Supplemental Online Material

MATERIALS AND METHODS

Study Setting

Liuzhou, China has a population of 3.6 million, with one million in four urban districts and 2.6 million in six surrounding rural counties. Liuzhou is known for a vibrant night life, with an estimated 5,000 women engaged in sex work. Guangxi, on the border with Vietnam, is among the provinces in China with the highest HIV and STI prevalence. In 2009, syphilis prevalence among women engaged in sex work in urban Liuzhou was estimated to be 11.0%[1].

PLACE Methods

The venue-based strategy used in this study, the Priorities for Local AIDS Control Efforts (PLACE), constructs a sampling frame of all public venues where people meet new sexual partners. This includes venues identified as places where sex workers solicit clients. The rationale for the more generous sampling frame is that it provides information on sexual mixing among a variety of groups; the sampling frame does not depend on the definition of sex work; and it can reveal previously unknown risk groups [2]. PLACE is similar to “time-location-sampling” which randomly samples venue-time periods from an inventory of time blocks developed by listing venues where the population of interest (e.g., female sex workers) is known to congregate, and segmenting the open hours of operation at each venue into discrete time blocks. Two differences are that PLACE does not restrict the mapping and listing to venues where particular types of most-at-risk populations can be found and PLACE does not sample time-periods but interviews people when the number of people at the venue is expected to be the largest.

The PLACE method follows five steps:

1) Planning

2) Community Informant Surveys to Identify Venues

3) Venue Visits to Characterize and Map Venues

4) Interviews (and Testing) of Workers and Patrons at Selected Venues

5) Feedback to the community and use of data to improve programs

These steps were modified slightly as necessary to accommodate the comparison of the two sampling methods. See Supplementary Figure 1.

Step 1. During the planning phase we identified the geographic boundaries of the study area (i.e., urban and rural Liuzhou); selected the rapid test for syphilis as the biomarker for comparing the characteristics of the populations sampled by two methods; selected an interviewing team of people knowledgeable of Liuzhou, not affiliated with the local Centers for Disease Control, willing to visit venues late at night, and able to gain the trust of sex workers as well as administer the rapid test and provide results; developed, pre-tested and translated the data collection instruments; obtained ethical review and approval of the study; and finalized the protocol. The decision to include the rural counties reflected the desire to avoid terminating an RDS recruitment chain if a person from the rural areas was recruited into the study. Based on information available about sex work in Liuzhou suggesting that sex workers were much more likely to be workers at the venue rather than patrons, we focused efforts on interviewing workers at venues rather than workers and patrons.

Step 2. The objective of the community informant surveys is to construct a sampling frame of venues where people from Liuzhou (including but not limited to sex workers) meet new sexual partners. The rationale for identifying venues in addition to known sex worker venues was to uncover venues unknown to be places where sex workers solicit and to gain insight into high risk sexual behavior among people who do not self-identify as sex workers. The target number of community informant interviews was 15 in each of the six counties and 310 across the four urban districts. Community informants were aged 18 an older. Selection of informants was done by convenience but based on targets for various types of informants, e.g. taxi drivers, police, health workers, out of school youth, migrant workers, business men and injecting drug users. Interviewers approached people appearing to meet the target type and requested verbal informed consent for a face to face interview. No unique identifiers were obtained and no testing was performed. Each informant was asked to name places where people go to meet new sexual partners. They were asked to describe each venue named including the type of venue and its name and address. Informants identified 971 venues.

Step 3. The objective of the venue visits is to obtain information about the characteristics of the venues identified by the community informants including the number of female workers at the venue, the estimated number of female workers who exchange sex for money at the venue, and whether female patrons included sex workers. It was not feasible to visit all of the 971 unique venues identified. We selected a stratified sample of 385 venues from three strata:

- Venue Group 1: 334 Urban Venues named by only one informant
- Venue Group 2: 317 Urban Venues named by two or more informants
- Venue Group 3: 320 Rural Venues

Within each stratum, venues were sorted by geographical district and type. Venues were selected for venue visits by interval from these ordered lists using a random start. In Group 1, every third venue was selected (33% selected=111 venues); in Group 2, every third venue was not selected (66% selected=211 venues); and in Group 3, 1 in 5 were selected (20% selected=63 venues). 395 venue visits were conducted between November 19,2009 and January 9, 2010. At each selected venue, interviewers asked the manager or some other knowledgeable person aged 18 and older about the characteristics of the venue including the total number of women working at the venue and the number of women who were sex workers. Verbal informed consent was requested. Analysis of venue characteristics is weighted to account for the differences in the probability of selection.

