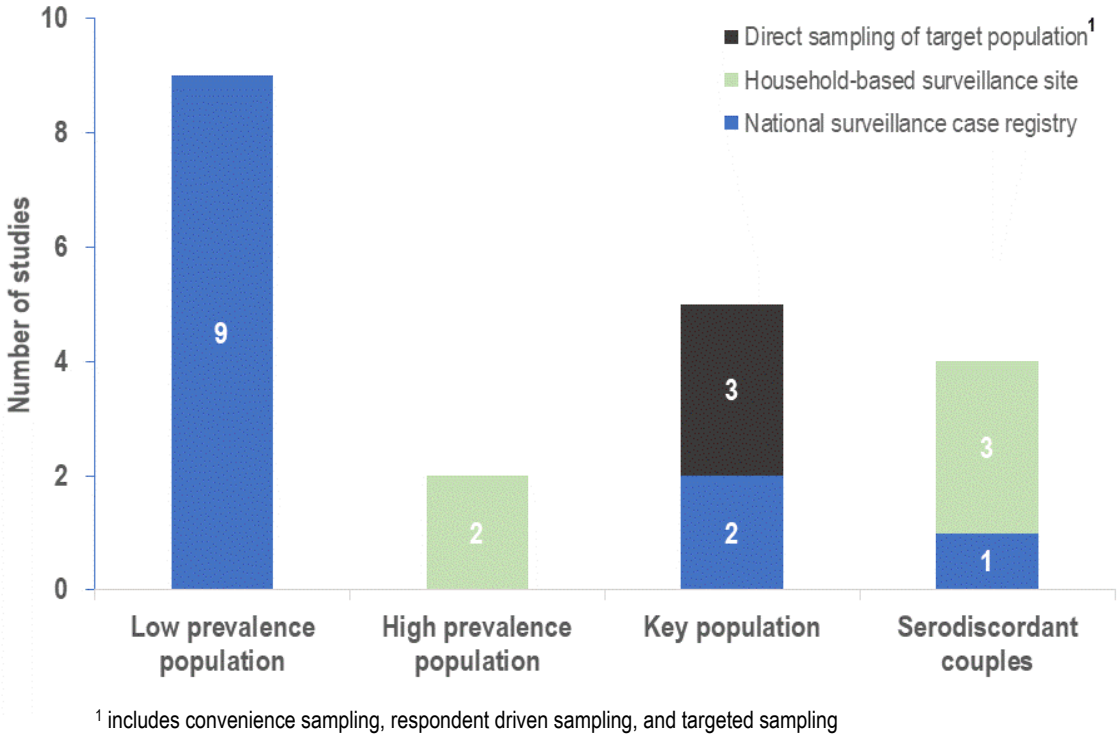
[11] PWID: person who injects drugs
[12] Recent infection determined using a multi-assay algorithm.
[13] Utilizes venue affiliation network analysis to describe the variability of HIV viral load across networks of sex partner meeting venues.

Study	Catchment area	Years	Sample characteristics	Study design	Data source	Outcome	Exposure
Biraro 2013 [20]	Kyamulibwa sub-county, Uganda	1989-2007	259 co-habitating HIV serodiscordant couples, 18-59 years	Retrospective cohort analysis; nested in prospective, population-based open cohort	General Population Cohort[14]	Observed seroconversions	**Numerator**: number of c/mL **Denominator**: HIV-positive persons, diagnosed
Quinn 2000 [19]	Rakai district, Uganda	Nov 1994-Oct 1998	415 HIV serodiscordant couples, 15-59 years	Prospective cohort analysis; nested in prospective, population-based open cohort	Rakai STD Control for AIDS Prevention Study[15]	Observed seroconversions	**Numerator**: number of c/mL, mean, mean log10, median, median log10 **Denominator**: HIV-positive persons, diagnosed
Kiwanuka 2009 [21]	Rakai district, Uganda	1997-2002	268 HIV serodiscordant couples, 15-59 years	Retrospective cohort analysis; nested in prospective, population-based open cohort	Rakai Community Cohort	Observed seroconversions	**Numerator**: log10 **Denominator**: HIV-positive persons, diagnosed

[14] Surveillance site of 20,000 residents enumerated through household census

[15] Predecessor of Rakai Community Cohort, a surveillance site of residents enumerated through household census.

***Figure 2.2**. Number of studies by data source and population type*



[1] includes convenience sampling, respondent driven sampling, and targeted sampling

***Table 2.2**. Summary of study findings*

Study	Population type	Sampling strategy	Quality score: Selection	Quality score: Comparability	Quality score: Outcome	Key measures of correlation/association
Assessing ecological trends						
Golden 2016 [23]	Low prevalence	Serial, cross-sectional cohorts via case surveillance registry	1	0 **Total score**: 2	1	From 2006 to 2013, the proportion with suppressed viral load (<200 c/mL) among reported HIV cases significantly increased from 45% to 86% (p<.0001). From 2009 to 2013, the duration when 84-86% of reported HIV cases had a recorded viral load test, the proportion with suppressed viral load significantly increased from 74% to 86% [p<.0001]. From 2004 to 2013, the number of new diagnoses declined 28%.
Chakraborty 2017 [25]	Low prevalence	Serial, cross-sectional cohorts via case surveillance registry	1	0 **Total score**: 2	1	From 2004 to 2013, mean log10 viral load among reported HIV cases significantly decreased from 3.02 to 2.14 [p=.0001]. During the same period, the mean annual number of new diagnoses significantly decreased from 342.29 to 41.68 [p=.0001].
Lohse 2005 [15]	Low prevalence	Prospective, population-based open cohort via case surveillance registry	1	0 **Total score**: 2	1	From 1995 to 2003, the proportion of reported HIV-positive persons with viral load >10,000 c/mL during the year of diagnosis was 77% [range: 71% -78%, trend p=.335]. From 1995 to 2002, the annual number of new diagnoses remained stable with a median of 231 per year [range: 192--258, p=.288].
Krentz 2013 [18]	Low prevalence	Prospective, population-based open cohort via case surveillance registry	1	0 **Total score**: 2	1	Among reported HIV-positive persons in HIV care from 2001 to 2010, the total viral load [range: 21,745,084-44,663,718 copies/mL] and annual mean participant viral load [range: 32,025-53,667 c/mL] were variable over time. The proportions with viral load ≤50 c/mL and ≤200 c/mL increased from 32% to 66% and from 49% to 72%, respectively. The proportion with viral load >200 c/mL and >100,000 c/mL reduced from 51% to 28% and 6.4% to 3.8%, respectively. The number of new diagnoses [range: 46 -88] and incidence per 100,000 [range: 4.4 - 7.1] were variable over time.

Study	Population type	Sampling strategy	Quality score: Selection	Quality score: Comparability	Quality score: Outcome	Key measures of correlation/association
Cowan 2012 [16]	Key population	Prospective, population-based open cohort [subsample] via case surveillance registry	1	0 **Total score**: 2	1	From 1995 to 2010, the number of MSM with viral load >400 showed a steady decrease over time of 75% from 1035 to 262. From 1995 to 2009, the number of new diagnoses in HIV care was variable (range: 71-137).
Brantley 2017 [27]	Key population	Population-based cross-sectional cohort via case surveillance registry	1	0 **Total score**: 2	1	Venue viral load was calculated as the geometric mean viral load of cases linked to a particular venue and was classified into low (<1500 c/mL), moderate (1500-50,000 c/mL) and high (>50,000 c/mL). Among the 23 venues, venue viral load was high in 4 venues, moderate in 18 venues, and low in 1 venue. Six of the venues accounted for 66% of reported sex meeting venues among cases. All six were moderate viral load venues.
Grabowski 2017 [22]	High prevalence	Prospective, population-based open cohort via household-based enumeration	4	3 **Total score**: 9	2	In unadjusted analysis, viral suppression (<1000 c/mL) among HIV-positive adults from 2009 to 2016 significantly increased from 42% to 75% [p<.001]. In adjusted analysis, the rate of observed HIV seroconversions significantly reduced 42%, from 1.17 cases per 100 person-years during 2002 to 2005 (i.e., the period before HIV service scale-up) to .66 cases per 100 person-years in 2016 [IRR: 0.58, CI 0.45, 0.76, p<.001].
Tanser 2017 [14]	High prevalence	Prospective, population-based open cohort via household-based enumeration	4	2 **Total score**: 8	2	When geometric mean viral load and the prevalence of detectable viremia (≥1550 c/mL) were measured among HIV-positive adults within a 3 kilometer radius of a HIV-negative adult, the number of seroconversions per 100 person-years plotted by quartile distribution of population viral load did not show an ecological relationship. Yet when geometric mean viral load and prevalence of detectable viremia was measured among adults, irrespective of HIV status, communities (i.e. areas within a 3 kilometer radius of a HIV-negative adult) with high and with unsuppressed viral loads showed a dose-response pattern to number of seroconversions.

Study	Population type	Sampling strategy	Quality score: Selection	Quality score: Comparability	Quality score: Outcome	Key measures of correlation/association
Evaluating ecological correlations						
Montaner 2010 [6]	Low prevalence	Serial, cross-sectional cohorts via case surveillance registry	1	0 **Total score**: 2	1	From 1996 to 2009, the annual number of new diagnoses decreased by 52% from 702 to 338 [p=.001]. Among reported HIV-positive cases , for every one unit log10 decrease in viral load, the annual number of new HIV diagnoses decreased by a factor of 0.86 [CI 0.75,-0.98]. IDU vs non-IDU: From 1999 to 2009, the annual number of new diagnoses among IDU decreased by 50% from 159 to 80 [p=.003]. Among non-IDU, annual new diagnoses remained stable [p=.623]. Among IDU, the number with highest annual viral load <500 c/mL increased 82% [p=.001] and with highest annual viral load <50 c/mL increased 86% [p=.002]. Among non-IDU, the number with highest annual viral load <500 c/mL increased 36% [p=.0002] and with highest annual viral load <50 c/mL increased 42% [p=.001].
Montaner 2014 [17]	Low prevalence	Serial, cross-sectional cohorts via case surveillance registry	1	0 **Total score**: 2	1	Among reported HIV-positive cases, from 1996 to 2012, the proportion with highest annual viral load <500 c/mL significantly increased from 8% to 74% [p.0001]. From 1999 to 2012, the proportion with highest annual viral load <50 c/mL significantly increased from 6% to 59% [p= .0001]. The annual number of new diagnoses from 1996 to 2012 significantly decreased 66% from 702 to 238 [p=.0004]. From 1996 to 2012, the number of incident cases significantly decreased 42% from 632 to 368 [p=.0003].For every 1% increase in the number of individuals with highest annual viral load <500 c/mL, incidence decreased by 1% (rate ratio: 0.99, CI 0.989,0.991].From 1996 to 2012, the number of reported HIV-positive IDU decreased 92% [p.0013] and the number of reported HIV-positive MSM decreased 22% [p.0046].

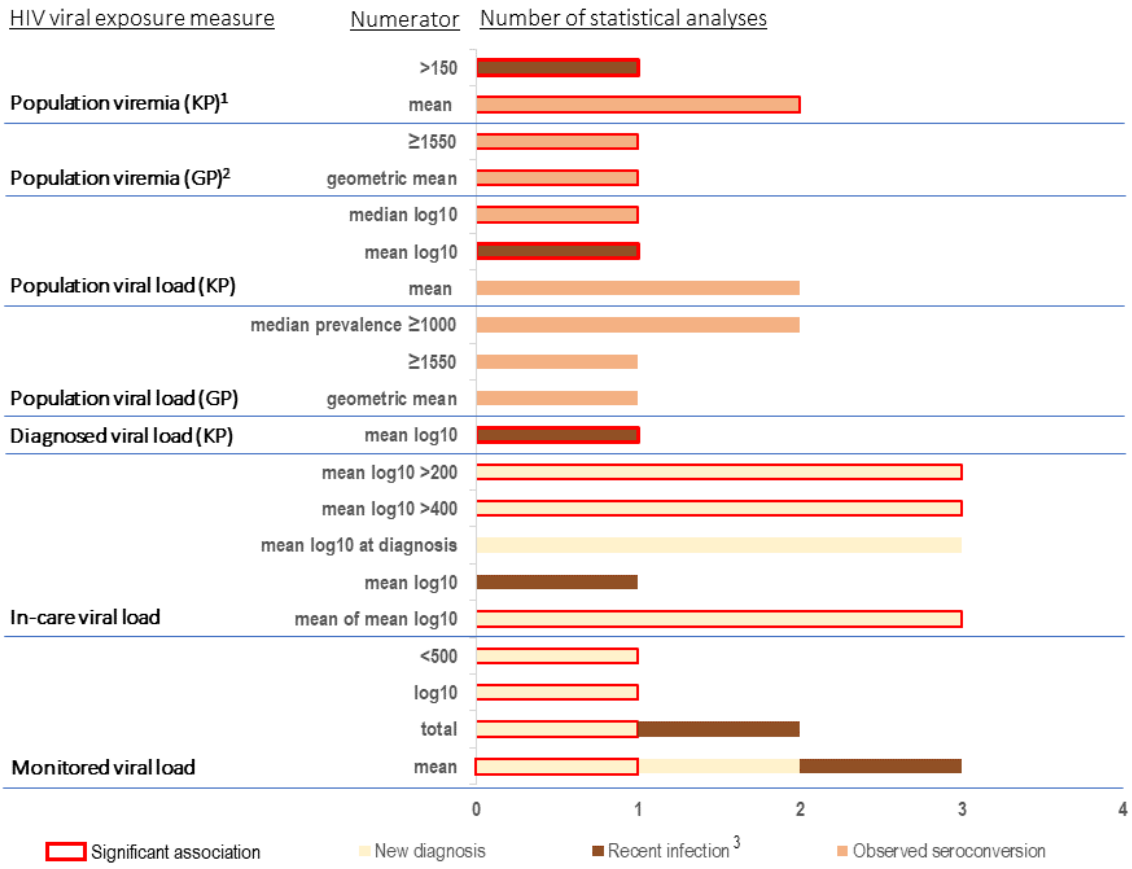
Study	Population type	Sampling strategy	Quality score: Selection	Quality score: Comparability	Quality score: Outcome	Key measures of correlation/association
Castel 2012 [24]	Low prevalence	Serial, cross-sectional cohorts via case surveillance registry	1	0 **Total score**: 2	1	Among reported HIV-positive cases, from 2004 to 2008, the annual proportion of reported HIV-positive persons with undetectable viral load (i.e. below assay's limit of detection) at their most recent viral load test increased from 15.4% to 57.7%. The annual mean viral load of the most recent viral load test significantly decreased [p=.0001]. In adjusted analysis, the decrease in mean annual viral load at the most recent viral load test was not associated with the number of annual new diagnoses [p=.11].
Das 2010 [5]	Low prevalence	Serial, cross-sectional cohorts via case surveillance registry	1	0 **Total score**: 2	1	Among reported HIV-positive cases, from 2004 to 2008, annual total viral load of the most recent viral load test significantly decreased [p=.021]. Annual mean viral load of the most recent test significantly decreased [p=.037]. During the same period, the number of new diagnoses significantly decreased from 798 to 434 [p<.005]. From 2006 to 2008, there was a non-significant reduction in HIV incidence from 935 [CI: 656, 1212] to 621 [CI: 462, 781], [p=0.29] The decrease in new diagnoses was significantly associated with reductions in mean [p=.003] and total [p=.002]viral load. Reductions in incidence were not associated with mean and total viral load [p>0.30].
Bolijn 2017 [13]	Low prevalence	Prospective, population-based open cohort via case surveillance registry	1	0 **Total score**: 2	1	Mean log viral load among reported HIV-positive persons in HIV care was significantly associated with the number of new diagnoses two [IRR 1.27, CI 1.11-1.46, p<.001], three [IRR 1.42, CI 1.19-1.70, p<.001]and four [IRR 1.61, CI 1.25-2.07, p<.001] years later. Mean log viral load at diagnosis among reported HIV-positive persons in HIV care was not associated with the number of new diagnoses one, three and four years later.The proportion with viral load <400 c/mL was significantly associated with number of new diagnoses two [IRR: 1.09, CI 1.03-1.14, p<.01], three [IRR: 1.12, CI 1.05-1.19, p<.001], and four [IRR: 1.15, CI 1.06-1.26, p<.01] years later.The proportion with viral load ≤200 c/mL was significantly associated with number of new diagnoses two [IRR: 0.92, CI 0.88-0.98, p<.01], three [IRR: 0.90, CI 0.85-0.96, p<.01],

Study	Population type	Sampling strategy	Quality score: Selection	Quality score: Comparability	Quality score: Outcome	Key measures of correlation/association
						and four [IRR: 0.88, CI 0.80-0.96, p<.01] years later.For every log unit decrease in mean viral load, the number of new diagnoses decreased 21% two years later. For every 10% increase in proportion of persons with viral load suppression, the number of new diagnoses decreased 8% two years later.
Grabowski 2017 [22]	High prevalence	Prospective, population-based open cohort via household-based enumeration	4	3 **Total score**: 9	2	In adjusted analysis, the median community prevalence of viremia (≥1000 c/mL) was not associated with the risk of seroconversion. Estimates and measures of association were not provided. Graphical display of estimates showed a trend of increasing incidence against increases in median prevalence of viremia.
Solomon 2016 [12]	Key population	Cross-sectional cohort via respondent-driven sampling	4	0 **Total score**: 7	3	Among HIV-positive MSM and IDU in care (i.e. physician visit in prior 6 months), mean log10 viral load was not correlated with incidence [r=0.29, CI -0.10-0.60, p=0.14]. Among HIV-positive MSM and IDU who were previously diagnosed, mean log10 viral load was correlated with HIV incidence [r=0.59, CI 0.27-0.79, p=.001]. Each standard deviation increase in mean log10 viral load was associated with 1.33% [CI 0.42-2.24] increase in HIV incidence. Among HIV-positive MSM and IDU regardless of prior diagnosis, the mean log10 viral load was correlated with HIV incidence [r=0.51, CI 0.16-0.74, p=.007]. Each standard deviation increase in mean log10 viral load was associated with 1.19% [CI 0.27-2.12] increase in HIV incidence. Among MSM and IDU regardless of HIV status, the prevalence of viremia (>150 c/mL) was most strongly correlated with HIV incidence [r=0.81, CI 0.62-0.91, p<.0001]. Each standard deviation increase in prevalence of viremia was associated with 1.51% [CI 0.70-2.32] increase in HIV incidence. The regression model also demonstrated the highest R2 and lowest AIC and BIC.

Study	Population type	Sampling strategy	Quality score: Selection	Quality score: Comparability	Quality score: Outcome	Key measures of correlation/association
Estimating population-level exposure-outcome associations						
He 2013 [11]	Serodiscordant couples	Prospective, population-based cohort via case surveillance registry	3	0 **Total score**: 5	2	In unadjusted analysis, risk of HIV seroconversion was higher in serodiscordant partnerships where the HIV-positive partner's viral load was greater than 400 c/mL. When compared to HIV-positive partners with <400 c/mL, the risk of HIV seroconversion was 12.89 [CI 1.59-104.78, p=.017], 27.29 [CI 2.47-301.00, p=.007]), and 27.49 [CI 3.38-223.39, p=.002] when the HIV-positive partner's viral load was 400-9,999 c/mL, 10,000-49,000 c/mL, and ≥50,000 c/mL, respectively.
Biraro 2013 [20]	Serodiscordant couples	Retrospective, population-based cohort via household-based enumeration	4	1 **Total score**: 6	1	In adjusted analysis, risk of HIV seroconversion in serodiscordant partnerships was not statistically different [p=0.25] across categories of the HIV-positive partner's viral load [<10,000 c/mL (ref), 10,000-49,999 c/mL (HR 2.18, CI 0.19-24.55) ≥50,000 c/mL (HR 5.85, CI 0.68-50.60)].
Quinn 2000 [19]	Serodiscordant couples	Prospective population-based cohort via household-based enumeration	4	1 **Total score**: 8	3	In adjusted analysis, the rate of HIV seroconversion in serodiscordant partnerships demonstrated a dose-response relationship to the HIV-positive partner's viral load. When compared to HIV-positive partners with <3500 c/mL, the risk of transmission was 5.80 [CI 2.26-17.80], 6.91 [CI 2.96-20.15], and 11.87 [CI 5.02-34.88] when the HIV-positive partner's viral load was 3500-9,999 c/mL, 10,000-49,000 c/mL, and ≥50,000 c/mL, respectively. With every log increase in viral load, the risk of transmission increased by a factor of 2.45.
Kiwanuka 2009 [21]	Serodiscordant couples	Prospective population-based cohort via household-based enumeration	4	2 **Total score**: 9	3	In adjusted analysis, the risk of HIV seroconversion in serodiscordant partnerships was significantly associated with the HIV-positive partner's viral load. For every log increase in viral load, the risk of seroconversion increased by a factor of 2.37 [CI (1.75-3.21), p<.001].

Study	Population type	Sampling strategy	Quality score: Selection	Quality score: Comparability	Quality score: Outcome	Key measures of correlation/association
Tanser 2017 [14]	High prevalence	Prospective, population-based open cohort via household-based enumeration	4	2 **Total score**: 8	2	In adjusted analysis, for every unit increase in the geometric mean viral load of the study population (irrespective of HIV status), there was a 9.1% [HR: 1.091, CI 1.045-1.138, p=.001] increased risk of HIV seroconversion. For every 1% increase in the prevalence of detectable viremia in the study population (irrespective of HIV status), there was a 6.3% [HR: 1.063, CI 1.025-1.103, p=.001] increased risk of HIV seroconversion. Analyses were not significant when geometric mean viral load and detectable viremia were measured only among HIV-positive persons.
Wood 2009 [4]	Key population	Prospective observational parallel cohorts via snowball sampling	3	3 **Total score**: 8	2	Median log10 viral load during the six months before each HIV-negative participant's follow-up visit was associated with HIV seroconversion [HR: 3.32, CI 1.82-6.08, p<.001]. When the median log10 concentration fell below 20,000 c/mL, median log10 viral load was no longer significantly associated with HIV seroconversion [HR: 1.7, CI .79-3.67, p=.175].
Glasheen 2018 [26]	Key population	Serial cross-sectional cohorts via targeted sampling	3	0 **Total score**: 5	2	Annual mean viral load among newly diagnosed HIV-positive IDU was not correlated with observed seroconversion in the same year [r=.32, p=.28] or following year [r=.20, p=.527]. Annual prevalence-adjusted mean viral load among both HIV-negative and newly diagnosed HIV-positive IDU was correlated with observed seroconversions in the same year [r=.69, p=.009] and following year [r=.762, p=.004]. In unadjusted analysis, the annual log10 viral load among newly diagnosed HIV-positive IDU was associated with an increase in risk of seroconversion [HR: 1.28, p=.038]. The annual prevalence-adjusted log10 viral load among both HIV-negative and newly diagnosed HIV-positive IDU was associated with an increase in risk of seroconversion [HR: 3.29, p<.001].

