Creating a frame: Random sampling in a non-homogeneously distributed urban migrant populations

Gayatri Singh, Brown University
Benjamin Clark, University of the Witwatersrand
Kennedy Otwombe, University of the Witwatersrand

Abstract

This paper details a research method for carrying out survey work in difficult-to-reach populations using a stratified sampling strategy as utilized in a multinational survey in Johannesburg (South Africa). The method was developed in response to a lack of baseline household distribution, unique (and often unsafe) terrain of inner city Johannesburg, need to collect methodologically sound data on social factors affecting HIV risk for urban migrants and past claims from researchers that random sampling in this context was impossible. Consequently, the success of the approach depended upon the several factors: utilization of geo-data to generate a sampling frame; mixed methods approach of the overall project; integration of the sampling strategy in questionnaire design and effective collaboration with community members. An enumeration of the method and challenges faced in the data collection is provided in this paper to demonstrate the feasibility of a random sampling approach within non-homogeneously distributed population groups.

The authors would like to acknowledge the support of the *Joint Mellon Node for Migration and HIV* for their financial support that made this project possible as well as Emthonjeni Center at the University of the Witwatersrand. Special thanks to Prof Mark Lurie, Prof Michael White and Dr. Georges Reniers for their valuable guidance on this project. Comments are welcome and should be directed to Gayatri Singh via email: **gayatri_singh@brown.edu**

*Please note that this is not a final version of the paper.

Section 1: Introduction

Over the last decade, processes of migration and their impact upon health outcomes of migrants and host populations have acquired a central place in both academic and policy discussions. In southern Africa, the spread of HIV in particular has generated much interest and important studies showing the relationship between migrancy and thespread of sexually transmitted diseases, HIV and other opportunistic infections have been carried out (Williams et al. 2002, Abdool Karim et al,1992). However, as Decosas and Adrien (1997) have pointed out, the association between migration and HIV is more likely to be a result of "the conditions and structures of the migration process than the actual dissemination of the virus along the corridors of migration." Yet very few studies have gone beyond demonstrating this connection to carrying out systematic investigation of the factors responsible for such increased vulnerability. Notable among these are those carried out by Lurie et al (2003), Campbell (2000) and Zuma et al (2003). But review of the relevant literature reveals two further gaps in research of this nature in sub-Saharan Africa: (1) the lack of methodologically robust studies focusing on the inner cities of rapidly urbanizing metros that are hubs of internal and international migration; and (2) an omission of the self-settled immigrants, particularly refugees and asylum seekers, from the research agenda on social stratification and vulnerability to infectious diseases. Given that these groups often occupy precarious social positions, especially in highly xenophobic environments, they tend live in places that are socially and physically difficult to navigate by outside researchers. Therefore, one of the biggest methodological challenges in researching such migrant populations is that of conceptualizing the relationship between migrants and the spaces they inhabit. The expectation of documented information on a population's physical distribution and orderly surveillance units for survey sampling purposes is often defied by the settlement patterns of migrant populations. Since the notion of a sampling frame is based upon some assumed knowledge about the characteristics and spatial distribution of a population, survey work with 'hidden' or difficult to access groups poses a challenge residence based sampling that requires creative solutions.

South African cities are particularly relevant to study in this regard given the massive urban population growth due to the international and internal migration since independence in early 1990s. At the same time, HIV prevalence in South Africa is among one of the highest in the world. Following the most recent antenatal survey, South Africa's Department of Health, in collaboration with UNAIDS and WHO has published an updated estimate of 18.34% prevalence in people aged 15-49 years old in 2006 and 29.2% among pregnant women. Population based Nelson Mandela/HSRC HIV prevalence survey (2002) showed that while the epidemic in South Africa is generalised throughout the population, the highest prevalence was found in those living in urban informal settlements with an HIV prevalence of 21.3%, followed by formal urban areas (12.1%), tribal areas (8.7%) and farms (7.9%). This report claimed that he higher rates of HIV prevalence in urban areas could be attributed to high levels of human mobility and repeated relocation *(Nelson Mandela/HSRC Survey, 2002)*. However, language barriers, issues of legal documentation, high levels of xenophobia and the reluctance of most research organizations to go into unsafe inner city areas results in an under-representation of foreign migrants in these surveys. Regardless of this, there has been a mushrooming of policy and advocacy oriented studies in South Africa around the issue of migrants (and recently, migrants and HIV). Unfortunately, many of these hastily formulated studies do not use very sound methodologies and have resulted in policy recommendations based on non-verifiable assumptions.[1] This is partly in response to the demands of policy makers to make available statistics on migration and partly to open doors of funding opportunities. These methodological limitations, coupled with a less than adequate engagement with the socio-economic factors underlying HIV risk and risk behaviors have led to a constrained understanding of the relationship

1 See in particular: A Minaar and M Hough, Who goes there? Perspectives on clandestine and illegal aliens in Southern Africa, Pretoria: HSRC Publishers, 1996, p. 127. This study by HSRC estimated that there were between 2.5 to 4.0 million Zimbabweans living illegally in South Africa. The figures were later recalled but continue to be cited in government documents and policy papers. Other policy directed research pertaining to migrants has come out of organizations like the South African Migration Project and International Organisation of Migration. While these organizations are more responsible about the methods used and the claims made, there is still a lack of methodological rigour in the work published.

between migration and HIV. This paper responds to both these challenges in the context of work carried out in the inner city Johannesburg.

We present a research strategy for carrying out survey work in difficult-to-reach populations using a multi-level stratified sampling strategy as utilized in a multinational survey in Johannesburg (South Africa). The survey was stratified into four migrant communities namely, Zimbabwean, Congolese (DRC), Somali and rural to urban South African migrants. This method was developed in response to a lack of baseline household distribution, unique (and often unsafe) terrain of inner city Johannesburg, need to methodologically sound data on social factors affecting HIV risk for urban migrants and past claims from researchers that random sampling in this context was impossible. The study was an eighteen month project funded by Joint Mellon Node on HIV and Migration. It took a mixed method approach carried out in two phases: qualitative data collection involving focus groups and semi structured interviews and a survey phase that collected demographic data on migration and social indicators of vulnerability to ill health in general and HIV in particular, including detailed information on sexual behavior. The prior use of qualitative methods was crucial to help us understand the relationship between migrants and the spaces they inhabited in the inner city. It played an important role in strengthening the construct validity of the concepts to be tested in the survey, especially since relatively little culturally specific information on people's perceptions and attitudes regarding health and specifically, sexual health. In the end, the success of the approach depended upon a combination of several factors, particularly the following: (i) the multi-phase, mixed methods approach of the overall project, (ii) attempts made to gain in-depth knowledge of the migrant residential patterns prior to generating a survey sample, (iii) utilization of spatial data layers to sample from and develop fieldworker maps, (iv) incorporation of the sampling strategy in the questionnaire design to ensure systematic procedures for carrying out random sampling at multiple levels (from selection of a 'residential unit' within a sampled building, to selection of households within sampled units), (v) effective

collaboration with community members and community based organizations, (vi) careful selection of a large number of surveyors who had pervious experience of working with migrant communities or who had previously lived in the areas to be sampled, (vii) iterative and participatory approach to questionnaire development with an emphasis on inputs from surveyors, (viii) intensive training of all surveyors in not only survey methods but also qualitative methods to prepare them for unexpected situations as far as possible. While each one of these considerations played an important role, here we mainly discuss the elements pertaining to the development and application of the sampling strategy, the challenges faced and some limitations of the study.