Step 4. Female Worker Surveys. The objective of the worker surveys is to interview a stratified random sample of female workers that would yield interviews with 400 sex workers. Data from the venue visits suggested that few female patrons were sex workers. In a small sample of 95 female patrons, only three reported sex work. Therefore we focused on interviews with female workers. A sample of venues visited in Step 3 were selected for worker interviews, taking into account differences in the probabilities of selection in Step 3 and the expected number of sex workers at each venue. Venues were selected from the three strata as described below:

- Stratum 1: Urban "High Volume Venues". This stratum includes 16 venues that were confirmed during a venue visit in Step 3 as being in operation and where half or more of the female workers exchanged sex for money and where interviewing all workers would yield interviews with at least five sex workers. (Specifically, at these sites, the venue informant reported: 1) that there are sex workers at the site (c21g=1); 2) the proportion of female staff who are sex workers is at least 50% (c29c/c29a > =.5) and 3) there are at least 5 sex workers at the site (c29c>4).) All female staff at these 16 venues were eligible for an interview.

- Stratum 2: 25 Other Urban venues. The 25 selected venues are a stratified random sample of the list of urban verified venues in operation such that all venues had an equal probability of selection. All workers at these venues were eligible for an interview.
- Stratum 3: 23 Rural Venues. Every third venue of the 63 rural operational verified venues was selected for interviews with workers. All workers were eligible up to 5 workers per venue. If more than 5 workers, the 5 were randomly selected.

Interviewers requested interviews with all workers at urban venues. Verbal informed consent was requested and no unique identifiers were obtained. Partial participation was not permitted. A payment of 100 yuan was given to all who agreed to participate in the survey and syphilis testing. Interviews were conducted from January 6-28, 2010.

Women who were age 15 and older who lived in Liuzhou were identified as sex workers based on their responses to multiple questions in a face-to-face interview eliciting whether they had exchanged sex for money in the past 4 weeks. There were two questionnaires administered to participants. The first confirmed age, willingness to participate and if the respondent had participated in the RDS arm. Female workers aged 15-17 were excluded from participation if they were at the venue with their parents or at the venue on a family errand. Parental consent was not requested for the other women age 15-17. The second questionnaire included 82 items including some with multiple questions. The second questionnaire had 6 sections that covered demographic information, venue employment, health issues, sexual history, contraceptive use, and sex work. The first question about sexual behavior followed 36 prior questions. The first question about ever exchanging sex for money was asked over half-way through the questionnaire (Q43e). Other participants in the study were not aware if someone reported sex work or not based on the length of interview. Everyone received the test for syphilis regardless of whether they reported having sex or exchanging sex for money.

Step 5. Feedback to Liuzhou

The preliminary results were presented in an informal meeting with public health professionals in Liuzhou in June, 2010.

Analysis of PLACE data

PLACE data analysis used SAS software. Responses were weighted based on the sampling strategy and participation rates both at the venue and individual level. Estimates adjust for stratum and sample weights. Characteristics were estimated using Proc Survey Freq in SAS [3].

RDS Methods

Assumptions of RDS Methods

Unlike in venue-based sampling, RDS sampling weights are not known to researchers. Participants' number of reciprocal ties with other members of the target population is used to approximate an individual's sample inclusion probability on the assumption that the chain referral process selects respondents with probability proportional to a respondent's personal network size. RDS theory rationalizes this assumption in terms of an idealized model of how sample subjects make referrals to new subjects and potential respondents are recruited into the sample, under the following stringent conditions: (a) all eligible members of a respondent's network have an equal probability of recruitment by the respondent; (b) there is reciprocity in recruitment (that is, if individual a recruits b, then b would recruit a); (c) respondents accurately report the number of members of their social networks who meet the study definition; and (d) the network is sufficiently large to allow sampling without replacement.

Threats to the representativeness of samples and validity of population estimates obtained through RDS may be introduced in two ways. The first is if the reported personal network size does not provide the correct sampling weight[4] [5, 6]. This could arise if RDS participants inaccurately report their network size; if they recruit preferentially from their immediate social network on a characteristic associated with infection; or if infected individuals have more contacts than their uninfected counterparts, leading to proportionally down-weighting high degree individuals, underrepresentation of the infected group and a negative bias in the proportion infected[6]. The second is if other factors, in addition to network size, influence the probability that infected members of the population are included in the sample. For example, infected sub-groups of the population might be missed because they are not linked through the peer network, or the RDS assumption of non-preferential recruitment of participants constrains researchers from guiding the referral process towards members of the population who are likely to be missed. The inability to redirect the chain referral can prove particularly problematic if recruitment chains become trapped in low infection venues which would limit the ability of a potentially infected respondent to be recruited into the sample; if incentives paid to recruited peers influence the likelihood that people who do not meet the study definition will be recruited; if the location of the interview is less accessible to those infected; and if participants who refuse to be tested are more likely to be infected than those who agree to the test.

