



Tobacco Control and Public Health in Eastern Europe

Pravosud, V. O., & Andreeva, T. I. (2016). Respondent-driven sampling: characteristics associated with productive and non-productive recruitment. *Tobacco Control and Public Health in Eastern Europe*, 6(1), 31–44. doi:<http://dx.doi.org/10.6084/m9.figshare.4495910>

Respondent-driven sampling: characteristics associated with productive and non-productive recruitment

Vira O Pravosud, Tatiana I Andreeva

BACKGROUND: While studying the population of female sex workers (FSWs) in Ukraine using respondent-driven sampling (RDS) methodology, we noticed that seeds could generate recruitment chains different with regards to their length. We presumed that homogeneity among respondents could be associated with prolonging the recruitment chains. Therefore, this study aimed to explore if the tendency of forming long and short chains was related to specific characteristics of FSWs.

METHODS: The study employed data from a bio-behavioral survey among FSWs, conducted in 2009, which used RDS method to recruit 975 participants in six Ukrainian cities with four or five initial seeds. Pearson chi-square test was used to evaluate differences in 29 parameters (demonstrating different ways of client search, social interaction, health and behavior of FSWs) between respondents who were included in either five productive (≥ 80 participants) or 20 non-productive chains. Stratified analysis with multivariable adjusted logistic regression model was carried out to show the association of examined covariates with recruitment productivity.

RESULTS: The findings revealed that affiliation to any non-governmental organization (NGO) for FSWs and injection drug users (IDUs) was a significant interaction factor. For those who were not members of the NGOs, while controlling for other covariates in the model, older age, higher level of education, being current or previous drug user, not appealing to state clinics for treatment of sexually transmitted infections (STIs), and involvement in escort services had larger adjusted estimated odds of productive recruitment of FSWs in the study. On the other hand, for members of any NGOs for FSWs/IDUs, street-based way of client search was strongly associated with productive recruitment.

CONCLUSIONS: The results of this study suggest that members and non-members of NGOs for FSWs/IDUs have different characteristics that allow them performing productive recruitment. This should be considered while developing a sampling technique for studies of hard-to-reach populations.

KEYWORDS: Ukraine; respondent-driven sampling; FSWs; wave; productive recruitment chains.

BACKGROUND

In Ukraine, respondent-driven sampling (RDS) is a relatively new methodology of obtaining data on bio-behavioral characteristics and disease prevalence among hidden populations such as injection drug users (IDUs), men who have sex with men (MSM), and female sex workers (FSWs), who are considered high-risk groups in regards

to poor health outcomes including high prevalence of sexually transmitted infections (STIs) (Berleva et al., 2010; Spreen, 1992).

The RDS methodology implies that the recruitment of participants is done by the respondents (starting from the initial seeds – those recruited by the investigator), where one can recruit up to three individuals with each recruiter and re-

cruitee getting a unique coupon number (Heckathorn, 1997). The process of recruitment lasts until everyone, or a sufficient number of respondents, is enrolled in the study (Salganik & Heckathorn, 2004).

The RDS methodology gives a possibility to form a probability sample if the equilibrium assumption is met. This assumption implies that

the final composition of the sample, according to the critical characteristics and behavioral factors, will not depend on the initially selected seeds if there were a sufficient number of recruitment waves. In fact, after a certain wave (usually at least after the fifth) the sample composition becomes stable, and network connections occur randomly, ensuring the representativeness of the sample for further analysis (Heckathorn, 2002; Salganik & Heckathorn, 2004). This can be explained by the increasing number of waves and therefore, the number of recruitees, which, in turn, minimizes the sampling error. Previous research showed that equilibrium could also be achieved within a limited number of waves (Wang et al., 2005), or in the sample where seeds had different levels of recruitment intensity, generating either long or short chains (Lachowsky et al., 2016; Saliuk & Andreeva, 2010). However, this brings us to the issue of obtaining a sample where characteristics of respondents from long recruitment chains can be more prevalent, which may lead to under- or overestimation of study outcomes associated with those characteristics. The existing literature does not address this question in depth; thus, further research is needed to answer, which characteristics can be more or less prevalent in a sample, leading to either over- or underestimation of corresponding health outcomes.

This study aimed to determine the factors that could be associated with recruitment intensity in different chains with which the sample of FSWs was formed using RDS technique. Specifically, the primary objective of this study was to determine if the tendency of forming long and short chains was associat-

ed with characteristics of FSWs. We assumed that long recruitment chains were more likely formed by more homogeneous groups of individuals with regards to demographic, behavioral, or social characteristics, with wider social networks as it was shown in the study from Canada, although among a different hidden population – MSM (Forrest et al., 2016), in the study from Kenya among MSM, FSWs, and IDUs (Okal et al., 2016), and among young men who have sex with men (YMSM) in the U.S. (Kuhns et al., 2015).

Although the RDS technique was introduced more than two decades ago, it is still considered a relatively new sampling methodology. This, in turn, generates many discussions about advantages and disadvantages of its implications, especially when the object of the research is a hidden population. Therefore, the clinical significance of this study lies in providing additional knowledge for investigators who focus on research among hard-to-reach populations, and would like to use an appropriate sampling technique, which would allow them to obtain a representative sample.