Section 2: Location and Case Selection

The choice of Johannesburg as the location of this study is not random. Of the five metropolitan cities in South Africa, it is arguably the most important a gateway city for both South African rural to urban migrants. Census South Africa 2001 figures show that 35.2% of the cities residents were born outside of the province where it is located and 6.7% of the cities population was born outside of South Africa with majority of these being from African countries. This marked a growth of 300,000 people between 1996-2001 in Johannesburg, most of which is attributed to forces of migration and urbanization (South African Cities Network 2004).

These statistics underestimate the number of foreign migrants as they are likely to count those with legal documents. These figures also fail to capture the dramatic transformation that inner city neighborhoods have been subject in the last 15 years of independence in South Africa. Indeed, the inner city Johannesburg is a unique socio-economic environment that has transitioned from being a predominantly white neighborhood prior to 1994 and undergoing swift transition into an international and internal migrant hub since 1994 (BW Ref). The residential units have deteriorated due to lack of government investment in migrant populated neighborhoods. There is a high level of overcrowding in residential

units, sex work hotspots are thriving and the prevalence of crime is high.[2] High levels of xenophobia towards foreign migrants prevail in Johannesburg in general and in the inner city in particular. HIV prevalence in Johannesburg was 29.5% in 2001 among women attending public sector antenatal clinics (National HIV and Syphilis sero-prevalence survey 2001).

The foreign nationalities selected in this survey were the top three sending countries of asylum seekers in the period of 2004-2005, when the study was being conceptualized. As can be seen from the figures in Appendix 1, numbers of asylum seekers have continued to rise steadily since 2001 with a sharp increase in the number of Zimbabweans. At the same time, South Africa has also seen a rapid increase in rural-urban migration from within its own borders. BW REF The South African internal migrant group was included for two reasons- one, as a control group of sorts to help disentangle the effects of nationality and legal status on behavioral patterns, perceptions and socio-economic factors influencing health; two, to study the difficulties faced by internal migrants in successfully completing the socio-economic aspects of urban transition. In addition, the focus of the studies with migrants in Johannesburg, especially self-settled foreign migrants, has been predominantly in livelihoods and xenophobia (Speigel, 2004, Landau, 2004). With an exception of a few studies, there is also an unfortunate tendency in the migration studies literature to speak about 'migrant communities' as homogeneous groups who, by virtue of being mobile populations are exposed to almost similar risks resulting in "one size fits all" type interventions (Fenigold, 2007, Gonzalez, 2002). While the qualitative work with migrants in Johannesburg (and elsewhere) has tried to contextualize socio-economic conditions and individual or community behaviors within cultural frameworks, small sample size of these qualitative studies seldom allows them to achieve the advantage of comparability to test whether the proposed argument is group specific or more generalizable. We sought such comparative

2 See Legget (2003) and Palmary et al. (2003) for more on crime in Johannesburg. Also see, Christopher (2005) on the question of ghettos developing in South Africa's inner cities.

advantage in our decision to look at these four migrant groups, which have diverse migration and settlement histories, language proficiencies, legal status and religious affiliations. Further, the application of this methodology on multiple groups highlights its potential to be applied to different social groups as well as its limitations in certain instances, as will be discussed later.

Section 3: Challenges to the development and implementation of survey strategy

Despite the ideal location to carry out the study from the perspective of the questions we were attempting to answer, designing an effective survey and its administration was going to be challenging for several reasons. Unlike many fieldwork situations, the development of this strategy was not as much based on what we knew about our sampling frame but instead the acute awareness of the lack of reliable knowledge about it. Mainly, we expected to face three challenges: one, the difficulties of accessing migrant communities included in the study- not only to get their agreement to respond to our questions but also the difficulties in visiting residences in a household based approach; two, the theme of the survey; and three, unique urban environment of the survey site (inner city Johannesburg). In this section, we discuss each of these in order to familiarize the reader with the specific context of our study that informed our choices.

(a) Problems of access:

While carrying out groundwork for this study, we found little concrete information on migration status of people living in the inner city Johannesburg. From the qualitative phase we knew that migrants often tended to live in communities close to each other, in overcrowded units inhabited by multiple households as well as in exploitative rental situations. In some cases access to residential units had to be negotiated with the building owner. An added problem was that a number of absentee building owners had hired a notorious security agency called "Bad Boy'z", known for their aggressive stance and suspicion of outsiders gathering information from residents of the buildings under their supervision.

Worse still, many of the migrant populated buildings in the inner city had been the target of the city administration's urban 'clean-up' zeal in the last few years. Determined to make Johannesburg a "World Class City" by 2030 and more immediately in preparation for the 2010 Soccer World Cup, municipal authorities have undertaken draconian eviction measures.[3] As a result, raids by the Johannesburg Metropolitan Police Department the South African Police Service (SAPS) (aimed at capturing undocumented immigrants), Home Affairs and municipally contracted Wozani Security (popularly known as the Red Ants due to their conspicuous red overalls and surreptitious early morning operations) had become commonplace.[4] Along with these circumstances, the existing socio-political context of urban South Africa, typified with high levels of xenophobia has made foreign migrants particularly reluctant to engage in the public sphere. Many of these migrants, especially new arrivals from Zimbabwe, are often undocumented and understandably apprehensive of giving out intimate details about their migration experiences. This climate of fear and paranoia about the motives of those approaching to "ask questions", carrying the survey research was undoubtedly going to be a challenge.

(b) Theme of the survey:

The survey contained detailed questions pertaining to migration information, living conditions in Johannesburg, sexual behavior and attitudes towards HIV/AIDS. High levels of stigma attached to HIV/AIDS coupled with an environment of mistrust as described above could prove detrimental to the successful completion of the survey. We were informed by key informants from community organizations during the grant proposal writing stage as well as in the qualitative phase that a survey planned by a prominent international

3 Cohere report

4 There have been several claims by organizations such as Center for Applied Legal Studies at the University of the Witwatersrand (Johannesburg) and South African investigative journalist teams (such as Carte Blance and Special Assignment) that Johannesburg city government is trying to flush out the inner core of the city of its poorer populations to encourage economic investment. See: http://www.law.wits.ac.za/cals, http://www.abahlali.org/node/110. Also, Special Assignment, South African Broadcasting Corporation, 25th April 2006, http://www.sabcnews.com/specialassignment/poorscript.html

organization in Johannesburg on the theme of HIV and Migration had to be pulled out from the field after rumors spread that its main purpose was to identify international migrants from Africa with HIV and deport them to their home countries.[5] This meant that not only did we have to be very careful in not gaining the reputation of "an HIV testing study" but also to dissociate ourselves from engagement with government bodies. The qualitative fieldwork and community mapping exercise (which we describe later as the backbone of our sampling strategy) became crucial for achieving this trust from the community.[6]

