**HIV infection and engagement in the care continuum among Venezuelan migrants and refugees: results of a biobehavioral survey in Colombia**

Andrea L. Wirtz, PhD, MHS,\*1,2 José Rafael Guillén,MSc,3 Megan Stevenson, MSPH,1 Jennifer Ortiz, MD, MSc, 3 Miguel Ángel Barriga Talero, MSc3 Kathleen R. Page, MD,1,2,4 Jhon Jairo López, MSc, 3 Damary Martínez Porras, MSc3 Jhon Fredy Ramirez Correa, Esp,3 Ricardo Luque Núñez, MD, MSc,5 Julian A. Fernandez-Niño, MD, Dr.Sc.,6 Prof Paul B. Spiegel, MD, MPH,2

Supplemental Material

**Table of Contents**

[Sampling methodology 2](#_Toc130288077)

[Statistical weighting to account for sampling 2](#_Toc130288078)

[RDS results 2](#_Toc130288079)

[Figure 1 Bogotá and Soacha RDS network graph, by city 3](#_Toc130288080)

[Figure 2 Barranquilla and Soledad RDS network graph, by city 4](#_Toc130288081)

[Figure 3 Expanded view of single RDS network in Bogotá and Soacha 5](#_Toc130288082)

[Table 1 Characteristics of participants with previously diagnosed and undiagnosed HIV infection 6](#_Toc130288083)

[Table 2 Multivariable logistic regression model: Variables associated with HIV infection among Venezuelan migrants and refugees in Colombia between July 30, 2021 through February 5, 2022 (N=6,221) 7](#_Toc130288084)

[References: 8](#_Toc130288085)

# **Sampling methodology**

Study methods have been previously described in detail.1 Respondent-driven sampling (RDS), a chain referral sampling method that employs limited referrals within peer networks,2 was used to accrue the study sample. Recruitment started with 20 “seeds” (~10 per site) selected from the target population. Seeds were purposively selected based on being influential among peers and diverse in characteristics (e.g., age, gender, geographic residence within each city). Seeds participated in all study activities and were asked to invite up to four peers (recruits) to participate in study activities, which produced the first sampling wave. Eligible and participating recruits were then asked to refer up to four more peers living in the study cities. This process continued until the target sample size was accrued. Participants had the option to use paper and/or e-coupons via SMS or WhatsApp to refer peers. Our in-house electronic study data management system was customized to trace recruitment linkages across participants and send automated study reminders to participants via SMS or WhatsApp, depending on technology access or preference.

## **Statistical weighting to account for sampling**

RDS diagnostic techniques were performed and RDS-Analyst platform.3 Diagnostic techniques included assessment of recruitment depth, bottlenecks, homophily, and convergence across select variables that may affect sampling and inferences.3 Population estimates were calculated by incorporating RDS survey weights, based on self-reported network size, to calculate population prevalence using the RDS-II (Volz-Heckathorn) estimator. Bootstrapping procedures were used to calculate associated 95% confidence intervals following the smoothing algorithm for RDS described by Schonlau and Liebau with 1,000 replications.4 For overall estimates (i.e., aggregated across sites), we incorporated complex survey design to account for clustering within sites. Analyses were performed within Stata Statistical software (version 17) using the RDS program.4

Some RDS investigators also weight by population size when combining across geographic locations when networks do not span geographic distances.5 Due to evidence of recruitment across sites and formative data that suggested networks across study sites, we calculated aggregated estimates across the two sites, using the RDS-II weighting while accounting for clustering within site. We conducted a sensitivity analysis by estimating overall population prevalence using both RDS-II weights and additional weighting for the estimated proportion of migrants and refugees and in each site.6 There was negligible difference in the results produced in the sensitivity analysis compared those with only RDS-II weights for the aggregated analysis. Because inclusion of population size estimates introduces new assumptions and potential error we utilized and present the aggregated estimates with only RDS-II weights. RDS weights are not recommended for multivariable regression analysis and so were not incorporated into the regression models.7

## **RDS results**

RDS was initiated by 20 seeds evenly distributed across study cities. Of those seeds, all but one successfully referred eligible and participating peers; an additional seed was then initiated to replace the unsuccessful seed. This produced a median and maximum recruitment depth of 9 and 17 waves, respectively. Convergence was met for key variables of interest (gender, migration status, HIV status, and syphilis infection). Figures 1 and 2 display RDS network graphs for each site, as well as cross-recruitments. Because the number of participants makes visualization of the networks difficult, we have included an expanded view of a single RDS network from Bogotá and Soacha (Figure 3).