RDS Sample recruitment

Between October 26, 2009 to January 29, 2010, RDS methods were used to recruit FSWs in Liuzhou. Eligible participants were women 15 years and older who had sex in exchange for money during the month prior to the interview and were currently working and living in Liuzhou (including rural counties). Eligibility for participation in the RDS arm was being at least 15 years old, a first time participant in the RDS arm, and self-identified as a sex worker in response to the question: “Have you exchanged sex for money in the past month?.”

Seven seeds, stratified by place where they solicited clients (massage parlour, hair salon, KTV-karaoke bar, sauna and park) were recruited with the help from experienced local outreach workers. All except one of the seven seeds recruited other participants. The six productive seeds generated between 9 and 20 waves of recruitment. 310 out of 583 respondents were recruiting participants, while the remaining 273 did not recruit any participant. Participants in the Liuzhou RDS were given from the start two coupons, a number which was further reduced to one after the 14th wave to control sample growth. The idea of systematic coupon reduction, which involves decreasing the number of coupons from two to one at a specific wave in the recruitment chains was discussed by Johnston et al.[7] Our decision to reduce the number of coupons for sample growth control was based on the theoretical consideration that this forces the sampling process to more closely resemble a non-branching random walk, a strategy that was shown to reduce the variance of RDS estimates[8].

As in the PLACE arm, interviews were administered face-to-face by trained nurses in Mandarin Chinese and Zhuang, the language spoken by the largest ethnic minority group in Liuzhou. However in the RDS arm, interviews were conducted at a single fixed interview site located in a local hospital in one of Liuzhou's urban districts, easily accessible by public transportation. Participants could drop-in or phone to make an appointment. Participants were screened upon arriving at the interview site.

Each participant received a primary incentive of 100 yuan (ca. US$ 14) for the interview and to cover transportation to the interview site and a secondary incentive of 50 yuan for each of her successful recruitments. The appropriate size of these incentives was determined after formative research was conducted prior to the fielding of the survey and in consultation with Liuzhou outreach workers who were familiar with the organization of sex work in the city.

To prevent potential respondent duplication, during screening of respondents for eligibility, biometrics (height, weight, and left/right forearm length and wrist circumference) were collected and entered in a password protected RDS coupon manager to identify potential duplicates. To minimize the threat presented by impersonators recruited into the sample by recruiters eager to fill their coupon quota or attracted by the size of the incentives, interviewers were trained to detect impersonators through a series of questions in the screening phase of the interview. Incentives were also judged not large enough as to encourage participation by imposters or coercive recruitment.

RDS Questionnaires

The questionnaires used to measure socio-demographic characteristics and HIV/STD-related risk behaviors were based on the RDS module provided by investigators who employed RDS to recruit female sex workers in Vietnam[9], on the FSW module of the Family Health International Behavioral Surveillance Surveys [10], on the standard PLACE protocol [2], and on a household survey of sexual behavior in Liuzhou implemented in 2008. The questionnaires included socioeconomic and demographic information, sexual behavior practices, drug use and STD symptoms. An additional set of questions on sex work history and reasons for sex work was introduced as part of this study. For RDS purposes, network size was measured using the question: "How many sex workers do you know in Liuzhou (including rural counties)? By knowing, I mean: you know their names and they know yours, and you have met or contacted them in the past month." When recruiting participants returned to the interview site to collect their secondary incentives, they were administered a brief follow-up questionnaire to collect information on basic socio-demographic characteristics, place of work, type and strength of relationship relative to all members of the participant's network, including those whom they recruited successfully, those who turned down the invitation to participate and, as a variation of the usual RDS practice, those whom they did not invite to participate. All respondents provided informed consent and anonymous interviews.