(c) Urban terrain of inner city Johannesburg

Finally, the difficulty to develop a sampling frame was exacerbated by the unique terrain of the survey site, inner city Johannesburg, where both foreign and South African migrants often inhabit dilapidated buildings and are usually missed out by the censuses or electoral rosters. While it is true that neighborhoods falling under the inner city boundaries are populated by foreign and internal migrants, their populations are not evenly distributed in a manner that it is easy to assign them to 'townships' or 'enumerator areas' that serve as city administration and census boundaries. Further, land utilization in the inner city in practice does not correspond with municipal residential or business use allocations, making it difficult to classify the buildings on the basis of their official zoning status. To add to the lack of information, prior surveys conducted with migrants in the same setting had left out one of the key suburbs - Hillbrow- that was deemed unsafe. In our case, it was impossible to avoid Hillbrow as this is where most Zimbabwean immigrants have recently located themselves. Badly maintained and at times dilapidated buildings and high levels of crime presented a challenge that required extra safety precautions for the surveyors.

5 While this cannot be confirmed by any written source, we were told that a UNHCR supported HIV KAP study had to be withdrawn from the field due to rumors regarding the institutional targeting of the immigrants by the government in cahoots with the refugee protection agency. This information provided by an executive committee member of Coordinating Body for Refugee Communities as well as a staff member of Zimbabwean Torture Victims Project, NGO organizations in Johannesburg.

6 Additionally, we were careful to embed within the survey, mechanisms for correct health service information and free psychological counseling service provision, graciously offered by the University of Witwatersrand's Emthonjeni Center for the migrant communities.

All this information about the difficulties we would face was valuable but it did not offer us any solutions for developing a sampling frame. Literature on research methodology also spoke of challenges in these contexts but did not offer good fixes. Conversations with those who had previously carried surveys in the inner city made us realized that none of the conventional approaches (e.g. using lists of residential units from the housing departments, sampling within enumerator areas etc) would worked (Jacobsen and Landau 2003, Vigneswaran, 2007). However, since random sampling is one of the few ways to eliminate biases that non-probability or other convenience samples are subject to, we wanted to pursue its feasibility. We did not want to go the route of randomly sampling from names in clinic records, NGO lists, church rosters etc. they are only capturing respondents exhibiting particular characteristics and interests(Carter-Edwards et al. 2002, Gritz et al 1992). Snowball sampling is suggested when working with difficult to access populations. But no matter how carefully and meticulously it is done, it is not easy to attain a fully unbiased sample as well as ensure that sensitive information about the respondent does not filter to other network members (Sommers, 2001, Jacobsen et al 2003). Finally, we decided to attempt 'creating' a sampling frame by undertaking a 'migrant community foot mapping' exercise.

Section 4: 'Creating' a Frame:

A Community Mapping Sheet was prepared and given to multiple fieldworkers who carried out mapping on foot based on their own knowledge, our inputs from the information generated during the qualitative phase as well as consultation with other community members and community organizations. On the sheet we asked the field workers not only to specify the streets and cross streets of the where migrants were 'usually observed', but also break it down by type of structure (residential, business and other). Each one of these three categories was further specified. If residential, then the field workers were asked to record whether the structure was a high rise building, a stand alone house or a

shack; if business, was it formal or informal; and if other, a description of the activity taking place. We also saw this as a good opportunity to inform the migrant communities about the survey being undertaken, in order to minimize hostility towards surveyors and suspicion regarding the survey motives.[7] The Community Mapping Sheets from different fieldworkers were reconciled with each other and further augmented using member residential lists from NGOs and churches (where possible).

The end product of this migrant community mapping exercise were printed maps of inner city Johannesburg with streets highlighted where each migrant group was 'known' to reside. Our main reason for collecting information on all the places where migrants could usually be found (and not just residential areas) was to have a back up strategy for finding respondents in case the random sampling method did not work for any number of unforeseen circumstances. Another advantage of this mapping was that we were not constrained by arbitrary politically or administratively defined boundaries while creating the sampling frame (such as wards, tracts, enumerator areas used in censuses etc.). This provided a more meaningful understanding of the social meaning of the space occupied by different migrant communities by allowing us a better knowledge of clustering, evenness of spread, residential segregation or overlap between various groups' residential patterns. Interestingly, the mapping information we received did not seem to suggest too much clustering (in three out of four communities) of the nature as had been lamented by other surveys in the inner city as a reason for the failure of their sampling strategies (Jacobsen and Landau 2003, Vigneswaran (2007).

During this exploratory phase, we found out about an inner city 'building footprints' geo-database from a Johannesburg based urban development consultancy group, on the agreement that we would only use this information for

7 This was important given that there was at least one previous instance, which we were aware of, where a UNHCR supported HIV KAP study had to be withdrawn from the field due to rumors regarding the institutional targeting of the immigrants by the government in cahoots with the refugee protection agency. Based on information provided by Coordinating Body for Refugee Communities.

purposes of research and no commercial activities.. This database had several layers of spatial data and included information on primary and secondary land usage for all structures in this area. A GIS was created from this database that consisted of a building footprint layer for the inner city neighborhoods in the Urban Development Zone from the above database (covering most of the inner city neighborhoods that we were focusing on)[8]; plot layers from City of Johannesburg corporate GIS (CoJ GIS) to map any portion of the area falling outside of the UDZ boundaries and an accompanying street layer. Migrant community specific sampling polygons were then developed based on the maps generated in the community mapping exercise. While we did not attempt to impose any administrative boundaries on the polygons, we found that in many cases, they roughly followed existing neighborhood boundaries. The 'building footprints' and plots that fell within these polygons were then extracted to form the population of residential structures that we could sample from[9].

Most of the area covered by the Kagiso geo-database allowed us to specify built structures whose primary or secondary land use was residential, in the case where 'plots' were used (instead of footprints, from CoJ GIS information), we only deleted plots that were very clearly not residential (typically, parks or other open areas). Each building footprint or plot was tagged with the migrant community of the polygon it fell into. This meant that it could be in multiple migrant community sampling polygons as some of them overlapped. A database routine was then used to randomly sample the layer for each migrant group with a 50% over sample. A 50% over sample was generated to account for any lacunae in accurate data on building or plot type and the inability to distinguish between residential and commercial structures in cases where the land use was both commercial and residential. We added significant buffer zones to the areas highlighted based on the migrant community foot mapping exercise, to minimize

8 There is no consensus on the areas that fall within or constitute the 'inner city'. UDZ boundaries is one of the ways in which the city administration defines the inner city boundaries. For this survey, we used the migrant community populated areas that fell within loosely defined boundaries of the inner city as our reference.

9 While the building footprint layer had attribute data that allowed for the exclusion of most type of non-residential building footprints the plot layers did not.

the key informant knowledge bias. All surveyors were further trained to ask for respondents belonging to any of the specified national groups in every sampled unit visited, irrespective of whether or not it was in a neighborhood identified by the foot mapping exercise as having migrants from a particular community.

