Using DHS to Identify Sites for HIV Prevention Trials

Elizabeth E Tolley, PhD[1]*; Mark A. Weaver, PhD[2]; Lisa M. Albert, MS, MPH[3]
[1] FHI 360, Durham, NC
[2] University of North Carolina-Chapel Hill, School of Medicine
[3] University of North Carolina-Chapel Hill, Gillings School of Global Public Health

Abstract

Recruiting high-risk, HIV-negative participants is challenging for HIV prevention trials. This study aimed to 1) identify and characterize subsets of high-risk women based on responses to demographic and psychosocial questions from the AIDS Indicator Survey (AIS) and 2) develop a rapid, inexpensive tool for site identification. We developed a latent class model (LCM), hypothesizing that AIS respondents could be grouped by responses to psychosocial indicators, and that these latent classes would vary by HIV status, socio-demographic, and other indicators. We tested our model on women respondents from the Tanzania 2003 AIS Survey, and replicated it in several other populations. LCM produced four classes of women who significantly varied by psychosocial indicators and HIV status. Geographic differences in HIV prevalence and class composition were observed. Our approach has the potential to provide a more systematic, inexpensive and rapid strategy to identify HIV prevention trial sites.

Keywords: HIV Prevention; Recruitment; Latent Class Models; Psychosocial Indicators; Clinical Trials

Résumé

Le recrutement de participants à haut risque de VIH, mais VIH-négatifs, est un défi pour les essais cliniques. Cette étude visait à 1) identifier et caractériser des sous-groupes de femmes á haut risque d'acquisition à partir de leurs réponses aux questions démographiques et psychosociales de l'Enquête sur les Indicateurs du SIDA (EIS) et 2) de développer un outil rapide et économique pour appuyer l'identification de lieux de recrutement. Nous avons développé un modèle à classes latentes (MCL) sous l'hypothèse que les participantes puissent être regroupées en fonction de leurs réponses aux questions psychosociales, et que les classes latentes varient en terme de statut VIH, d'indicateurs sociodémographiques, et d'autres indicateurs. Nous avons testé notre modèle en utilisant les données de l'enquête EIS conduite en Tanzanie en 2003, et l'avons ensuite répliqué sur d'autres populations. Le modèle MCL a produit 4 classes de femmes différant de manière significative en termes des indicateurs psychosociaux et du statut VIH. Des différences ont été constatées au niveau de la prévalence du VIH et de la composition des classes entre les différentes zones géographiques. Notre approche pourrait offrir une stratégie systématique, économique, et rapide pour l'identification de sites de recrutement pour les essais cliniques.

Keywords: Age distribution; marriage; fertility; mortality; net migration

---

* Elizabeth E. Tolley, Senior Scientist, Social and Behavioral Health Sciences, FHI 360, 359 Blackwell Street, Suite 200 Durham, North Carolina 27701, Tel: 919-544-7040 ext 11334, Email: btolley@fhi360.org

Introduction

Efforts to develop new HIV prevention technologies, including vaginal microbicide gels and films, oral pre-exposure prophylaxis (PrEP), and, more recently, longer-acting vaginal rings and injectables, have spanned almost three decades and included approximately 50 different pharmaceutical products formulated and/or delivered through a range of approaches (Stone and Harrison 2010). While more than a dozen products have moved from preclinical/ laboratory evaluation into human clinical trials to evaluate longer-term safety, acceptability and effectiveness, only three HIV prevention approaches, thus far, have demonstrated evidence of effectiveness. Three clinical trials of either tenofovir alone, or Truvada (tenofovir in combination with emtricitibine), concluded that the pre-exposure prophylactic (PrEP) use of ARVs taken orally could reduce HIV transmission in heterosexual discordant couples and among men-who-have-sex-with men (Grant, Lama et al. 2010, Baeten, Donnell et al. 2012). In addition, the CAPRISA 004 clinical trial in South Africa produced a proof of concept for the vaginal use of tenofovir 1% gel when used peri-coitally in reducing HIV transmission to heterosexual women (Abdool Karim, Abdool Karim et al. 2010).

Despite these recent successes, HIV prevention development efforts have faced a number of challenges along the way. One major challenge has been the ability of clinical trials implementers to recruit sufficiently high-risk but HIV-negative women (Skoler 2006). The difficulty is four-fold. First, methods for identifying geographic localities with high HIV incidence are problematic. Clinical trials may rely on pre-trial roll-in studies to calculate incidence (Ramjee, Kapiga et al. 2008, Feldblum, Latka et al. 2012) – an approach that is both costly and time-consuming. Newer and more sensitive blood tests or testing algorithms are promising (Hecht, Wellman et al. 2011, Combes, A et al. 2013), but may be expensive and pose logistical problems if individual diagnosis requires that more than one test is given, or ethical hurdles if the expectation of inflated estimates precludes individual-level diagnosis. Second, researchers' efforts to target "high-risk" populations may miss the mark. For example, while several first generation microbicide trials focused recruitment on women engaged in commercial sex work, incidence rates realized within the trials did not provide for a sensitive test of product effectiveness (Feldblum, Adeiga et al. 2008, Halpern, Obunge et al. 2009).