METHODS

Study sample and design

This was an observational study with exploratory analysis employing data from a bio-behavioral survey among FSWs, conducted in 2009 by the Kyiv International Institute of Sociology and commissioned by the International Charitable Foundation (ICF) Alliance for Public Health (formerly “International HIV/AIDS Alliance in Ukraine”) in the frame of the program “Tackling the HIV/AIDS epidemic in Ukraine” (supported by the

Global Fund to Fight AIDS, Tuberculosis and Malaria) (Grushetsky, 2010). Overall, there were 975 participants recruited using RDS in six Ukrainian cities (Donetsk, Cherkassy, Chernihiv, Kyiv, Kharkiv, and Zaporizhia). The number of primary respondents (seeds) was five for Kharkiv and four for each of the remaining five cities. The primary respondents had to meet several requirements: to be 14 or older, have at least ten friends among other FSWs; use multiple ways of client search; and live in different parts of the city (Grushetsky, 2010).

The investigators were interested to see if specific characteristics of the respondents were associated with productivity of recruitment (i.e. the length of the recruitment chain) generated from different seeds. We operationalized a recruitment chain as “productive” or “non-productive” based on the number of people included in it. The cutoff was decided based on the variable distribution and with the aim to maximize the power of the analysis. Hence, a chain was considered productive if it included 80 or more respondents. The number of respondents in non-productive chains was significantly lower ranging from 2 to 43 recruitees. There were 495 (50.8%) respondents in the first group consisting of productive chains from five seeds representing FSWs from Donetsk, Zaporizhia, Kharkiv, and Kyiv; and 480 (49.2%) FSWs in the second group including twenty non-productive chains from all six cities.

Characteristics of the respondents

Overall, we examined a set of 29 characteristics of respondents which could be associated with the

intensity of recruitment, including demographic characteristics, different ways of client search, social interaction, health outcomes, and several behavioral risk factors of FSWs. In the survey, the respondents were asked about their age, the language they usually spoke (Russian or Ukrainian) and the level of education they had (elementary, secondary, high school or uncompleted high education, technical school, or university). The FSWs were asked to confirm or deny whether they were searching for their clients on streets, highways, railway stations, in hotels, saunas, online or by phone, in entertainment facilities, and if the subjects were involved in escort-services, or preferred other ways of client search. The respondents were also asked if they were diagnosed during the 12 months prior to the date of the interview with syphilis, hepatitis C, gonorrhoea, genital herpes, chlamydia, trichomoniasis, candidiasis, tuberculosis, and if they were aware of their HIV status, and whether they appealed to state clinics in case of STIs. In addition, the recruits were questioned about their behavioral risk factors such as the use of protection while having physical intercourse, daily alcohol consumption, and usage of any substance. Also, the respondents were asked if they were members of any NGO for FSWs/IDUs, if they appealed to the associations of people living with HIV or to self-support groups for FSWs. Finally, either positive or negative HIV status was recorded based on the result of the rapid test.

Statistical analyses

Statistical analyses were carried out using SPSS version 23. Exploratory analyses included the use of descriptive statistics presented in

Table 1. Specifically, means, standard deviations, quartiles, and ranges were examined for continuous variables, whereas frequencies and percentages were examined for categorical variables. Pearson χ^2 and Fisher's Exact Tests (if the expected cell was ≤ 5) were used to determine if differences in characteristics of respondents in two groups from productive and non-productive chains were statistically significant on the five percent significance level.

Listwise deletion of missing data was used in the regression procedures steps as there were only five missing values not related to the outcome variable. Some other variables had additional level corresponding to "No response" value. However, those cases were not excluded from the analysis in order not to lose power.

This study assessed the unadjusted (Table 2) as well as stratified and adjusted (Table 3) associations between the independent covariates and the outcome – productivity of the recruitment chains. Therefore, the bivariate and multivariable logistic regression models were carried out with the productivity of chains as a dependent variable. As it was mentioned earlier, a chain was considered productive if it had included 80 participants or more. Other characteristics were treated as independent covariates.

Variables presented in Table 1, for which we observed significant differences between participants from productive and non-productive chains, were then included in the bivariate logistic regression models. To control for potential confounding factors, an adjusted multivariate logistic regression analysis was performed. Backward elimination was applied for this model,

starting with the main effects of all potential predictors which were significant in bivariate analysis. Comparing to the model with all the covariates, a slight decrease in AIC value of the reduced model was observed (1205.595 vs. 1202.076, respectively).

Moreover, due to the nature of the data, we had to consider a possibility of multicollinearity for a number of independent covariates. Specifically, the descriptive analysis revealed that the FSWs were mentioning several ways of searching for clients. Therefore, to avoid inflation of parameter estimates in the adjusted logistic regression model, we performed a collinearity analysis for categorical variables (Table 4 in Appendices) judging from the significance and the magnitude of the phi-coefficients (Ekström, 2011; Muir, Berg, Chesworth, Klar, & Speechley, 2010). The results showed that street-based client search was moderately positively correlated with searching for clients on highways and railway stations, and a large share of women who were involved in escort services (almost 85%) also confirmed that they could be searching for clients online or by phone demonstrating a highly significant correlation of .147.