After the sample was drawn, a unique identifier was given to each sampled structure. Finally, survey maps were developed, containing the building foot prints (and in a few cases, plots), street names and sample buildings that were highlighted and marked with an identifying number (Figure 1 below). The unique identifiers were to be recorded on the cover sheet (location cover as well as roster sheets- see Appendix 2) of each interview collected in order to keep track of the sampling process such that the teams did not over-sample a building. These maps played a key role in guiding the surveyors to the sampled structures and also to monitor the progress of the survey. Most importantly, they became a tool for the field team leaders of each team to coordinate the members and create a method of checks to ensure their safety at all points.[10] The maps essentially broke down the neighborhoods into manageable and easy to navigate sub-sections for the fieldworkers instead of having them work with census maps of enumerator areas that did not show street names or structures. Field workers teams were divided by nationality and as a starting point sent into sampled areas which had been identified by the community mapping as areas with higher likelihood of finding migrants of particular nationalities to boost the morale by increasing the possibility of securing successful interviews. This was important in our case as the fieldworkers were being paid per interview collected due to budget constraints. So not only did we have to come up with a method for keeping them motivated but at the same time ensuring that the quality of the surveys collected was not compromised due to the nature of payment. Section 6 speaks about the latter concern a little more. Related to this, we also demonstrate the weakness of this methodology with respect to the Somali community arising from budget constraints (particularly, the payment method) and highly clustered nature of their settlement.

10 Many measures were taken to ensure the safety of fieldworkers that are not discussed in detail in this version of the paper.

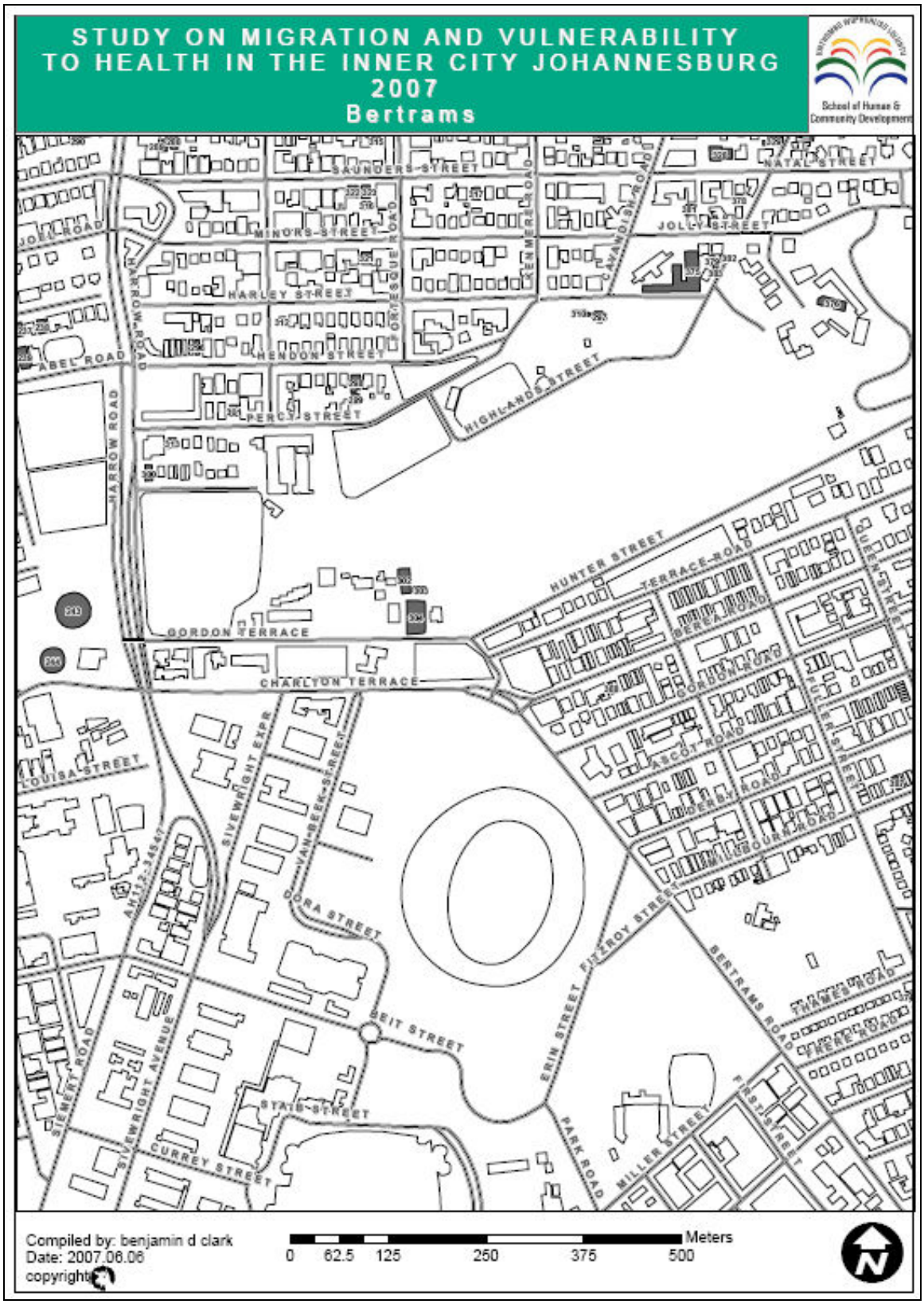


Figure 1: Field Map showing numbered and shaded samples

Section 5: Multiple types of structure and problem of multiple households

The structures sampled in the manner described above represented the first stage of the random sampling. However, these structures were not all of the same kind. We expected to encounter a combination high rise buildings, free standing houses, houses with semi-detached or detached backyard cottages and shelters or boarding houses (single or multi-storied). While a free standing house could be considered a single residential unit (but not a household unit), a high rise building necessarily contained many sub-residential units. In the same way, free standing house with a rental backyard cottage represented the presence of two units on a sampled structure. Hence, the second stage involved sampling within the identified structure to get to a smaller sampling unit.

In addition to this, each of these residential units was likely to have multiple households (families, independent unrelated individuals or a combination thereof) living in the same unit. For instance, in the neighborhood of Hillbrow, we typically had high rise buildings with 5-20 floors. On each floor, there were multiple residential units (apartments, in this case) and within each of these units, multiple families or single persons or a combination of the thereof. In many cases each room of the unit (apartment) was rented out to an individual or family or be sub-divided into several temporary partitions of ply-wood or curtains. As a result, we needed to come up with a strategy that could adapt to this kind of structural and residential complexity.