Some research had suggested that HIV risk may be less determined by the number of sex partners or frequency of sex, than by characteristics of the sexual relationship (Nagot, Ouedraogo et al. 2005, Karim, Kharsany et al. 2011, Ramjee 2011, Dunkle and Decker 2013). Indeed, women who acknowledge sex work may have greater ability to negotiate condom use, at least with clients and casual partners, than other women. Third, by focusing recruitment efforts on women who feel capable of adhering to the trial protocol – in other words, women who perceive greater self-efficacy - microbicide trials may recruit women who are at lower risk of HIV. Finally, trial implementers have increasingly conducted their studies in a small number of geographic areas, especially in regions of South Africa with continued high incidence rates – potentially leading to the emergence of a class of professional trial participants who may enrol in one or even multiple clinical trials without a real commitment to adhere to product use (Karim, Kharsany et al. 2011).

Literature Review and Theoretical Framework

In the past decade, Measure/Demographic and Health Survey (DHS) has introduced an HIV biomarker component into over 30 countries, including Tanzania and Côte d'Ivoire (Measure DHS 2013). We set out to determine whether use of DHS data could assist in identifying and characterizing sub-populations within a given country who might be at greater risk of HIV infection for the purpose of identifying sites for future prevention trials. Our approach is based on "market segmentation" analyses that were originally developed by marketing and business firms, but have been adopted by health communications specialists. In short, this approach uses individual data on psychosocial indicators to develop class profiles towards which to tailor their marketing or communication activities (Barry and Weinstein 2009).

The goals of this analysis were to 1) identify and characterize subsets of women at a country, region or sub-regional level who are at high risk of HIV, based on their responses to demographic and psychosocial questions currently collected by DHS; and 2) develop a rapid and inexpensive tool for identifying population segments in specified countries at high risk of HIV.

Conceptual Model

Behavioural theories identify how psychosocial factors predict health behaviours which in turn influence health status (Glanz, Rimer et al. 2008). Several theories have been applied to examining HIV risk and adoption of prevention behaviours, including the AIDS Risk Reduction Model (ARRM) (Catania, Kegeles et al. 1990), the Information, Motivation and Behaviour Change theory (Fisher and Fisher 1992) and the Risk Perception Attitudes Framework (Rimal, Brown et al. 2009). They share similar constructs. The ARRM posits that, prior to sustaining a complex behaviour like consistent condom use, individuals must first perceive or "label" themselves to be at risk, commit to changing their risky behaviour and then enact the new behaviour. Commitment requires both knowledge of prevention options and belief that the individual has the ability to engage in these new behaviours. An individual's perceived ability, or self-efficacy, to adopt new prevention behaviours is essential to enacting, and eventually sustaining them. The IMB suggests that an individual must have appropriate information about HIV transmission and prevention behaviours, strong motivation and a range of behavioural skills to successfully carry out HIV prevention behaviours like condom use. Both perceived vulnerability to HIV and normative attitudes towards HIV and prevention behaviours influence an individual's motivation for behaviour change. According to the IMB, AIDS-related information and motivation work directly through a number of behavioural skills (for example, acquiring new information, communicating with a partner or purchasing condoms), which in turn enable an individual to reduce his/her risk for HIV. As in the ARRM, self-efficacy is a critical component of behavioural skills. Finally, the Risk Perception Attitudes Framework focuses on two of the above constructs, suggesting that the strength of an individual's efficacy beliefs will influence whether or how his or her risk perceptions affect the uptake of prevention behaviours (Rimal, Brown et al. 2009). For example, when a person's efficacy beliefs and risk perception are low, s/he will not be adequately motivated to engage in new prevention behaviours, while low risk perception but strong efficacy beliefs will cause a person to engage in proactive behaviours. In contrast, a person with high risk perception but low efficacy beliefs is likely to engage in "avoidance" behaviours.

The Risk Perception Attitudes Framework has been used to identify different population segments, based on their HIV-related perceived risk and efficacy beliefs, in order to develop and disseminate communication materials tailored to the needs of each audience segment (Rimal, Brown et al. 2009). We adapt and extend this approach, but using psychosocial constructs to identify population segments with potentially different predispositions to HIV prevention. Our assumption is that the different psychosocial profiles of our population segments will be related to underlying risk profiles that predict HIV status – and can be used to better tailor recruitment strategies for HIV prevention trials.

The diagram below presents our conceptual model for this study. It suggests that population sub-groups (or, latent classes) will be characterized by different levels of HIV knowledge, perceived and demonstrated efficacy for HIV or related health behaviours, attitudes towards HIV norms and HIV risk perception, and that these classes will differ with respect to actual HIV risk. Further, socio-demographic, geographic, and household characteristics, as well as HIV risk indicators, will predict latent class membership.