***Figure 2.3**. Number of statistical analyses by population HIV viral exposure measure, disaggregated by new HIV infection measure*



[1] KP: key population, [2] GP: general population, [3] STARHS or multi-assay algorithm

Supplements

Supplement One: search strategy

Supplemental Table S2.1*. Search strategies used in PubMed and Embase*

Search	Query	Results
PubMed		
#1	Search ("population viral load" OR "population-level viral load" OR "population level viral load")	24
#2	Search ("population viral suppression" OR "population-level viral suppression" OR "population level viral suppression")	284
#3	Search ("population HIV viral load" OR "population-level HIV viral load" OR "population level HIV viral load")	3998
#4	Search ("population HIV viral suppression" OR "population-level HIV viral suppression" OR "population level HIV viral suppression")	1140
#5	Search (#1 OR #2 OR #3 OR #4)	4513
#6	Search ("community viral load" OR "community-level viral load" OR "community level viral load")	236
#7	Search ("community viral suppression" OR "community-level viral suppression" OR "community level viral suppression")	430
#8	Search ("community HIV viral load" OR "community-level HIV viral load" OR "community level HIV viral load")	977
#9	Search ("community HIV viral suppression" OR "community-level HIV viral suppression" OR "community level HIV viral suppression")	382
#10	Search (#6 OR #7 OR #8 OR #9)	1217
#11	Search (HIV AND ("detectable viremia" OR "uncontrolled viremia" OR "viremia")	5021
#12	Search (#5 OR #10 OR #11)	10,037
#13	Search (HIV AND ("incidence" OR "new infections" OR "recency" OR "new diagnos*"))	34,001
#14	Search ("HIV incidence" OR "new HIV infections" OR "HIV recency" OR "new HIV diagnos*")	16,009
#15	Search (#13 OR #14)	34,547
#16	Search (#12 AND #15)	1146
#17	Search viral load [MeSH Terms]	30,238
#18	Search incidence [MeSH Terms]	239,497
#19	Search (#17 AND #18)	617
#20	Search (#16 OR #19)	1600
Embase		
	('hiv'/exp OR 'hiv') AND ('new diagnos*s' OR 'new infection' OR 'incidence'/exp OR 'incidence') AND ('virus load' OR 'virus transmission' OR 'viremia') NOT ('hbv' OR 'hbc' OR 'hepatitis' OR 'cardiovascular' OR 'diabetes' OR 'ebv' OR 'epstein barr' OR 'myocardial' OR 'tuberculosis' OR 'cervical' OR 'cytomegalovirus' OR 'immune reconstitution' OR 'cancer' OR 'malignancy' OR 'hpv' OR 'pneumonia')	2579

Supplement Two: study quality

Study quality scores were influenced by sampling strategies, which were often dictated by the study population of interest (Supplemental Table S2.2). Nine of the 20 studies were conducted in low HIV prevalence populations and, having been sampled from case surveillance registries, scored low on quality at 2.[5,6,13,15,17,18,23-25] Two studies were conducted in high HIV prevalence settings, scoring high (8 and 9) since they were household-based samples that measured directly observed HIV seroconversions as the study outcome and that statistically induced exposure comparability by adjusting for risk factors.[14,22]

Five studies in key populations ranged in quality score from 2 to 8 [mean: 4.8], due to varied sampling strategies and study designs (Supplemental Table S2.2).[4,12,16,26,27] For example, two studies, each scoring 2, sampled cross-sectional cohorts from case surveillance registries. Another study, scoring 7, used respondent-driven sampling to develop a probability sample and a HIV recency assay to detect recent HIV infections from a cross-sectional sample, thus mitigating the temporal directionality limitation of cross-sectional designs.

The four serodiscordant couples studies, with a mean score of 7 [range: 5–9], were generally higher quality studies (Supplemental Table S2.2).[11,19-21] Most used direct observation of HIV seroconversions and sampled from household-based surveillance sites. These factors better assured representative samples that allowed for exposure comparisons, adjusted for risk factors, within the population.

***Supplemental Table S2.2**. Newman-Ottawa risk bias scores*

	Low prevalence									High prevalence	
	Montaner 2010	Montaner 2014	Lohse 2005	Castel 2012	Golden 2016	Chakraborty 2017	Das 2010	Krentz 2013	Boljin 2017	Grabowski 2017	Tanser 2017
Selection (maximum 4)	1	1	1	1	1	1	1	1	1	4	4
1) Representativeness of the exposed cohort											
a) truly representative of the average population viral load in community										X	X
b) somewhat representative of average population viral load in community											
c) selected group: surveillance registry of reported HIV cases	X	X	X	X	X	X	X	X	X		
d) non-probability sampling method (i.e. snowball or targeted sampling)											
e) no description of the derivation of the cohort											
2) Selection of the non exposed cohort											
a) drawn from the same community as the exposed cohort										X	X
b) drawn from a different source											
c) unexposed comparison cohort not included	X	X	X	X	X	X	X	X	X		
d) no description of the derivation of the non exposed cohort											
3) Ascertainment of exposure											
a) secure record: surveillance registry	X	X	X	X	X	X	X	X	X		
b) direct biomarker testing for HIV viral load										X	X
c) no description											
4) Demonstration that outcome of interest was not present at start of study											
a) yes										X	X
b) no	X	X	X	X	X	X	X	X	X		
Comparability (maximum 3)	0	0	0	0	0	0	0	0	0	3	2
1) Comparability of cohorts on the basis of the design or analysis											
a) study controls for key factor relevant to study population										a	a
b) study controls for other important factors relevant to study population										b	b
c) study controls for factors unique to study population										c	
d) no unexposed comparison cohort for comparability	X	X	X	X	X	X	X	X	X		
e) no control for confounding factors (i.e. correlation analysis)											
Outcome (maximum 3)	1	1	1	1	1	1	1	1	1	2	2
1) Assessment of outcome											
a) independent blind assessment											
b) linked record: surveillance registry	X	X	X	X	X	X	X	X	X		
c) direct biomarker testing for HIV status										X	X
d) no description											
2) Was follow-up long enough for outcomes to occur											
a) yes										X	X
b) no											
c) study design did not include follow up period, limiting interpretation	X	X	X	X	X	X	X	X	X		
3) Adequacy of follow up of cohorts											
a) complete follow up - all subjects accounted for											
b) loss is small, based on description is unlikely to introduce bias											
c) follow up unnecessary due to outcome (i.e. recency assay for incidence)											
d) follow up rate inadequate										X	
e) study design did not include follow-up, limits temporal interpretation	X	X	X	X	X	X	X	X	X		
f) no statement											X
Total score (maximum 10)	2	2	2	2	2	2	2	2	2	9	8

a: demographics; b: sexual behaviors; c: extramarital relationships; d: sharing syringes, other drug use; e: frequency of sex; f: circumcision status; g: condom use
Note: cells with hash pattern indicate point value of 1; cells without a pattern indicate point value of 0

Supplemental Table S2.2 Newman-Ottawa risk bias scores (continued)

	Key populations					Serodiscordant couples			
	Cowan 2012	Wood 2009	Glasheen 2018	Solomon 2016	Brantley 2017	Biraro 2013	Quinn 2000	Kiwanuka 2009	He 2013
Selection (maximum 4)	1	3	3	4	1	4	4	4	3
1) Representativeness of the exposed cohort									
a) truly representative of the average population viral load in community									
b) somewhat representative of average population viral load in community				X		X	X	X	
c) selected group: surveillance registry of reported HIV cases	X				X				X
d) non-probability sampling method (i.e. snowball or targeted sampling)		X	X						
e) no description of the derivation of the cohort									
2) Selection of the non exposed cohort									
a) drawn from the same community as the exposed cohort		X	X	X		X	X	X	X
b) drawn from a different source									
c) unexposed comparison cohort not included	X				X				
d) no description of the derivation of the non exposed cohort									
3) Ascertainment of exposure									
a) secure record: surveillance registry	X	X			X				
b) direct biomarker testing for HIV viral load			X	X		X	X	X	X
c) no description									
4) Demonstration that outcome of interest was not present at start of study									
a) yes		X	X	X		X	X	X	X
b) no	X				X				
Comparability (maximum 3)	0	3	0	0	0	1	1	2	0
1) Comparability of cohorts on the basis of the design or analysis									
a) study controls for key factor relevant to study population		a				e		e	
b) study controls for other important factors relevant to study population		b						g	
c) study controls for factors unique to study population		d					f		
d) no unexposed comparison cohort for comparability	X				X				
e) no control for confounding factors (i.e. correlation analysis)									
Outcome (maximum 3)	1	2	2	3	1	1	3	3	2
1) Assessment of outcome									
a) independent blind assessment									
b) linked record: surveillance registry	X				X				
c) direct biomarker testing for HIV status		X	X	X		X	X	X	X
d) no description									
2) Was follow-up long enough for outcomes to occur									
a) yes		X	X	X			X	X	X
b) no						X			
c) study design did not include follow up period, limiting interpretation	X				X				
3) Adequacy of follow up of cohorts									
a) complete follow up - all subjects accounted for				X			X	X	
b) loss is small, based on description is unlikely to introduce bias									
c) follow up unnecessary due to outcome (i.e. recency assay for incidence)									
d) follow up rate inadequate						X			X
e) study design did not include follow-up, limits temporal interpretation	X		X		X				
f) no statement		X							
Total score (maximum 10)	2	8	5	7	2	6	8	9	5

a: demographics; b: sexual behaviors; c: extramarital relationships; d: sharing syringes, other drug use; e: frequency of sex; f: circumcision status; g: condom use
Note: cells with hash pattern indicate point value of 1; cells without a pattern indicate point value of 0

References

1. Centers for Disease Control and Preention. Guidance on Community Viral Load: A Family of Measures, Definitions, and Method for Calculation. Atlanta: CDC; 2011.

2. Millett GA, Peterson JL, Wolitski RJ, Stall R. Greater risk for HIV infection of black men who have sex with men: a critical literature review. Am J Public Health 2006;96:1007-19.

3. Stall R. What's driving the U.S. epidemic in men who have sex with men. Oral abstract presented at: 15th Conference on Retroviruses and Opportunistic Infections; 2008 Feb 3-6; Boston, United States.

4. Wood E, Kerr T, Marshall BD, et al. Longitudinal community plasma HIV-1 RNA concentrations and incidence of HIV-1 among injecting drug users: prospective cohort study. BMJ 2009;338:b1649.

5. Das M, Chu PL, Santos GM, et al. Decreases in community viral load are accompanied by reductions in new HIV infections in San Francisco. PLoS One 2010;5:e11068.

6. Montaner JS, Lima VD, Barrios R, et al. Association of highly active antiretroviral therapy coverage, population viral load, and yearly new HIV diagnoses in British Columbia, Canada: a population-based study. Lancet 2010;376:532-9.

7. Cohen MS, Chen YQ, McCauley M, et al. Prevention of HIV-1 infection with early antiretroviral therapy. N Engl J Med 2011;365:493-505.

8. Moher D, Liberati A, Tetzlaff J, Altman DG, et al. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. PLoS Med 2009;6:e1000097.

9. The Newcastle-Ottawa Scale (NOS) for assessing the quality of nonrandomised studies in meta-analyses [Internet]. [cited 15 Jun 2019]. Available from: http://www.ohri.ca/programs/clinical_epidemiology/oxford.asp.)

10. Higgins JP, Altman DG, Gotzsche PC, et al. The Cochrane collaboration's tool for assessing risk of bias in randomised trials. BMJ 2011;343:d5928.

11. He N, Duan S, Ding Y, et al. Antiretroviral therapy reduces HIV transmission in discordant couples in rural Yunnan, China. PLoS One 2013;8:e77981.

12. Solomon SS, Mehta SH, McFall AM, et al. Community viral load, antiretroviral therapy coverage, and HIV incidence in India: a cross-sectional, comparative study. Lancet HIV 2016;3:e183-90.

13. Bolijn R, Op de Coul ELM, van Sighem A, et al. Comparing viral load metrics and evaluating their use for HIV surveillance. J Infect 2017;75:169-78.

14. Tanser F, Vandormael A, Cuadros D, et al. Effect of population viral load on prospective HIV incidence in a hyperendemic rural African community. Sci Transl Med 2017;9: eaam8012.

15. Lohse N, Hansen ABE, Jensen-Fangel S, et al. Demographics of HIV-1 infection in Denmark: Results from the Danish HIV cohort study. Scand J Infect Dis 2005;37:338-43.

16. Cowan SA, Gerstoft J, Haff J, Christiansen AH, Nielsen J, Obel N. Stable incidence of HIV diagnoses among Danish MSM despite increased engagement in unsafe sex. J Acquir Immune Defic Syndr 2012;61:106-11.

17. Montaner JSG, Lima VD, Harrigan PR, et al. Expansion of HAART coverage Is associated with sustained decreases in HIV/AIDS morbidity, mortality and HIV transmission: the "HIV treatment as prevention'' experience in a Canadian setting. PLoS One 2014;9.

18. Krentz HB, Gill MJ. The effect of churn on "community viral load" in a well-defined regional population. J Acquir Immune Defic Syndr 2013;64:190-6.

19. Quinn TC, Wawer MJ, Sewankambo N, et al. Viral load and heterosexual transmission of human immunodeficiency virus type 1. Rakai project study group. N Engl J Med 2000;342:921-9.

20. Biraro S, Ruzagira E, Kamali A, Whitworth J, Grosskurth H, Weiss HA. HIV-1 transmission within marriage in rural Uganda: a longitudinal study. PLoS One 2013;8:e55060.

21. Kiwanuka N, Laeyendecker O, Quinn TC, et al. HIV-1 subtypes and differences in heterosexual HIV transmission among HIV-discordant couples in Rakai, Uganda. J Acquir Immune Defic Syndr 2009;23:2479-84.

22. Grabowski MK, Serwadda DM, Gray RH, et al. HIV prevention efforts and incidence of HIV in Uganda. N Engl J Med 2017;377:2154-66.

23. Golden MR, Bennett AB, Dombrowski JC, Buskin SE. Achieving the goals of the national HIV/AIDS strategy: declining HIV diagnoses, improving clinical outcomes, and diminishing racial/ethnic disparities in King County, WA (2004-2013). Sex Transm Infect 2016;43:269-76.

24. Castel AD, Befus M, Willis S, et al. Use of the community viral load as a population-based biomarker of HIV burden. AIDS 2012;26:345-53.

25. Chakraborty H, Weissman S, Duffus WA, et al. HIV community viral load trends in South Carolina. Int J STD AIDS 2017;28:265-76.

26. Glasheen C, Johnson EO, Lorvick J, Kral AH. Measures of human immunodeficiency virus (HIV) community viral load and HIV incidence among people who inject drugs. Ann Epidemiol 2018;28:8-12.

27. Brantley M, Schumacher C, Fields EL, et al. The network structure of sex partner meeting places reported by HIV-infected MSM: Opportunities for HIV targeted control. Soc Sci Med 2017;182:20-9.

28. Rothman KJ, Greenland S, Lash TL. Modern epidemiology. 3rd ed. Philadelphia: Wolters Kluwer Health/Lippincott Williams & Wilkins; 2008.

29. Miller WC, Powers KA, Smith MK, Cohen MS. Community viral load as a measure for assessment of HIV treatment as prevention. Lancet Infect Dis 2013;13:459-64.

30. Kelley CF, Rosenberg ES, O'Hara BM, et al. Measuring population transmission risk for HIV: An alternative metric of exposure risk in men who have sex with men (MSM) in the US. PLoS One 2012;7.

31. Eisinger RW, Dieffenbach CW, Fauci AS. HIV viral load and transmissibility of HIV infection: Undetectable equals untransmittable. JAMA 2019;321:451-2.

32. WHO. Consolidated Guidelines on the Use of Antiretroviral Drugs for Treating and Preventing HIV Infection: Recommendations for a Public Health Approach. Geneva: World Health Organization; 2013.

33. WHO. Consolidated Guidelines on the Use of Antiretroviral Drugs for Treating and Preventing HIV Infection: Recommendations for a Public Health Approach - Second Edition. Geneva: World Health Organization; 2016.

34. Havlir DV, Balzer LB, Charlebois ED, et al. HIV testing and treatment with the use of a community health approach in rural Africa. N Engl J Med 2019;381:219-29.

35. Hayes RJ, Donnell D, Floyd S, et al. Effect of universal testing and treatment on HIV incidence - HPTN 071 (PopART). N Engl J Med 2019;381:207-18.

36. Makhema J, Wirth KE, Pretorius Holme M, et al. Universal testing, expanded treatment, and incidence of HIV infection in Botswana. N Engl J Med 2019;381:230-42.

37. Abdool Karim SS. HIV-1 epidemic control - insights from test-and-treat trials. N Engl J Med 2019;381:286-8.

38. Olawore O, Tobian AAR, Kagaayi J, et al. Migration and risk of HIV acquisition in Rakai, Uganda: a population-based cohort study. Lancet HIV 2018;5:e181-e9.

39. Baral S, Rao A, Sullivan P, et al. The disconnect between individual-level and population-level HIV prevention benefits of antiretroviral treatment. Lancet HIV 2019;6:e632-e8.

40. Thomas JC, Tucker MJ. The development and use of the concept of a sexually transmitted disease core. J Infect Dis 1996;174 Suppl 2:S134-43.

41. Fichtenberg CM, Ellen JM. Editorial response: moving from core groups to risk spaces. Sex Transm Infect 2003;30:825-6.

42. Skaathun B, Khanna AS, Morgan E, Friedman SR, Schneider JA. Network viral load: a critical metric for HIV elimination. J Acquir Immune Defic Syndr 2018;77:167-74.

43. Rhodes T, Simic M. Transition and the HIV risk environment. BMJ 2005;331:220-3.

44. Tanser F, de Oliveira T, Maheu-Giroux M, Baernighausen T. Concentrated HIV subepidemics in generalized epidemic settings. Curr Opin HIV AIDS 2014;9:115-25.

45. Tanser F, Barnighausen T, Dobra A, Sartorius B. Identifying 'corridors of HIV transmission' in a severely affected rural South African population: a case for a shift toward targeted prevention strategies. Int J Epidemiol 2018;47:537-49.

46. Cuadros DFG T, de Oliveira T, Bärnighausen T, Tanser F. Assessing the role of geographical HIV hot-spots in the spread of the epidemic. Oral abstract presented at: 25th Conference on Retroviruses and Opportunistic Infections; 2018 Mar 4-7; Boston, United States.

Chapter Three. Small area HIV risk profiles and population-level HIV viral measures: a national-level, cross-sectional analysis in Eswatini

Summary

Background: HIV core areas, places with high levels of uncontrolled HIV infection, may promote ongoing HIV transmission and sustain micro-epidemics. While individual uptake of HIV treatment is known to reduce aggregate population HIV viral measures, environmental factors that produce core areas is unclear. Understanding the contextual factors that promote core areas is needed.

Methods: Using a two-stage cluster-based design, a nationally representative, household-based sample of adults, ages 18-49 years was enrolled from December 2010 to June 2011 in Eswatini. Consenting adults completed an interview and received home-based rapid HIV testing and counseling. All seropositive samples were tested for HIV viral load using the COBAS AmpliPrep/Taqman HIV-1 Test, v 2.0. Multi-level latent class modeling was performed to identify statistically significant combinations of HIV risk factors and classify the combinations into small area risk profiles, categorized by HIV risk types. Linear regression with robust standard errors was conducted to assess the correlation between area profiles and area prevalence of detectable virus (≥20 copies/milliliter) among only HIV-positive adults (viral load) and among all adults, regardless of HIV status (viremia).

Results: A total of 18,172 adults were surveyed of which about half [54%] were female and over a third [36%] was ages 18 to 24 years. All adults were categorized as one of six HIV risk types, each showing a unique pattern of multiple risk factors that conveyed their propensity for HIV transmission and/or acquisition risk. The three most frequent composite prevalences of HIV risk types was categorized into area profiles: low-moderate acquisition, moderate acquisition/transmission, and high acquisition/transmission profiles. The prevalence of detectable viremia progressively increased from low-moderate acquisition, moderate acquisition/transmission, and high acquisition/transmission profiles [17.7%, 25.4%, and 35.1%, respectively]. On average, the prevalence of detectable viremia was 7.4% [p<.001] and 17.1% [p<.001] higher, respectively, when compared with low-moderate acquisition areas. The prevalence of detectable viral load significantly decreased from low-moderate acquisition to moderate acquisition/transmission areas [76.6% versus 68.7%, p<.001], and was significantly higher in high acquisition/transmission areas by 7.3% [p<.001], when compared with low-moderate acquisition areas. The high acquisition/transmission areas comprised the largest proportions of the highest transmission/acquisition risk types.

Conclusion: An area typology of HIV risk can explain variation in area-level HIV viral measures. The co-location of higher transmission and acquisition risk types in larger frequencies within small areas may enable elevated levels of uncontrolled HIV viremia with the potential to progress into geographic sources of ongoing HIV transmission.

Introduction

Core theory has been a fundamental tenet of sexually transmitted infection (STI) epidemiology for decades, positing that a disproportionate number of infections, when transmitted through sexual networks, results from a small sub-group of infected network members.[1] This core group of members is widely accepted for its strategic importance in STI transmission, and its role in driving and sustaining epidemics. Intervening with this small sub-group is presumed to produce greater decreases in community-level infections than compared to intervening with those less often affected by STIs.[2,3]

Members of the STI core are often difficult to identify and show changing membership over time. Yet the groups themselves remain geographically stable. The consistency in location of cores, coupled with its fluid membership, implies that places could be more important than people when defining the core. [2] Such "core areas" would involve geographies with high concentrations of infection; and core group members would entail all area residents, regardless of infection status, rather than a sub-group of infected members.[4]

In HIV epidemiology, geographic areas with high levels of HIV viremia are likely sources of ongoing HIV transmission,[5,6] and therefore, a spatially derived core would comprise geographies with high concentrations of uncontrolled HIV infection. While individual-level uptake of HIV antiretroviral treatment (ART) is a well-established predictor of population levels of HIV viral load,[7-9] the evidence is less clear on contextual predictors of HIV core areas, or places with high area HIV viral quantities. Individual factors only partially account for risk behavior variation in populations, and contextual domains, such as availability of services, social norms, and the physical environment, may more comprehensively explain the factors shaping individual-level risk behaviors and their population-level outcomes.[10] An analytic approach inclusive of environmental influences promoting HIV risk behaviors and rendering environments susceptible to HIV is warranted. A risk environment approach is one proposed framework.[11,12]

A risk environment is a socially or physically defined space where several factors interact to increase the likelihood of health-related harm. Factors manifest according to the type of environment (social, policy, economic, and physical) and the level of influence (macro and micro).[11] HIV core areas may accordingly be viewed as geographically defined environments that enable HIV related risk behaviors resulting in geographically concentrated, high area viral quantities. As HIV core areas may operate as a

population mechanism for new HIV infections, understanding the factors that produce core areas is necessary to design interventions that alter risk environments and maximize population-level HIV risk reduction.

In 2011, Eswatini, formerly known as Swaziland, experienced the highest national HIV prevalence (32%) and incidence (2.4%) among adults ages 18-49 years.[13,14] HIV and viral suppression prevalences appeared uniform across regions in the country[15] but may have masked sub-regional variation in HIV viral quantities, thus hindering the detection of potential core areas. This analysis applied a risk environment approach using multi-level latent class modeling to categorize the density of individual-level HIV risk factors into small area, geographic risk profiles. The area risk profiles were then evaluated for a relationship with area HIV viral quantities to infer the potential for predicting HIV core areas. Analyses were hypothesized to show that environments with higher numbers of adults practicing multiple high-risk behaviors would lead to HIV core areas characterized by high levels of uncontrolled HIV infection.