Thus, due to backward selection method and inferences of the analysis of multicollinearity, the final adjusted multivariable logistic regression model was reduced to seven variables (Table 5 in Appendices). Specifically, it included such variables as age groups, level of education, searching for clients in streets, involvement in escort services, use of any substance, affiliation to any NGO for FSWs and/or IDUs, and appeal to state clinics for treating STIs. Results

Table 1. Descriptive statistics for recruits in productive and non-productive chains

Variable	Values	Indicators	Productive (495, 50.8%)	Non-Productive (480, 49.2%)	p-value
Age groups					
	14–18	n, (%)	26 (5.3)	60 (12.5)	<.001
	19–25	n, (%)	117 (23.7)	170 (35.4)	
	26–55	n, (%)	350 (71.0)	250 (52.1)	
Age (years)					
		Mean (SD)	29.92 (7.405)	27.09 (7.539)	<.001
		Median (Q1, Q3)	30 (25, 35)	26 (21, 32)	
		(Min, Max)	(14, 54)	(14, 55)	
Language					
	Ukrainian	n, (%)	5 (1)	104 (21.7)	<.001
	Russian	n, (%)	490 (99)	376 (78.3)	
Education					
	Elementary education (<9 grades)		20 (4.1)	11 (2.3)	.003
	Secondary education (9 grades)		73 (14.9)	90 (18.8)	
	High school/ uncompleted high education		275 (56)	264 (55)	
	Technical/vocational school		69 (14.1)	88 (18.3)	
	University		54 (11)	27 (5.6)	
<i>The ways of searching for clients:</i>					
In streets					
	Yes		200 (40.4)	143 (29.8)	.001
On highways					
	Yes		169 (34.1)	120 (25)	.002
In hotels					
	Yes		86 (17.4)	104 (21.7)	.091
In saunas					
	Yes		104 (21.0)	116 (24.2)	.238
On railway station					
	Yes		42 (8.5)	18 (3.8)	.002
Online/phone					
	Yes		290 (58.6)	274 (57.1)	.635
In bars, clubs, casinos, discos, etc.					
	Yes		204 (41.2)	200 (41.7)	.885
Escort-services					
	Yes		44 (8.9)	22 (4.6)	.007
Other ways of client search					
	Yes		45 (9.1)	28 (5.8)	.053
<i>Behavioral risk factors:</i>					
Condom Use					
	No		23 (4.6)	17 (3.5)	.385
Alcohol use					
	Daily alcohol use		53 (12.4)	43 (10.9)	.520
Any drug use					
	Yes		147 (29.7)	70 (14.6)	<.001
	Previous experience of drug use		81 (16.4)	35 (7.3)	
	No		267 (53.9)	370 (77.1)	
	Doubts in response		0 (.0)	5 (1.0)	
Affiliation to any NGO for FSWs/IDUs					
	Yes		199 (40.2)	78 (16.3)	<.001
	No		293 (59.2)	401 (83.5)	
	No response		3 (.6)	1 (.2)	

Table 1. Descriptive statistics for recruits in productive and non-productive chains (continued)

Variable	Values	Indicators	Productive (495, 50.8%)	Non-Productive (480, 49.2%)	p-value
Appeal to a state clinic for treatment of STIs					
	Yes		68 (13.7)	126 (26.3)	<.001
	No		423 (85.5)	353 (73.5)	
	No response		4 (.8)	1 (.2)	
Appeal to the associations of people who live with HIV					
	Yes		32 (6.5)	21 (4.4)	.216
	No		462 (93.3)	459 (95.6)	
	No response		1 (.2)	0 (.0)	
Appeal to self-support groups for FSWs					
	Yes		18 (3.6)	14 (2.9)	.820
	No		476 (96.2)	465 (96.9)	
	No response		1 (.2)	1 (.2)	
<i>Diseases occurred during the last 12 months</i>					
Syphilis					
	Yes		7 (1.4)	9 (1.9)	.571
Hepatitis C					
	Yes		59 (11.9)	21 (4.4)	<.001
Gonorrhea					
	Yes		19 (3.8)	11 (2.3)	.162
Genital Herpes					
	Yes		26 (5.3)	29 (6.0)	.593
Chlamydia					
	Yes		30 (6.1)	45 (9.4)	.052
Trichomoniasis					
	Yes		16 (3.2)	15 (3.1)	.924
Candidiasis					
	Yes		95 (19.2)	111 (23.1)	.133
TB					
	Yes		5 (1.0)	7 (1.5)	.526
<i>HIV status</i>					
Self-reported HIV status known on the moment of interview (n=514)					
	Positive		56 (17.6)	26 (13.3)	.191
	Negative		262 (82.4)	170 (86.7)	
HIV positive test					
	Yes		105 (21.2)	68 (14.2)	.004
	No		390 (78.8)	412 (85.8)	

Means and standard deviations (in parentheses) are given for continuous variables, whereas frequencies and percentages (in parentheses) are given for categorical variables. The two-sided p-values $\leq .05$ for Wilcoxon-Mann-Whitney, Pearson chi-square, and Fisher's Exact Tests (if the expected cell was ≤ 5) were considered statistically significant (N=975).

The Wilcoxon-Mann-Whitney test was used to determine if two groups were different with regards to age. The Wilcoxon-Mann-Whitney test was carried out, since age treated as a continuous variable failed the normality tests (Kolmogorov-Smirnov test: $p_{age} < .001$, and Shapiro-Wilk: $p_{age} < .001$).

from bivariate logistic regressions demonstrated unadjusted statistically significant association at .05 level for age (treated as both continuous and categorical variables) as a predictor. However, treating age variable as a categorical rather than a continuous one was more informative due to lower AIC value (1314.46 vs. 1316.35). HIV test result and hepatitis C status were no longer significantly associated with the outcome in the multivariate adjusted logistic regression model.