Diagram 1: Latent Class Model for Population Segmentation
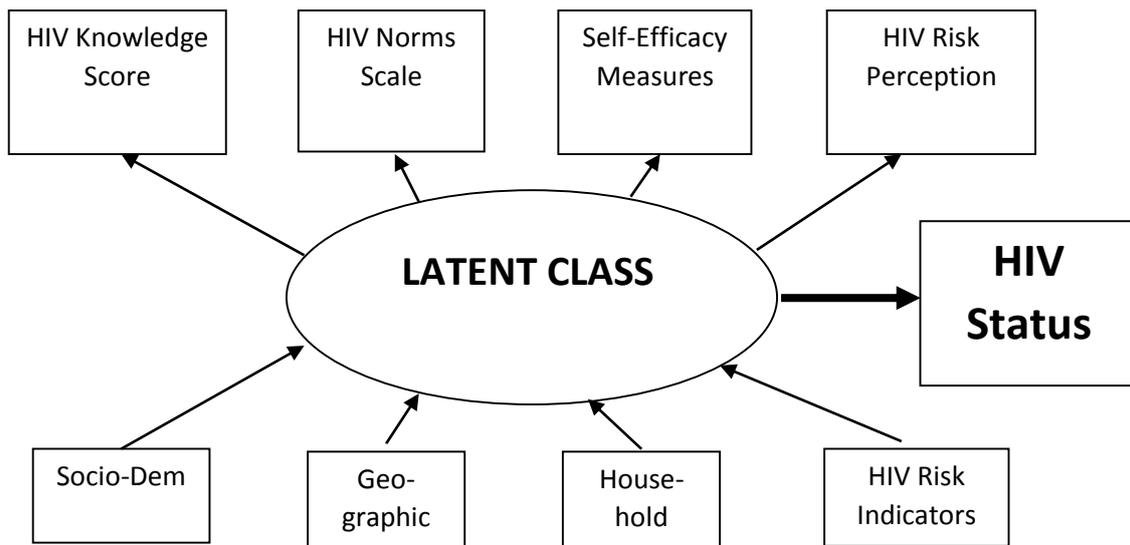
Data and Methods

Our original model was developed and tested on data from female respondents to the Tanzania 2003-04 HIV/AIDS Indicator Survey (AIS) DHS IV Release Version 0, and further replicated using Tanzania 2003-04 male data and on a second female dataset from the Côte d'Ivoire 2005 AIS DHS V Release Version 0. (Data for Tanzania males and Cote d'Ivoire females are not presented in this paper.) We attempted to further replicate the model on more recent data from the 2011-12 AIS DHS VI Release Version 1 in Tanzanian females. Below, we describe the samples and measures for the Tanzania AIS 2003-04 and 2011-12 models in women.

Tanzania AIS 2003-2004 Sample: The total sample of women in the Tanzanian DHS/AIS was 6863. A total of 1733 women were excluded for at least one of three reasons: their HIV test result was missing; they reported either never having had sex; or, they reported having been previously tested for HIV, were diagnosed as HIV-positive and knew their results. This last category of women was excluded from our sample in order to reduce the likely effect that such knowledge would have on subsequent reporting of HIV-related knowledge, attitudes and behaviours; in a sense, we were hoping to better approximate HIV incidence, as opposed to prevalence, using this restriction. Finally, 562 women who were missing data for at least one predictor variable were excluded from the model. Our final sample included 4,568 sexually active women for whom an HIV test result existed and who were either negative or did not know the result of their positive, anonymous test.

Tanzania 2011-2012 AIS Sample: We used the same algorithm to determine our 2011-12 sample. In this sample, a total of 10,299 women had both DHS survey data and AIS biomarker data. However, 2037 women were excluded because they reported either never having had sex, had a missing HIV result, or they reported having been previously tested for HIV, were diagnosed as HIV-positive and knew their test result. An additional 941 women were excluded from the model because they were missing data on at least one predictor variable. Our final model included data for 7,321 sexually active women for whom an HIV test result existed and who were either HIV-negative or did not know the result of their positive, anonymous test.

Measures:

Table 1 summarizes the variables used in our models, by year. The main study outcome was HIV status, either positive or negative. (HIV status was determined by a central laboratory, based on a standard algorithm using two different HIV antibody ELISA tests of dried blood spots collected through finger prick (Mishra, Medley et al. 2009).)

Four sets of psychosocial variables were used as indicators for the latent classes: HIV Knowledge, HIV Norms, Self-Efficacy, and Risk Perception. An HIV Knowledge score was computed by summing correct responses to 15 items in the 2003-04 AIS, but only 10 items in the 2011-12 AIS, assessing knowledge about potential modes of HIV transmission. Some examples of items include: Can people get the AIDS virus from mosquito bites? and Can people reduce their chances of getting the AIDS virus by using a condom every time they have sex?

Each correct response received one point; incorrect responses, including a response of "don't know", received 0 points.

We used exploratory factor analysis to derive a summary HIV Norms score for each respondent, using questions measuring tolerance towards HIV – infected individuals (9 items in 2003 and 6 in 2012). We extracted a single factor using principal factor methods (Hatcher 1994). Examples of HIV Norms items include: Would you shake hands with someone who is infected with the virus that causes AIDS? and If a male teacher has the AIDS virus but is not sick, should he be allowed to continue teaching in the school? Scores were standardized to range between 0 and 1 with 0 indicating no agreement with any of the items and 1 indicating agreement with all of the items.