Methods

Study population

The Swaziland HIV Incidence Measurement Survey (SHIMS) is a household-based, nationally representative sample of adults, ages 18-49 years. Using a two-stage, cluster-based design, 14,891 randomly selected households from 575 enumeration areas (i.e., census tracts) were approached from December 2010 to June 2011; 12,571 [94.3%, n=13,335] households participated, and 24,462 potential participants were identified. Persons who resided or had slept in the household the prior night and who were willing to receive home-based HIV testing, answer questions in siSwati or English, and provide written informed consent were eligible for participation. A total of 18,172 [74%, n=24,462] eligible adults consented to participate. The survey was approved by ethical review boards at Columbia University Irving Medical Center, the Eswatini Ministry of Health, and the US Centers for Disease Control and Prevention.

Study procedures

Selected households were approached by study staff trained in Good Clinical Practice and survey procedures.[16] Field staff conducted informed consent for an interview with HIV testing. Consenting participants completed a paper-based, face-to-face interview about sexual behaviors, HIV service uptake, and demographics in a private location around or within the home.

Trained staff performed venipuncture and conducted home-based HIV testing and counseling. Rapid HIV testing was conducted with Determine HIV-1/2 Ag/Ab Combo (Alere, Japan). Determine-reactive samples were confirmed with Uni-Gold HIV Test (Trinity Biotech, Ireland). Laboratory-based confirmatory testing and discrepancy resolution were done according to a survey-specific test algorithm (Supplemental Figure S3.1). HIV-seropositive participants were provided referrals to HIV care. All seropositive samples were tested for plasma viral load using the COBAS AmpliPrep/TaqMan HIV-1 Test, v2.0 automated platform (Roche, USA), according to manufacturers' instructions.

Study measures

Outcome

Area HIV viral quantity is an observed, macro-level, continuous outcome describing the amount of HIV within a study enumeration area (EA). Two definitions of HIV viral quantity were assessed. First, a viremia quantity was measured as the prevalence of detectable virus (≥20 c/mL) among all adults within an EA, regardless of HIV status; and second, a viral load quantity was measured as the prevalence of detectable virus (≥20 c/mL) among only HIV-positive adults within an EA.

Exposure

Area profile, the exposure of interest, is an unobservable HIV risk profile of the EA. It is a latent, macro-level variable summarizing the area environment of HIV transmission and acquisition risks affecting an EA-specific population. It is a nominal variable defined by the composite mean prevalences of adult HIV risk types across a sub-set of similar EA populations.

Adult HIV risk type indicators

HIV risk type is an unobservable adult HIV risk profile that describes a person's propensity for HIV transmission and/or acquisition. It is a latent, micro-level variable summarizing a set of indicators, specifically five micro-level observed HIV risk factors: 1) HIV and ART status (HIV-seronegative; HIV-seropositive, reported on ART; HIV-seropositive, reported not on ART; HIV-seropositive, reported no prior diagnosis); 2) lifetime HIV testing frequency (tested, three or more times; tested, one or two times; never tested), 3) number of sexual partners in the prior six months (no partners; one partner; two or more partners), 4) knowledge of those partners' HIV status (all HIV-negative; at least one HIV-positive; at least one with unknown status), and 5) condom use in the prior six months (always; sometimes; never).

Covariates

HIV risk type was disaggregated by two micro-level demographic variables: gender (female or male) and age group (18-24 years; 25-34 years; 35-49 years). When assessing the association between area HIV viral quantity and area profile, four covariates were included in the model as potential confounders: area gender ratio (expressed as a percentage), area mean age of all participants in an EA, area marital status, defined as the prevalence of persons who were married and living with a partner in the EA, and area residence, defined as either urban or rural.

Statistical analysis

Multi-level latent class modeling using maximum likelihood estimation was conducted to perform a micro-macro analysis. Model assessment and final selection was based upon a combination of statistical criteria, parsimony, and model interpretability (Supplement Two).

First, the five observed micro-level risk factor indicators were regressed on HIV risk type, a latent micro-level variable, to identify similar combinations of HIV risk factors and group them into risk type categories. Each indicator was tested on whether it significantly differentiated the risk type categories, using the Wald test statistic and assuming a chi-square distribution. After HIV risk type was identified, it was stratified by micro-level gender and age to qualitatively assess the distribution of risk types across gender and age groups. This assessment helped to validate the latent class structure.

Once the micro-level HIV risk types were identified, they were used as indicators to identify the latent macro-level area profiles. A macro-level latent variable model was estimated to identify similar combinations of micro-level HIV risk types and group them into area profile categories. Each area profile category represents a homogenous, qualitatively unique group of EAs in this study population with similar distributions of HIV risk types. Using the Wald test statistic with an assumed chi-square distribution, each area profile was tested on whether it significantly differed in proportion of each HIV risk type.

Lastly, linear regression analysis with robust standard error estimation determined whether area profiles could explain differences in the observed area HIV viral quantity. Analyses were weighted, including adjustments for design weights, nonresponse, and post-stratification adjustments. Descriptive statistics were estimated by calculating a conditional probability table of the indicator categories of the latent micro-level HIV risk type and latent macro-level area profile. The analyses described the prevalence of each

indicator across each HIV risk type and the prevalence of each HIV risk type across each area profile. Latent class modeling was conducted using Latent Gold version 5.1, and linear regression was conducted using STATA version 16.

Results

Participant characteristics

A total of 18,172 adults, ages 18 to 49 years, were surveyed (Table 3.1). Overall, 36% of the population was ages 18 to 24 years, with the majority [64%] having completed secondary education or higher and living in a rural area [70%]. A minority [7%] had never attended school. While most [45%] adults were unmarried and sexually active, the proportion of males reported as having never had sex was nearly threefold greater than among females [17% versus 6%]. Females were almost twice as likely as males [21% versus 11%] to be married with their partner living elsewhere.

Almost a third [32%] of adults tested HIV-seropositive, with HIV prevalence higher among females than males [39% versus 24%] (Table 3.1). Nearly equal proportions of HIV-positive adults reported either as diagnosed and currently on ART [10%], diagnosed and not on ART [9%], or having no prior diagnosis [12%]. Diagnosis of HIV infection was more than twice as common among females than males [26% versus 12%]. In fact, the probability among females of reporting three or more HIV tests in their lifetime was double that of males [42% versus 21%], and close to half [45%] of males had never tested for HIV, compared with 16% of females.

In the prior six months, males were notably more likely than females [18% versus 4%] to report multiple partners but also to not having any partners [27% versus 18%] (Table 3.1). A quarter [25%] of adults reported having at least one partner with unknown HIV status. Less than a third [32%] reported always using a condom, with more females than males [37% versus 30%] having never used a condom.

HIV risk types

Latent class modeling identified six HIV risk types. The probabilities of each risk factor, conditioned on risk type membership, and the gender and age distribution across risk types is presented in Table 3.2. Each risk type was examined for patterns in risk factor probabilities, labelled to convey the potential propensity for HIV transmission and/or acquisition risk implied by the pattern, and organized along a qualitative sequence of risk to facilitate interpretation.

Acquisition types – low, moderate, and high – were adults who tested HIV-seronegative and comprised nearly two-thirds [65%] of reproductive age adults (Table 3.2):

- **Low acquisition adults** were typically HIV-negative persons [100%] who were not sexually active in the prior six months [73%]. Among those who were sexually active [27%], all were monogamous and reported only HIV-negative partners [100%]. They did not demonstrate high risk behaviors like multiple partners and partners with unknown HIV status. Low acquisition risk type adults were generally young males [64% ages 18 to 24 years and 67% male].
- **Moderate acquisition adults** were sexually active [87%], monogamous [78%], and HIV-negative [100%]. They were more likely than low acquisition adults to have a partner with unknown HIV status [16% versus 0%] and had generally never used a condom in the prior six months [46%]. All [100%] moderate acquisition adults had tested for HIV at least once in their lifetime, with the majority [60%] having tested at least three times. Moderate acquisition adults were mostly female [64%].
- **High acquisition adults** were sexually active [97%] and, when compared to all risk types, demonstrated the highest likelihood of multiple partners [30%] and nearly the highest likelihood of having a partner with unknown HIV status [36%]. All [100%] were HIV-negative who typically never tested for HIV [74%]. They were male [70%] and generally ages 18-24 years [44%].

The very high, high, and moderate transmission types predominantly tested HIV-seropositive. Very high transmission adults were the least common type at 6% (Table 3.2):

- **Very high transmission adults** were sexually active [93%] and, compared to all risk types, demonstrated the highest likelihood of having a partner with unknown HIV status [37%] and the second highest likelihood of multiple partners [16%]. All [100%] were HIV-positive adults with no prior diagnosis and typically had never tested for HIV [83%] – the highest likelihood of no prior test history among all risk types. Very high transmission adults were most likely among the three transmission risk types to have never used a condom in the prior six months [43%]. Very high transmission adults were generally male [57%] and ages 25 to 34 years [43%].

- **High transmission adults** were sexually active [90%], HIV-positive [100%], and not on ART [89%]. They generally had all HIV-negative partners [47%] or at least one partner with unknown HIV status [34%], along with inconsistent use of condoms [73%].
- **Moderate transmission adults** were mainly monogamous [75%] or not sexually active [20%]. Most were HIV-positive [77%] and generally on ART [55%]. Moderate transmission adults were the most likely of all risk types to report always using a condom in the prior six months [69%]. They typically were female [67%] and older [43% ages 35 to 49 years].

Area profiles

Latent class modeling identified three area profiles, and the prevalence of each HIV risk type is presented by area profile in Figure 3.1. Each area profile was reviewed for meaningful differences in prevalence patterns and labelled to describe the potential area-level environment of HIV transmission and acquisition risk according to HIV risk type. Low-moderate acquisition areas were predominated [58%] by low acquisition [23%] and moderate acquisition [35%] risk type adults. Moderate acquisition/transmission areas were mostly [55%] populated by moderate acquisition [33%] and moderate transmission [22%] adults. And, high acquisition/transmission areas had the highest proportions of high acquisition [18%], very high transmission [11%], and high transmission [19%] adults. The prevalences of low-moderate acquisition areas and moderate acquisition/transmission areas were equivalent [41%]. High acquisition/transmission areas represented nearly a fifth [19%] of all EAs in Eswatini.

When comparing low-moderate acquisition, moderate acquisition/transmission, and high acquisition/ transmission areas, a progressive decrease occurred in the proportion of low acquisition [23%, 18%, and 10%, respectively] and moderate acquisition [35%, 33%, and 28%, respectively] adults (Figure 3.1). This pattern changed among high acquisition adults, with an initial decrease from low-moderate acquisition to moderate acquisition/transmission areas [15% versus 10%] followed by an 80% increase in high acquisition/transmission areas [18%]. Moderate transmission adults were notably more prevalent in moderate acquisition/transmission areas [22%] compared with low-moderate acquisition [11%] and high acquisition/transmission areas [13%].

Area HIV viral quantity by area profile

The prevalence of detectable viremia progressively increased from low-moderate acquisition, moderate acquisition/transmission, and high acquisition/transmission profiles [17.7%, 25.4%, and 35.1%, respectively] (Table 3.3). Detectable viremia was significantly higher in moderate acquisition/transmission and high acquisition/transmission areas. On average, the prevalence of detectable viremia was 7.4% [p<.001] and 17.1% [p<.001] higher, respectively, when compared with low-moderate acquisition areas.

Viral load quantities demonstrated a different pattern (Table 3.3). The prevalence of detectable viral load significantly decreased from low-moderate acquisition to moderate acquisition/transmission areas [76.6% versus 68.7%, p<.001]. Detectable viral load then increased among high acquisition/transmission areas to the highest proportion [83.3%] across area profiles. When compared with low-moderate acquisition areas, the prevalence of detectable viral load was significantly higher in high acquisition/transmission areas by 7.3% [p<.001]. Area geometric mean and high viral quantities were also assessed and showed similar results (Supplemental Table S3.10).

Risk factor distribution across area profiles

The prevalence of HIV-positive adults was highest in high acquisition/transmission areas at 40% (Figure 3.2a). The high acquisition/transmission profile, however, had the lowest proportions of HIV-positive adults who engaged in HIV-related services. Only half [50%, 20/40] of HIV-positive adults were previously diagnosed, and less than a quarter [23%, 9/40] of HIV-positive adults were on ART. Moderate acquisition/transmission areas, showing a 12% decrease in prevalence of HIV-positive adults compared with high acquisition/transmission areas [35% versus 40%], had the highest uptake of HIV-related services among HIV-positive adults. Two-thirds [69%, 24/35] of HIV-positive adults were previously diagnosed, and 40% [14/35] were on ART.

The moderate acquisition/transmission profile also demonstrated greater uptake of HIV preventive behaviors, including the lowest prevalence of having two or more partners [9% versus 10% in low-moderate areas and 13% in high acquisition/transmission areas] and the highest prevalence of having always used a condom [31% versus 26% in low-moderate areas and 29% in high acquisition/transmission areas] (Figure 3.2c and 3.2e).

The low-moderate acquisition profile had the lowest prevalence of HIV-positive adults [23%] (Figure 3.2a). It also showed the highest proportion of having never tested for HIV [32%], having all HIV-negative partners [67%], and having never used a condom [40%] (Figures 3.2b, 3.2d, and 3.2e).

Discussion

This study provides evidence that an area typology of HIV risk can explain variation in population-level HIV viral measures. High acquisition/transmission areas – locations with the largest proportions of high acquisition, very high transmission, and high transmission risk type adults – were more likely to have the highest prevalences of HIV viral measures in Eswatini. The high acquisition/transmission profile may consequently signal geographic risk environments that enable high risk behaviors and produce potential HIV core areas. This nationally representative finding is supported through a large, household-based sample with high participation rates and rigorous biomarker testing.

The correlation between population-level HIV viral measures and area prevalences of HIV risk types is likely due to the clustering of adults with a specific combination of factors affecting higher transmission and acquisition risks. High acquisition and very high transmission adults, apart from the four other risk types in Eswatini, had the highest likelihoods of multiple partnerships, of partners with unknown HIV status, and, among risk types characterized as sexually active, of having never tested for HIV. High transmission adults (i.e., sexually active HIV-positive persons) had the second highest likelihood of not being on ART, enabling high transmission risk. These three highest risk types combined in the largest concentrations in high acquisition/transmission areas and illustrate two possible mechanisms affecting area HIV viral quantities: interdependence of risk factors and crossover of highest risk types.

Using multi-level latent class analysis, the most prevalent, statistically significant combinations of risk type clusters were identified; and the high acquisition/transmission area profile showed the largest population sub-groups of highest transmission and acquisition risks. Given the geographic stability of core areas,[2,3] this sub-group combination may imply an interdependence of risk types. The co-location of the largest concentrations of high acquisition and high transmission adults may have fueled HIV prevalence over time; and this contextual effect of prevalent HIV risk types may explain the durability of HIV prevalence “hotspots” in endemic settings.[17]

Risk factor interdependence also manifested at the individual level. HIV risk types represented statistically significant combinations of risk factors. The highest risk types (i.e., high acquisition and very high transmission) denoted adults who were more likely to be sexually active with multiple partners, inconsistent condom use, and a partner of unknown HIV status. They also were very unlikely to test for HIV. This risk combination amplified the chances not only of incident HIV infection but also of undiagnosed infection promoting HIV transmission.[18-24]

The "crossover" between highest risk types may also promote area HIV viral quantities. High acquisition and very high transmission types are analogous profiles, mainly distinguished by HIV status. Their similarity may indicate an invisible shift in membership between the highest risk population sub-groups simply upon HIV acquisition. This transition between high risk types may not only support the geographic stability of core membership[2,3] but also amplify area HIV viral quantities towards formation of HIV core areas, given the higher likelihood of undiagnosed infection and HIV transmission among these adults. This intra-individual transition could suggest a risk type-specific interface to potentially exploit through interventions.

This study also provides evidence that clustering of HIV risk types that utilize HIV services is associated with lower area viral load quantities. The moderate acquisition/transmission area profile had the second highest proportion of HIV-positive persons and, accordingly, the second highest prevalence of detectable viremia. These areas, however, showed the lowest prevalence of detectable viral load. The predominance of moderate acquisition and moderate transmission types, who commonly tested three or more times for HIV, plausibly led to the highest area profile proportion of previously diagnosed adults at 24%. Moreover, moderate transmission adults had the greatest probability of being on ART, correlating with the lowest prevalence of detectable viral load across area profiles. Nationwide expansion of prevention of mother-to-child transmission (PMTCT) services with opt-out HIV testing and ART initiation in antenatal care probably facilitated the high testing and ART uptake in these areas,[25] since the moderate acquisition and transmission risk types were commonly women.

Findings also suggest that HIV testing and treatment may promote the clustering of preventive factors among risk types. HIV testing is associated with partner reduction,[19] and risk types most likely to test for HIV, moderate acquisition and transmission adults, were also most likely to be abstinent and least

likely to have multiple partners. Evidence also shows that ART initiation leads to uptake of HIV preventive behaviors,[26-29] and accordingly, moderate transmission adults showed not only the highest proportions of testing three or more times and of being on ART but also the highest probability of always using a condom (notably at 69%) and the lowest prevalence of unknown partner HIV status, a potential indicator of HIV disclosure. In fact, the prevalence of no prior diagnosis, having two or more partners, having a partner of unknown HIV status, and having never used a condom continually decreased from very high to high to moderate transmission adults; and the proportion of being diagnosed but not on ART reduced by more than half from high to moderate transmission adults. Similarly, the probability of having tested three or more times and always using a condom increased from very high to high to moderate transmission adults. This consistent pattern in transmission risk reduction across the three types implies a continuum of preventive behavior uptake among HIV-positive persons that could be leveraged in the design of area-focused HIV prevention intervention packages.

There are several strengths to this study. The results are drawn from a large, nationally representative sample with high participation and rigorous biomarker testing, lending to strong outcome and external validity. Strict methods for latent variable estimation were used and provided results with very low risk of misclassification of risk types and area profiles, good distinction between HIV risk types, modeled local dependencies between indicators, and strong assurance of identifying the global maximum solution of the model (Supplement Two). However, the analysis is a cross-sectional assessment and therefore hinders the ability to confidently assess temporal causality of area profiles preceding area HIV viral quantities. This limitation is somewhat mitigated by the necessity of HIV testing to diagnose HIV and initiate ART and the biological role of ART in reducing HIV viral load.[30] Also, HIV stigma may have biased the self-report of prior HIV diagnosis and ART status, which could have under-reported known HIV status and either over or under-reported ART prevalence.

In conclusion, this study shows that profiling the HIV risk of places can explain variations in area-level HIV viral measures. Areas that enable the clustering of higher risk behaviors are associated with increased population-level detectable quantities of HIV, suggesting that the co-location of higher transmission and acquisition risk types in larger frequencies may operate as a population mechanism of HIV infection. Moreover, HIV service interventions may alter area risk environments, enabling situational

and/or structural support for community risk reduction and thereby easing area HIV viral quantities with the potential to mitigate HIV core areas.

Prevention interventions targeting individual factors may only partially reduce transmission;[31] a more effective, resilient strategy aims towards shifts in the social and structural environments in which risk is produced.[11] Given the correlation between areas characterized by risk types that utilize HIV services and lower area HIV viral measures, overall findings support the possibility that area viral load quantities could be reversed by altering the prevalence patterns of HIV risk types towards those that utilize a combination of HIV services. Area interventions, designed according to the risk type prevalence pattern, could be implemented to facilitate the crossover of adults into lower transmission and acquisition risk groups. For example, a high acquisition/transmission area intervention may implement a HIV testing service aimed toward repeat testing and tailored to identify very high transmission and acquisition risk adults. Such an intervention may include an integrated strategy (i.e., integrating a non-HIV specific intervention within the intervention package) to enhance the likelihood of HIV testing uptake particular to the needs of high transmission and acquisition risk types. Further research identifying the contextual risk factors that cultivate HIV risk types and promote area-level risk type clustering could improve the prediction of area profiles of HIV risk and better tailor area interventions towards prevention of HIV core areas and eventual control of micro-epidemics. Area profiles defined by HIV risk type prevalences may be an informative tool for HIV programming, particularly in hyperendemic settings like Eswatini.

Tables and figures

***Table 3.1**. Participant characteristics*

	Males*		Females*		Total*	
	%	n	%	n	%	N
	46	8329	54	9843	100	18,172
Age [N=18,172]						
18-24 years	37	3088	35	3481	36	6569
25-34 years	35	2949	33	3284	34	6233
35-49 years	28	2291	31	3078	30	5369
Education [N=18,089]						
Never attended	7	540	7	633	7	1173
Primary	28	2335	30	2912	29	5247
Secondary	49	4067	51	5000	50	9066
Tertiary	16	1349	13	1253	14	2602
Residence [N=18,172]						
Rural	70	5869	70	6879	70	12,747
Urban	30	2460	30	2965	30	5425
Marital status [N=17,436]						
Not married, never had sex	17	1327	6	563	11	1890
Not married, ever had sex	46	3708	44	4105	45	7813
Married, living with partner	26	2112	30	2811	28	4923
Married, partner elsewhere	11	854	21	1956	16	2810
HIV and ART status [N=18,168]						
HIV-negative	76	6321	61	6023	68	12,343
HIV-positive, on ART	8	625	13	1259	10	1885
HIV-positive, not on ART	4	369	13	1327	9	1695
HIV-positive, no prior diagnosis	12	1013	13	1231	12	2244
HIV testing frequency [N=18,172]						
Tested, 3 or more times	21	1743	42	4147	32	5891
Tested, 1-2 times	34	2852	42	4155	39	7007
Never tested	45	3733	16	1541	29	5274
Number of partners† [N=18,034]						
No partners	27	2268	18	1795	23	4063
One partner	54	4466	78	7618	67	12,084
Two or more partners	18	1515	4	373	10	1887
Knowledge of partner HIV status† [N=13,779]						
All HIV-negative	60	3535	56	4440	58	7975
At least one HIV-positive	14	822	19	1473	17	2296
At least one with unknown HIV status‡	26	1536	25	1972	25	3508

Condom use† [N=13,334]						
Always	34	1962	30	2268	32	4230
Sometimes	36	2052	34	2552	35	4604
Never	30	1708	37	2793	34	4500

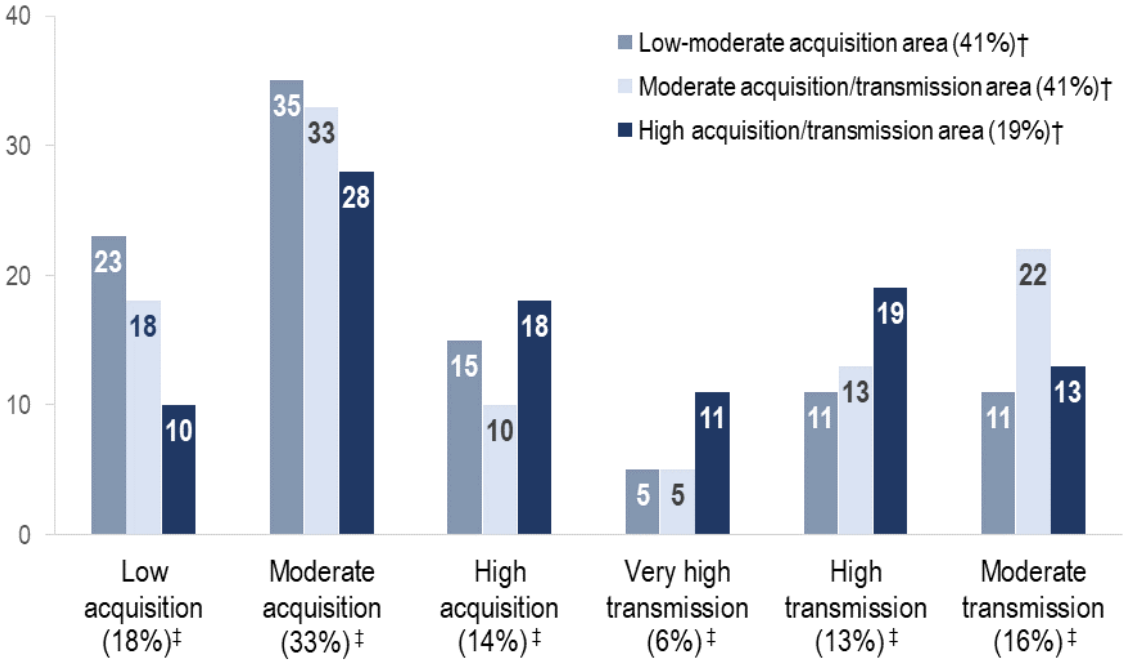
* Column totals may not total to 100% due to rounding error
† Reported in the prior six months
‡ No reported HIV-positive partners