RDS Syphilis Testing

The RDS and PLACE arms used the same rapid syphilis test. Blood samples for rapid syphilis testing were collected by finger prick (Wantai anti-TP Antibody Rapid Test, Wantai Biological Pharmacy Enterprise, Beijing, China) on a voluntary basis. 47 out of 583 eligible RDS respondents did not agree to the test on grounds that they were tested in the past. Participants were administered the questionnaires regardless of their participation in the rapid syphilis test. (In the PLACE arm, individuals were required to participate in both the test and interview.) Test results were provided to those who agreed to take it at the end of the interview. If RDS participants tested positive, they were invited to immediately provide a blood sample using a needle for confirmatory testing (Toulidine Red Unheated Serum Test, Rongsheng Biotechnical Company, Shanghai, China). However, to increase confidentiality of results, positive-testing participants on the rapid test had the option to return on another day for the confirmatory test. Of the 40 participants who tested positive on the rapid test, 36 agreed to the confirmatory test. Of these, 20 tested positive for active syphilis. All tests were performed by trained doctors and lab technician at the same hospital of the interview site. Free treatment was offered to those with confirmed active syphilis. PLACE participants who agreed to the confirmatory test were required to visit a hospital to obtain the confirmatory testing.

RDS data analysis

RDS data were analyzed using the RDS Analysis Tool (RDSAT) Version 5.6.0 [11] and the RDS-2 estimator[12] implemented using R software by one of our co-authors (Neely). We used both RDSAT and RDS-2 estimators [12]to generate estimates of population proportions and their confidence intervals. The validity of estimates from RDSAT requires that the chain referral process samples individuals with probability proportional to their network sizes and that it represents a first-order Markov process that has achieved equilibrium in the sense that the overall composition of the sample has become stable with increasing numbers of recruitment waves[13, 14]. RDSAT estimates of group proportions are calculated based on cross-group recruitments, tracked through the unique codes on the recruitment coupons and used to calculate equilibrium estimates from the Markov model, and a weighted-average of participants' personal networks.

The options used in the RDSAT software were:

- Network size estimation: dual component (default)
- Mean Cell Size: 12 (default)

- Number of re-samples required for bootstrap: 2500 (default)
- Confidence interval (2-tailed alpha): 0.025 per tail
- Did not pull in outliers of network sizes. (default)
- Algorithm type: Data smoothing (default)

The RDS-2 estimator, on the other hand, requires only the assumption of sampling probability proportional to network size and computes population proportions based on a weighted average of participants' network sizes. Our decision to present RDS-2 estimates together with the more conventionally used estimates from RDSAT was motivated by the following reasons: First, simulation and empirical studies have found RDS-2 may yield estimates with less bias [12, 15]. Second, reliance on RDS-2 estimates is greater when equilibrium estimates cannot be computed and RDSAT estimates cannot be generated. This was the case for proportion estimates for a few selected socio-demographic characteristics.

RDSAT relies on a bootstrap computations presented in Salganik[16] to estimate confidence intervals. To estimate confidence intervals for proportions computed using the RDS-2 estimator, we used a variant of this bootstrap methodology. Volz and Heckathorn [12] also described an algebraic variance estimation procedure designed specifically for use with the RDS-2 estimator. However that procedure (at least in its published form) was derived under the assumption that the recruitment chain was linear (that is, there was no branching). Furthermore the Volz and Heckathorn variance estimation formula [12]has only been published in a form suitable for binary variables, while in our analyses we also deal with multi-category variables. As a result we decided to implement a simple model-based bootstrap estimator consistent with the underlying statistical model that is used throughout the RDS literature. This estimator is described in full below. All estimates were independently derived using the formulas described in Volz and Heckathorn[2008] [12] and Salganik[16] using the R statistical language **(**www.r-project.org**)**. Pre-processing of the data was performed using Stata[17].

The figure in the showing the chains of recruitment in RDS were produced by Netdraw with data output from RDSAT using the following settings: 1,000,000,000 iterations, spring embedding, distance between components:30, proximities: geodesic distances.

RDS-2 Confidence Intervals Estimation

In this procedure one defines a resampling scheme that can be used to create a collection of samples that, one hopes, have a distribution similar to the actual sampling process that generated the data. To make this concrete, suppose we wish to estimate the proportion of the population who are in some subgroup A (say, for example that A represents members of the population who test positive for HIV, or some other sexually transmitted disease). In this situation we wish to estimate both the proportion of the population in A, and (in order to examine the statistical significance of the result) we need to compute a confidence interval. Salganik's procedure provides a recipe for computing such a confidence interval as follows. First, construct the following algorithm for computing a single bootstrap estimate:

1. Divide the sample into two subsets: A[rec] consisting of individuals recruited by members of A, and B[rec] consisting of individuals recruited by members in the complement of A.

2. Select a “bootstrap seed” from the data by selecting an observation of the sample at random (i.e. make a random draw from the sample by selecting from amongst the observations so that each observation has equal probability of being drawn).