Added to this was challenge of the in conceptualizing a 'household' when carrying out interviews. The problem of conceptualizing households under one definition is an obstinate theoretical problem and not a new one (Messer 1983). As Gurney and Omolalu noted as far back as 1971, defining co-residence itself may be challenging when many independent units fall in one structure. This issue is even more exacerbated when dealing with multi-national and multi-cultural

communities due to subjective cultural understandings of what a household (or even a family) means. Different durations and patterns of settlements of these migrant communities complicated the matter even more. For instance, in Somali community, households could not only be multi-generational and incorporating multiple extended family ties but could also include unrelated individuals belonging to the same tribe who may have recently migrated to South Africa. Within a single residential unit (an apartment or a free standing house) more than one of such group could exist. Or there could be other single people renting a room or a bed in the same unit. This meant that in one residential unit, one could have multiple nuclear or joint families with intergenerational co-residence, second and third order relatives as well as one or more unrelated individuals living with the family who may be considered part of the 'family', share rental costs, food and other resources. At the same time, there could be collections of related individuals living together but not sharing resources, or in some cases, unrelated individuals living together and sharing resources or large groups of single men and/or women sharing a rental apartment and at times, living in half day shifts. As a result, we needed to come up with a working definition of "household". We decided to define it as "a group of people who regularly stay together in the same residential unit, who are related to each other and who share resources with each other; or a single person living independently or living with other unrelated people but not sharing everyday resources."

Section 5: Securing Interviews

To accommodate the multiplicity of structure types and multiple co-residing households described above, a method for random sampling within the selected buildings as well as within a selected household was developed. A Location Cover and a Household Roster were developed (Appendix 2) that were attached to each questionnaire to aid the fieldworkers in randomly selecting a household with in the sampled location as well as randomly selecting an individual respondent within a selected household. This was accomplished by generating three unique random tables for the location cover sheets and one

unique random table for the Household Roster sheet. These random tables were produced using database routines and then merged onto the Location Cover and Household Roster sheets during printing.

The Location Cover had one random table for selecting the floor to be sampled if the selected structure was a multi-storey building and accounted for building with up to 20 floors.[11] The next random table allowed for the random selection of a residential unit (particularly, apartment units on the selected floor of a multi-storey structure). It could also be used to pick out a unit on a plot of land, if the surveyors found more than one residential unit built on the plot (for instance, a backyard detached cottage or separate rooms constructed outside of the main property). The last table was used in the case where more then one household was present within the selected residential unit. This was particularly useful in the case of the Somali community where the overall number of Somali occupied structures were relatively small, but several independent households could be found living under one roof. The location cover sheet allowed between two to four respondents to be interviewed per location (the initial sampled structure), depending on whether it was a high rise building or a free standing house.

Once a household was selected, the Roster Sheet was used to determine eligibility of the each of the household members and select a respondent out of the list of all household members listed in terms of the relationship with the 'person opening the door'. The reason for this additional step was to ensure that the sample was not systematically biased by always interviewing those who opened the door or allowing the fieldworkers discretion of selecting a 'convenient' respondent. For instance, it was likely that we would find more unemployed people or women (especially in the Somali community) since the survey took place predominantly in the day time. We did not want to use the usual relationship with the head of the household as we felt that in the given context this may be a difficult concept to explain to the respondents. For an

11 If a building had more than 20 floors, it could be thought of as two buildings and the fieldworkers could re-sample it after excluding the first 20 floors.

individual to be included in the survey, they had to be between 18-50 years of age for women and 18-60 for men, usually living in one of the pre-identified structures (i.e. not be a visitor), identify themselves (or are reported by the initial respondent to the roster) as Zimbabwean, Congolese, Somali or South African, and finally must have been born outside Johannesburg (for South Africans) and outside South Africa (for non-nationals). Following this, a random selection of the eligible household members was done using the unique random number table pre-printed on the roster sheet (See appendix 2 for the complete sheet).

HHR1	HHR2	HHR3	HHR4	HHR5	HHR6	HHR7	HHR8	HHR9	HHR10	HHR11	HHR12
1	1	1	**4**	5	4	4	3	6	8	1	12
RAS_012	**WHAT IS THE SERIAL NUMBER OF THE PERSON TO BE INTERVIEWED:**										1

Figure 2: Random number table taken from the roster sheet.

The fieldworkers were asked to enter the serial number of the person they were selecting for an interview on the space provided in the table (Figure 2) so that we could monitor the proper application of the random selection procedure. If the selected respondent was not available at that time to be interviewed, the interviewers were asked to set up an appointment. The survey allowed up to two follow up visits for an appointment, after which the interview was abandoned. In the event of complete refusals by the selected respondent, the surveyors could delete the person from the eligibility list and re-calculate an eligible respondent from the random table within the same household. The appointment strategy also helped in overcoming some of the language difficulties of doing a multilingual survey. For instance, if an English speaking interviewer's sampled a residential unit that had French speaking Congolese migrants, he or she could make an appointment to send another colleague from the Congolese team. Since men were allowed to interview only men but women could interview men as well as women (based on the experience of the qualitative phase), the appointment system was also used in such cases where the roster taken by a male interviewer selected a female respondent.

During the initial week of the survey it was determined that the Somali population, which was highly concentrated in a relatively few buildings within Mayfair neighborhood, was not being captured by the initial sample. This population was distributed across Mayfair (predominantly) but divided into clusters that were not adjacent to each other. This was also significantly smaller in number as compared to the other three groups (See Appendix 1) such that the random sampling of plots was likely to miss them and thus require multiple rounds of re-sampling. When the strategy did not work even with the second resample, we knew that not only were we wasting precious time but that the frustration of fieldworkers was rapidly mounting as the payment structure required them to be securing interviews. In addition to this, most people with whom appointments were made would default, leading to the interview being abandoned. Due to the small size of the Mayfair area it was concluded that it would be more efficient to perform a census of the entire area and access all plots that contained Somali residents at least once.

The community mapping sheets gave a good starting point for this exercise but we could not be sure that our key informants had knowledge of all the Somalis in the area. At this stage, we began employing some active community members of the Somali community to further support the survey by not only supporting the mini census effort but also by soliciting participation in the survey. In the end, even though we collected the smallest number of interviews from this community, it took us the longest time to complete. This also demonstrates the limitations of our sampling strategy in situations where multiple clusters exist in a small population. We could perhaps have continued to generate re-samples until we managed to get the required number of respondents but it would have required us to change the payment structure of interviews and requesting more funds from the donor organization, both of which were unrealistic options for us.[12] However, despite the problems, the Somali experience still highlighted the

12 It should be noted here that the authors do not think that the payment strategy that used in this survey is an optimal one. We had to ascribe to it due to the work being carried out in a very limited budget. But we suggest that donor organizations should be cognizant of such issues.

importance of the community foot mapping exercise, which had even initially indicated the existence of such multiple, fragmented clusters and the particular structure of Somali households that supported the development of Location Covers and Roster Sheets.

Section 6: Concluding Remarks

At the end of a five week period, a total of 1067 completed interviews were collected.[13] Despite the challenges faced in the data collection we were able to successfully execute the sampling strategy in three out of four migrant groups. While we did not abandon aim to achieve a statistically unbiased sample of respondents in the Somali group, the challenges faced in collecting interviews with this group highlight the limitation of the approach used for the other three groups. However, these limitations themselves point to a crucial need for close collaboration with community based organizations and influential members of the community as well as a meticulous pre-planning of the survey. Given the space constraints and the focus of this paper, we have not been able to provide anecdotes from the fieldwork that highlight the manner in which this strategy was successful in dealing with challenging situations arising in the field. Most importantly, we have demonstrated here that residence based random sampling is possible within a non-homogeneously distributed population context and have made arguments for using a mixed methods approach, especially in the context of work with migrants.