***Table 3.2**. Characteristics of HIV risk types*

	HIV risk type (%)					
	Acquisition risk			Transmission risk		
	Low (n=3271)	Moderate (n=5997)	High (n=2544)	Very high (n=1090)	High (n=2362)	Moderate (n=2908)
Risk type prevalence	18	33	14	6	13	16
HIV and ART status*						
HIV-negative	100	100	100	0	0	23
HIV-positive, on ART	0	0	0	0	11	55
HIV-positive, not on ART	0	0	0	0	46	20
HIV-positive, no prior diagnosis	0	0	0	100	43	2
HIV testing frequency*						
Tested, 3 or more times	0	60	0	0	40	46
Tested, 1-2 times	24	40	26	17	60	54
Never tested	76	0	74	83	0	0
Number of partners*†						
No partners	73	13	3	7	10	20
One partner	27	78	67	77	78	75
Two or more partners	0	9	30	16	12	5
Knowledge of partner HIV status*†						
All HIV-negative	100	69	45	44	47	52
At least one HIV-positive	0	16	19	19	19	19
At least one unknown HIV status‡	0	16	36	37	34	29
Condom use*†						
Always	13	18	35	21	28	69
Sometimes	32	35	38	36	38	25
Never	55	46	28	43	35	6
Gender						
Female	33	64	30	43	72	67
Male	67	36	70	57	28	33
Age						
18-24 years	64	34	44	21	23	19
25-34 years	19	36	32	43	45	38
35-49 years	17	30	24	37	32	43

* Risk factor was a significant predictor (p<.001) of HIV risk type (Supplement Two)
† Reported in the prior six months
‡ No reported HIV-positive partners

Figure 3.1. *Prevalence of HIV risk type by area profiles*[*]



* Risk type was a significant predictor of area profiles [<.001] (Supplement Two)

† Prevalence of area profile among 575 enumeration areas

‡ National-level prevalence of risk type among 18,172 adults, ages 18-49 years

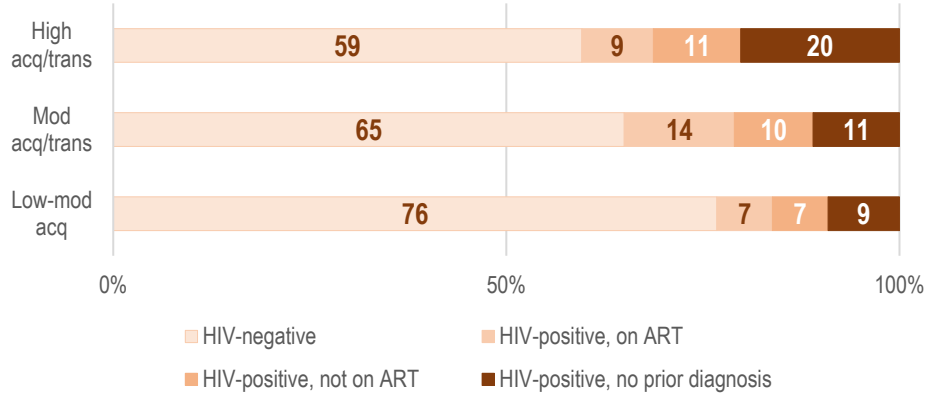
***Table 3.3**. Prevalence of detectable viral quantity by area profile*

	Area profiles							
	Low-moderate acquisition (n=233)		Moderate acquisition/transmission (n=235)			High acquisition/transmission (n=107)		
	Prevalence	β (95% CI*)	Prevalence	β (95% CI*)	p-value	Prevalence	β (95% CI*)	p-value
Detectable viremia (%)†	17.7	ref	25.4	7.42 (5.99-8.86)	<.001	35.1	17.09 (15.25-18.94)	<.001
Detectable viral load (%)‡	76.6	ref	68.7	-7.56 (-10.60- -4.53)	<.001	83.3	7.25 (4.10-10.39)	<.001

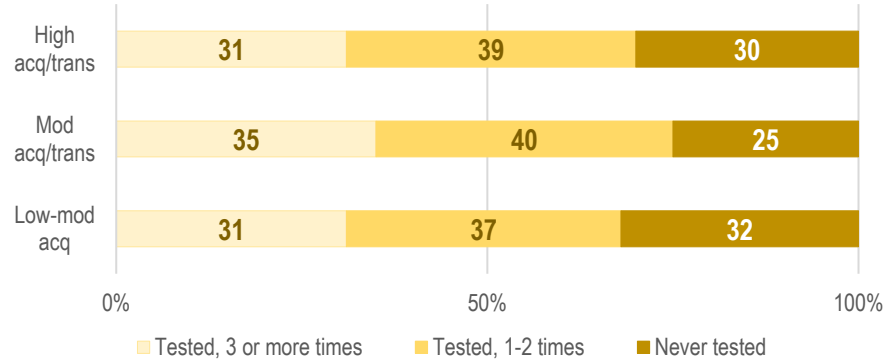
*: Regression coefficient with 95% confidence interval using robust standard errors, adjusted for area gender ratio, area mean age, area marital status, and area residence
†: Detectable viral quantity threshold is ≥ 20 viral copies/milliliter among all adults
‡: Detectable viral quantity threshold is ≥ 20 viral copies/milliliter among HIV-positive adults

***Figure 3.2**. Risk factor distribution across area profiles*

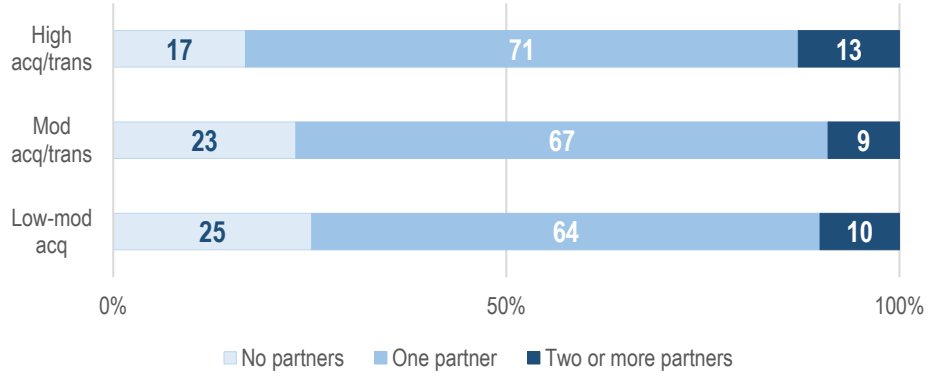
3.2a. HIV and ART status



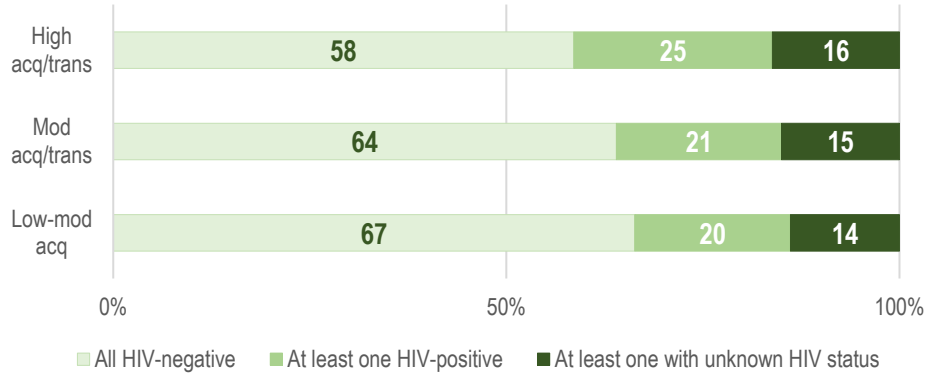
3.2b. HIV testing frequency



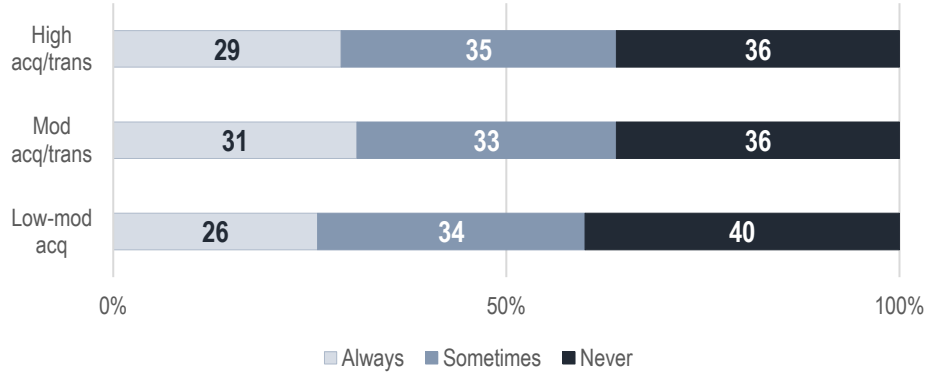
3.2c. Number of partners in the prior six months



3.2d. Knowledge of partner HIV status in prior six months



3.2e. Condom use in the prior six months

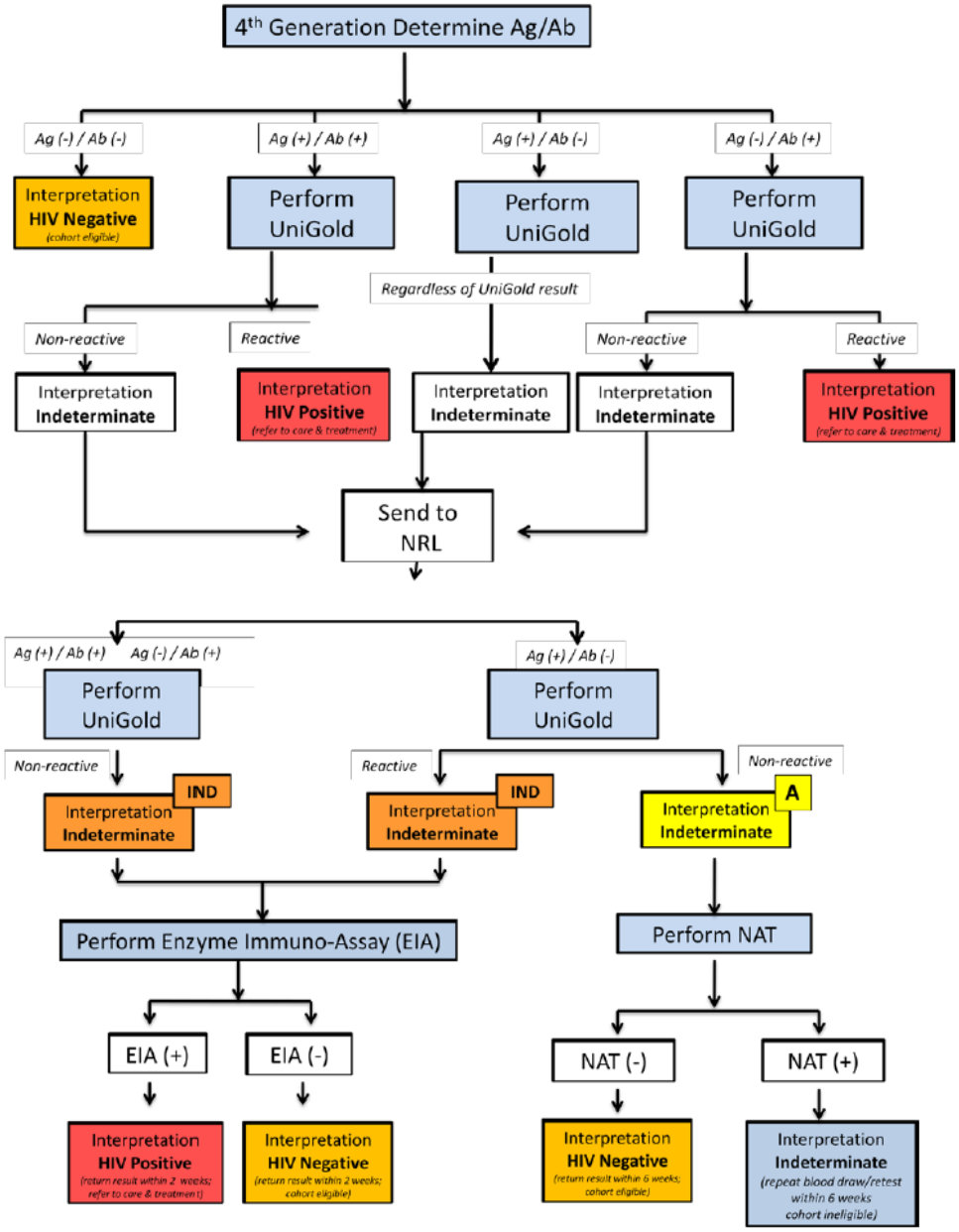


Supplements

Supplement One

This supplement provides the HIV test algorithm for determining final HIV status of participants.

Supplemental Figure S3.1*. Swaziland HIV incidence measurement survey (SHIMS): serial HIV test algorithm*



Supplement Two

This supplement provides additional methodology considerations used to estimate the area profile latent class model.

Latent model estimation

Maximum likelihood (ML) estimation was used to identify the most likely model parameter values, given the observed data. Two iterative procedures – the expectation-maximization (EM) algorithm and Newton-Raphson algorithm – were used. Estimation was initiated with the EM algorithm with a maximum number of 750 iterations and a tolerance/convergence index (i.e. a stopping rule when estimates maximize or nearly maximize the likelihood function) of 0.1. If 750 iterations with a convergence criterion of 0.1 was reached without an identified model, estimation continued with the Newton-Raphson algorithm, using a maximum of 50 iterations and tolerance of 1.0×10^{-8}, a conservative criterion.[32]

To improve the likelihood of estimating the global maximum solution, 100 random sets of start values were used to initiate model estimation.[33] The number of iterations per start values set was increased to 50 to maximize the chances of avoiding a local maximum solution[32] and a common threshold convergence criterion of maximum absolute deviation (MAD) ≤.000001 to define when the change between successive iterations was sufficiently small and therefore near the ML solution.[33]

ML estimation assumed data as missing at random and computed the average effect from available data.[34] This approach to "ignore" missingness was appropriate given the infrequency of missing values: 4% or less of demographic data, 9% of number of partners, and <1% of all other variables.

Supplemental Table S3.1*. Model fit statistics for latent class models*

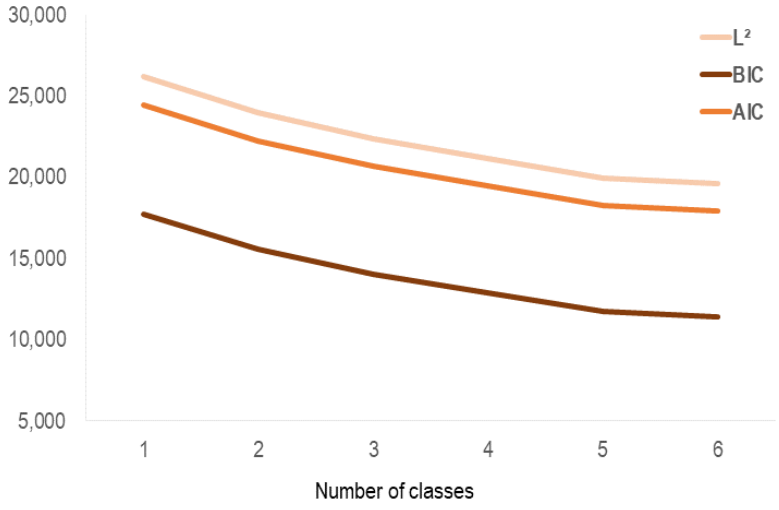
	Parameters	L^2 *	Degrees of freedom	BIC	AIC	LL
1 class	11	26,135	863	17,671	24,409	-80,053
2 class	17	23,934	857	15,529	22,220	-78,952
3 class	23	22,353	851	14,007	20,651	-78,162
4 class	29	21,126	845	12,839	19,436	-77,548
5 class	35	19,944	839	11,715	18,266	-76,957
6 class	41	19,580	833	11,410	17,914	-76,775
7 class	Model not identified					
8 class	Model not identified					

* Degrees of freedom are large and therefore, the reference distribution for L^2 statistic is not known and p-values are not presented.

Assessing latent model fitness

Model assessment and final selection was based upon a combination of statistical criteria, parsimony, and model interpretability.[33] Six micro-level models were estimated, ranging from models with one to six classes (note: models with seven and eight classes could not be identified) (Supplemental Table S3.1). In less complex models, the L^2 statistic is commonly used to measure the absolute model fitness by estimating the degree of association between variables that remain unexplained after model estimation. A lower value indicates a better model fit. To test the statistical significance of the model L^2 statistic, a chi-square distribution is assumed for testing, and a p-value greater than .05 indicates a good model fit. However, the distribution of the L^2 statistic is unknown when the degrees of freedom are large; and consequently, the Bayesian information criterion (BIC)[35] and Akaike information criterion (AIC)[36] were used for model selection in this analysis.[33]

***Supplemental Figure S3.2**. Model fit statistics curves for latent class models*



BIC and AIC show the relative model fitness across different models, where the lowest value depicts the best fitting model. BIC and AIC values progressively decreased with higher numbers of latent classes, with the six-class model showing the lowest values (Supplemental Figure S3.2). Given the smaller relative decrease in BIC and AIC between the 5-class and 6-class models, the interpretability of the two models were compared. The 6-class model showed potentially useful distinctions in sexual risk factors and

therefore was selected. The entropy R-squared ($R^2_{entropy}$) value was .77, indicating good separation of classes. $R^2_{entropy}$ ranges from 0, indicating indiscernible separation, to 1, indicating distinct separation. A good value is greater than .70.[34]

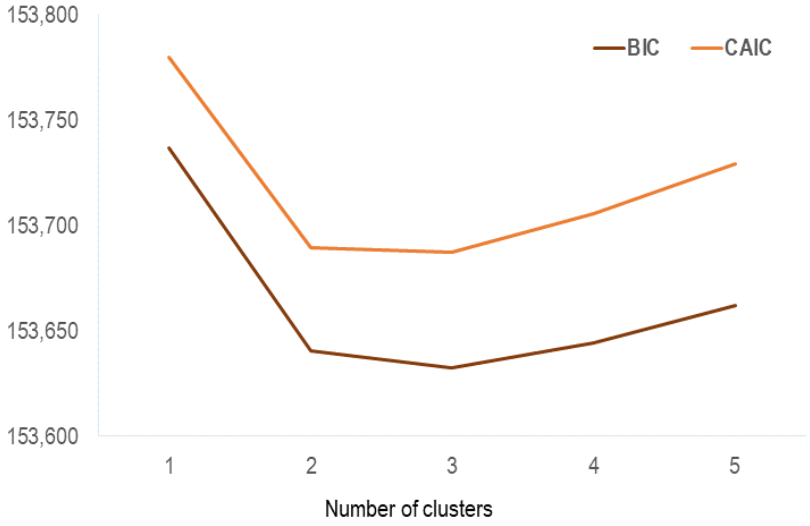
Using the 6-class micro-level latent class model, macro-level models ranging from one to five clusters were estimated. Among various information criteria, BIC is most suitable for identifying the number of latent groups at two levels and is applicable when using a simultaneous or stepwise approach to determining a latent variable structure.[37] The three cluster model showed the lowest BIC value and was chosen (Supplemental Table S3.2, Supplemental Figure S3.3). The $R^2_{entropy}$ value for the macro-level model was .42 and .74 for the micro-level model on which it was regressed.

***Supplemental Table S3.2**. Model fit statistics for latent cluster models*

	Parameters	BIC	CAIC	LL
1 cluster	43	153,736	153,779	-76,732
2 cluster	49	153,640	153,689	-76,665
3 cluster	55	153,632	153,687	-76,641
4 cluster	61	153,644	153,705	-76,628
5 cluster	67	153,662	153,729	-76,618

Note: BIC and CAIC based on log likelihood (LL) and number of clusters

***Supplemental Figure S3.3**. Model fit statistics curves for latent cluster models*



Testing prediction of the latent model's dependent variable

For the final 6-class micro-level model, the R^2 statistic was estimated for each indicator to determine the proportion of the model explained by the indicator (Supplemental Table S3.3). HIV and ART status explained a large proportion (88%) of the model. HIV testing frequency informed a good proportion (64%) of the model, as well. Each indicator was tested on whether it significantly differentiated the classes (i.e. the adult risk type), using the Wald test statistic and assuming a chi-square distribution. All indicators were significant (<.001) and were included in the model. Similarly, using the Wald test statistic with an assumed chi-square distribution, each small area risk profile significantly differentiated the HIV risk type (<.001) (Supplemental Table S3.4).

***Supplemental Table S3.3**. Multinomial logistic regression of HIV risk factor indicators on HIV risk type*

	Risk type beta coefficient						Wald statistic	p-value	R^2
	1	2	3	4	5	6			
HIV and ART status	-5.15	-8.13	2.76	-7.81	5.23	13.11	122.92	<.001	0.88
HIV testing frequency	5.6	-5.04	5.05	-4.9	4.76	-5.46	184.2	<.001	0.64
Number of partners	0.28	-2.42	-0.54	1.48	0.43	0.77	1531.61	<.001	0.30
Knowledge of partner HIV status	0.26	-3.49	0.74	0.81	0.82	0.88	329.58	<.001	0.15
Condom use	0.43	0.48	-1.30	-0.07	0.09	0.37	202.33	<.001	0.18

Key: (1) moderate acquisition, (2) low acquisition, (3) moderate transmission, (4) high acquisition, (5) high transmission, (6) very high transmission. Note: risk type numbers do not correlate to increasing risk pattern as described by risk type labels.

***Supplemental Table S3.4**. Multinomial logistic regression of area profiles on HIV risk type*

	Risk type beta coefficient						Wald statistic	p-value
Area profile	1	2	3	4	5	6		
Low-moderate acquisition	0.94	0.51	-0.21	0.09	-0.23	-1.1	445.48	<.001
Moderate acquisition/transmission	0.83	0.22	0.44	-0.39	-0.13	-0.98	267.9	<.001
High acquisition/transmission	0.59	-0.46	-0.16	0.15	0.21	-0.33	133.49	<.001
R^2	.88							

Key: (1) moderate acquisition, (2) low acquisition, (3) moderate transmission, (4) high acquisition, (5) high transmission, (6) very high transmission. Note: risk type numbers do not correlate to increasing risk pattern as described by risk type labels.

Analysis plan

The analysis plan comprised a three-step estimation procedure.[38] In step one, a measurement model for HIV risk type and area profile was estimated. Nonparametric random-effects multinomial logistic regression modelling was conducted to allow latent variable sizes to differ.