Both bivariate and multivariate analysis revealed that affiliation to any NGO for FSWs/IDUs was strongly associated with recruitment productivity. In the survey description, it was mentioned that the investigators consulted abovementioned NGOs about recruitment of the initial respondents, and this was how the seeds were reached (Grushetsky, 2010). After adjusting for other covariates in the final multivariable model, we could observe more than 10 percent differences in the estimated ORs for a number of independent predictors as compared to their bivariate results (Table 2 vs. Table 5). Thus, we were concerned about affiliation to any NGO for FSWs/IDUs being a strong effect modifier or a confounding factor. To check that, we firstly assessed significance of the effect of interaction terms, by including all possible two-way interactions with NGO membership status to the abovementioned reduced final multivariate logistic regression. The results revealed several statistically significant interactions of NGO membership status with drug use ($p < .001$), appeal to state clinics for treating STIs ($p = .007$), and having technical school or college education ($p = .034$); and with having university degree ($p = .090$). Further on, we performed a strati-

fied analysis (Table 3a and Table 3b) based on the interaction factor – affiliation to any NGO for FSWs/IDUs. As the result, strata-specific point estimates of odds ratios were sufficiently different for members and non-members of any NGOs for FSWs/IDUs implying the presence of effect modification. Thus, we adjusted final multivariable logistic regression model accordingly and reported strata-specific results.

RESULTS

Univariate descriptive statistics

Table 1 presents descriptive statistics for women in each of the two groups of interest based on the outcome variable defined as recruitment productivity observed in five chains with 80 or more respondents. According to these results, a number of variables seemed to be distributed approximately evenly between the two groups. In particular, we refer to such variables as searching for clients in hotels, saunas, entertainment facilities, not using condoms while having physical intercourse, daily alcohol use, affiliation with associations for people living with HIV or self-support groups for FSWs, having experienced tuberculosis or such STIs as syphilis, gonorrhoea, genital herpes, chlamydia, trichomoniasis, and candidiasis.

However, we observed significant differences for a number of examined characteristics of the study respondents. Specifically, higher percentages of recruits in the group from productive chains were Russian-speaking women of older age who had complete or incomplete higher education, those who searched for clients in streets, on

highways, railway stations, who were involved in escort services, had positive HIV and/or hepatitis C status, were current drug users, were involved in any FSW/IDUs NGOs, and did not choose state clinics for treatment of STIs as their preferable medical care facility.

Bivariate findings: unadjusted analysis using the logistic regression model

Table 2 presents results from the bivariate logistic regression using productivity of recruitment as the outcome of interest. As it was expected, older age, university education, searching for clients in streets, highways, railway stations, involvement in escort services, current or previous drug usage, affiliation with any NGO for FSWs or IDUs, positive HIV and/or hepatitis C status as well as appealing to public clinics for treating STIs appeared to be statistically significantly associated with recruitment productivity.

Multivariable findings: stratified analysis using adjusted logistic regression model

Table 3a presents results from our final reduced logistic regression model using the recruitment productivity as the outcome of interest for the first stratum, i.e. for those respondents who were members of any NGO for FSWs/IDUs.

Only street-based client search predictor remained statistically significant after adjusting for other covariates in the model ($p = .012$), with estimated adjusted odds [odds] of recruitment productivity of 2.268 (95% CI: 1.199–4.292).

The stratum-specific results revealed that FSWs in the youngest

Table 2. Bivariate logistic regression results using the event of productive recruitment as the outcome of interest, (N=970). Estimate corresponds to the respective non-intercept parameter estimate(s) in the given model

Variable	Values	OR	95% CI	p-value
Age groups	14–18	.310	.190–.504	<.001
	19–25	.492	.369–.654	<.001
	26–55	reference		
Age (years)		1.052	1.034–1.071	<.001
Education	Elementary education (<9 grades)	1.745	.820–3.713	.148
	Secondary education (9 grades)	.779	.548–1.107	.163
	High school/uncompleted high education	reference		
	Technical or vocational school	.753	.526–1.076	.119
	University	1.920	1.174–3.140	.009
Searching for clients in streets	Yes	1.598	1.225–2.084	.001
	No	reference		
Searching for clients on highways	Yes	1.555	1.178–2.053	.002
	No	reference		
Searching for clients on railway station	Yes	2.380	1.350–4.196	.003
	No	reference		
Providing escort-services	Yes	2.031	1.198–3.444	.009
	No	reference		
Drug use	Yes	2.910	2.102–4.029	.696
	Previous experience of drug use	3.207	2.093–4.913	.000
	No	reference		
Affiliation to any NGO for FSWs/IDUs	Yes	3.492	2.581–4.723	<.001
	No	reference		
	No response	4.106	.425–39.669	.222
Appeal to a state clinic for treatment of STIs	Yes	.450	.325–.625	<.001
	No	reference		
	No response	3.338	.371–30.001	.282
Hepatitis C	Yes	2.958	1.767–4.950	<.001
	No	reference		
HIV positive test	Yes	1.631	1.167–2.280	.004
	No	reference		

Table 3. Logistic regression results for the final model using the productive recruitment as the outcome of interest for Strata 1 and 2. Estimate corresponds to the respective non-intercept parameter estimate(s) in the given model