3. Starting with the bootstrap seed select a new observation by selecting at random (with equal probability) from either A[rec] or B[rec] depending on the group membership of the bootstrap seed. Continue this process recursively until one has selected a sequence of observations of the same size as the original sample. One then has a single bootstrap sample whose size is identical to the original data set. The data in the sample consist of an observed group membership and self-reported network size. A set of data constructed in this manner will be called a bootstrapped data set, or bootstrapped data for short.

Next we use 1-3 above to create many bootstrapped data sets:

4. Repeat 1-3 above until one has many versions of the bootstrapped data (i.e one has many artificial data sets, each constructed by the procedure above).

5. To compute a confidence interval for an estimator of p[A] (in principal any estimator, though at the time that Salganik was writing Volz's RDS-2 had not be developed yet), apply the estimator to each member of the collection of data sets above to yield a collection of bootstrapped estimates. This collection can be used to compute confidence intervals by either computing the variance of the bootstrapped variance (which is what Salganik does) or by reporting 95% central quantiles of the bootstrapped estimates. The latter strategy has the advantages that it is a standard approach in the bootstrap literature[18]and because it automatically provides intervals that are constrained to be within the interval [0,1] as is natural for a proportion estimate.

In order to examine the implications of using this procedure to compute confidence intervals, we recast this algorithm in a mathematically equivalent form that describes the model under which groups and degrees are simulated for the bootstrapped data. First, in order to simulate group memberships, we approximate the probability of transitions within and between the groups A and B by constructing a Markov transition matrix

$$\hat{P}=\begin{bmatrix} n_{AA}/(n_{AA}+n_{AB}) & n_{AB}/(n_{AA}+n_{AB}) \\ n_{BA}/(n_{BA}+n_{BB}) & n_{BB}/(n_{BA}+n_{BB}) \end{bmatrix}$$

where $n_{A,B}$ is the number of observed transitions from group A to group B seen in the original data. The terms n_{AA}, n_{BA} and n_{BB} are similarly the number of observed transitions from the group indicated by the first subscript to the group indicated in the second subscript. This matrix gives the exact probabilities of the transitions between groups under Salganik's bootstrap. In terms of statistical modeling, is the maximum likelihood estimate for the transition probabilities under a first order Markov model for the group transitions observed in the sample (see Anderson[19] for classical material on inference under this model, see either Volz[12] or Goel and Salganik[20] for detailed discussions of in the context of RDS). In Salganik's bootstrap, once groups have been simulated, we can simulate degrees by making a random draw from the observed degrees for the appropriate group. The entire process of creating a single bootstrapped data set can be described as follows.

1. Select a seed, y_0^{boot} by making a random draw from the observed sample, thus y_0^{boot} will be in A with probability $n_A/(n_A+n_B)$ and in B with probability $n_B/(n_A+n_B)$ where n_A and n_B are the number of observations in the sample from groups A and B respectively.

2. Select y_1^{boot}through y_n^{boot}iteratively by using the transition probabilities determined by.

In other terms, if $y_i^{boot}$is an A, theny_{i+1}^{boot}will be an A with probability $n_{AA}/(n_{AA}+n_{AB})$and in B with probability $n_{AB}/(n_{AA}+n_{AB})$.

3. After selecting y_1^{boot}through y_n^{boot}, select bootstrapped degrees d_1^{boot}through $d_n^{boot}$by selecting d_i^{boot}randomly from the observed degrees in the group corresponding too y_i^{boot}. Again, to be concrete, if$y_i^{boot}$is A then we select $d_i^{boot}$by making a random draw from the observed degrees for group A. If$y_i^{boot}$is B then we select$d_i^{boot}$by making a random draw from the observed degrees for group B.

The description above can be summarized quite concisely by saying that the Salganik bootstrap (i) models group membership (in A or B) as a linear first order Markov chain of length $n$with transition probabilities and (ii) models degrees as conditionally independent given group membership. Before we discuss the potential shortcomings of this approach we briefly describe the modeling assumptions behind the other variance estimation approach currently in use with RDS data.

There are two features of the above bootstrap procedure that are worth noting. First, the above procedure can be used to estimate the sampling distribution of any RDS estimator. This is because the bootstrap method is primarily a procedure for creating bootstrapped data. Thus, one uses the procedure to construct a large number of synthetic data sets whose distribution, one hopes, matches the sampling distribution of the actual RDS process. Then, in order to estimate the sampling distribution of a population estimator, one applies that estimator to each of the bootstrapped data sets in turn in order to create a large sample of bootstrapped population estimates. Consequently, one can apply this approach to any RDS estimator, including the RDS-2 estimator developed by Volz.