Latent class analysis assumes independence between indicators within each estimated risk type/area profile. Given the likely correlations between several indicators (e.g., HIV testing history and HIV and ART status where no prior test history correlates with no prior diagnosis), the assumption of local independence between indicators when conditioned on HIV risk type category was checked by assessing for bivariate residuals (BVRs). When BVRs were larger than 1, the association between those two indicators was accommodated in the model by relaxing the assumption of local independence.

Pairwise comparisons of the latent classes using the Wald test statistic was done to test for the statistical significance of each indicator in discriminating between HIV risk types (Supplemental Table S3.5). Fifteen pairwise comparisons were conducted for each indicator, and the majority (80%, 60/75) were significant. Analyses indicated that HIV and ART status did not distinguish between the three acquisition risk types (low, moderate and high acquisition types), which were all HIV-negative adults. HIV testing frequency probabilities were similar across low and high acquisition and high transmission types, which were characterized by no probability of three or more lifetime tests and high proportions of having never tested; and low and moderate acquisition and very high transmission adults did not differ in condom use, where generally half (55%, 46%, and 43%, respectively) had never used a condom. Knowledge of partner HIV status demonstrated the least ability to discriminate between HIV risk types. High acquisition and high and very high transmission adults did not differ, nor did high acquisition and moderate and high transmission adults. The probability distribution of partner HIV status, however, did differ between moderate and very high transmission adults. Low and moderate acquisition adults also differed in partner status distribution with all other risk types.

***Supplemental Table S3.5**. Pairwise comparison of risk types by risk factor indicator*

Risk type comparison		p-value	Risk type comparison		p-value	Risk type comparison		p-value	Risk type comparison		p-value	Risk type comparison		p-value
HIV status			HIV testing frequency			Number of partners			Partner's HIV status			Condom use		
1	2	**NS**	1	2	<.001	1	2	<.001	1	2	0.008	1	2	**NS**
1	3	0.007	1	3	<.001	1	3	<.001	1	3	<.001	1	3	<.001
1	4	**NS**	1	4	<.001	1	4	<.001	1	4	<.001	1	4	<.001
1	5	<.001	1	5	<.001	1	5	**NS**	1	5	<.001	1	5	<.001
1	6	<.001	1	6	<.001	1	6	<.001	1	6	<.001	1	6	**NS**
2	3	<.001	2	3	<.001	2	3	<.001	2	3	0.003	2	3	<.001
2	4	**NS**	2	4	**NS**	2	4	<.001	2	4	0.002	2	4	<.001
2	5	<.001	2	5	<.001	2	5	<.001	2	5	0.002	2	5	0.002
2	6	<.001	2	6	**NS**	2	6	<.001	2	6	0.002	2	6	**NS**
3	4	<.001	3	4	<.001	3	4	<.001	3	4	**NS**	3	4	<.001
3	5	<.001	3	5	0.009	3	5	<.001	3	5	**NS**	3	5	<.001
3	6	<.001	3	6	<.001	3	6	<.001	3	6	0.004	3	6	<.001
4	5	<.001	4	5	<.001	4	5	<.001	4	5	**NS**	4	5	0.03
4	6	<.001	4	6	**NS**	4	6	<.001	4	6	**NS**	4	6	<.001
5	6	<.001	5	6	<.001	5	6	0.001	5	6	**NS**	5	6	<.001

Key: (1) moderate acquisition, (2) low acquisition, (3) moderate transmission, (4) high acquisition, (5) high transmission, (6) very high transmission. Note: risk type numbers do not correlate to increasing risk pattern as described by risk type labels.

Pairwise comparisons of area profiles indicate that the HIV risk type distribution significantly differed only between low-moderate acquisition and high acquisition/transmission area profiles (p=.047) (Supplemental Table S3.6). Area profile comparisons also showed that each pair significantly differed (<.001) in each of the six viral quantity measures (Supplemental Table S3.7)

***Supplemental Table S3.6**. Pairwise comparison of area profiles*

Area profile comparison		Wald	df	p-value
1	2	0.03	1	**NS**
1	3	2.26	1	**NS**
2	3	3.94	1	0.047

Key: (1) moderate acquisition/transmission, (2) low-moderate acquisition profile, (3) high acquisition/ transmission profile.

***Supplemental Table S3.7**. Pairwise comparison of area profiles by HIV viral quantity*

Area profile comparison		p-value	Area profile comparison		p-value	Area profile comparison		p-value
Area geometric mean			Area prevalence ≥ 20 c/mL (%)			Area prevalence ≥ 100,000 c/mL (%)		
Viral load*			Viral load*			Viral load*		
1	2	<.001	1	2	<.001	1	2	<.001
1	3	<.001	1	3	<.001	1	3	<.001
2	3	<.001	2	3	<.001	2	3	<.001
Viremia			Viremia			Viremia		
1	2	<.001	1	2	<.001	1	2	<.001
1	3	<.001	1	3	<.001	1	3	<.001
2	3	<.001	2	3	<.001	2	3	<.001

Key: (1) low-moderate acquisition profile, (2) moderate acquisition/transmission profile, (3) high acquisition/ transmission profile.

* Among HIV-positive adults

† Among all adults regardless of HIV status

In step two, using the final selected HIV risk type and area profile measurement models, the observed survey data was assigned using posterior class membership probabilities so that risk types were assigned to classes and EAs were assigned to clusters using ***Z*** scores for micro-units and macro-units. Clusters and classes were assigned using a modal rule where a macro-unit or micro-unit was assigned to the cluster or class, respectively, with the highest frequency. The amount of misclassification error was assessed by conducting a cross-tabulation of the modal clusters/classes by the probability clusters/classes. The largest

potential error in using the modal rule for individual assignment versus the probability rule was 4.1% [n=743] of adults being misclassified as low acquisition type (Supplemental Table S3.8). Similarly, the largest error in area profiles may have misclassified 4.2% (24/575) of EAs as low-moderate acquisition areas (Supplemental Table S3.9).

Supplemental Table S3.8. *Cross tabulation of individual assignment rules for HIV risk types*

Number of individuals assigned to risk types using modal rule							
	1	2	3	4	5	6	Total
1	5542	294	31	70	1	0	5938
2	220	2830	0	209	0	0	3259
3	557	70	1930	19	337	0	2913
4	488	360	20	1597	0	0	2465
5	9	1	450	0	1981	0	2441
6	0	0	0	0	195	961	1156
Total	6815	3555	2430	1896	2515	961	18172
Number of individuals assigned to risk types using proportional rule							
	1	2	3	4	5	6	Total
1	4799	329	438	363	9	0	5938
2	329	2486	54	390	1	0	3259
3	438	54	1812	64	541	4	2913
4	363	390	64	1647	1	0	2465
5	9	1	541	1	1746	143	2441
6	0	0	4	0	143	1008	1156
Total	5938	3259	2913	2465	2441	1156	18172
Difference in number of individuals assigned between modal and proportional rules [# (% of 18172)]							
	1	2	3	4	5	6	
1	743 (4.1)	-34 (0.2)	-408 (2.2)	-293 (1.6)	-7 (<.01)	0	
2	-109 (0.6)	344 (1.9)	-54 (0.3)	-181 (1.0)	-1 (<.01)	0	
3	118 (0.7)	16 (0.1)	118 (0.6)	-44 (0.2)	-204 (1.1)	-4 (<.01)	
4	125 (0.7)	-30 (0.2)	-43 (0.2)	-50 (0.3)	-1 (<.01)	0	
5	0	0	-91 (0.5)	-1 (<.01)	235 (1.3)	-143 (0.8)	
6	0	0	-4 (<.01)	0 (<.01)	51 (0.3)	-47 (0.3)	

Key: (1) moderate acquisition, (2) low acquisition, (3) moderate transmission, (4) high acquisition, (5) high transmission, (6) very high transmission. Note: risk type numbers do not correlate to increasing risk pattern as described by risk type labels.

***Supplemental Table S3.9**. Cross tabulation of EA assignment rules to area profiles*

Number of assigned EAs using modal rule				
	1	2	3	Total
1	167	46	18	232
2	45	163	10	218
3	24	15	87	125
Total	236	224	115	575

Number of assigned EAs using proportional rule				
	1	2	3	Total
1	144	59	29	232
2	59	140	18	218
3	29	18	78	125
Total	232	218	125	575

Difference in number of EA assignments between modal and proportional rules [# (% of 575)]			
	1	2	3
1	24 (4.2)	-13 (2.3)	-11 (1.9)
2	-14 (2.4)	23 (4.0)	-9 (1.6)
3	-5 (.9)	-4 (.7)	9 (1.6)

Key: (1) low-moderate acquisition profile, (2) moderate acquisition/transmission profile, (3) high acquisition/ transmission profile.

Lastly in step three, ordinary least squares regression was used to estimate model parameters of the relationship between area profiles and area HIV viral quantity.

Supplement Three

This supplement provides the prevalence of area HIV viral quantity, when measured by geometric mean, detectable virus, and high virus levels, and its statistical association with area profile.

***Supplemental Table S3.10**. Prevalence of area HIV viral quantity by area profile*

	Area profiles							
	Low-moderate acquisition (n=233)		Moderate acquisition/transmission (n=235)			High acquisition/transmission (n=107)		
	Viral quantity	β (95% CI*)	Viral quantity	β (95% CI*)	p-value	Viral quantity	β (95% CI*)	p-value
Viremia (c/mL among all adults)								
Geometric mean (mean)	8.3	ref	22.1	12.61 (8.23-17.00)	<.001	71.1	60.75 (43.43-78.06)	<.001
Detectable (%)†	17.7	ref	25.4	7.42 (5.99-8.86)	<.001	35.1	17.09 (15.25-18.94)	<.001
High (%)‡	15.6	ref	21.4	5.59 (4.25-6.92)	<.001	31.6	15.74 (14.02-17.46)	<.001
Viral load (c/mL among HIV-positive adults)								
Geometric mean (mean)	9328.9	ref	3840.5	-5194.81 (-7280.10- -3109.52)	<.001	13,122.3	4095.35 (576.84-7613.86)	0.023
Detectable (%)†	76.6	ref	68.7	-7.56 (-10.60- -4.53)	<.001	83.3	7.25 (4.10-10.39)	<.001
High (%)‡,	19.5	ref	18.6	-0.37 (-3.32-2.59)	**NS**	26.7	7.62 (3.98-11.26)	<.001

*: Regression coefficient with 95% confidence interval using robust standard errors, adjusted for area gender ratio, area mean age, area marital status, and area residence
†: Detectable viral quantity threshold is ≥ 20 c/mL
‡: High viral quantity threshold is ≥ 100,000 c/mL

References

1. Yorke JA, Hethcote HW, Nold A. Dynamics and control of the transmission of gonorrhea. Sex Transm Infect 1978;5:51-6.

2. Fichtenberg CM, Ellen JM. Editorial response: moving from core groups to risk spaces. Sex Transm Infect 2003;30:825-6.

3. Thomas JC, Tucker MJ. The development and use of the concept of a sexually transmitted disease core. J Infect Dis 1996;174 Suppl 2:S134-43.

4. Gesink DC, Sullivan AB, Miller WC, Bernstein KT. Sexually transmitted disease core theory: roles of person, place, and time. Am J Epidemiol 2011;174:81-9.

5. Jain V, Petersen ML, Liegler T, et al. Population levels and geographical distribution of HIV RNA in rural Ugandan and Kenyan communities, including serodiscordant couples: a cross-sectional analysis. Lancet HIV 2017;4:e122-e33.

6. Vandormael A, Newell M-L, Baernighausen T, Tanser F. Use of antiretroviral therapy in households and risk of HIV acquisition in rural KwaZulu-Natal, South Africa, 2004-12: a prospective cohort study. Lancet Glob Health 2014;2:E209-E15.

7. Montaner JSG, Lima VD, Barrios R, et al. Association of highly active antiretroviral therapy coverage, population viral load, and yearly new HIV diagnoses in British Columbia, Canada: a population-based study. Lancet 2010;376:532-9.

8. Barnabas RV, van Rooyen H, Tumwesigye E, et al. Initiation of antiretroviral therapy and viral suppression after home HIV testing and counselling in KwaZulu-Natal, South Africa, and Mbarara district, Uganda: a prospective, observational intervention study. Lancet HIV 2014;1:e68-e76.

9. Jain V, Byonanebye DM, Liegler T, et al. Changes in population HIV RNA levels in Mbarara, Uganda, during scale-up of HIV antiretroviral therapy access. J Acquir Immune Defic Syndr 2014;65:327-32.

10. Galea S, Ahern J, Vlahov D. Contextual determinants of drug use risk behavior: a theoretic framework. J Urban Health 2003;80:iii50-8.

11. Rhodes T. The 'risk environment': a framework for understanding and reducing drug-related harm. Int J Drug Policy 2002;13:85-94.

12. Rhodes T, Simic M. Transition and the HIV risk environment. BMJ 2005;331:220-3.

13. Bicego GT, Nkambule R, Peterson I, et al. Recent patterns in population-based HIV prevalence in Swaziland. PLoS One 2013;8:e77101.

14. Justman J, Reed JB, Bicego G, et al. Swaziland HIV incidence measurement survey (SHIMS): a prospective national cohort study. Lancet HIV 2017;4:e83-e92.

15. Justman J, Ellman, T.M., Donnell, D., Duong, Y.T., Reed, J., Bicego, G., Ehrenkranz, P., Chang, J., Wang, L., Bock, N., Nkambule, R. for the SHIMS Study Team. Population HIV viral load in Swaziland: Assessing ART program effectiveness and transmission potential. Oral abstract presented at: 20th Conference on Retroviruses and Opportunistic Infections; 2013 Mar 3-6; Atlanta, United States.

16. Dixon JR. The International Conference on Harmonization Good Clinical Practice Guideline. Quality Assurance 1999;6:65-74.

17. Cuadros DFG T, de Oliveira T, Bärnighausen T, Tanser F. Assessing the role of geographical HIV hot-spots in the spread of the epidemic. Oral abstract presented at: 25th Conference on Retroviruses and Opportunistic Infections; 2018 Mar 4-7; Boston, United States.

18. Mwangi M, Bunnell R, Nyoka R, et al. Unsafe sex among HIV-infected adults in Kenya: results of a nationally representative survey. J Acquir Immune Defic Syndr 2011;58:80-8.

19. Fonner VA, Denison J, Kennedy CE, O'Reilly K, Sweat M. Voluntary counseling and testing (VCT) for changing HIV-related risk behavior in developing countries. Cochrane Database Syst Rev 2012:CD001224.

20. Coates TJ, Kulich M, Celentano DD, et al. Effect of community-based voluntary counselling and testing on HIV incidence and social and behavioural outcomes (NIMH Project Accept; HPTN 043): a cluster-randomised trial. Lancet Glob Health 2014;2:e267-77.

21. Wong LH, Rooyen HV, Modiba P, et al. Test and tell: correlates and consequences of testing and disclosure of HIV status in South Africa (HPTN 043 Project Accept). J Acquir Immune Defic Syndr 2009;50:215-22.

22. Abdool Karim Q, Dellar RC, Bearnot B, et al. HIV-positive status disclosure in patients in care in rural South Africa: implications for scaling up treatment and prevention interventions. AIDS Behav 2015;19:322-9.

23. Simbayi LC, Kalichman SC, Strebel A, Cloete A, Henda N, Mqeketo A. Disclosure of HIV status to sex partners and sexual risk behaviours among HIV-positive men and women, Cape Town, South Africa. Sex Transm Infect 2007;83:29-34.

24. Kalichman SC, Kalichman MO, Cherry C, Grebler T. HIV disclosure and transmission risks to sex partners among HIV-positive men. AIDS Patient Care STDS 2016;30:221-8.

25. UNAIDS. Start Free Stay Free AIDS Free. Geneva: UNAIDS; 2019.

26. Mujugira A, Celum C, Ngure K, et al. Antiretroviral therapy initiation is not associated with risky sexual behavior among heterosexual human immunodeficiency virus-infected persons in serodiscordant partnerships. Sex Transm Infect 2017;44:57-61.

27. Eisele TP, Mathews C, Chopra M, et al. Changes in risk behavior among HIV-positive patients during their first year of antiretroviral therapy in Cape Town South Africa. AIDS Behav 2009;13:1097-105.

28. Doyle JS, Degenhardt L, Pedrana AE, et al. Effects of HIV antiretroviral therapy on sexual and injecting risk-taking behavior: a systematic review and meta-analysis. Clin Infect Dis 2014;59:1483-94.

29. Berhan A, Berhan Y. Is the sexual behaviour of HIV patients on antiretroviral therapy safe or risky in Sub-Saharan Africa? Meta-analysis and meta-regression. AIDS Res Ther 2012;9:14.

30. Cohen MS, Chen YQ, McCauley M, et al. Antiretroviral therapy for the prevention of HIV-1 transmission. N Engl J Med 2016;375:830-9.

31. Gupta GR, Parkhurst JO, Ogden JA, Aggleton P, Mahal A. Structural approaches to HIV prevention. Lancet 2008;372:764-75.

32. Vermunt JKaMJ. Latent Gold 4.0 User's Guide. Belmont, Massachusetts: Statistical Innovations; 2005.

33. Collins LMaLST. Latent Class and Latent Transition Analysis: With Applications in the Social, Behavioral, and Health Sciences. Boston, Massachusetts: John Wiley and Sons, Inc; 2009.

34. Vermunt JKaMJ. Technical Guide for Latent Fold 5.1: Basic, Advanced, and Syntax. Belmont, Massachusetts: Statistical Innovations, Inc.; 2016.

35. Schwartz G. Estimating the dimension of a model. Ann Stat 1978;6:461-4.

36. Akaike H. Factor analysis and AIC. Psychometrika 1987;52:317–32.

37. Lukociene O, Varriale R, Vermunt JK. The simultaneous decision(s) about the number of lower- and higher-level classes in multilevel latent class analysis. Sociol Methodol 2010;40:247-83.

38. Bennink M, Croon MA, Keuning J, Vermunt JK. Measuring student ability, classifying schools, and detecting item bias at school level, based on student-level dichotomous items. J Educ Behav Stat 2014;39:180-202.

Chapter Four. Area risk profiles and HIV susceptible areas: detecting the places most prone to incident HIV in a national hyperendemic setting

Summary

Background: Contextual factors, like gender inequality and stigma, can influence area environments to enable HIV risk factors. Places susceptible to HIV may result from a risk environment of community norms and practices that produce HIV acquisition risk. Understanding the factors that create them may help pinpoint the next incident infection and support strategies for population risk reduction and epidemic control.

Methods: Using a two-stage cluster-based design, a nationally representative, household-based sample of adults, ages 18-49 years was enrolled from December 2010 to June 2011 in Eswatini. Consenting adults completed an interview and received home-based rapid HIV testing and counseling. All seropositive samples were tested for HIV viral load using the COBAS AmpliPrep/Taqman HIV-1 Test, v 2.0. Adults testing HIV-seronegative were enrolled in a prospective cohort for the direct observation of HIV seroconversion and completed an interview and received home-based rapid HIV testing and counseling six months later. Multi-level latent class modeling was performed to identify statistically significant combinations of HIV risk factors and classify the combinations into small area risk profiles. Generalized linear regression was conducted to assess the association between area profiles and area-level HIV seroconversion. The observed average area-level prevalences of each risk factor were assessed for bivariable and multivariable associations with HIV susceptible areas.

Results: A total of 18,172 adults were surveyed of which 4396 [24%] had detectable HIV viremia (≥20 copies/milliliter) among all adults, regardless of HIV status. 11,880 [96%; n=12,357] HIV-seronegative adults enrolled in the prospective cohort and 11,155 (94%) of them completed an endline visit. Four area profiles were identified. The proportion of HIV susceptible areas progressively increased from Profiles A, B, C, and D [14.3%, 21.8%, 24.6%, and 30.8%, respectively]. HIV susceptible areas were more than twice as likely to occur in Profile D than Profile A risk environments [RR 2.13, 95% confidence interval (CI) (1.13, 4.00); p=0.02]. Profile D areas had prevalences of unknown partner HIV status and detectable viremia at 28% and 24%, respectively. In contrast, Profile A areas had prevalences of only 8% with unknown HIV status and 31% with detectable viremia. When micro-level risk factor indicators were aggregated at the area level, none of the risk factor area prevalences were independently associated with HIV susceptible areas.

Conclusion: An area typology of HIV acquisition risk can help predict the places most susceptible to HIV acquisition. While uncontrolled HIV infection is necessary for areas to produce incident HIV, the concurrent levels of other risk factors may have more influence in reducing population-level incidence than previously considered.

Introduction

Treatment as prevention (TasP) is the cornerstone of the global strategy to control the HIV epidemic, based on the premise that treatment-induced sustained HIV viral suppression virtually eliminates ongoing transmission.[1] Centered on this approach, UNAIDS announced the Fast-Track targets for 2030, aiming to reduce the number of new infections worldwide to less than 200,000.[2] Global efforts, however, are currently falling short of this goal. New HIV infections since 2010 have decreased only 16% from 2.1 million to 1.7 million in 2018, revealing a trajectory well above the 2020 interim target of 500,000 new infections.[3,4]

Coupled with this sobering progress report is the mixed results across several well-conducted TasP trials, indicating that treatment coverage is necessary but insufficient to restraining new HIV infections.[5-9] While some have suggested the fatal flaw of these studies may have been an inadequate focus on hard to reach populations (e.g., migrants, young persons, and men)[6] and core groups with higher likelihood for ongoing HIV transmission,[10] an alternative explanation may be linked to the spatial dynamics of epidemics. Given that TasP was designed to mitigate individual risk factors and has not shown the population-level HIV incidence reductions needed to achieve epidemic control, a shift in the unit of analysis from susceptible individuals to susceptible areas may be instructive. Reorienting to environmental risk factors that are exogenous to people[11] and that encourage the circumstances for contracting HIV may clarify how treatment can optimally prevent populations from acquiring new HIV infections.

The geographic diffusion of HIV is shaped by population behavior patterns constrained through social, economic, cultural, and political circumstances;[11] describing the environment of HIV risk created by these conditions is a potential analytic approach to explain population HIV outcomes. A risk environment is a social or physical space where various factors combine and promote the likelihood of health-related harm.[12] In this framework, areas prone to new HIV infections, or HIV susceptible areas, may result from a risk environment of community norms and practices that produce HIV acquisition risk. As HIV susceptible areas may potentially pinpoint the next incident infection, understanding the factors that create them may be key to designing interventions that change risk environments, augment risk reduction, and attain epidemic control.

In 2011, Eswatini, formerly known as Swaziland, experienced the highest national HIV incidence (2.4%) among adults ages 18-49 years. Individual-level analyses revealed few acquisition risk factors, namely marital status and knowledge of partner's HIV status;[13] to build on these findings, this analysis applied a risk environment approach to describe the environmental conditions affecting sexual HIV acquisition risk among HIV-seronegative adults in Eswatini. Multi-level latent class modeling of longitudinal data was used to categorize the combined prevalences of contextual HIV risk factors, specifically area-specific sexual behavior norms and probability of viral exposure of HIV-seronegative persons, into small area, geographic risk profiles. The area risk profiles were then assessed for their prediction of HIV susceptible areas measured through directly observed HIV seroconversions. Analyses were hypothesized to show that HIV-seronegative adults living in environments with higher HIV viremia, lower knowledge of partner HIV status and greater frequency of multiple partnerships would result in areas experiencing a new HIV infection.