Table 3a. Stratum 1 (NGO members, N=274)

Variable	Values	Adjusted OR (NGO+)	95% CI	p-value
Age groups				
	14–18	.202	.033–1.227	.082
	19–25	1.062	.524–2.151	.868
	26–55	reference		
Education				
	Elementary education (<9 grades)	1.091	.248–4.793	.908
	Secondary education (9 grades)	.796	.373–1.699	.555
	High school/uncompleted high education	reference		
	Technical or vocational school	1.444	.658–3.170	.360
	University	.875	.336–2.283	.786
Searching for clients in streets				
	Yes	2.268	1.199–4.292	.012
	No	reference		
Providing escort-services				
	Yes	1.349	.532–3.419	.528
	No	reference		
Drug use				
	Yes	.660	.345–1.265	.211
	Previous experience of drug use	1.347	.572–3.170	.495
	No	reference		
Appeal to a state clinic for treatment of STIs				
	Yes	1.614	.658–3.959	.295
	No	reference		

Table 3b. Stratum 2 (NGO non-members, N=696)

Variable	Values	Adjusted OR (NGO–)	95% CI	p-value
Age groups				
	14–18	.689	.383–1.241	.215
	19–25	.592	.406–.862	.006
	26–55	reference		
Education				
	Elementary education (<9 grades)	1.839	.671–5.034	.236
	Secondary education (9 grades)	.527	.322–.864	.011
	High school/uncompleted high education	reference		
	Technical or vocational school	.519	.311–.867	.012
	University	2.362	1.240–4.498	.009
Searching for clients in streets				
	Yes	1.309	.902–1.901	.157
	No	reference		
Providing escort-services				
	Yes	2.353	1.084–5.108	.030
	No	reference		
Drug use				
	Yes	3.690	2.343–5.813	<.001
	Previous experience of drug use	2.679	1.524–4.712	.001
	No	reference		
Appeal to a state clinic for treatment of STIs				
	Yes	.399	.254–.626	<.001
	No	reference		
	No response	3.138	.308–32.014	.334

group had 80% lower odds to generate long recruitment chains comparing to those women aged between 26 and 55 years. Albeit, those results were significant only on a ten percent significance level ($p=.082$) due to the probable lack of power as there were only six NGO members in this age group.

Other covariates presented in Table 3a appeared to be insignificant, although the direction of the association remained similar as in the bivariate logistic regression models. However, unexpectedly, some predictors became not only insignificant but also were characterized with an opposite direction as compared to the results from their bivariate models.

Table 3b presents results for the second stratum, i.e. for those respondents who were non-members of any NGO for FSWs/IDUs. Moreover, consistent with the bivariate analysis for the pooled data (Table 2), the stratum-specific results revealed that FSWs aged 19–25 had significantly lower odds of .592 ($p=.006$) to generate long recruitment chains comparing to those women aged between 26 and 55 years.

Women who graduated from a university had 2.362 times larger odds of recruitment productivity comparing to women who still were at high school or had completed high school education (95% CI: 1.240–4.498, $p=.009$); whereas those FSWs with secondary education had 47.3% lower odds of generating long recruitment chains comparing to women with completed or just started high school education (95% CI: .322–.864, $p=.011$). Finally, those with technical or vocational school education had .519 times lower adjusted odds of recruitment productivity as compared

to women who were at high school or had completed high school education (95% CI: .311–.867, $p=.012$).

Searching for clients in streets was no longer significant ($p=.157$); albeit likewise in the bivariate model, we observed larger odds of recruitment productivity (OR:1.309, 95% CI: .902–1.901) as compared to non-street-based FSWs.

FSWs who were involved in escort services had 2.353 larger adjusted odds of recruitment productivity (95% CI: 1.084–5.108, $p=.030$) as compared to those who were not.

Furthermore, current and previous drug users had, respectively, more than three- and twofold larger adjusted odds of productive recruitment comparing to non-drug users (95% CI: 2.343–5.813, $p<.001$; 95% CI: 1.524–4.712, $p<.001$, respectively). While controlling for other covariates, those participants who went to public clinics for treating STIs had 60% lower odds (95% CI: .254–.626, $p<.001$) of forming productive recruitment chains comparing to those women who did not prefer public medical facilities.

DISCUSSION

The aim of this study was to determine which characteristics of respondents were associated with successful recruitment of FSWs. Initially, we assumed that the ability to create long and short recruitment chains depends on whether participants belong to wide or isolated networks of FSWs, which differ based on demographic, social, behavioral, or health characteristics. For instance, street-based FSWs may know many FSWs who also prefer similar ways of client search and thus, will have more op-

portunities to recruit other FSWs in a study. At the same time, women of commercial sex who mainly used Internet or phone contacts to find their clients may not know other FSWs who use the same methods. Therefore, these recruits will more likely form short recruitment chains.

The results from univariate analysis and the non-stratified adjusted logistic regression model indicated that, indeed, respondents who may belong to specific, although different, social networks of FSWs, had significantly larger odds of productive recruitment. In particular, those were street-based FSWs, who were probably drug users, and women who were involved in escort services.