Several studies with migrants have demonstrated the use of mixed methods approaches to achieve similar aims. For instance, Durand and Massey's ethnosurvey methodology looking at Mexican migrants at communities or origin and destination (Massey 1987); Stepick and Stepick's (1990) work with Haitians in Miami and socio-medical surveys that combine ethnographic work to get a

13The sample was drawn assuming a power of 80.6%, significance level of 0.05, standard deviation of 10, a size effect of 1 and a two-sided test, we get a sample of 800 (@ 200 per group) was determined to give statistically valid results. Calculated thus, this study should generate results with a power of at least 80%. We then decided to top this figure up by over-sampling each population.

more culturally sensitive understanding of health seeking behaviors such as Freidenberg, Mulvihill, and Caraballo 1993; Manton 1993; Stratford et al. 2003. The closest to our study is the work carried out by Parrado et al (2005) with Mexican migrants in United States in both approach and purpose where they highlighted the importance of community collaboration. However, our study is unique in several respects. One, it is one of its kind in South Africa so far, where the urban inner city environment posed challenges that many similar studies, such as those cited above, did not face. Two, the use of spatial sampling techniques has demonstrated many advantages in streamlining field work, and these suggestions can be made use of by other studies in similar environments. Three, the multi-phase approach of the project ensured that we did not intertwine the qualitative data collection with close ended questions in survey work but instead used the former to build on to the latter in content as well as method. Four, unlike Parrado et al (2005) we did not use surveyors known closely to communities to be interviewed, given the sensitive nature of the questions we were asking. While it has not be discussed adequately in this paper, the training of surveyors was crucial for the success of this survey. Five, the use of community foot mapping exercise with intense involvement from key informants in the migrant communities as well as migrant organizations was rewarding at several levels. Not only did it allow us to insight that led to the development of our sampling frame but it also helped us better understand the residential patterns of migrant households, give an indication of the conditions to expect in the fieldwork and most importantly gave us a chance gain the acceptance of migrant communities. Finally, substantively this project makes an important sociological contribution to work on migrants (especially, self-settled foreign migrants in urban areas) and socio-economic conditions impacting upon vulnerability to HIV risk.

Bibliography

Abdool Karim, Q, Abdool Karim, S, Singh, B, Short, R. and Ngxongo, S (1992) Seroprevalence of HIV infection in Rural South Africa, *AIDS* 6(12): 1,535-9.
Africa', Demographic Research Special Collection 1.

Campbell, C. 2000. 'Selling sex in the time of AIDS: the psycho-social context of condom use by sex workers on a Southern African mine.' Social Science & Medicine Vol 50: 479-494.

Christopher, A. 2005. 'Does South Africa Have Ghettos?' Tijdschrift voor Economische en Sociale Geografie, Vol. 96, No. 3, pp. 241–252.

Carter-Edwards, Lori, John T. Fisher, Benjamin J. Vaughn, and Laura P. Svetkey. 2002. ''Church Rosters: Is This a Viable Mechanism for Effectively Recruiting African Americans for a Community-Based Study?'' Ethnicity and Health 7:41-55.

Crush, Jonathan. 2000. "The Dark Side of Democracy: Migration, Xenophobia and Human Rights in South Africa." International Migration 38: 103–31.

Decosas J, and Adrien, A. (1997) 'Migration and AIDS', *AIDS* Vol. 11 (Suppl. A) S77-S84.

Department of Health, South Africa. 2007. "National HIV and Syphilis Antenatal Sero-prevalence Survey in South Africa", Department of Health, 2002-2006

Fenigold, D. 2007 ONE SIZE DOES NOT FIT ALL: ETHNO-LINGUISTIC MINORITIES AND 'INCLUSIVE GROWTH', Presentation at Inclusive growth and poverty reduction in New Asia and Pacific Symposium, October 2007.

Flaskerud, Jacquelyn and Betty Winslow. 1998. ''Conceptualizing Vulnerable Populations in Health-Related Research.'' Nursing Research 47:69-78.

Freidenberg, Judith, Michael Mulvihill, and Louis Caraballo. 1993. ''From Ethnography to Survey: Some Methodological Issues in Research on Health Seeking in East Harlem.'' Human Organization 52:151-61.

Gonzalez M. (2002) Promoting HIV/AIDS prevention among migrant farmworkers and their families: a replicable model. Int Conf AIDS. 2002 Jul 7-12.

Gritz, E. R., B. A. Berman, and G. Bennett. 1992. ''Recruitment Through Schools and Churches.'' Pp. 79-90 in Health Behavior Research in Minority Populations: Access, Design and Implementation, edited by D. M. Becker, D.

R. Hill, J. S. Jackson, D. M. Levine, F. A. Stillman, and S. M. Weiss. NIH Pub. No. 92-2965.

Gurney, I.M., and Omolalu, A. 1971. 'A Nutritional Survey in Southwestern Nigeria: The Anthropometric and Clinical Findings'. Journal of Tropical Pediatrics, 17(2): 50-61.

Jacobsen, K. and Landau, L. and. 'The Dual Imperative in Refugee Research: Some Methodological and Ethical Considerations in Social Science Research on Forced Migration' *Disasters*. Vol. 27(2): 95-116.

Landau, L. (2005) "Immigration and the State of Exception: Nativism, Security and Sovereignty in Refugee-Affected Africa", available on http://migration.wits.ac.za/StateofException.pdf.

Landau, L., Ramjathan-Keogh, K. and Singh, G. (2004) 'Xenophobia in South Africa and Problems Related to it', Background Paper prepared for Open Hearings hosted by South African Human Rights Commission with the Portfolio Committee of the Department of Foreign and Home Affairs. Available on www.pmg.org.za

Leggett, Ted. 2003. "Rainbow Tenement: Crime and Policing in Inner Johannesburg." Monograph no. 78. Pretoria: Institute for Security Studies.

Lurie, M., Williams, B., Zuma, K., Mkaya-Mwamburi, D., Garnett, G., Sweat, M., Gittelsohn, J., and Abdool Karim, S (2003) 'Who infects whom? HIV-1 concordance and discordance among migrant and non-migrant couples in South Africa', *AIDS*. Vol.17:2245–2252.

Manton, Kenneth I. 1993. ''A Bridge Between Cultures: Linked Ethnographic-Empirical Methodology for Culture-Anchored Research.'' American Journal of Community Psychology 21:747-73.

Massey, Douglas S. 1987a. ''The Ethnosurvey in Theory and Practice.'' International Migration Review 21:1498-1522.

Messer, E. 1983. 'The Household Focus in Nutritional Anthropology: An Overview'. Food Nutrition Bulletin. Vol 5(4): 2-12.

Palmary, Ingrid, Janine Rauch, and Graeme Simpson. 2003. "Violent Crime in Johannesburg." In R. Tomlinson et al., eds., *Emerging Johannesburg: Perspectives on the Postapartheid City*. London: Routledge.