Methods

Study population

The Swaziland HIV Incidence Measurement Survey (SHIMS) is a household-based, nationally representative cross-sectional sample of adults, ages 18-49 years. The design and methods are described elsewhere.[14] A total of 18,172 adults from 12,571 households across 575 enumeration areas (i.e., census tracts) completed a baseline survey from December 2010 to June 2011. Among them, 12,357 participants tested HIV-seronegative and were asked to participate in a six-month cohort, designed to directly observe the rate of HIV seroconversions. Eligibility criteria included eligibility criteria for the cross-sectional cohort,[14] confirmed HIV-seronegative status in the baseline survey, willingness to undergo all study visits and procedures, and ability to provide informed consent and answer interview questions in English or siSwati. Persons who stated an intention to leave Swaziland indefinitely in the upcoming six months were excluded. A total of 11,880 (96%) persons consented to participate, and at the endline visit six months later, 11,155 (94%) participants completed a follow-up interview and received HIV rapid testing. The study was approved by ethical review boards at Columbia University Irving Medical Center, the Eswatini Ministry of Health, and the US Centers for Disease Control and Prevention.

Study procedures

Selected households were approached by study staff trained in Good Clinical Practice and survey procedures.[15] Field staff conducted informed consent for completing an interview with HIV testing. At both baseline and endline visits, consenting participants completed a paper-based, face-to-face interview about sexual behaviors, HIV service uptake, and demographics and home-based HIV testing and counseling in a private location around or within the home.

Trained staff performed venipuncture and HIV testing and counseling. Rapid HIV testing was conducted with Determine HIV-1/2 Ag/Ab Combo (Alere, Japan). Determine-reactive samples were confirmed with Uni-Gold HIV Test (Trinity Biotech, Ireland). Laboratory-based confirmatory testing and discrepancy resolution were done according to a survey-specific test algorithm (Supplemental Figure S4.1). HIV-seropositive participants were provided referrals to HIV care.

To measure the biological environment of HIV exposure, the HIV viral quantity in the baseline sample of 18,172 adults was included in this analysis. All seropositive samples in the cross-sectional cohort were tested for plasma viral load using the COBAS AmpliPrep/TaqMan HIV-1 Test, v2.0 automated platform (Roche, USA), according to manufacturers' instructions.

Study measures

Outcome

HIV susceptible area is an observed, macro-level, dichotomous outcome indicating if at least one HIV seroconversion was observed in the study enumeration area (EA) during the six-month follow-up period.

Exposure

Area profile, the exposure of interest, is an unobservable profile of HIV acquisition risk in the EA. It is a latent, macro-level variable summarizing the area-level risk environment of social norms and HIV viral quantity affecting seroconversion in the HIV-seronegative, EA-specific population. It is a nominal variable defined by the composite mean prevalences of seven indicators across a sub-set of similar HIV-seronegative EA populations.

Area profile indicators

Area profile comprises seven micro-level observed risk factor indicators: 1) sexually active (yes, no); 2) number of sexual partners in the prior six months (two or more partners, one partner, no partners); 3) casual partner in the prior six months (yes, no); 4) knowledge of HIV status of partners in the prior six months (at least one with unknown status, at least one HIV-positive, all HIV-negative); 5) condom use in the prior six months (never, sometimes, always); 6) gender-age demographic (female, 18-24 years; female, 25-34 years; female, 35-49 years; male, 18-24 years; male, 25-34 years; male, 35-49 years); and 7) HIV viremia (detectable virus at ≥20 copies/mL [c/mL], undetectable virus at <20 c/mL).

HIV viremia was measured at baseline among the 18,172 cross-sectional cohort participants. It includes all EA residents, regardless of HIV status. When estimated as a profile indicator, it describes the area-level context of uncontrolled HIV viremia where HIV-seronegative participants reside. Gender-age demographic was measured at baseline only among the 11,880 six-month cohort participants. All other risk factor indicators were measured among the 11,155 six-month cohort participants who completed the endline visit.

Covariates

Three covariates were included as potential confounders in the model assessing association between area profile and HIV susceptible areas: area gender ratio (expressed as a percentage), area marital status, defined as the prevalence of persons who were married and living with a partner in the EA, and area residence, defined as either urban or rural. Area mean age of all participants in an EA was not included in the final model since age distribution was reflected in the area profile. Covariates were measured at baseline among the 18,172 cross-sectional cohort participants.

Statistical analysis

Multi-level latent class modeling using maximum likelihood estimation was conducted to perform a micro-macro analysis. Model assessment and final selection was based upon a combination of statistical criteria, parsimony, and model interpretability (Supplement Two).

The seven observed micro-level risk factor indicators were regressed on area profile, a latent macro-level variable, to identify similar combinations of contextual HIV risk factors and group them into area profile categories. Each category represents a qualitatively unique, homogenous group of EAs with similar

distributions of contextual HIV risk factors. Using the Wald test statistic with an assumed chi-square distribution, each risk factor was tested for a statistically significant association with area profile.

After estimating area profile, generalized linear regression analysis, using the log function to obtain risk ratios, determined whether area profiles could predict the observed distribution of HIV susceptible areas. Analyses were weighted, including adjustments for design weights, nonresponse, and post-stratification adjustments. Descriptive statistics were estimated by calculating a conditional probability table of the indicator categories of contextual risk factors, describing the prevalence of each indicator across each area profile. Latent class modeling was conducted using Latent Gold version 5.1, and generalized linear modeling was conducted using STATA version 16.

Results

Participant characteristics

A total of 11,856 adults, ages 18 to 49 years, were surveyed (Table 4.1). Nearly half [45%] of the population was ages 18 to 24 years, with the majority [70%] living in a rural area. Over two-thirds [68%] had completed secondary education or higher, and only 5% having never attended school. While most [44%] adults were unmarried and sexually active, the proportion of males reported as having never had sex was more than double that of females [22% versus 9%]. Females were more than twice as likely as males [21% versus 9%] to be married with their partner living elsewhere.

The vast majority [87%] of adults reported as sexually debuted, with a larger proportion among females than males [92% versus 82%] (Table 4.1). While about a tenth of sexually debuted adults reported having two or more partners [10%] or having a casual partner [8%] in the prior six months, males were notably more likely than females to report multiple partners [19% versus 2%] and casual partners [14% versus 2%]. Around a tenth [11%] of adults reported at least one partner with unknown HIV status, with more females than males [42% versus 30%] having never used a condom.

To quantify the context of uncontrolled HIV infection where HIV-seronegative adults reside, the prevalence of detectable viremia was estimated among all residents in study areas, regardless of HIV status (Table 4.1). Overall, nearly a quarter [24%] of adults had detectable HIV, which was higher in females [29%] than males [19%].

Area profiles

Latent class modeling detected four area profiles. The prevalence of each contextual risk factor, disaggregated by area profile, is presented in Table 4.2. Each profile was assessed for patterns in risk factor probabilities to identify meaningful differences between profiles that infer the potential environmental risk for HIV acquisition.

Profile A reflected area-level risk environments with the highest proportion [94%] of sexually active HIV-seronegative adults, yet also some of the lowest levels of contextual sexual risk: having two or more partners [10%], any casual partner [5%], and a partner with unknown HIV status [8%] (Table 4.2). Having never used a condom, however, was at the greatest [42%] level, compared to other profiles. Areas under this profile showed the largest prevalences of detectable viremia [31%] and HIV [38%] (Figure 4.1).

Profile B had the lowest proportion [84%] of sexually active HIV-seronegative adults but presented a similar sexual risk behavior pattern as Profile A (Table 4.2). Profiles A and B had statistically equivalent prevalences of two or more partners [8%], any casual partner [6%], a partner with unknown HIV status [8%], and never having used a condom [39%] (Table 4.2, Supplemental Table S4.3). Area risk environments under Profile B, however, showed the lowest prevalences of detectable viremia [21%] and HIV [29%] (Figure 4.1).

Sexual behavior norms in Profile C risk environments showed elevated risk-taking. With the highest levels across profiles, Profile C had a 70% and 112% increase, when compared to the national average, in having two or more partners [17% versus 10%] and having casual partners [17% versus 8%] (Table 4.2). These sexual risk “hotspots,” however, were complemented by the highest likelihood [38%] of always using a condom. Detectable viremia prevalence was statistically equivalent to the national average [27% versus 24%] (Supplemental Table S4.3) and was intermediate in level to Profiles A and B [38%, 27%, and 21% in Profiles A, C, and B, respectively], as was levels of HIV prevalence [38%, 36%, and 29% in Profiles A, C, and B, respectively] (Figure 4.1). Additionally, males ages 18-34 years [51%] was notably larger than other profiles [32%, 41%, and 42% in Profiles A, B, and D, respectively].

Risk environments summarized in Profile D depicted a comparatively lower [85%] prevalence of sexually active HIV-seronegative adults, with most contextual risk factors (i.e., number of partners, casual partnerships, and condom use) similar to the national average (Table 4.2). Except, having a partner with

unknown HIV status was more than twice the proportion nationwide [28% versus 12%], showing more than a doubled and tripled increase compared to other profiles [28%, 13%, 8%, and 8% in Profiles D, C, B, and A, respectively]. Detectable viremia prevalence [24%] was statistically equivalent to Profile C [27%] and the national average [24%]. Average area HIV prevalence was 33% (Figure 4.1). Additionally, females ages 18-24 years [24%] was somewhat greater than other profiles [19%, 22%, and 17% in Profiles A, B, and C, respectively].

Differences in ART prevalence across all profiles was negligible [10%, 10%, 11%, and 10% in Profiles A, B, C, and D, respectively] (Figure 4.1).

HIV susceptible area by area profile

Table 4.3 describes the frequencies and relative risk (RR) comparisons of HIV susceptible areas by area profile. The proportion of HIV susceptible areas progressively increased from Profiles A, B, C, and D [14.3%, 21.8%, 24.6%, and 30.8%, respectively]. To determine whether an area profile was more likely to result in a prospectively observed HIV susceptible area, two approaches were used.

The first approach examined whether a HIV susceptible area was more common under Profiles B, C, or D, when compared to Profile A, the risk environment showing the lowest proportion of HIV susceptible areas. Adjusted analysis suggests that HIV susceptible areas were more than twice as likely to occur in Profile D than Profile A risk environments [RR 2.13, 95% confidence interval (CI) (1.13, 4.00); p=0.02]. A relative risk ranging from a 13% to a fourfold comparative increase was also reasonably compatible with the data. Profiles B and C showed a 51% [p=0.15] and 55% [p=0.17] increase, respectively, though the range in possible relative risk values included both negative and positive associations (Table 4.3).

The second approach explored whether an area profile could identify risk environments with potential for micro-epidemics of new HIV infections. It compared the difference in proportion of HIV susceptible areas between an area profile and the grand mean proportion of HIV susceptible areas across all profiles (i.e., the weighted national-level area average). Using the deviation of Profile A from the grand mean as reference, the risk of an HIV susceptible area was 43% greater in Profile D risk environments in adjusted analysis [RR 1.43, 95% CI (1.03, 1.97); p=0.03]. Increased risk, ranging from 3 to 97% were also compatible with the data. Profiles B and C showed only 1% [p=0.92] and 4% [p=0.83] increases in risk, respectively, with potential risk comparisons showing both negative and positive associations (Table 4.3).

Observed area prevalences of risk factors and HIV susceptible areas

In Figure 4.2, aggregation of the observed micro-level risk factor indicators was done to calculate the average area-level prevalences of each risk factor. The aggregate measures were assessed for bivariable and multivariable associations with HIV susceptible areas. Of the seven risk factors (i.e., sexual debut, two or more partners, casual partner, unknown partner HIV status, never having used a condom, detectable viremia, and being a female ages 18-24 years), only one aggregate risk factor (prevalence of unknown partner HIV status) was significant [p=0.045] in bivariate analysis, albeit showing a marginal relative risk with HIV susceptible areas [RR 1.01, 95% CI (1.00, 1.02)]. The other six aggregate risk factors demonstrated statistically null relative risk values, ranging from 0.99 to 1.01.

In analysis adjusting for baseline aggregated mean area gender ratio, area marital status, and area residence, no aggregate risk factor showed a significant association with HIV susceptible areas, with relative risk values ranging from 0.99 to 1.01 (Figure 4.2). Results were similar when adding baseline aggregated mean area age to the model and when in a separate model adjusting for analogous aggregated demographics measured among only HIV-seronegative adults.

Discussion

This study provides evidence that the concurrence of risk factors– and not levels of uncontrolled HIV infection alone – in small areas may differentiate risk in population-level HIV incidence. Using an area typology of HIV acquisition risk, places with the highest detectable viremia combined with lower sexual risk had the least likelihood of HIV seroconversions. Yet the risk of a new infection was doubled in areas with average detectable viremia and sexual risk, when coupled with the highest levels of having a partner with unknown or HIV-positive status. This finding was shown in the most severe national HIV epidemic through a large, nationally representative sample with rigorous biomarker testing, high participation, and exceptional retention.

The association between a unique HIV risk environment profile and HIV susceptible areas may be due to the geographic clustering of known individual-level risk factors in this study population. The high risk partner status-moderate viremia profile (i.e., Profile D) described places with more than double and triple the prevalences of unknown partner HIV status, when compared to other profiles. These places were also higher in proportions of HIV-seronegative persons reporting a HIV-positive partner. The clustering of these

individual-level risk factors is meaningful, given that knowledge of partner HIV status was the sole predictor of HIV seroconversion in men and (besides marital status) women, in a multivariable analysis of demographic and sexual risk factors.[13] Profile D appeared to pinpoint places with social norms that hamper HIV-seronegative adults from learning the status of their partners or from protecting themselves against HIV from partners with unsuppressed viremia. Stigma, discrimination, and gender inequality may have differentially precluded disclosure in these social environments or limited access to HIV testing services.[16-21] Given the doubled risk of new infections when compared to lower sexual risk-highest viremia areas (i.e., Profile A) and the context of comparatively average – yet high at 24% – viremia, Profile D areas appeared less likely to enable HIV disclosure discussions and, perhaps, viral suppression among known HIV-positive partners, resulting in significantly greater susceptibility to HIV acquisition.

Alternatively, population mobility may also explain the differences in HIV susceptible areas across profiles. Mobility and migration are common in southern Africa[22] and have long been putative drivers of HIV in the region.[23-25] They are shown to limit HIV diagnosis, ART adherence, retention in care, and prevention of HIV acquisition.[26-28] Since migration and mobility were not measured or included in analysis models, they may independently explain incidence variation across areas. Mobility may, however, predict risk environments, and if so, could be reflected in the area profiles. For example, the proportion of men ages 35-49 years in Profile D was half that of the national average and less than a third of Profile A. Population mobility may have reduced the availability of this group of men in communities, creating a social environment less conducive for couples to discuss HIV testing and status disclosure. While mobility is surmised to increase risk of ongoing transmission among migrants, some evidence also indicates that high risk behaviors, such as extra-marital sex, is less associated with labor-related mobility among men and more likely among women whose partners live elsewhere.[29,30] As married women whose partners lived elsewhere independently predicted HIV seroconversion in this study population,[13] the absence of men ages 35-49 years may have facilitated the likelihood of extra-marital sex among wives, thereby enabling an environment susceptible to incident HIV. While the role of mobility is unclear, the clustering of absent men in risk environments possibly predisposed to new infections may imply value in understanding its affect on HIV risk environments, as population mobility patterns may independently predict HIV risk environment profiles, if not also HIV susceptible areas overall.

This study also found that when micro-level risk factor indicators were aggregated at the area level, none of the risk factor area prevalences were independently associated with HIV susceptible areas. Thus, no informative pattern in incident infection could be discerned from the variability of individual area mean risk estimates. Yet underlying the area-specific distributions of risk factors, a meaningful pattern in HIV incidence was observed when investigating sets of risk factors. The seven observed micro-level risk factor indicators were significant predictors of area profile, resulting in four indicator combinations occurring beyond expectations of chance. These combinations revealed statistically significant clustering of multiple and casual partnerships in Profile C, of unknown partner HIV status in Profile D, and of detectable viremia in Profile A; and the grouping of unknown partner HIV status in tandem with high detectable viremia, average sexual risk levels, and select demographic shifts appeared associated with areas experiencing HIV seroconversions. Overall, these findings suggest that aggregate area-level risk factors in this population may simply summarize the individual risks of aggregated adults – rather than the constructs of environmental HIV risk – without providing much new understanding to HIV incidence.[31] Methods detecting statistically significant combinations of risk factors, on the other hand, may more adequately capture the heterogeneity in risk due to area-level clustering, with greater precision in determining where the next incident HIV infection may occur.

This study has several strengths. The outcome is based on directly observed HIV seroconversion, the gold standard in incidence measurement, with rigorous biomarker testing. The findings are from a large, nationally representative sample of HIV-seronegative adults characterized by high participation and strong retention, lending to strong external validity. Latent variable estimation followed strict parameters, providing results with very low risk of misclassification of and moderate distinction between area profiles, modeled local dependencies between indicators, and strong assurance of identifying the global maximum solution of the model (Supplement Two). The method also estimated sampling and measurement errors, which cannot be accounted for in summary aggregation, thereby strengthening the statistical validity of these results. However, the prevalence of detectable viremia was a single baseline assessment and therefore assumed a stable measure during the six-month follow-up period.

In conclusion, area HIV risk profiles suggest that the clustering of risk factors, in tandem with high levels of detectable viremia, may heighten the susceptibility of places to incident HIV. Area-level mean

estimates of risk factors were not independently associated with area-level incidence; and places with the highest levels of detectable viremia did not have the highest risk of prospectively observed HIV seroconversion. Instead, places with 28% of HIV-seronegative adults reporting a partner with unknown HIV status had more than twice the risk of an incident infection, when contrasted to areas with only 8% reporting partners with unknown status yet also a 25% elevation in detectable viremia. This implies that in similar settings with moderate differences in uncontrolled HIV infection, treatment-induced viral suppression may not independently trigger large reductions in population-level incidence.

Demographic or high risk behavior sub-populations differentially at risk of ongoing transmission may hinder the success of TasP interventions.[6,10] The notion that select sub-groups could maintain epidemics, plus the evidence that area environments can distinguish HIV acquisition risk, may mean that higher risk sub-populations driving epidemic trends could be defined by place, rather than people. Contextual determinants, such as economic disadvantage and gender inequality, create the social and physical environments that enable individual risk behavior and modulate how persons experience risk and access risk management resources.[20,32] Thus, the risk embedded in environments of places, rather than in individuals, may be a more appropriate target for interventions. A greater understanding of the contextual determinants of high risk environments, as may be defined through area risk profiles, could be particularly informative.

Additionally, given the finding that geographic clustering of less HIV disclosure and less detectable viremia appeared to render places more susceptible to new HIV infections, it may be that optimal treatment requires the augmenting of other prevention interventions, in order to achieve the population-level equivalent of undetectable equals untransmittable.[1] Since substantial incidence reductions were not achieved in TasP trials, despite high ART coverage and viral suppression, a latent pattern of other HIV risk factors may help explain the observed pattern in incidence. Future research on how best to tailor comprehensive combination prevention interventions for populations at risk may prove fruitful.

This analysis used an innovative approach to describing the HIV risk of places and its relationship to variation of incidence in populations. By distinguishing geographic risk profiles that may predict incident HIV infection, HIV programs may be better positioned to develop interventions interfacing with places, not

simply people, at risk of HIV. This shift in approach may help formulate new strategies to achieving target reductions in new infections with eventual control of the global epidemic.

Tables and figures

***Table 4.1**. Participant characteristics*

	Males*		Females*		Total*	
	%	n	%	n	%	N
HIV-seronegative adults [N=11,856]						
	51	6026	49	5830	100	11,856
Age [N=11,856]						
18-24 years	47	2821	42	2475	45	5296
25-34 years	33	2018	27	1598	31	3616
35-49 years	20	1187	30	1757	25	2944
Marital status [N=11,460]						
Not married, never had sex	22	1268	9	519	16	1787
Not married, ever had sex	48	2783	40	2244	44	5027
Married, living with partner	22	1254	30	1710	26	2964
Married, partner elsewhere	9	515	21	1167	15	1682
Education [N=11,802]						
Never attended	5	300	5	314	5	614
Primary	26	1550	26	1527	26	3077
Secondary	51	3085	53	3081	52	6166
Tertiary	18	1065	15	880	16	1945
Residence [N=11,735]						
Rural	70	4184	70	4015	70	8199
Urban	30	1776	30	1760	30	3536
Ever had sex? [N=11,087]						
Yes	82	4577	92	5017	87	9593
No	18	1036	8	457	13	1494
Number of partners† [N=9573]						
Two or more partners	19	859	2	93	10	952
One partner	64	2902	83	4163	74	7065
No partners	18	803	15	753	16	1555
Any casual partner? † [N=8686]						
Yes	14	570	2	110	8	680
No	86	3551	98	4455	92	8006
Knowledge of partner HIV status† [N=7974]						
At least one with unknown HIV status‡	12	459	11	458	11	916
At least one HIV-positive	10	387	12	518	11	906
All HIV-negative	77	2896	77	3257	77	6152

Condom use† [N=7851]						
Never	30	1125	42	1751	37	2876
Sometimes	34	1245	34	1397	34	2641
Always	36	1330	24	1004	30	2334
HIV-seronegative and HIV-seropositive adults [N=18,172]						
HIV viremia prevalence¥ [N=18,172]						
Detectable (≥20 viral copies/milliliter)	19	1544	29	2852	24	4396
Undetectable	81	6785	71	6991	76	13,776

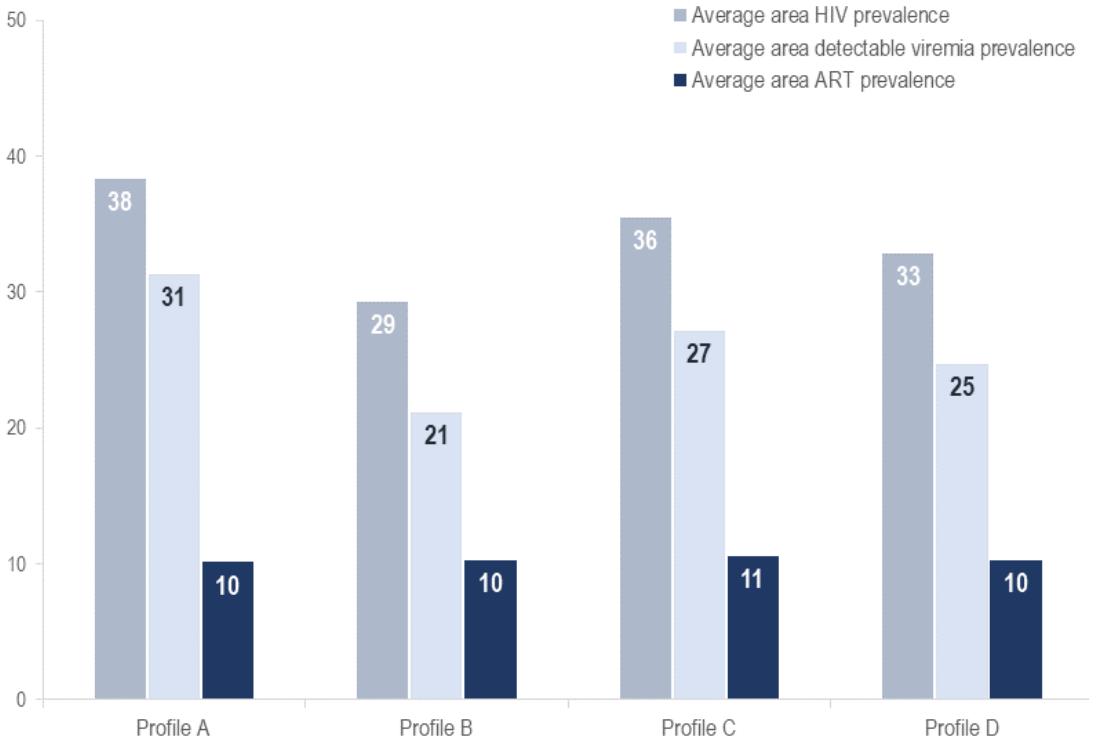
* Column totals may not total to 100% due to rounding error
† Reported in the prior six months
‡ No reported HIV-positive partners
¥ Baseline prevalence among HIV-seropositive and HIV-seronegative persons in sampled study areas