However, this exploratory analysis revealed a strong effect modification based on the affiliation to any NGO, which focuses mainly on FSWs and/or IDUs. Having been highly associated with productive recruitment in both univariate and multivariate logistic regression models for the pooled data, this predictor raised some doubts, first of all, since survey investigators were in contact with abovementioned NGOs to get access to initial seeds and involve them in the study (Grushetsky, 2010). Therefore, there might have been a possibility that other members of those NGOs were very likely to be recruited as well. We assumed that overall participants, who were members of abovementioned NGOs, were more likely to form well-networked groups of FSWs and IDUs, who were also members of same or similar NGOs, and therefore, had higher chances to perform successful recruitment of their acquaintances.

Secondly, we were concerned that significantly larger odds of recruit-

ment productivity associated with NGO membership in the pooled dataset could be explained by easier access of prevention programs in Ukraine to street-based FSWs with greater risks of STIs and HIV. Additionally, we observed that the majority of respondents in productive chains belonged to the older age group (26–55 years). This factor could explain a higher adherence of older women to participate in an NGO as it was demonstrated in a study from Nepal, where older and better-educated FSWs were more involved in those NGOs that provided health services (Ghimire, Smith, & van Teijlingen, 2011), and also in a study from China (Yi et al., 2010) showing that street-based female sex workers of older age were in better contact with organizations providing prevention services.

Thus, we presumed that because of having more network connections due to high possibility of NGO membership, street-based FSWs should have performed a successful recruitment of other street-based sex workers. Hence, we expected to see larger percentages of women who preferred to search for clients in streets or highways as compared to other ways of looking for clients. Although the study participants could mention several ways of client search, it was unexpected to observe sufficiently high frequencies of searching for clients via phone and Internet, or in entertainment establishments such as saunas, bars, restaurants, cafes, casinos, clubs and discos, and low negative correlations of these ways of client search with street-based work (Table 4). However, there were no differences between groups in productive and non-productive chains with regards to these places of searching for clients.

Finally, the findings of the stratified analysis and assessment of interaction suggested that affiliation to NGOs for FSWs/IDUs was a significant effect modifier. For women who were members of any NGO for FSWs/IDUs, the only covariate that remained significant even after stratification and adjustment for other covariates in the model was street-based client search showing more than twofold larger adjusted odds of recruitment productivity. For non-members of NGOs, street-based work also showed larger odds of productive recruitment, although highly insignificant.

As for the second stratum, the results indicated that women from the older age group, and also those with university education had larger odds of recruitment productivity. Thus, the observed effect of age and education on the recruitment process could no longer be questionable due to confounding factor of NGO membership as the results of pooled data may have suggested.

As expected, after stratification, drug usage was no longer a significant predictor among NGO members. The results for the second stratum, including non-members of NGOs for FSWs/IDUs, were in line with other research (Reisner et al., 2010), and revealed that current and previous drug users had larger odds of generating long recruitment chains comparing to non-users. In fact, it can be assumed that FSWs-IDUs may form a separate subgroup of commercial sex workers, members of which would be very likely to recruit their peers. A study among Mexican FSWs revealed that injection drug usage was common among female sex workers of younger age, and those who prefer to consume drugs be-

fore providing sex services (Strathdee et al., 2008). Also, injection drug usage was associated with longer time involved in sex business and thus, knowing other women who had substance abuse, higher prevalence of STIs, and were engaged in other risky behaviors including unprotected physical intercourse (Strathdee et al., 2008).

Based on the results of this study, we may assume that older and better-educated FSWs regardless of their NGO membership and those who probably have issues with substance abuse form a separate network with a high homogeneity level due to which they had a high likelihood of successful peer recruitment. The same inference can be made with regards to FSWs who were involved in escort services even if they did not have NGO membership status, as well as a separate group of NGO members for FSWs/IDUs could be formed by those women who prefer street-based client search.

Our findings also suggest that participants with abovementioned characteristics might have been overrepresented in the sample. Therefore, further research is recommended to validate heterogeneity of the sample, and also to examine whether high levels of respondents' homogeneity could have overestimated health-related outcomes which were reported based on the sample collected by RDS methodology (Grushetsky, 2010).

This study had several limitations including the exploratory nature of performed analysis and usage of secondary data. Moreover, we lacked assessment of socio-economic status (SES) of the respondents, which, in fact, could be related to the frequency of providing commercial sex services, and thus,

increased risks of STI transmission as it was shown in previous research (Deering et al., 2013; Griensven, Limanonda, Chongwatana, Tirasawat, & Coutinho, 1995; Reed, Gupta, Biradavolu, Devireddy, & Blankenship, 2010; Saggurti et al., 2012; Yi et al., 2010). However, we observed that women, although only in the second stratum, who preferred to receive medical care of STIs in public clinics had much lower odds of forming larger recruitment chains. With regards to the health care system in Ukraine, where not every citizen may afford private medical care, this variable may be considered a proxy for socio-economic status (SES).

Despite the mentioned limitations, this study had several strengths. First of all, we excluded variables that could introduce multicollinearity and inflation of parameter estimates. Secondly, the adjusted multivariable logistic regression model was carried out allowing control for potential confounding factors. Thirdly, we assessed interaction effect of a potential effect modifier and conducted an appropriate stratified analysis. Moreover, although it is difficult to consider all possible factors which may bias the final inferences, the investigators assume these results are free of residual confounding as complete case analysis was carried out. Finally, this study determined several characteristics of respondents, albeit different for two strata, which were strongly associated with productivity of recruitment even after stratification and adjusting for other covariates in the model. Hence, these findings may be considered by other investigators who plan to design a study of hard-to-reach populations of FSWs using RDS methodology.