Speigel, P.B. (2004) "HIV/AIDS in Conflict-affected and Displaced Populations: Dispelling Myths and Taking Action" Disasters 28(3): 322-339.

Stepick, Alex and Carol Dutton Stepick. 1990. ''People in the Shadows: Survey Research Among Haitians in Miami.'' Human Organization 49:64-77.

Stratford, Dale, Sandra Chamblee, Tedd V. Ellerbrock, Jonny W. Johnson, Denny Abbott, C. Fordham von Reyn, and C. Robert Horsburgh Jr. 2003. ''Integration of a Participatory Research Strategy Into a Rural Health Survey.'' Journal of General Internal Medicine 18:586–8.

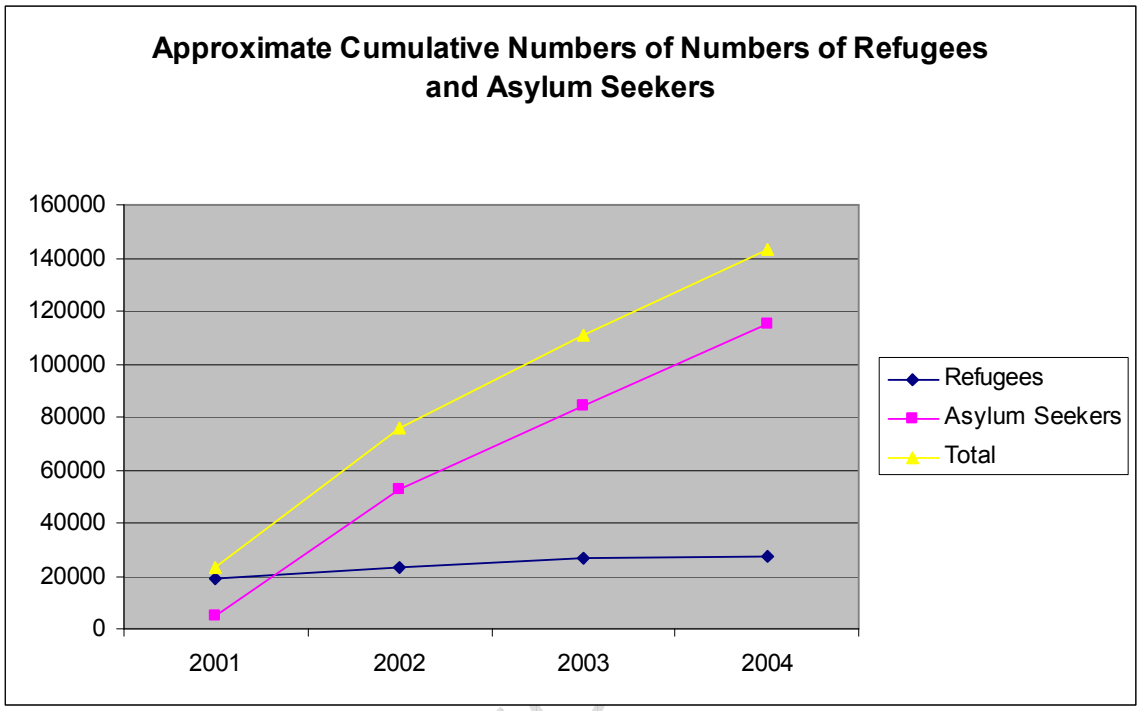
UNAIDS, (2007) *AIDS Epidemic Update*. Available on http://data.unaids.org/pub/EPISlides/2007/2007_epiupdate_en.pdf. Accessed on 4th December, 2007.

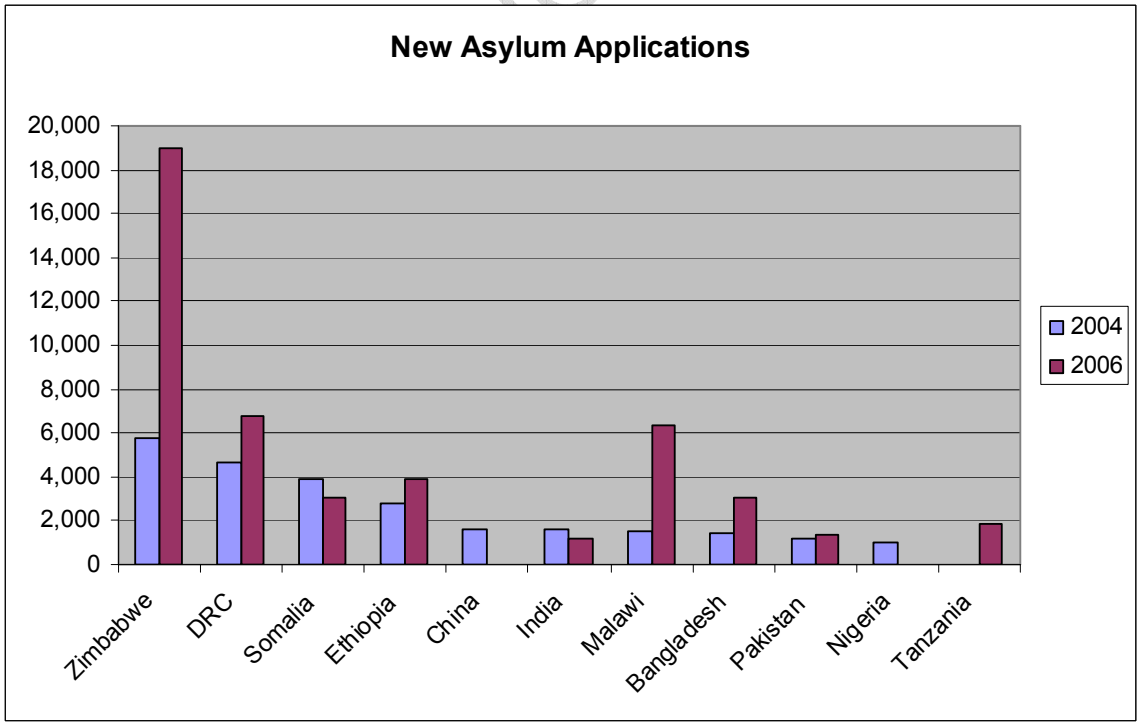
Williams, B, Gouws, E, Lurie, M, Crush, J (2002) 'Spaces of Vulnerability: Migration and HIV/AIDS In South Africa', *Migration Policy Series* No. 24 South African Migration Project.

World Health Organization; 2007.Towards Universal Access: Scaling up Priority HIV/ AIDS Interventions in the Health Sector. Progress Report. Geneva, Switzerland:

Zuma, K., Gouws, E., Williams, B. and Lurie, M. (2003) 'Risk factors for HIV infection among women in Carletonville, South Africa: migration, demography and sexually transmitted diseases', *International Journal of STD & AIDS* 2003; 14: 814–817.