***Table 4.2**. Characteristics of area profiles*

	Profile A		Profile B		Profile C		Profile D	
	%	95% CI*	%	95% CI*	%	95% CI*	%	95% CI*
Area profile prevalence	15	(12-18)	56	(52-60)	17	(14-20)	12	(10-15)
Ever had sex?								
Yes	94	(93-96)	84	(83-85)	91	(89-93)	85	(82-87)
No	6	(4-7)	16	(15-17)	9	(7-11)	15	(13-18)
Number of partners†								
Two or more partners	10	(9-11)	8	(7-9)	17	(15-20)	9	(8-10)
One partner	75	(74-76)	73	(72-74)	74	(73-75)	74	(73-75)
No partners	15	(14-17)	19	(17-20)	9	(7-10)	17	(15-19)
Any casual partner? †								
Yes	5	(3-6)	6	(5-7)	17	(15-20)	9	(6-11)
No	96	(94-97)	94	(93-95)	83	(80-85)	91	(89-94)
Knowledge of partner HIV status†								
At least one unknown HIV status‡	8	(6-9)	8	(7-9)	13	(11-15)	28	(23-32)
At least one HIV-positive	10	(9-11)	10	(10-11)	12	(11-13)	16	(15-17)
All HIV-negative	82	(80-85)	82	(80-83)	75	(72-78)	57	(51-62)
Condom use†								
Never	42	(38-45)	39	(37-41)	28	(25-31)	33	(29-37)
Sometimes	33	(32-34)	34	(33-35)	34	(33-35)	34	(33-35)
Always	25	(22-28)	28	(26-29)	38	(35-41)	33	(29-37)
HIV viremia prevalence¥								
Detectable (≥20 copies/milliliter)	31	(28-34)	21	(20-22)	27	(25-29)	24	(22-27)
Undetectable	69	(66-72)	79	(78-80)	73	(71-75)	76	(73-78)
Gender-age group								
Female, 18-24y	19	(16-21)	22	(21-23)	17	(14-19)	24	(22-27)
Female, 25-34y	16	(14-18)	13	(12-14)	13	(11-15)	14	(12-16)
Female, 35-49y	15	(12-17)	16	(15-17)	9	(8-11)	15	(12-17)
Male, 18-24y	13	(11-15)	26	(24-27)	26	(23-28)	26	(23-29)
Male, 25-34y	19	(17-22)	15	(14-16)	25	(22-28)	16	(13-18)
Male, 35-49y	18	(15-20)	9	(8-10)	11	(9-13)	5	(4-7)

* Confidence interval
† Reported in the prior six months
‡ No reported HIV-positive partners
¥ Baseline prevalence among HIV-seropositive and HIV-seronegative persons in sampled study areas

***Figure 4.1**. Average prevalence of HIV, detectable viremia,* and antiretroviral treatment (ART) by area profile*



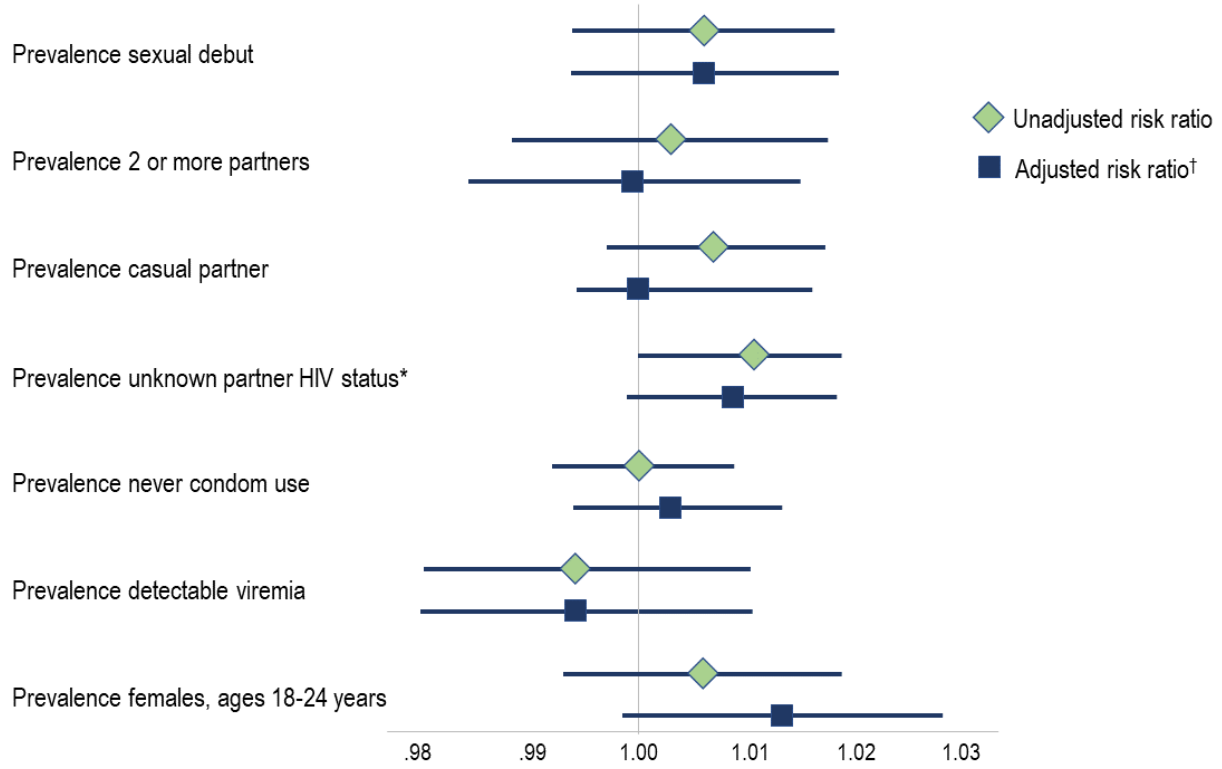
* Prevalence of ≥20 HIV viral copies/milliliter among all persons, regardless of HIV status

***Table 4.3**. Generalized linear regression of HIV susceptible areas on area profile*

	Total areas (N=575)	HIV susceptible areas	Comparisons in relative magnitude of risk				Comparisons of deviations from the grand mean			
Profile	# (%)	# (%)	Unadjusted risk ratio (95% CI*)	p-value	Adjusted risk ratio (95% CI*)	p-value	Unadjusted risk ratio (95% CI*)	p-value	Adjusted risk ratio (95% CI*)	p-value
A	86 (15.0)	14 (14.3)	ref	ref	ref	ref	ref	ref	ref	ref
B	323 (56.1)	70 (21.8)	1.53 (0.88-2.64)	0.13	1.51 (0.86-2.65)	0.15	0.99 (0.78-1.26)	0.93	1.01 (0.79-1.30)	0.92
C	96 (16.7)	25 (24.6)	1.73 (0.94-3.18)	0.08	1.55 (0.83-2.87)	0.17	1.12 (0.82-1.52)	0.48	1.04 (0.75-1.42)	0.83
D	70 (12.1)	22 (30.8)	2.15 (1.16-4.00)	**0.02**	2.13 (1.13-4.00)	**0.02**	1.40 (1.02-1.92)	**0.04**	1.43 (1.03-1.97)	**0.03**

* Confidence interval

***Figure 4.2**. Unadjusted and adjusted risk ratios of the association between observed area-level prevalences of risk factors and HIV susceptible areas*



* Unadjusted risk ratio, p=0.045

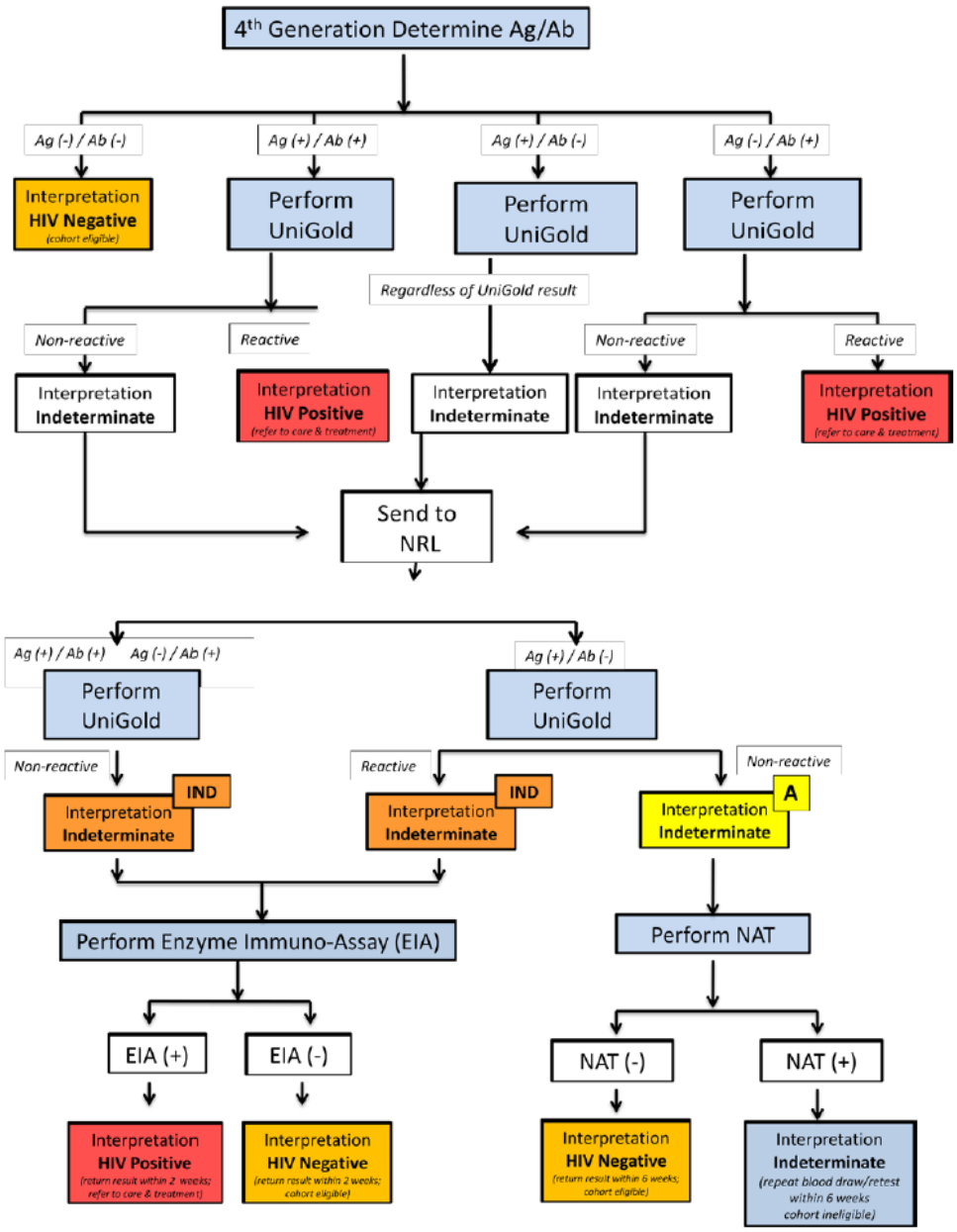
† Adjusted for area gender ratio, area marital status, and area residence

Supplements

Supplement One

This supplement provides the HIV test algorithm for determining final HIV status of participants at baseline and endline measures.

Supplemental Figure S4.1*. Swaziland HIV incidence measurement survey (SHIMS): serial HIV test algorithm*



Supplement Two

This supplement provides additional methodology considerations used to estimate the area profile latent class model.

Latent model estimation

Maximum likelihood (ML) estimation was conducted using the expectation-maximization (EM) algorithm and Newton-Raphson algorithm to identify the most likely model parameter values. Estimation was initiated with the EM algorithm with a maximum number of 250 iterations and a tolerance/convergence index (i.e. a stopping rule when estimates maximize or nearly maximize the likelihood function) of 0.1. If the EM algorithm failed to identify a model, estimation continued with the Newton-Raphson algorithm, using a maximum of 50 iterations and tolerance of 1.0×10^{-8}, a conservative criterion.[33]

To improve the likelihood of estimating the global maximum solution, 50 random sets of start values were used to initiate model estimation.[34] The number of iterations per start values set was increased to 50 to maximize the chances of avoiding a local maximum solution[33] and a common threshold convergence criterion of maximum absolute deviation (MAD) ≤.000001 to define when the change between successive iterations was sufficiently small and therefore near the ML solution.[34]

ML estimation assumed data as missing at random and computed the average effect from available data.[35] This approach to "ignore" missingness was appropriate given the infrequency of missing values: less than 2% of demographic data; 1.5% or less of HIV viremia, gender-age, and casual partner; and 6-7% of all other indicators.

Assessing latent model fitness

Model assessment and final selection was based upon statistical criteria, parsimony, and model interpretability.[34] Macro-level models ranging from one to six clusters (i.e., area profiles) were estimated. Bayesian information criterion (BIC) is appropriate for identifying the number of latent groups.[36] BIC values progressively decreased from one to four cluster models and subsequently increased among five and six cluster models. The four cluster model showed the lowest BIC value and potentially useful distinctions between indicators and therefore was selected (Supplemental Table S4.1). The entropy R-squared ($R^2_{entropy}$) value for the model was 0.62, indicating adequate separation of clusters. A good value is greater than 0.70.[35]

***Supplemental Table S4.1**. Model fit statistics for latent cluster models*

	Parameters	BIC	LL
1 cluster	14	117,777	-58,820
2 cluster	26	117,530	-58,638
3 cluster	38	117,419	-58,523
4 cluster	50	117,410	-58,460
5 cluster	62	117,433	-58,413
6 cluster	74	117,475	-58,374

Note: BIC is based on log likelihood (LL) and number of clusters

Testing prediction of the latent model's dependent variable

For the final 4-cluster macro-level model, each indicator was tested on whether it was a statistically significant predictor of the clusters (i.e. area profiles), using the Wald test statistic and assuming a chi-square distribution (Supplemental Table S4.2). Of note, when estimating the models, Latent Gold used the largest sized cluster, Profile B, as the reference cluster. All indicators were significant (<.001) and were included in the model.

***Supplemental Table S4.2**. Multinomial logistic regression of HIV risk factor indicators on area profile*

	Profile B	Profile A		Profile C		Profile D		Wald statistic	p-value
		OR*	95% CI	OR*	95% CI	OR*	95% CI		
Ever had sex?	ref	3.27	(2.44-4.38)	1.93	(1.54-2.41)	1.07	(0.84-1.38)	95.76	<.001
Number of partners	ref	1.23	(1.07-1.41)	2.17	(1.85-2.55)	1.12	(0.95-1.32)	89.95	<.001
Any casual partner?	ref	0.78	(0.53-1.14)	3.42	(2.70-4.34)	1.57	(1.06-2.31)	135.25	<.001
Knowledge of partner HIV status	ref	0.96	(0.81-1.14)	1.30	(1.15-1.47)	2.22	(1.97-2.51)	177.34	<.001
Condom use	ref	1.09	(0.96-1.23)	0.72	(0.65-0.80)	0.85	(0.74-0.98)	54.84	<.001
Viremia prevalence (≥20 c/mL†)	ref	1.69	(1.46-1.95)	1.38	(1.20-1.57)	1.21	(1.02-1.44)	62.02	<.001
Gender-age groups									
Female, 18-24y		ref		ref		ref		225.15	<.001
Female, 25-34y	ref	1.49	(1.16-1.90)	1.32	(1.05-1.66)	0.97	(0.77-1.23)		
Female, 35-49y	ref	1.03	(0.82-1.29)	0.76	(0.60-0.96)	0.80	(0.63-1.00)		
Male, 18-24y	ref	0.60	(0.46-0.77)	1.32	(1.07-1.63)	0.91	(0.73-1.12)		
Male, 25-34y	ref	1.55	(1.20-1.99)	2.25	(1.80-2.81)	0.98	(0.78-1.23)		
Male, 35-49y	ref	2.33	(1.79-3.02)	1.62	(1.21-2.16)	0.55	(0.38-0.79)		

* OR: odds ratio

† c/mL: copies per milliliter

Note: Latent Gold software by default uses the latent class (i.e. area profile) with the greatest prevalence as the reference category.

***Supplemental Table S4.3**. Pairwise comparisons of area profiles by risk factor indicator*

Area Profile	A	B	C	D
A		Ever had sex: <.001 Number partners: 0.004 Casual partner: **NS** Knowledge partner HIV status: **NS** Condom use: **NS** Detectable viremia: <.001 Gender-age group: <.001	Ever had sex: 0.005 Number partners: <.001 Casual partner: <.001 Knowledge partner HIV status: <.001 Condom use: <.001 Detectable viremia: 0.022 Gender-age group: <.001	Ever had sex: <.001 Number partners: **NS** Casual partner: 0.003 Knowledge partner HIV status: <.001 Condom use: 0.001 Detectable viremia: 0.001 Gender-age group: <.001
B	B<A, ever had sex B<A, 2+ partners Casual partner: **NS** Knowledge partner HIV status: **NS** Condom use: **NS** B<A, detectable viremia B>A, female 18-24y		Ever had sex: <.001 Number partners: <.001 Casual partner: <.001 Knowledge partner HIV status: <.001 Condom use: **NS** Detectable viremia: <.001 Gender-age group: <.001	Ever had sex: **NS** Number partners: **NS** Casual partner: 0.024 Knowledge partner HIV status: <.001 Condom use: 0.021 Detectable viremia: 0.031 Gender-age group: 0.022
C	C<A, ever had sex C>A, 2+ partners C>A, casual partner C>A, unknown partner HIV status C<A, never use condom C<A, detectable viremia C<A, female 18-24y	C>B, ever had sex C>B, 2+ partners C>B, casual partner C>B, unknown partner HIV status Condom use: **NS** C<B, detectable viremia C<B, female 18-24y		Ever had sex: <.001 Number partners: <.001 Casual partner: <.001 Knowledge partner HIV status: <.001 Condom use: 0.035 Detectable viremia: **NS** Gender-age group: <.001
D	D<A, ever had sex Number partners: **NS** D>A, casual partner D>A, unknown partner HIV status D<A, never use condom D<A, detectable viremia D>A, female 18-24y	Ever had sex: **NS** Number partners: **NS** D>B, casual partner D>B, unknown partner HIV status D<B, never condom use D>B, detectable viremia D>B, female 18-24y	D<C, ever had sex D<C, 2+ partners D<C, casual partner D>C, unknown partner HIV status D>C, never use condom Detectable viremia: **NS** D>C, female 18-24y	

* c/mL: copies per milliliter

Analysis plan

Analysis entailed a three-step estimation procedure.[37] First, a measurement model for area profile was estimated using nonparametric random-effects multinomial logistic regression to allow latent variable sizes to differ.

Latent class analysis is based on the assumption of local independence, where independence between indicators is assumed when conditioned on area profile. The assumption was checked by assessing for whether bivariate residuals (BVRs) were larger than 1. Residuals ranged from 0.93 to 1.4 indicating independence between indicators within area profiles was a reasonable assumption.

Pairwise comparisons of the area profiles were conducted using the Wald test statistic to test for the statistical significance of each indicator in discriminating between area profiles (Supplemental Table S4.3). A total of 42 pairwise comparisons were conducted, including six for each indicator, with the majority (83%, 35/42) showing a statistically significant differentiation between area profiles. Comparisons showed that Profiles A and B did not statistically differ in probabilities of casual partnerships, knowledge of partner HIV status, and condom use. Similarly, Profiles A and D did not statistically differ in probabilities of sexual debut and number of partners; Profiles B and D did not differ in probabilities of number of partners; and Profiles C and D did not differ in prevalence of detectable viremia.

In step two, the observed survey data was assigned to the final selected area profile measurement model using posterior class membership probabilities. Enumeration areas (EAs) were assigned to area profiles using $\boldsymbol{Z}$ scores for macro-units. Using a modal rule, EAs were assigned to profiles with the highest frequency. Misclassification error was assessed by conducting a cross-tabulation of the modal by the probability assignments. The largest error in area profiles may have misclassified 4.4% (25/575) of EAs as Profile B areas (Supplemental Table S4.4).

Lastly in step three, generalized linear regression was used to estimate model parameters of the relationship between area profiles and HIV susceptible areas.

***Supplemental Table S4.4**. Cross tabulation of enumeration area (EA) assignment rules to area profiles*

Number of assigned EAs using modal rule					
	B	A	C	D	Total
B	262	12	9	10	293
A	18	75	6	1	100
C	11	5	77	4	97
D	22	2	5	54	84
Total	313	95	97	70	575

Number of assigned EAs using proportional rule					
	B	A	C	D	Total
B	236	21	13	23	293
A	21	69	7	3	100
C	13	7	69	7	97
D	23	3	7	51	84
Total	293	100	97	84	575

Difference in number of EA assignments between modal and proportional rules [# (% of 575)]				
	B	A	C	D
B	25 (4.4)	-9 (-1.5)	-4 (-0.7)	-12 (-2.2)
A	-3 (-0.5)	7 (1.2)	-2 (-0.3)	-2 (-0.3)
C	-2 (-0.4)	-2 (-0.4)	7 (1.3)	-3 (-0.5)
D	0 (-0.1)	-1 (-0.2)	-1 (-0.2)	3 (0.6)

References

1. Eisinger RW, Dieffenbach CW, Fauci AS. HIV viral load and transmissibility of HIV infection: undetectable equals untransmittable. JAMA 2019;321:451-2.

2. UNAIDS. 90-90-90 an Ambitious Treatment Target to Help End the AIDS Epidemic. Geneva: Joint United Nations Programme on HIV/AIDS; 2014.

3. UNAIDS. Fact sheet - World AIDS Day 2019. Geneva: Journal United Nations Programme on HIV/AIDS; 2019.

4. Global Health Collaborators. Global, regional, and national incidence, prevalence, and mortality of HIV, 1980-2017, and forecasts to 2030, for 195 countries and territories: a systematic analysis for the global burden of diseases, injuries, and risk factors study 2017. Lancet HIV 2019;6:e831-e59.

5. Hayes RJ, Donnell D, Floyd S, et al. Effect of universal testing and treatment on HIV incidence - HPTN 071 (PopART). N Engl J Med 2019;381:207-18.

6. Abdool Karim SS. HIV-1 epidemic control - insights from test-and-treat trials. N Engl J Med 2019;381:286-8.

7. Makhema J, Wirth KE, Pretorius Holme M, et al. Universal testing, expanded treatment, and incidence of HIV infection in Botswana. N Engl J Med 2019;381:230-42.

8. Havlir DV, Balzer LB, Charlebois ED, et al. HIV testing and treatment with the use of a community health approach in rural Africa. N Engl J Med 2019;381:219-29.