The second feature that is worth noting is that in Salganik's procedure there are two factors that clearly influence the ability of the bootstrap to approximate the actual sampling distribution. The first of these is that one replaces the branching observations of the RDS sampling process with a linear chain. One would expect therefore that a bootstrap method that uses the same branching procedure as the data collection process would do a better job of replicating the sampling distribution of RDS. The danger that a linear chain runs the risk of underestimating variance has been observed previously by Goel and Salganik (2009)[20]. As a result in our implementation we have used the observed branching structure of our sample, rather than a linear structure. The second feature is that Salganik samples the entire data set when selecting seeds. We believe that this approach is contrary to the very motivation for RDS: the seeds are drawn from an accessible stratum of the target population and are surely not distributed in the same manner as the actual RDS sample. As a result our bootstrap procedure treats seeds as a fixed aspect of the sampling design since they are selected by the researcher. Thus our bootstrap algorithm can be described as follows:

1. Select bootstrapped seeds, y_0^{boot} by making as identical to the observed seeds.

2. For each wave in the data set, select member y_i^{boot} by using the transition probabilities in

and the value of the recruiter $y_j^{boot}$where $j$is the recruiter of $i$in the original data. In other terms, if $y_j^{boot}$is an A, theny_i^{boot}will be an A with probability $n_{AA}/(n_{AA}+n_{AB})$and in B with probability $n_{AB}/(n_{AA}+n_{AB})$.

3. After selecting y_1^{boot}through y_n^{boot}, select bootstrapped degrees d_1^{boot}through d_n^{boot}as before. Thus we select d_i^{boot}randomly from the observed degrees in the group corresponding too y_i^{boot}. Again, to be concrete, if$y_i^{boot}$is A then we select $d_i^{boot}$by making a random draw from the observed degrees for group A. If$y_i^{boot}$is B then we select$d_i^{boot}$by making a random draw from the observed degrees for group B.

In order to obtain confidence intervals we use the above method to simulate 100000 bootstrapped data sets and apply the RDS-2 estimator to each of these yielding 100000 bootstrapped population estimates. The confidence intervals reported as thus the central 0.025% to 0.975% quantiles of the bootstrapped population estimates.

Methods to Compare Characteristics of PLACE and RDS Sampled Populations

We expected that 15% of women engaged in sex work would have a positive rapid test. Assuming that the population of sex workers in Liuzhou is approximately 5,000, a sample size of 380 sex workers in each group would have 80% power to detect a difference of +/- 5% , assuming alpha=0.05 and the design effect is 2. We aimed for at least 400 sex workers in each group and planned on a larger sample for the RDS arm. Characteristics of the two samples and the corresponding confidence intervals were done separately. For the PLACE estimates, SAS

Proc Survey Freq was used with appropriate sampling weight and clustering by strata. We compared RDS estimates as calculated by RDSAT software and by RDS-2 estimates as described above.

For the multivariable analysis, the two data sets were combined and analyzed using Proc Genmod in SAS using the method as recommended by Cole[21]. The models were run with RDSAT weights for the RDS participants and separately with RDS-2 weights. Weights for the RDS sample were estimated using two different methods with similar results except that prevalence ratios could not be estimated using weights output from RDSAT software. Data from each PLACE worker were weighted based on the probability that the venue where the worker was interviewed was selected for a venue visit, the probability that the venue was selected for worker visits, the proportion of venues willing to participate, and the proportion of workers who participated.

We assumed an independent correlation matrix. The models were run using PROC GENMOD [21] with the weights estimated separately for persons recruited by PLACE and RDS. The natural log of the probability of the jth person in the ith cluster having a positive rapid test for syphilis was modeled as a linear function:

$$\mathbf{ln}[E(Y_{ij})] = \beta_0 + \beta_1 \delta_{ij} + \beta_2 Age_{ij}$$

where $[E(Y_{ij})]$ is the expected value of the probability of having a positive rapid test for the jth person in the ith cluster; δ_{ij} indicates whether the person is in the PLACE arm or not; and age is the respondent's age.

The code for the model follows. For RDS, class is the recruitment chain; For PLACE, class is the stratum from which the participant was sampled.

```
proc genmod  data=china.modeldata4 descending;
class cluster; model syphilis2 = place   / d=b link=log;
scwgt normalw1; repeated subject = cluster / type=ind;
title ratio place  all weighted;  estimate "prevalence ratio" place  1 / exp;
output out=results5 p=prevalence;
```

Supplementary Figure 2 was constructed using the predicted prevalence of a positive test from the binomial regression model similar to the model indicated above. The model-predicted probability of a positive screening test by age was lowest for RDS sex workers and highest for rural PLACE sex workers.