Self-Efficacy was measured by three items, including the questions: Can you say no if you don't want to have sex?, Can you ask your partner to use a condom? (2011-12 only), and Do you know a place where you could get a HIV test if you wanted to? We also derived a score to assess the level of efficacy a respondent reported in relation to obtaining an HIV test. In 2003-04, this score was derived from two questions:: I don't want to know the results, but have you ever been tested to see if you have the AIDS virus?, and The last time you had the test, did you yourself ask for the test, was it offer to you and you accepted, or was it required? Those who had never been tested received a score of 0; if testing was required, they received a score of 1; being offered and accepting a test was scored 2; and requesting a test was scored 3. In 2011-12, the second question was replaced with On your most recent HIV test, did you get the results of your test? Those who had never been tested again scored 0, those who had been tested but did not receive the results of their most recent test scored 1, while those who did receive their most recent results scored 2.

HIV Risk Perception was measured by the question, scored 0-3, Do you think your chances of getting AIDS are small, moderate, great or no risk at all? This question was not included in the Tanzania 2011-12 AIS. Finally, standard socio-demographic, household, geographic and HIV risk indicators were drawn from the AIS dataset and are shown in Table 1.

Data Analysis

We used latent class analysis (LCA) with a distal outcome (HIV) to fit the model depicted in Figure 1, using Mplus, version 7 (Muthén and Muthén 2012). This is an exploratory modelling approach that assumes the population is comprised of an unknown number of distinct HIV risk classes. We first determined the number of latent classes using simplified models that included only the four sets of psychosocial variables and HIV status. We fit models with two to five classes and used the Bayesian information criterion, Lo-Mendell-Rubin likelihood ratio tests, and a subjective assessment of overall interpretability to select the final number of classes. Finally, we added the predictor variables to the model with the selected number of classes. Latent class analysis does not explicitly assign women to classes but rather predicts posterior class membership probabilities for each woman. To facilitate comparisons across classes, we assigned women to classes based on their highest posterior probabilities. We used analysis of variance or chi-squared tests, respectively, to compare continuous or categorical characteristics across classes. All analyses other than the LCA were conducted using SAS, version 9.2 (SAS Institute, Cary, NC).

Results

In the original analysis of AIS 2003-04 data, four classes emerged from the latent class modelling based on women's responses to HIV knowledge, norms, self-efficacy and risk perception constructs (Table 2). All six variables showed significant differences between classes (p<0.001 for each). Class membership also varied by the socio-demographic, household, geographic and risk variables that were included in the model as predictors.

Examination of the latent class and predictor variables helped us develop labels for the four classes. For example, the first and largest, or "Typical", class reported relatively high levels of HIV knowledge and tolerance towards HIV, but lower levels of self-efficacy. A second class appeared to be comprised of more "Rural, Traditional" women. They showed the lowest levels of HIV knowledge, least tolerant views towards HIV, and lowest levels of self-efficacy and risk perception. Two other classes consisted of younger, more educated, urban

women. Women in an "Urban, Traditional" group (class 3) were more likely to be married than women in class 4, labelled "Urban, Empowered". In addition, women in class 4 had fewer children, were more likely to be using contraception and to have used a condom at last sex, and all had requested an HIV test in the past.

HIV status significantly varied across the classes (p<.0001). The "Typical" class had the highest HIV prevalence rate at 8%. In contrast, less than 1% of members associated with classes 3 and 4 were infected with HIV. Our initial model verifications, using male data from the 2003-04 Tanzania AIS and the female data from 2005 Côte d'Ivoire AIS, found similar patterns, in which a more urban, empowered group had significantly lower HIV status than the typical/referent group or a more traditional rural group (data not shown).

Table 1: Model Variables

	DHS 2003	DHS 2012

HIV Knowledge: 1 composite variable
Summary score for questions assessing knowledge about potential modes of HIV transmission. Correct responses received 1 point; incorrect responses received 0 points; a response of "don't know" was scored as an incorrect response. Examples of questions:
    Can people get the AIDS virus from mosquito bites?
    Can people reduce their chances of getting the AIDS virus by using a condom every time they have sex?
	1 variable based on 15 items; response range 0-15	1 variable based on 10 items; response range 0-10

HIV Norms: 1 composite variable
Multi-item scale produced through exploratory factor analysis of questions assessing attitudes towards people living with HIV/AIDS. Examples of questions:
    Would you shake hands with someone who is infected with the virus that causes AIDS?
    If a male teacher has the AIDS virus but is not sick, shoe he be allowed to continue teaching in the school?
	1 variable based on factor analysis of 9 items; 5 items retained; α=.64	1 variable based on factor analysis of 6 items; 5 items retained; α=.49

Perceived Efficacy: 2-3 variables
    You know a place to get an AIDS test.
    You can say NO if you don't want to have sex.
    You can ask your partner to use a condom.
	2 variables	3 variables