CONCLUSIONS

The aim of this study was to explore and determine which characteristics of the respondents in the sample obtained by RDS methodology were associated with productive recruitment. The former was defined as an event of forming recruitment chains with 80 or more respondents.

The findings from our analysis revealed that affiliation to any NGO for FSWs/IDUs was a significant interaction factor. However, several factors remained significantly associated with recruitment productivity even after stratification and adjustment for other covariates in the model.

For those who were not members of NGOs, older age, higher level of education, being current or previous drug user, not appealing to a state clinic for treatment of STIs, and involvement in escort services were characterized by larger adjusted estimated odds of productive recruitment of FSWs in the study. For members of any NGOs for FSWs/IDUs, street-based way of client search showed more than twice larger odds of recruitment productivity while controlling for other covariates.

These results have methodological implications for development of the design of future studies of hard-to-reach population which would use the RDS technique.

ACKNOWLEDGMENT

We thank Alliance for Public Health (Ukraine) for providing the data collected by Kiev International Institute of Sociology (KIIS). We thank K. Fedorenko for proofread-

ing and feedback. We also thank S. Rivera Polanco for his support during the manuscript preparation.

ABOUT THE AUTHORS

Vira O Pravosud, MPH, is affiliated to the University of Kentucky, USA, and National University of "Kyiv-Mohyla Academy", Ukraine;
ORCID 0000-0002-5762-6632
Email: vira.pravosud@uky.edu

Tatiana I Andreeva, MD, Ph.D., at the time when this study was started, was affiliated with the School of Public Health, National University of "Kyiv-Mohyla Academy", Ukraine;
ORCID: 0000-0001-5663-6225
Email: tatianandreeva@gmail.com

This paper was received: November 5, 2016; accepted: December 21, 2016; published: December 30, 2016.

REFERENCES

- Berleva, G. O., Dumchev, K. V., Kobyshcha, Y. V., Paniotto, V. I., Petrenko, T. V., Saliuk, T. O., & Shvab, I. A. (2010). Analytical report based on sociological study results "Estimation of the Size of Populations Most-at-Risk for HIV Infection in Ukraine in 2009" from http://www.aidsalliance.org.ua/ru/library/our/monitoring/pdf/indd_en.pdf.
- Deering, K. N., Lyons, T., Feng, C. X., Nosyk, B., Strathdee, S. A., Montaner, J. S. G., & Shannon, K. (2013). Client demands for unsafe sex: the socio-economic risk environment for HIV among street and off-street sex workers. *J Acquir Immune Defic Syndr*, 63(4), 522-531. doi:10.1097/QAI.0b013e3182968d39
- Ekström, J. (2011). The phi-coefficient, the tetrachoric correlation coefficient, and the Pearson-Yule Debate. *Department of Statistics, UCLA*.