The model used for the graph included two indicator variables to represent the three groups on the graph (RDS, PLACE urban, and PLACE rural), and included interaction between group and age. Age was modeled using restricted cubic splines with knots at 16, 20, and 30 years. The macro "RCSPLINE" was used[22]. The figure was constructed from a model containing only individuals age 35 and younger and does not present the probability of infection at older ages. A comparison of the probability of infection by age among older women is not presented because the probability estimates were unstable at older ages among PLACE women because there were few older sex workers in the PLACE arm (n=26). In the RDS arm, women older than 35 comprised 12% of the sample and 35% had a positive rapid test compared to 3.8% of younger women. Of the 64 infections among the RDS arm, 42 were older than 35. In the PLACE arm, the proportion of women with a positive test was similar for women older and younger than 35 (26% vs 23%), although the confidence intervals were much wider around the estimate for the older women (0.0,53.5 vs. 11.9,35.0) due to the small sample size in the PLACE arm.

Overall, there was not a statistically significant difference in the RDS and PLACE estimates of the percentage of sex workers recruiting in urban Liuzhou who had a positive rapid test for syphilis (8.2 vs 17.8). The study was not powered to assess differences between the urban and rural sub-groups.

The table below compares RDSAT and RDS –II estimates.

Comparison of Estimates from RDSAT and RDS II

Table 1. Sociodemographic and Behavioral Characteristics of Female Sex Workers (FSW) in Liuzhou Recruited by Respondent-driven sampling (RDS): Comparison of RDSAT and RDS II estimates.

	RDS			
	RDSAT		RDS-II	
	%	95%CI	%	95%CI
total	576		576	
Age group				
15-19	13.1	8.1, 17.8	11.42	8.0,20.2
20-24	32.1	24.6, 40.4	31.35	23.0,42.6
25-29	17.4	13.0, 22.4	20.23	12.7,23.2
30-34	16.5	10.6, 21.7	19.5	11.1,22.3
35-39	10.9	6.1, 16.9	10.70	4.8, 16.8
40+	10.1	3.8, 18.5	6.81	2.1,19.8
Residence				
District of RDS Office	47.0	40.7, 53.0	46.21	40.8, 53.4
Other Urban districts	50.5	44.6, 57.0	51.28	44.0, 56.6

	RDS			
	RDSAT		RDS-II	
	%	95%CI	%	95%CI
Rural Counties	2.5	1.2, 4.0	2.51	1.3, 4.1
Marital status				
Never Married	62.9	55.1, 69.8	62.20	54.8, 70.7
Divorced/Widowed	23.5	17.9, 29.8	24.01	17.4, 29.6
< Junior High Education	25.8	19.7, 31.8	25.34	19.8, 31.5
Monthly Income RMB	n/a		4,888	
Aged<15 at first sex	2.2	0.9, 3.9	2.09	0.9, 3.9
Ever Arrested	10.2	6.9, 13.9	10.64	7.2, 14.0
> weekly Alcohol use	37.7	30.1, 45.8	36.6	28.1, 45.6
Ever injected drugs	2.4	0.4, 5.1	1.97	0.4,5.4
Sex work past 4 weeks	100		100	
>10 Partners past 4 weeks	56.7	50.0, 64.0	55.73	55.8,72.2
Condom use at last sex	71.0	64.9, 76.8	71.5	65.3, 76.6
Solicited past year in:				
urban Liuzhou	n/a	n/a	99.39	83.4, 1.0
rural Liuzhou	3.8	2.0, 6.2	3.88	1.9, 6.0
outside Liuzhou	12.6	8.7, 16.9	12.17	8.5 16.9
Solicited past 6 months				
Outdoors	6.4	1.5, 13.3	4.1	0.9,14.0
Phone/Internet	32.0	26.3, 37.8	31.6	26.4,37.7
KTV, Karaoke	29.0	17.0, 43.4	22.4	11.9, 43.8

	RDS			
	RDSAT		RDS-II	
	%	95%CI	%	95%CI
Hair Salon	12.4	6.5, 18.7	12.6	7.4,19.7
Massage	35.2	26.6, 44.5	33.9	24.0, 47.1
HIV tested, know results	28.3	22.6, 34.1	28.9	22.9, 35.6
Syphilis test past year	7.3	4.5, 10.3	7.59	4.8, 10.5

Note: RDSAT, Respondent-driven sampling analysis tool; CI, confidence interval; n/a, not available since all RDS respondents worked in urban area. The 7 RDS seeds were excluded from RDSAT and RDS-II estimates. The 47 RDS FSW who participated in the survey but refused the rapid syphilis test were included in this table. The 161 sex workers are a subset of the 680 workers recruited through PLACE. RDSAT does not calculate means. RDS-II does not calculate medians. Income is mean monthly income.