HIV Testing Efficacy: Ordinal variable created from several questions. For DHS 2003-04, respondents received 0 if they had never tested before survey participation; 1 if testing was required, 2 if testing offered and they accepted; 3 if they requested a test on their own. For DHS 2011-12, respondents received 0 if they had never tested, 1 if tested but hadn't received results, and 2 if tested and received results.	Variable scored 0-3	Variable scored 0-2
HIV Risk Perception: Single item *Do you think your chances of getting AIDS are small, moderate, great or no risk at all?*	1 variable	Question not asked
Socio-demographic variables: Age; education; religion; marital status; total children ever born; use of effective contraception; current pregnancy status	6 variables	Religion, contra-ception missing
Household variables: Wealth index; female-headed household; head of household age; number household members	2 variables	4 variables
Geographic variables: Urban/rural	1 variable	1 variable
HIV risk indicators: Age at 1st sex; total lifetime sexual partners; sexual partners last 12 months; received any injections last 12 months; condom use last sex	5 variables	5 variables

Table 2: Latent Class Models for Tanzania AIS 2003-04 and 2011-12 Female Respondents (* Between-class comparison (but within each DHS, separately) significant with p<0.001.)

CLASS NAME	AIS 2003-04				AIS 2011-12		
	Typical (referent)	Rural, Traditional	Urban, Traditional	Urban, Empowered	Typical (referent)	Rural High Risk	Rural Low Risk
Latent Model – Psychosocial Variables	Class 1 N=2795	Class2 N=1153	Class 3 N=400	Class 4 N=220	Class 1 N=5574	Class 2 N=651	Class 3 N=1096
HIV Knowledge Score, mean	12.2 *	7.8	12.1	12.5	8.5 *	4.0	6.9
HIV Norms, mean	0.8 *	0.4	0.7	0.8	0.9 *	0.8	0.4
Know a place to get AIDS test, %	69.4 *	20.1	100	100	98.7 *	69.4	88.4
You can say NO to sex , %	77.4 *	58.2	70.0	82.2	77.3 *	44.4	55.3
Can ask a partner to use a condom, %					61.8 *	23.4	32.4
Demonstrated Efficacy – HIV testing, mean	0*	0	1.8	3	1.6 *	0.6	1.2
Chances of getting AIDS are, mean	1.1 *	0.8	1.0	1.0			
Predictor Variables							
Mean age	29.8 *	29.5	27.6	27.0	30.4	31.1	30.8
Mean education (years)	5.8*	2.6	6.3	7.3	6.6 *	2.5	3.4
Religion, %:							
Moslem	37.3 *	20.9	41.2	37.2			
Christian	59.7	57.2	56.0	60.9			
Currently married %	78.8 *	85.1	78.2	66.8	80.3 *	79.4	89.1
Use of effective contraception, %	28.3 *	5.8	26.0	28.6			
Total children ever born, mean	3.4 *	4.0	3.0	2.2	3.4 *	4.2	4.4
Age at first intercourse, mean	17.0 *	16.2	17.9	17.6	17.5 *	16.7	16.9
Total lifetime partners, mean	2.5 *	2.2	2.2	2.3	2.2 *	1.8	2.1
Used condom last sex, %	10.6 *	4.3	12.7	25.9	14.8 *	5.6	7.0
Female headed household, %	19.7 *	14.3	20.5	23.6	20.2 *	16.2	12.3
Wealth index, %:	*				*		

Poorest	15.3	41.1	13.5	6.8	12.7	38.4	30.3
Middle (3 categories)	64.2	57.0	48.0	43.2	60.0	58.8	66.8
Richest	20.5	1.9	38.5	50.0	27.3	2.8	2.9
Urban, %	26.7 *	3.9	45.0	54.0	27.1	4.6	4.5
Distal Outcome							
HIV positive, %	7.9 *	3.5	0.8	0.9	1.5 *	4.6	1.3

Tanzania 2011-12 AIS Replication

Similar to our 2003-04 analysis, the LCA identified subsets of women who differed significantly (p<0.001) by responses to all psychosocial variables. Unlike 2003-04, however, a three-class, rather than four-class, solution seemed to fit the data best. In addition, unlike 2003-04, the typical/referent class was significantly less likely to be HIV-positive compared to one predominantly rural class in which almost 5% of members were HIV- infected. The third class, which was also predominantly rural, was also lower risk, with about 1% infected.

Geographic Variation in Class and HIV Status

Table 3 compares the HIV prevalence and class affiliation of survey respondents in a subset of regions in Tanzania for the 2003-04 and 2011-2012 analyses. In 2003-04, three regions had HIV prevalence rates of 10% or higher. However, survey respondents' class affiliation varied considerably. In particular, almost 40% of respondents in the Dar es Salaam sample were affiliated with a low risk class (the Urban Traditional or Urban Empowered groups), compared to just 9% in Iringa and 13% in Mbeya. Drawing on HIV prevalence, class and regional data would suggest that Iringa and Mbeya might be better options as HIV prevention clinical trial sites than Dar es Salaam.