Appendix 1: Refugee and asylum seekers entering South Africa by year and nationality. Source: Department of Home Affairs, South Africa, 2005





Appendix 2: Location Cover sheet and Household Roster with unique random tables

School of Human & Community Development

Location Cover

MVH-LocationCover2007-v1

Location: 1 1 1 1
Fieldworker: A A
Date (yyyy/mm/dd): Y Y Y Y M M D D

LOC_000	Location Contact		
Appointment made (or attempted) by (surname)	LOC_001	A A A A A A A A A A A A	
Interview carried out by (surname):	LOC_002	A A A A A A A A A A A A	
Date on which appointment was made:	LOC_003	Y Y Y Y - M M - D D	
Date on which interview was carried out:	LOC_004	Y Y Y Y - M M - D D	
Outcome of appointments and interviews at this location: INSTRUCTIONS: FILL FOR BOTH SAMPLES	1. Interview initiated and completed 2. Interview initiated but not completed due to refusal to carry on by the respondent 3. Appointment could not be made as an entry to the building could not be secured 4. Appointment was made but not honored by respondent after 3 attempts and thus abandoned	LOC_005	1 2 1 1

LOC_100	Location Information		
Was this a safe or a dangerous building to be accesses?	1. Safe 2. Dangerous	LOC_101	1
Can the location be included in re-sampling, if needed? i.e. Does it have unsampled household units (individuals or families unrelated to people who were interviewed) remaining that can be accessed in the second wave?	1. Yes 2. No	LOC_102	1
Was it easy to access the location or was permission from an external authority was needed to do interviews?	1. Easy Access 2. Permission was needed 3. Other difficulty (specify)	LOC_103	1 _____ _____
From whom was the permission needed? Please be as specific about the person as you can and include contact details, if possible.			
LOC_104:			

LOC_200	Sample Tables																			
Floor	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
	1	1	2	2	1	4	5	2	1	6	1	4	13	3	3	15	9	6	5	11
	First Sample										Second Sample									
	1	2	3	4	5	6	7	8	9	10	1	2	3	4	5	6	7	8	9	10
Structures	1	1	2	4	2	2	3	7	4	2	1	2	1	3	1	3	1	6	4	4
Households	1	1	2	2	4	4	5	3	1	6	1	1	3	4	2	2	6	4	7	8

C:\Documents and Settings\User\Desktop\GayatriSurvey\LocationCover.doc
STUDY ON MIGRATION AND VULNERABILITY TO HEALTH IN THE INNER CITY JOHANNESBURG 2007
email: gayatrimail@gmail.com
2008.03.09 1 of 100

Figure 3: Loction cover with unique random tables at bottom.

Location: | | | | |

Household Roster and Appointment Sheet:

Hello! Does anyone who identifies as Zimbabwean, Somali, from DRC or South African live here? (If Yes, give more information as printed below).

Good Day! My name is _____ and I am here from the University of the Witwatersrand, Emthonjeni Community Center. We are currently carrying out a survey to find out more about the experiences of people who migrate to Johannesburg. In this survey, we want to know how these experiences impact upon the health and well being people coming to live in Johannesburg. This also includes the difficulties faced in accessing health services as well as other social conditions that may affect a person's health in any way. Neither me, nor any one else involved in this research is working for or on behalf of any government or any aid organization. This study is **only** for purposes of research in order to know your views in this matter. Your participation will help us a lot in completing this study as in assisting your community living in Johannesburg. Participation in this study is completely voluntary. The information you give me and your identity will be kept strictly confidential. Anyone taking part in this study can refuse to any question that they don't want to at anytime in the interview. I am happy to answer any question you may have for me. I invite you to be a part of our research.

RAS_001	**With your permission, may I go ahead and ask a few questions that will help me decide who I can interview in your household as we have to follow the research method?** These couple of questions prior to the interview will take about 5 minutes of your time.	Yes = 2 No = 1	1
RAS_002	Including yourself how many people in total (related and unrelated) usually live with you here in this structure? (**ENTER NUMBER.** PROBE FOR ALL THE PEOPLE IN THE STRUCTURE, EVEN WHERE DIFFERENT FAMILIES ARE LIVING IN VARIOUS ROOMS OF ONE HOUSE.)		
RAS_003	Including your family/household, how many family units or single people live here? (WHEN I SAY HOUSEHOLD, I MEAN PEOPLE WHO REGULARLY STAY TOGETHER IN THE SAME STRUCTURE AND WHO ARE RELATED TO EACH OTHER AND WHO SHARE RESOURCES WITH EACH OTHER)	NO OF FAMILIES NO OF SINGLE PEOPLE	1 1 1 1
RAS_004	Apart from these members of your household, who are the other people you are living with now? (CHECK ALL APPLICABLE. DO NOT READ)	South Africans People from community/country of origin Other Foreigners	✓ ✓ ✓
RAS_005	How would you classify the type of dwelling you live in? Is it: (ASK **RELATIVE CATEGORIES** BASED ON **OBSERVATION** OF HOUSING TYPE.)	Free Standing House (Single Family) = 7 Free Standing House (Multi Family) = 6 Apartment (Single Family/Self) = 5 Apartment (Multi-Family/Sharing with others non-related) = 4 Semi-Detached House (cottage etc.) = 3 Self-Built/Informal Housing/Shack/ Squatter Settlement = 2 Hostel, Dormitory, Boarding House, Shelter = 1 Other = 0	1
RAS_005.1	Other House type (Specify)		

1

Location:

SPEAK: Please tell me the of everyone who you consider a part of your household/family. Could we start

RAS_6._	RAS_7._	RAS_8._	RAS_9.__	RAS_10._	RAS_11.1	SS No.
Relations-hip to the person opening the door?	Gender FEMALE = 1 MALE = 2	AGE IN YEARS	Nationality SA = 1 ZIM = 2 DRC = 3 SOM= 4 OTHER = 0	Non SA: Was 'NAME' born outside South Africa? South Africans: Was 'NAME' born outside Johannesburg? (YES = 1; NO = 2)	Eligibility (IF RAS_9._ > 0 AND AGE > 15 THEN 2 ELSE 1) YES = 2 NO = 1	
SELF						1
Now, could you please list other members of your household starting from the youngest member going to till the oldest member.						
						2
						3
						4
						5
						6
						7
						8
						9
						10
						11
						12

with your details first please? Then, we can go on from the youngest to the oldest person.

Now **SELECT RESPONDENT** based on the interview selection matrix below. Mark a big X in eligibility box to show has been selected according to the selection matrix.

HHR1	HHR2	HHR3	HHR4	HHR5	HHR6	HHR7	HHR8	HHR9	HHR10	HHR11	HHR12
1	1	1	4	5	4	4	3	6	8	1	12
RAS_012	WHAT IS THE SERIAL NUMBER OF THE PERSON TO BE INTERVIEWED:										

If the person selected for the interview is different from the person who opened the door, respectfully inform the person who you have selected and request to speak that person to interview them? If it is a new person explain the purpose of the interview again and make an appointment to do the interview.

RAS_013	SAY: The questions we ask involve very brief answers and it will take about 1 hour of your time. Are you willing to let me interview you?	YES = 2 NO = 1	
Appointment details:			
RAS_014	Appointment Date:	Y Y Y Y - M M - D D	
RAS_015	Appointment Time:	H H : M M	
RAS_016	Any contact details taken to follow up:		

2

Figure 4: Household roster with unique random table on page 2