FIGURE TITLES

Supplementary Figure 1. PLACE Study flow chart, Liuzhou, China, 2009 – 2010

Supplementary Figure 2 Predicted probability of a positive rapid test for syphilis among women age 15-35 by age and urban district-county status

REFERENCES

1. Lu, F., et al., *Prevalence of HIV infection and predictors for syphilis infection among female sex workers in southern China.* Southeast Asian J Trop Med Public Health, 2009. **40**(2): p. 263-72.

2. Weir SS, Tate J, Hileman SB, et al. Priorities for Local AIDS Control Efforts: A Manual for Implementing the PLACE Method. Chapel Hill (NC): MEASURE Evaluation Project; 2005.

3. SAS Institute Inc. SAS/STAT ® 9.2 User's Guide. Cary (NC): SAS Institute Inc.; 2008. Chapter 83, The SURVEYFREQ Procedure; p. 6286-6362.

4. Goel S, Salganik MJ. Assessing respondent-driven sampling. *Proc Natl Acad Sci U S A.* 2010;**107**:6743-6747.

5. Abdul-Quader AS, Heckathorn DD, Sabin K, et al. Implementation and analysis of respondent driven sampling: lessons learned from the field. *J Urban Health*. 2006;**83**:i1-i5.

6. Handcock MS, Gile KJ. Modeling Social Networks from Sampled Data. *Ann Appl Stat.* 2010;**4**:5-25.

7. Johnston LG, Khanam R, Reza M, et al. The effectiveness of respondent driven sampling for recruiting males who have sex with males in Dhaka, Bangladesh. *AIDS Behav*. 2008;**12**:294-304.

8. Goel S, Salganik MJ. Respondent-driven sampling as Markov chain Monte Carlo. *Stat Med*. 2009;**28**: 2202-2229.

9. Johnston LG, Sabin K, Mai TH, et al. Assessment of respondent driven sampling for recruiting female sex workers in two Vietnamese cities: reaching the unseen sex worker. *J Urban Health*. 2006;**83**:16-28.

10. Family Health International. Behavioral Surveillance Surveys: Guidelines for Repeated Behavioral Surveys in Populations at Risk of HIV. Arlington (VA): Family Health International; 2000.358 p.

11. Volz E, Wejnert C, Degani I, et al. Respondent-Driven Sampling Analysis Tool (RDSAT) Version 5.6. Ithaca (NY): Cornell University; 2007.

12. Volz E, Heckathorn DD. Probability Based Estimation Theory for Respondent Driven Sampling. *Journal of Official Statistics.* 2008.**24**:79-97.

13. Salganik MJ, Heckathorn DD. Sampling and estimation in hidden populations using respondent-driven sampling. *Sociol Methodol.* 2004;**34**:193-239.

14. Heckathorn DD. Respondent-driven sampling: A new approach to the study of hidden populations. *Soc Probl*. 1997;**44**:174-199.

15. Wejnert C. An Empirical Test of Respondent-Driven Sampling: Point Estimates, Variance, Degree Measures, and out-of-Equilibrium Data. *Sociol Methodol.* 2009;**9**:73-116.

16. Salganik MJ. Variance estimation, design effects, and sample size calculations for respondent-driven sampling. *J Urban Health*. 2006;**83**:i98-i112.

17. Stata Statistical Software. College Station (TX): StataCorp LP; 2007.

18. Davison AC, Hinkley DV. Bootstrap methods and their application. New York: Cambridge University Press; 1997.

19. Anderson TW, Goodman LA. Statistical-Inference About Markov-Chains. *Annals of Mathematical Statistics*. 1957;**28**:89-110.

20. Goel S, Salganik MJ. Respondent-driven sampling as Markov chain Monte Carlo. *Stat Med*. 2009;**28**:2202-2229.

21. Cole SR. Analysis of complex survey data using SAS. *Comput Methods Programs Biomed*. 2001; **64**:65-69.

22. Devlin TF, Weeks BJ. Spline functions for logistic regression modeling. Proceedings of the 11th Annual SAS Users Group International; 1986 February 9-12; Atlanta, Georgia. Cary (NC): SAS Institute, Inc.