In the 2011-2012 analysis, Iringa and Mbeya remain two of the regions with the highest HIV prevalence rates, but Mbeya has a higher proportion of respondents affiliated with the "Rural High Risk" group than Iringa (12% versus 1%). Several other regions with relatively higher HIV prevalence rates and high risk class composition include Rukwa, Myanza, Morogoro and Njombe (locations bolded in Table 3).

Table 3: Comparisons of HIV Prevalence and Class Affiliation of Survey Respondents for Selected Tanzania Regions, for AIS 2003-04 and 2011-12 Analyses

			AIS 2003-04 Classes						AIS 2011-12 Classes			
	N	% HIV+	Typical N=2795	Rural Traditional N=1153	Urban Traditional N=400	Urban, Empowered N=220		N	% HIV+	Typical N=5574	Rural High Risk N=651	Rural Low Risk N=1096
Iringa	153	**12.4**	70%	21%	**4%**	**5%**		163	**3.7**	94%	1%	5%
Mbeya	216	**12.0**	63%	24%	**8%**	**5%**		221	**4.1**	72%	**12%**	15%
Dar Es Salaam	314	10.2	61%	2%	25%	12%		407	1.5	98%	1%	1%
Pwani	197	7.6	73%	12%	11%	4%		171	2.3	92%	4%	4%
Kilimanjaro	160	7.5	69%	12%	15%	4%		174	1.7	94%	3%	3%
Tabora	266	6.4	44%	47%	6%	3%		343	0.6	75%	7%	18%
Shinyanga	204	6.4	19%	73%	6%	2%		239	2.9	64%	12%	24%
Lindi	205	6.3	77%	13%	6%	3%		236	0.9	89%	5%	6%
Mtwara	199	6.0	88%	8%	2%	2%		216	1.4	90%	5%	5%
Rukwa	237	**5.9**	52%	43%	**2%**	**3%**		237	**3.4**	66%	**15%**	19%
Mwanza	261	**5.8**	58%	30%	7%	5%		308	**2.6**	72%	**13%**	15%
Tanga	204	5.4	61%	26%	8%	5%		225	0.9	72%	14%	14%
Arusha	193	5.2	58%	20%	13%	8%		223	2.2	70%	12%	18%
Morogoro	191	**4.7**	72%	12%	12%	4%		253	2.4	83%	**11%**	6%
Njombe	n/a	n/a	n/a	n/a	n/a	n/a		181	**3.9**	90%	6%	4%

Discussion

An important challenge in recruitment for clinical trials of new HIV prevention technologies is the need to identify populations who are at high risk of HIV – that is, likely to be exposed at some point during trial participation, but who are HIV negative. In some South African sites involved in recent microbicide trials, as many as two-fifths of prospective trial participants were not eligible because they were determined to be HIV-infected at screening (Parikh, Kiepiela et al. 2013). Consequently, the challenge is to develop recruitment mechanisms that identify high incidence but avoid high prevalence. By removing AIS respondents who had been previously diagnosed with HIV and knew their status, we believe our approach to sample selection approximates HIV incidence, thereby refining our ability to characterize population groups who are either most or least at risk due to more recent trends.

Our use of LCA for population segmentation is a highly exploratory approach. Nevertheless, our analyses were able to identify and characterize subpopulations either at lower HIV risk than the norm in 2003-04, or at higher risk in 2011-12. Despite seeming differences in the 2003-04 and 2011-12 models, there are some interesting similarities in findings. For example, lower risk classes (classes 3 and 4 in 2003-04 and class 1 in 2011-12) had high HIV knowledge and HIV norms scores; despite changes in how the testing efficacy variable was measured, lower risk groups were much more likely to know a place where one could get tested and to have gotten tested, than higher risk groups. Indeed, in the 2003-04 women's and men's samples, having requested a test oneself, rather than having just accepted a test that was offered as part of a healthcare visit, differentiated the "Urban Empowered" class from other classes. In contrast, women in classes with a higher proportion of HIV-positive members tended to be older, poorer, less educated, and more likely to reside in rural areas.

Theoretically, minimizing recruitment of individuals who fit a lower risk profile or increasing recruitment of those with higher risk profiles would improve the ability of HIV prevention trials to determine effectiveness of new products. This could potentially be accomplished by refining trial eligibility criteria to include psychosocial indicators (HIV knowledge, norms, risk perception and/or self-efficacy), in addition to the socio-demographic and risk indicators typically included on screening forms.

However, political sensitivities would also need to be considered. A more politic approach would be to identify regions with high proportions of individuals in classes associated with higher prevalence. Our LCA analysis of DHS/AIS data provides a relatively low-cost, rapid approach to augment the site selection process.