- Forrest, J. I., Lachowsky, N. J., Lal, A., Cui, Z., Sereda, P., Raymond, H. F., . . . Hogg, R. S. (2016). Factors Associated with Productive Recruiting in a Respondent-Driven Sample of Men who Have Sex with Men in Vancouver, Canada. *Journal of Urban Health*, 93(2), 379-387. doi:10.1007/s11524-016-0032-2
- Ghimire, L., Smith, W. C. S., & van Teijlingen, E. R. (2011). Utilisation of sexual health services by female sex workers in Nepal. *BMC health services research*, 11(1), 1.
- Griensven, G. J. P., Limanonda, B., Chongwatana, N., Tirasawat, P., & Coutinho, R. A. (1995). Socio-economic and demographic characteristics and HIV-1 infection among female commercial sex workers in Thailand. *AIDS Care*, 7(5), 557-566. doi:10.1080/09540129550126218
- Grushetsky, A. (2010). Analytic report of the survey among female sex workers in 2009 "Monitoring of behaviors and prevalence of HIV among FSW as a component of HIV surveillance of second generation".
- Heckathorn, D. D. (1997). Respondent-driven sampling: a new approach to the study of hidden populations. *Social Problems*, 44(2), 174-199.
- Heckathorn, D. D. (2002). Respondent-driven sampling II: deriving valid population estimates from chain-referral samples of hidden populations. *Social Problems*, 49(1), 11-34.
- Kuhns, L. M., Kwon, S., Ryan, D. T., Garofalo, R., Phillips, G., & Mustanski, B. S. (2015). Evaluation of Respondent-Driven Sampling in a Study of Urban Young Men Who Have Sex with Men. *Journal of Urban Health*, 92(1), 151-167. doi:10.1007/s11524-014-9897-0
- Lachowsky, N. J., Sorge, J. T., Raymond, H. F., Cui, Z., Sereda, P., Rich, A., . . . Moore, D. M. (2016). Does size really matter? A sensitivity analysis of number of seeds in a respondent-driven sampling study of gay, bisexual and other men who have sex with men in Vancouver, Canada. *BMC Medical Research Methodology*, 16(1), 157. doi:10.1186/s12874-016-0258-4
- Muir, S. W., Berg, K., Chesworth, B., Klar, N., & Speechley, M. (2010). Balance impairment as a risk factor for falls in community-dwelling older adults who are high functioning: a prospective study. *Phys Ther*, 90(3), 338-347. doi:10.2522/ptj.20090163
- Okal, J., Raymond, H. F., Tun, W., Musyoki, H., Dadabhai, S., Broz, D., . . . Geibel, S. (2016). Lessons learned from respondent-driven sampling recruitment in Nairobi: experiences from the field. *BMC Research Notes*, 9(1), 158. doi:10.1186/s13104-016-1965-y
- Reed, E., Gupta, J., Biradavolu, M., Devireddy, V., & Blankenship, K. M. (2010). The Context of Economic Insecurity and Its Relation to Violence and Risk Factors for HIV Among Female Sex Workers in Andhra Pradesh, India. *Public Health Reports*, 125(Suppl 4), 81-89.
- Reisner, S. L., Mimiaga, M. J., Johnson, C. V., Bland, S., Case, P., Safren, S. A., & Mayer, K. H. (2010). What makes a respondent-driven sampling "seed" productive? Example of finding at-risk Massachusetts men who have sex with men. *Journal of Urban Health*, 87(3), 467-479.
- Saggurti, N., Jain, A. K., Sebastian, M. P., Singh, R., Modugu, H. R., Halli, S. S., & Verma, R. K. (2012). Indicators of mobility, socio-economic vulnerabilities and HIV risk behaviours among mobile female sex workers in India. *AIDS Behav*, 16(4), 952-959.
- Salganik, M. J., & Heckathorn, D. D. (2004). Sampling and estimation in hidden populations using respondent-driven sampling. *Sociological methodology*, 34(1), 193-240.
- Saliuk, T., & Andreeva, T. (2010). Evaluation of the necessary number of waves in RDS recruiting of hard-to-reach populations as exemplified by female sex workers surveyed in 2009. *Grani*, 6 (74), 124-128.
- Spreen, M. (1992). Rare populations, hidden populations, and link-tracing designs: What and why? *Bulletin de Methodologie Sociologique*, 36(1), 34-58.
- Strathdee, S. A., Philbin, M. M., Semple, S. J., Pu, M., Orozovich, P., Martinez, G., . . . Patterson, T. L. (2008). Correlates of Injection Drug Use among Female Sex Workers in Two Mexico-U.S. Border Cities. *Drug Alcohol Depend*, 92(1-3), 132-140. doi:10.1016/j.drugalcdep.2007.07.001
- Wang, J., Carlson, R. G., Falck, R. S., Siegal, H. A., Rahman, A., & Li, L. (2005). Respondent-driven sampling to recruit MDMA users: a methodological assessment. *Drug Alcohol Depend*, 78(2), 147-157.
- Yi, H., Mantell, J. E., Wu, R., Lu, Z., Zeng, J., & Wan, Y. (2010). A Profile of HIV Risk Factors in the Context of Sex Work Environments among Migrant Female Sex Workers in Beijing, China. *Psychol Health Med*, 15(2), 172-187. doi:10.1080/13548501003623914.

APPENDICES

Table 4. Assessing the associations for categorical dichotomous variables

The ways of searching for clients	On highways	In hotels	In saunas	On railway station	Online/ phone	In bars, clubs, casinos, discos, etc.	Escort-services	Other ways of client search
In streets (N=343) (Column %)	197 (68.2)	55 (28.9)	85 (38.6)	55 (91.7)	147 (26.1)	103 (25.5)	6 (9.1)	12 (16.4)
Phi-coefficient (p-value)	.448 (p<.001)	-.064 (p=.045)	.039 (p=.222)	.303 (p<.001)	-.224 (p<.001)	-.171 (p<.001)	-.147 (p<.001)	-.112 (p<.001)
The ways of searching for clients	On highways	In hotels	In saunas	On railway station	Online/ phone	In bars, clubs, casinos, discos, etc.	In streets	Other ways of client search
Escort-services (N=66) (Column %)	5 (1.7)	18 (9.5)	14 (6.4)	1 (1.7)	56 (9.9)	37 (9.2)	6 (1.7)	3 (4.1)
Phi-coefficient (p-value)	-.130 (p<.001)	.053 (p=.098)	-.009 (p=.786)	-.052 (p=.104)	.147 (p<.001)	.080 (p=.012)	-.147 (p<.001)	-.030 (p=.347)

Table 5. Logistic regression results for the final model of the pooled data using the productive recruitment as the outcome of interest, (N=970). Estimate corresponds to the respective non-intercept parameter estimate(s) in the given model

Variable	Values	Adjusted OR	95% CI	p-value
Age groups				
	14–18	.567	.330–.975	.040
	19–25	.663	.483–.910	.011
	26–55	reference		
Education				
	Elementary education (<9 grades)	1.529	.658–3.555	.323
	Secondary education (9 grades)	.629	.423–.937	.023
	High school/uncompleted high education	reference		
	Technical or vocational school	.719	.483–1.071	.104
	University	1.793	1.044–3.079	.034
Searching for clients in streets				
	Yes	1.511	1.108–2.061	.009
	No	reference		
Providing escort-services				
	Yes	1.912	1.062–3.442	.031
	No	reference		
Drug use				
	Yes	2.229	1.554–3.197	<.001
	Previous experience of drug use	2.201	1.388–3.489	.001
	No	reference		
Affiliation to any NGO for FSWs/IDUs				
	Yes	2.676	1.926–3.718	<.001
	No	reference		
	No response	2.129	.184–24.609	.555
Appeal to a state clinic for treatment of STIs				
	Yes	.546	.380–.785	.001
	No	reference		
	No response	3.317	.349–31.546	.297