9. Iwuji CC, Orne-Gliemann J, Larmarange J, et al. Universal test and treat and the HIV epidemic in rural South Africa: a phase 4, open-label, community cluster randomised trial. Lancet HIV 2018;5:e116-e25.

10. Baral S, Rao A, Sullivan P, et al. The disconnect between individual-level and population-level HIV prevention benefits of antiretroviral treatment. Lancet HIV 2019;6:e632-e8.

11. Rhodes T, Simic M. Transition and the HIV risk environment. BMJ 2005;331:220-3.

12. Rhodes T. The 'risk environment': a framework for understanding and reducing drug-related harm. Int J Drug Policy 2002;13:85-94.

13. Justman J, Reed JB, Bicego G, et al. Swaziland HIV incidence measurement survey (SHIMS): a prospective national cohort study. Lancet HIV 2017;4:e83-e92.

14. Bicego GT, Nkambule R, Peterson I, et al. Recent patterns in population-based HIV prevalence in Swaziland. PLoS One 2013;8:e77101.

15. Dixon JR. The International Conference on Harmonization Good Clinical Practice Guideline. Quality Assurance 1999;6:65-74.

16. Antelman G, Smith Fawzi MC, Kaaya S, et al. Predictors of HIV-1 serostatus disclosure: a prospective study among HIV-infected pregnant women in Dar es Salaam, Tanzania. AIDS 2001;15:1865-74.

17. Medley A, Garcia-Moreno C, McGill S, Maman S. Rates, barriers and outcomes of HIV serostatus disclosure among women in developing countries: implications for prevention of mother-to-child transmission programmes. Bull World Health Organ 2004;82:299-307.

18. Chaudoir SR, Fisher JD. The disclosure processes model: understanding disclosure decision making and postdisclosure outcomes among people living with a concealable stigmatized identity. Psychol Bull 2010;136:236-56.

19. UNAIDS. Miles to Go: Closing Gaps, Breaking Barriers, Righting Injustices. Global AIDS Update 2018. Geneva: Joint United Nations Programmes on HIV/AIDS; 2018.

20. Gupta GR, Parkhurst JO, Ogden JA, Aggleton P, Mahal A. Structural approaches to HIV prevention. Lancet 2008;372:764-75.

21. Kalichman SC, Simbayi LC. HIV testing attitudes, AIDS stigma, and voluntary HIV counselling and testing in a black township in Cape Town, South Africa. Sex Transm Infect 2003;79:442-7.

22. International Organization for Migration. World Migration Report 2020. Geneva: International Organization for Migration; 2019.

23. Coffee MP, Garnett GP, Mlilo M, Voeten HA, Chandiwana S, Gregson S. Patterns of movement and risk of HIV infection in rural Zimbabwe. J Infect Dis 2005;191 Suppl 1:S159-67.

24. Glynn JR, Ponnighaus J, Crampin AC, et al. The development of the HIV epidemic in Karonga District, Malawi. AIDS 2001;15:2025-9.

25. Jochelson K, Mothibeli M, Leger JP. Human immunodeficiency virus and migrant labor in South Africa. Int J Health Serv 1991;21:157-73.

26. Olawore O, Tobian AAR, Kagaayi J, et al. Migration and risk of HIV acquisition in Rakai, Uganda: a population-based cohort study. Lancet HIV 2018;5:e181-e9.

27. Tanser F, Barnighausen T, Vandormael A, Dobra A. HIV treatment cascade in migrants and mobile populations. Curr Opin HIV AIDS 2015;10:430-8.

28. McGrath N, Eaton JW, Newell ML, Hosegood V. Migration, sexual behaviour, and HIV risk: a general population cohort in rural South Africa. Lancet HIV 2015;2:e252-9.

29. Vissers DC, Voeten HA, Urassa M, et al. Separation of spouses due to travel and living apart raises HIV risk in Tanzanian couples. Sex Transm Dis 2008;35:714-20.

30. Camlin CS, Akullian A, Neilands TB, et al. Population mobility associated with higher risk sexual behaviour in eastern African communities participating in a universal testing and treatment trial. J Int AIDS Soc 2018;21 Suppl 4:e25115.

31. Diez-Roux AV. Bringing context back into epidemiology: variables and fallacies in multilevel analysis. Amer J Public Health 1998;88:216-22.

32. Blankenship KM, Friedman SR, Dworkin S, Mantell JE. Structural interventions: concepts, challenges and opportunities for research. J Urban Health 2006;83:59-72.

33. Vermunt JKaMJ. Latent Gold 4.0 User's Guide. Belmont, Massachusetts: Statistical Innovations; 2005.

34. Collins LMaLST. Latent Class and Latent Transition Analysis: With Applications in the Social, Behavioral, and Health Sciences Boston, Massachusetts: John Wiley and Sons, Inc; 2009.

35. Vermunt JKaMJ. Technical Guide for Latent Fold 5.1: Basic, Advanced, and Syntax. Belmont, Massachusetts: Statistical Innovations, Inc.; 2016.

36. Lukociene O, Varriale R, Vermunt JK. The simultaneous decision(s) about the number of lower- and higher-level classes in multilevel latent class analysis. Sociol Methodol 2010;40:247-83.

37. Bennink M, Croon MA, Keuning J, Vermunt JK. Measuring student ability, classifying schools, and detecting item bias at school level, based on student-level dichotomous items. J Educ Behav Stat 2014;39:180-202.

Chapter Five. Conclusion

This dissertation provides evidence that an area typology of HIV risk can explain variations in population-level HIV outcomes. Risk factors spatially cluster in predictable, meaningful combinations that can inform the ranking of small area geographies into risk categories. Using a risk environment framework to define risk categories, area profiles describing the contextual risk of HIV transmission and acquisition in small areas correlated with area prevalences of detectable HIV viremia. Also, area profiles of the contextual HIV acquisition risk among seronegative populations predicted areas susceptible to prospectively observed HIV seroconversions. These findings were evaluated in a large, nationally-representative sample from 2011 in Eswatini, a country marked with the highest HIV prevalence and incidence. High levels of participation and retention, robust biomarker quality, and the direct observation of HIV seroconversions lend greater validity to these results.

A review of the literature in Chapter 2 showed an evidence base mainly of lower quality studies and inconsistent HIV viral exposure measures. The overall findings, however, supported a relationship between population-level HIV viral exposure and new HIV infections. Increasing levels of suppressed HIV in HIV-infected populations were associated with fewer new infections over time. Although population exposure was measured using 14 numerators among five different groups, the most robust findings were determined in studies of household-based and directly sampled populations, where population viremia (i.e. a summary measure of HIV viral quantity among all persons, regardless of HIV status) in high prevalence populations was consistently associated with HIV incidence; population viral load (i.e., a summary measure of HIV viral quantity among only HIV-positive persons) did not show an association in high prevalence, general populations and was inconsistent in key populations.[1-3]

The review highlighted a gap where only one low quality study incorporated a spatial aspect to HIV viral exposure.[4] Additionally, randomized cluster trials conducted after the review failed to definitively show a population-level impact of treatment as prevention in reducing population HIV incidence.[5-7] Core group hypotheses were posited that a disproportionate number of HIV infections in the trials were due to a small sub-group of hard to reach and high-risk populations and that the observed population viral load reductions obscured differences in HIV transmission and acquisition risks across these sub-groups.[8,9] Given the inadequate knowledge in the heterogeneity and spatial elements of population HIV risk, Chapters 3 and 4 applied a spatial framework to investigating area-level differences in HIV viral quantities and new infections.

Conceptual model of environmental influences on population HIV outcomes

Figure 5.1 illustrates a spatial model of small area risk environments in Eswatini in 2011. The model comprises three levels of environmental influence. First, the supermacro environment encompasses superstructural factors that are distal to the biologic transmission and acquisition of HIV. It encompasses broad structural factors, such as cultural values and economic development that affect large populations (e.g., nations or global regions) and policies and laws that affect portions of those populations. Collectively, these factors can selectively subject sub-populations to gender inequality, for example, and potentially determine levels of geographically bound environments of HIV risk. Second, the macroenvironment encompasses small area factors that are less distal yet still distant from biologic HIV infection. It affects small areas and shapes the conditions and resources available to individuals residing in those areas. And third, the microenvironment encompasses individual factors that are more proximal to biologic HIV infection. It affects individuals and encompasses how the macroenvironmental factors are experienced.[10-14]

Chapter 3 focused on small area, macroenvironments of HIV transmission-acquisition risk (Figure 5.1). The macroenvironments were categorized into three area profiles that likely evolved from the prevailing superstructural conditions. Each profiled macroenvironment fostered unique patterns in HIV behavioral norms and service access, resulting in six HIV risk types defined by distinct combinations of individual-level HIV risk factors (i.e., HIV testing and ART uptake, number of sexual partnerships, knowledge of partner HIV status, and condom use). At the microenvironmental level, the HIV risk types characterized how adults, ages 18-49 years, engaged in the macroenvironmental conditions of where they lived and suggest an individual's propensity for HIV transmission and/or acquisition risk.

Chapter 3 analyses evaluated whether an area profile of HIV transmission-acquisition risk environments, when categorized by the composite prevalences of HIV risk types, correlated with area levels of HIV viral quantities. Evidence of such a relationship could clarify the macroenvironmental conditions that promote HIV core areas, or areas with high concentrations of uncontrolled HIV infection. HIV core areas, the spatial analogue of core groups, are likely sources of ongoing HIV transmission[15,16] and could operate as a population mechanism for new HIV infections.

High acquisition/transmission areas – locations with the largest proportions of high acquisition, very high transmission, and high transmission risk type adults – were more likely to have the highest prevalences

of HIV viral measures in Eswatini. The high acquisition/transmission profile may indicate geographic risk environments that enable high risk behaviors and produce potential HIV core areas. The emergence of HIV core areas, in relation to the proportional frequency of HIV risk types, may be due to two possible population mechanisms between micro and macro levels of HIV infection.

First is risk type clustering (Figure 5.1). The behavioral norms and HIV service access characterizing high acquisition/transmission area environments may have promoted not only the individual-level clustering of risk factors among HIV risk types but also the area-level clustering of risk types among profiled macroenvironments. The three highest risk types (i.e., high acquisition, very high transmission, and high transmission) combined in the greatest proportions in areas profiled as high acquisition/transmission areas; and the co-location of the largest concentrations of high acquisition and high transmission adults may have fueled HIV prevalence over time. This contextual effect of prevalent HIV risk types could explain the durability of HIV prevalence “hotspots” in endemic settings.[17]

Second is risk type crossover between highest risk types (Figure 5.1). High acquisition and very high transmission types are analogous profiles, mainly separated by HIV status. Their similarity may suggest an unseen shift in membership between the highest risk population sub-groups upon HIV acquisition. This transition between high risk types may not only support the geographic stability of core membership[18,19] but also amplify area HIV viral quantities towards formation of HIV core areas, because of the greater possibility of undiagnosed infection and HIV transmission characterizing these risk types.

Chapter 4 focused on small area, macroenvironments of HIV acquisition risk that were characterized into four area profiles. The risk environments of HIV acquisition and of HIV transmission-acquisition are related, since the former is inherent to the latter (Figure 5.1). However, each HIV acquisition area profile reflected macroenvironmental conditions that shaped the demographic distribution and behavioral norms on sexual behavior and disclosure, specifically among HIV-seronegative adults residing in the small areas. They also shaped the area viral environments where HIV-seronegative adults lived, as measured through the area prevalence of detectable HIV viremia among all area residents, regardless of HIV status. Since the area-specific measurement of HIV core areas and of area viral environments are the same in this analysis, HIV core areas could represent a subset of area viral environments.

Chapter 4 analyses evaluated whether an area profile of HIV acquisition risk environments predicted places with a HIV seroconversion, or HIV susceptible areas. The profiles described the most common composites of contextual, macroenvironmental risk factors, specifically the demographic distribution and prevalences of sexual risk behaviors and disclosure among HIV-seronegative adults and the area prevalence of detectable HIV viremia. Although uncontrolled HIV infection is a necessary factor for areas to produce incident HIV, results showed that the concurrent levels of other risk factors may have a greater role in population-level incidence than previously considered. Places with the highest detectable viremia when combined with lower sexual risk had the least likelihood of HIV seroconversions. Yet the risk of a new infection was doubled in areas with average detectable viremia and sexual risk, when coupled with the highest levels of having a partner with unknown or HIV-positive status. The relationship between a unique HIV risk environment profile and HIV susceptible areas may relate to the geographic clustering of known individual-level risk factors in the study population, namely knowledge of partner HIV status. Also, population mobility was not measured in this study, but results imply a possible role in decreasing the presence of older men in the community, which may have influenced the social norms promoting protection against HIV acquisition.

Program implications

Dissertation findings show that area environments can distinguish population-level HIV outcomes, suggesting that sub-populations driving epidemic trends could be defined by place. The risk embedded in the environments of places, rather than in individuals, may be a more appropriate target for interventions. Such a shift in program strategies would recognize not only individuals but also area populations as the unit of intervention and places as an interface to stemming disease.

Since places susceptible to incident HIV were characterized by lower detectable viremia and norms hindering HIV disclosure, it may be that optimal treatment requires better implementation of prevention interventions. Amplifying the full spectrum of combination HIV prevention may be needed to achieve the population-level equivalent of undetectable equals untransmittable.[20] In fact, results in Chapter 3 indicate that risk environments could be altered through HIV service delivery. Findings showed that clustering of HIV risk types that utilized any HIV service, whether prevention or treatment, correlated with lower area viral load quantities; and, the uptake of HIV testing and treatment appeared to promote the clustering of

preventive factors among risk types. A consistent pattern in transmission risk reduction occurred across the very high, high, and moderate transmission risk types, alluding to a continuum of preventive behavior uptake among HIV-positive persons that could be leveraged in the design of area-focused combination prevention packages.

The prominence of macroenvironmental risk factors – and the supermacro conditions that may determine them – is underscored through the HIV-related community norms and practices captured by the area profiles. The demonstrated relationships between profiled risk environments and population HIV outcomes reveal a potential for area-focused structural interventions. Community sensitization campaigns designed to deter harmful community norms and practices and promote those that are protective could play a larger role; and similarly, integrating strategies and interventions focused on area-specific structural factors, such as gender and economic inequalities, may also warrant greater attention. Combination HIV prevention approaches augmented with structural interventions matched to area risk environments may increase the specificity of current public health efforts by targeting the spatial, rather than individual, risk patterns associated with geographic clusters of HIV viremia and incident HIV.

Future directions

Measuring area profiles of HIV risk environments may explain how micro and macro level risk factors produced population mechanisms of HIV disease in Eswatini. Further investigation is warranted to learn whether similar area profiles and HIV risk types are prevalent in other endemic settings and if they exhibit comparable relationships to HIV core and susceptible areas. Recent data across several sub-Saharan African countries[21] could help validate these findings in generalized epidemics and clarify whether risk factors other than knowledge of partner HIV status, in tandem with viremia levels, can differentially predict where the next HIV infection may occur. Also, a follow-on survey in Eswatini was conducted in 2016,[22] lending an opportunity to assess whether area profiles of composite prevalences of HIV risk types varies over time. Similarly, the transportability of the identified area risk profiles and HIV risk types from generalized to concentrated epidemics requires investigation.

Phylogenetic data could enhance the interpretation of area risk profiles. For example, the mapping of phylogenetic strains against HIV risk types could help understand whether high HIV transmission risk types are more likely to transmit HIV to very high or high HIV acquisition risk types. Similarly, phylogenetic

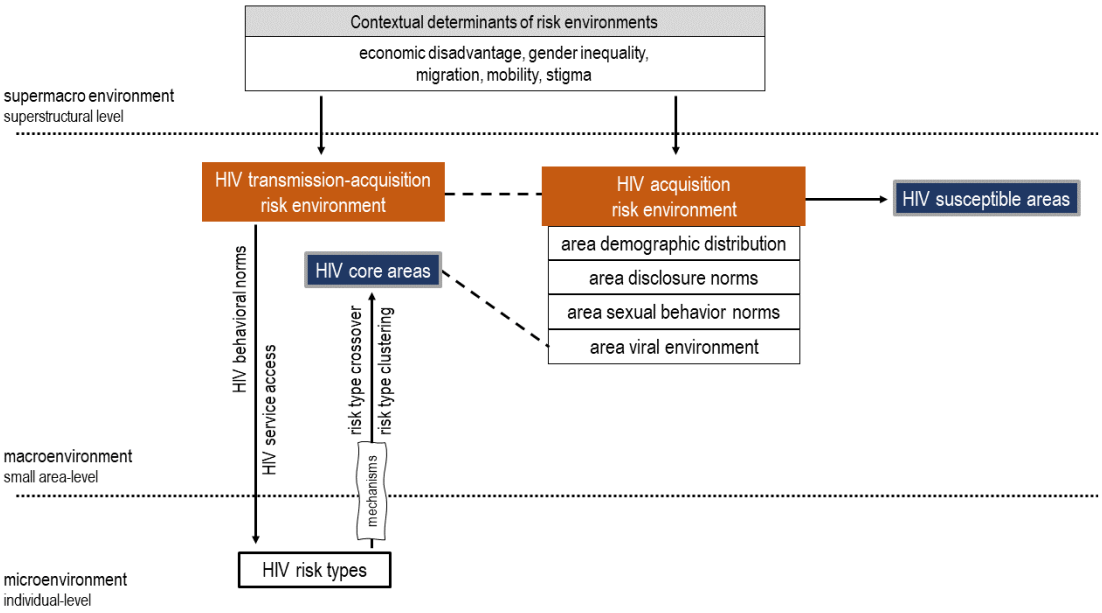
data could inform whether certain area risk profiles are more or less likely to be the source of ongoing HIV transmission.

Additional research can also explore if other indicators better capture the heterogeneity in risk associated with area-level clustering. Since micro-level risk factors vary by setting, other individual-level indicators may carry greater predictive ability when estimated into composite prevalences. Superstructural or macroenvironmental level factors may also be ideal indicators to better understand the determinants of geographic clustering of HIV risk types and of micro-level clustering of risk behaviors.

The focus of combination HIV prevention on multiple risk domains and high-transmission geographies[23] is an area-based approach to intervention targeting. Yet, recent interventions were predominantly aimed at individual-level factors more proximal to infection, rather than at population patterns in HIV risk.[5-7,24,25] Implementation science research can help identify methods to translate area profiles into intervention packages across behavioral, biomedical, and structural domains that contour against spatial risk patterns. Non-HIV specific strategies addressing contextual determinants of those risk patterns could also be integrated to complement and enhance the effectiveness of HIV-specific interventions. Applying a spatial framework to area-based combination prevention design could shift the precision in prevention strategies from individuals to the spatial risk of populations and may potentially unlock greater progress in efforts towards epidemic control.

Tables and figures

Figure 5.1. *Conceptual model of environmental influences on population HIV outcomes*



References

1. Tanser F, Vandormael A, Cuadros D, et al. Effect of population viral load on prospective HIV incidence in a hyperendemic rural African community. Sci Transl Med 2017;9.

2. Glasheen C, Johnson EO, Lorvick J, Kral AH. Measures of human immunodeficiency virus (HIV) community viral load and HIV incidence among people who inject drugs. Ann Epidemiol 2018;28:8-12.

3. Solomon SS, Mehta SH, McFall AM, et al. Community viral load, antiretroviral therapy coverage, and HIV incidence in India: a cross-sectional, comparative study. Lancet HIV 2016;3:e183-90.

4. Brantley M, Schumacher C, Fields EL, et al. The network structure of sex partner meeting places reported by HIV-infected MSM: opportunities for HIV targeted control. Soc Sci Med 2017;182:20-9.

5. Havlir DV, Balzer LB, Charlebois ED, et al. HIV testing and treatment with the use of a community health approach in rural Africa. N Engl J Med 2019;381:219-29.

6. Hayes RJ, Donnell D, Floyd S, et al. Effect of universal testing and treatment on HIV incidence - HPTN071 (PopART). N Engl J Med 2019;381:230-42.

7. Makhema J, Wirth KE, Pretorius Holme M, et al. Universal testing, expanded treatment, and incidence of HIV infection in Botswana. N Engl J Med 2019;381:230-42.

8. Baral S, Rao A, Sullivan P, et al. The disconnect between individual-level and population-level HIV prevention benefits of antiretroviral treatment. Lancet HIV 2019;6:e632-e8.

9. Abdool Karim SS. HIV-1 epidemic control - insights from test-and-treat trials. N Engl J Med 2019;381:286-8.

10. Gupta GR, Parkhurst JO, Ogden JA, Aggleton P, Mahal A. Structural approaches to HIV prevention. Lancet 2008;372:764-75.

11. Sweat MD, Denison JA. Reducing HIV incidence in developing countries with structural and environmental interventions. AIDS 1995;9 Suppl A:S251-7.

12. Barnett T WA. AIDS in the Twenty-first Century: Disease and Globalisation. Basingstoke: Palgrave Macmillan; 2002.

13. Rhodes T. The 'risk environment': a framework for understanding and reducing drug-related harm. Intl J Drug Policy 2002;13:85-94.

14. Rhodes T, Simic M. Transition and the HIV risk environment. BMJ 2005;331:220-3.

15. Jain V, Petersen ML, Liegler T, et al. Population levels and geographical distribution of HIV RNA in rural Ugandan and Kenyan communities, including serodiscordant couples: a cross-sectional analysis. Lancet HIV 2017;4:e122-e33.

16. Vandormael A, Newell M-L, Baernighausen T, Tanser F. Use of antiretroviral therapy in households and risk of HIV acquisition in rural KwaZulu-Natal, South Africa, 2004-12: a prospective cohort study. Lancet Glob Health 2014;2:E209-E15.

17. Cuadros DFG T, de Oliveira T, Bärnighausen T, Tanser F. Assessing the role of geographical HIV hot-spots in the spread of the epidemic. Oral abstract presented at: 25th Conference on Retroviruses and Opportunistic Infections. 2018 Mar 4-7; Boston, United States.

18. Thomas JC, Tucker MJ. The development and use of the concept of a sexually transmitted disease core. J Infect Dis 1996;174 Suppl 2:S134-43.

19. Fichtenberg CM, Ellen JM. Editorial response: moving from core groups to risk spaces. Sex Transm Infect 2003;30:825-6.

20. Eisinger RW, Dieffenbach CW, Fauci AS. HIV viral load and transmissibility of HIV infection: undetectable equals untransmittable. JAMA 2019;321:451-2.

21. PHIA Project [Internet]. New York: ICAP at Columbia University; 2020. [cited: 1 Mar 2020] Available from: https://phia.icap.columbia.edu

22. Nkambule R, Nuwagaba-Biribonwoha H, Mnisi Z, Ao TT, Ginindza C, Duong YT, et al. Substantial progress in confronting the HIV epidemic in Swaziland: first evidence of national impact. Oral abstract presented at: 9th IAS Conference on HIV Science; 2017 Jul 8-13; Paris, France.

23. Kurth AE, Celum C, Baeten JM, Vermund SH, Wasserheit JN. Combination HIV prevention: significance, challenges, and opportunities. Curr HIV/AIDS Rep 2011;8:62-72.

24. Iwuji CC, Orne-Gliemann J, Larmarange J, et al. Universal test and treat and the HIV epidemic in rural South Africa: a phase 4, open-label, community cluster randomised trial. Lancet HIV 2018;5:e116-e25.

25. Grabowski MK, Serwadda DM, Gray RH, et al. HIV prevention efforts and incidence of HIV in Uganda. N Engl J Med 2017;377:2154-66.