Differences between 2003-04 and 2011-12 Models

The differences in our 2003-04 and 2011-12 models may be a result both of changes in the HIV epidemic itself, as well as changes in some of the key measured variables on which the models are based. (Indeed, changes in the DHS/AIS surveys between these two periods may also be a reflection of changes in the epidemic.) Between 2002 and 2012, the number of HIV-infected persons living in Africa who were receiving antiretroviral therapy increased 150-fold. Women – particularly adult women of reproductive age were most likely to have benefited from expanded access to HIV treatment and care, given efforts to scale up prevention of mother-to-child treatment services. In Tanzania, an estimated 68% of pregnant women were tested for HIV in 2012 (World Health Organization 2013). HIV testing questions within the AIS 2011-12 women's survey were primarily anchored to testing within the context of antenatal care, likely reflecting structural changes in how HIV testing is currently implemented in Tanzania.

Despite unprecedented expansion in HIV testing and treatment access, the majority of HIV-infected individuals have never been diagnosed. Structural factors, logistical and operational issues and social factors, including stigma, continue to act as barriers to HIV-related services (World Health Organization 2013). The NIH has called for innovative strategies to identify persons unaware of their HIV infection in specific high-risk populations and settings on an on-going basis. Our 2011-12 analysis found that women with a higher risk for HIV continued to have lower HIV testing knowledge and efficacy, suggesting that some continue to be missed by provider-initiated testing and remain at risk. Finally, the most recent AIS did not include a question on HIV risk perception. Yet, individual perception of risk plays an important (albeit complicated) role in determining initiation and maintenance of prevention behaviours (de Zoysa, Sweat et al. 1996, Misovich, Fisher et al. 1997, Khawcharoenporn, Kendrick et al. 2012, Van Damme, Corneli et al. 2012, Van Damme and Corneli 2013, Corneli, Wong et al. 2013), and may also affect retention and adherence within HIV prevention clinical trials.

Unfortunately, because of these differences in the measurement of some key variables, we cannot easily explore time effects across the models. Nevertheless, there are some remarkable similarities between the results of the two models. Notably, despite changes in the epidemic, less knowledge of where to get an HIV test, lower HIV testing efficacy, and HIV knowledge in general seem to identify a rural sub-population that may be at increased risk of HIV. These lower scores on HIV knowledge and efficacy measures may reflect pervasive social and structural factors that continue to undermine HIV prevention behaviours and perpetuate risk among some populations.

Limitations

This analysis has several limitations. As we pointed out, the use of LCA for population segmentation is a very exploratory process involving relatively subjective decisions about which indicators to include, how many classes to retain and how each class should be labelled. Therefore, our analysis of HIV prevalence, class affiliation and region provides some useful information to consider when identifying potential trial sites, but should not be considered conclusive. For example, clinical trial site selection must also consider the existence and quality of physical and human resources needed to implement trials. In addition, our use of a modified HIV prevalence indicator, by removing individuals from our analysis who knew of their HIV positive status prior to DHS interview, may not be the best indicator of incidence. However, the modification does increase the likelihood that the latent classes we identified reflect the psychosocial characteristics of those who may be "at-risk" but not currently living with HIV. Finally, the analysis also has several strengths; it makes low cost use of existing data and is relatively straight-forward and easily replicated for other countries, as and when country-level biomarker data become available.

Conclusions

Given that most site identification activities currently rely on personal networks and/or intensive – but limited -- primary data collection efforts, this analysis has the potential of providing a more systematic, inexpensive, and relatively rapid approach to identifying countries, regions and sub-districts for HIV prevention trials. By comparing the models developed from different country or regional-level DHS data, it may also provide some new insights into the role that culture plays in shaping HIV risk – leading to better tailoring of recruitment strategies to the local contexts in which a trial is situated.

References

Abdool Karim, Q., Abdool Karim, S. S., Frohlich, J. A. et. al. (2010). "Effectiveness and safety of tenofovir gel, an antiretroviral microbicide, for the prevention of HIV infection in women." Science **329**(5996): 1168-1174.

Baeten, J. M., Donnell, D., Ndase, P. et. al. (2012). "Antiretroviral prophylaxis for HIV prevention in heterosexual men and women." N Engl J Med **367**(5): 399-410.

Barry, J. and Weinstein, A. (2009). "Business psychographics revisited: from segmentation theory to successful marketing practice." Journal of Marketing Management **25**(3-4): 315-340.

Catania, J. A., Kegeles, S. M. and T. J. Coates (1990). "Towards an understanding of risk behavior: an AIDS risk reduction model (ARRM)." Health Educ Q 17(1): 53-72.

Combes, S. L., G-Yohannes, A., Kidane, A. et. al. (2013). "HIV prevalence and incidence among women at higher risk of infection in Addis Ababa, Ethiopia." AIDS Res Hum Retroviruses **29**(3): 535-540.

Corneli, A., Wong, C., Agot. K. et. al. (2013). "The association between risk perception and adherence in the FEM-PrEP clinical trial [MOLBPE28]." The 7th IAS Conference on HIV Pathogenesis, Treatment and Prevention. Kuala Lumpur.

de Zoysa, I., Sweat, M. D. and Denison, J. A. (1996). "Faithful but fearful: reducing HIV transmission in stable relationships." AIDS **10**(Supplement A): S197-S203.

Dunkle, K. L. and Decker, M. R. (2013). "Gender-based violence and HIV: reviewing the evidence for links and causal pathways in the general population and high-risk groups." Am J Reprod Immunol **69 Suppl 1**: 20-26.

Feldblum, P. Adeiga, J., A., Bakare, R. et. al. (2008). "SAVVY vaginal gel (C31G) for prevention of HIV infection: a randomized controlled trial in Nigeria." PLoS One **3**(1): e1474.

Feldblum, P. J., Latka, M. H., Lombaard, J. et. al. (2012). "HIV incidence and prevalence among cohorts of women with higher risk behaviour in Bloemfontein and Rustenburg, South Africa: a prospective study." BMJ Open **2**(1): e000626.

Fisher, J. D. and Fisher, W. A. (1992). "Changing AIDS-risk behavior." Psychol Bull **111**(3): 455-474.

Glanz, K., Rimer, B. K. and Viswanath, K. (2008). Health behavior and health education : theory, research, and practice. San Francisco, CA, Jossey-Bass.

Grant, R. M., Lama, J. R., Anderson, P. L. et. al. (2010). "Preexposure chemoprophylaxis for HIV prevention in men who have sex with men." N Engl J Med 363(27): 2587-2599.

Halpern, V., Obunge, O., Ogunsola, F. et. al. (2009). "Interim data monitoring to enroll higher-risk participants in HIV prevention trials." BMC Med Res Methodol 9: 44.

Hatcher, L. (1994). A Step-by-Step Approach to Using SAS for Factor Analysis and Structural Equation Modeling. Cary, NC, SAS Institute, Inc.

Hecht, F. M., Wellman, R., Busch, M. P. et. al. (2011). "Identifying the early post-HIV antibody seroconversion period." J Infect Dis 204(4): 526-533.

Karim, Q. A., Kharsany, A. B., Frohlich, J. A. et. al. (2011). "Recruitment of high risk women for HIV prevention trials: baseline HIV prevalence and sexual behavior in the CAPRISA 004 tenofovir gel trial." Trials 12: 67.

Karim, Q. A., Kharsany, A. B., Naidoo, K. et. al. (2011). "Co-enrollment in multiple HIV prevention trials - experiences from the CAPRISA 004 Tenofovir gel trial." Contemp Clin Trials 32(3): 333-338.

Khawcharoenporn, T., Kendrick, S. and Smith, K. (2012). "HIV risk perception and preexposure prophylaxis interest among a heterosexual population visiting a sexually transmitted infection clinic." AIDS Patient Care STDS 26(4): 222-233.

Measure DHS (2013). Survey Search Results. Calverton, MD, Measure DHS.

Mishra, V., Medley, A., Hong, R. et al. (2009). Levels and Spread of HIV Seroprevalence and Associated Factors: Evidence from National Household Surveys. DHS Comparative Reports. DHS. Calverton, MD. 22: 142.

Misovich, S. J., Fisher, J. D. and Fisher, W. A. (1997). "Close Relationships and elevated HIV Risk Behavior: Evidence and Possible Underlying Psychological Processes." Review of General Psychology 1(1): 72-107.

Muthén, L. K. and Muthén, B. O. (2012). MPlus. Los Angeles, CA.

Nagot, N., Ouedraogo, A., Ouangre, et. al. (2005). "Is sexually transmitted infection management among sex workers still able to mitigate the spread of HIV infection in West Africa?" J Acquir Immune Defic Syndr 39(4): 454-458.

Parikh, U. M., Kiepiela, P., Ganesh, S. et al. (2013). "Prevalence of HIV-1 drug resistance among women screening for HIV prevention trials in KwaZulu-Natal, South Africa (MTN-009)." PLoS One 8(4): e59787.

Ramjee, G. (2011). "Antiretroviral-based Microbicides and Oral Pre-exposure Prophylaxis for HIV Prevention." Future Virology 6(8): 997-1014.

Ramjee, G., Kapiga, S., Weiss, S. et. al. (2008). "The value of site preparedness studies for future implementation of phase 2/IIb/III HIV prevention trials: experience from the HPTN 055 study." J Acquir Immune Defic Syndr 47(1): 93-100.

Rimal, R. N., Brown, J., Mkandawire, G. et. al. (2009). "Audience segmentation as a social-marketing tool in health promotion: use of the risk perception attitude framework in HIV prevention in Malawi." Am J Public Health 99(12): 2224-2229.

Skoler, S. (2006). Our Current Microbicide Trials: Lessons Learned And To Be Learned. The Microbicide Quarterly, Alliance for Microbicide Development. 4: 1-6.

Stone, A. and Harrison, P. F. (2010). Microbicides - Ways Forward. A. f. M. Development. Silver Spring, MC: 52.

Van Damme, L. and Corneli, A. (2013). "Antiretroviral preexposure prophylaxis for HIV prevention." N Engl J Med 368(1): 84.

Van Damme, L., Corneli, A., Ahmed, K. et. al. (2012). "Preexposure prophylaxis for HIV infection among African women." N Engl J Med 367(5): 411-422.

The World Health Organization. (2013). Global update on HIV treatment 2013: Results, impact and opportunities. Geneva, Switzerland.