**Figure 1 Bogotá and Soacha RDS network graph, by city**

A picture containing chart

Description automatically generated

**Figure 2 Barranquilla and Soledad RDS network graph, by city**

Chart

Description automatically generated with low confidence

## **Figure 3 Expanded view of single RDS network in Bogotá and Soacha**

A picture containing outdoor object

Description automatically generated

Note: The seed is represented by the large central node. City of residence is represented by color: green, Bogotá, and orange, Soacha. This seed had a recruitment depth of 14 waves and was the source of referrals for 1,459 recruited participants. All RDS recruitment networks are displayed in the Supplemental Appendix.

**Table 1 Characteristics of participants with previously diagnosed and undiagnosed HIV infection**

|  | **HIV diagnostic status** | | | | | | |  | |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | | **Diagnosed  (n=34)** | | **Undiagnosed (n=37)** | | **Total  (n=71)** | | |  |
|  | | **n** | **Col%** | **n** | **Col%** | **n** | **Col%** | | ***p-value*** |
| **Site** | |  |  |  |  |  |  | | *0.242* |
| Bogotá and Soacha | | 11 | 32 | 17 | 46 | 28 | 39 | |  |
| Barranquilla and Soledad | | 23 | 68 | 20 | 54 | 43 | 61 | |  |
| **Age (years)** | |  |  |  |  |  |  | | *0.397* |
| 18 to 29 | | 11 | 32 | 18 | 49 | 29 | 41 | |  |
| 30 to 39 | | 13 | 38 | 13 | 35 | 26 | 37 | |  |
| 40 to 49 | | 5 | 15 | 4 | 11 | 9 | 13 | |  |
| 50+ | | 5 | 15 | 2 | 5 | 7 | 10 | |  |
| **Gender** | |  |  |  |  |  |  | | *0.773* |
| Man | | 21 | 62 | 20 | 54 | 41 | 58 | |  |
| Woman | | 11 | 32 | 15 | 41 | 26 | 37 | |  |
| Transgender or Nonbinary | | 2 | 6 | 2 | 5 | 4 | 6 | |  |
| **Migration status** | | |  |  |  |  |  | | *0.025* |
| Regular | | 17 | 50 | 9 | 24 | 26 | 37 | |  |
| Irregular | | 17 | 50 | 28 | 76 | 45 | 63 | |  |
| **Men who have sex with men** (ref: no; among MSM LHIV, n=41) | | 14 | 67 | 9 | 45 | 23 | 56 | | *0.162* |
| **Ever paid for sex (ref: no)** | | 2 | 6 | 1 | 3 | 3 | 4 | | *0.506* |
| **Transactional sex (ref: no)** | | 4 | 12 | 3 | 8 | 7 | 10 | | *0.606* |
| **Ever injected drugs (ref: no)** | | 2 | 6 | 1 | 3 | 3 | 4 | | *0.506* |
| **Key population** | | 16 | 47 | 11 | 30 | 27 | 38 | | *0.133* |
| **Lifetime HIV test** (self-reported; ref: never) | | 29 | 85 | 16 | 43 | 45 | 63 | | *p<0.001* |
| **Country of last HIV Test** (among those ever tested, n=45) | |  |  |  |  |  |  | | *p<0.001* |
| Colombia | | 3 | 10 | 10 | 63 | 13 | 29 | |  |
| Venezuela | | 26 | 90 | 6 | 38 | 32 | 71 | |  |
| **Laboratory results:** | |  |  |  |  |  |  | |  |
| **Syphilis infection (ref: non-reactive)** | | 9 | 27 | 8 | 22 | 17 | 24 | | *0.632* |
| **CD4 count (cells per µL, n=70)** | |  |  |  |  |  |  | |  |
| Less than 200 | | 7 | 21 | 6 | 16 | 13 | 19 | | *0.165* |
| 200 to 499 | | 8 | 24 | 17 | 46 | 25 | 36 | |  |
| 500 and above | | 18 | 55 | 14 | 38 | 32 | 46 | |  |
| **Viral load (copies per mL, n=70)** | |  |  |  |  |  |  | | *p<0.001* |
| <=50 | | 20 | 61 | 0 | 0 | 20 | 29 | |  |
| 51 - 1000 | | 5 | 15 | 0 | 0 | 5 | 7 | |  |
| >1000 | | 8 | 24 | 37 | 100 | 45 | 64 | |  |
| **Notes:** n: denominator for variable; ref: reference group not displayed; CD4 and viral load data are missing for one participant who had previously been diagnosed and declined further laboratory testing. Only sample estimates are provided given small number of events and subgroup analysis. | | | | | | | | | |

# **Table** **2 Multivariable logistic regression model: Variables associated with HIV infection among Venezuelan migrants and refugees in Colombia between July 30, 2021 through February 5, 2022 (N=6,221)**

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | **OR** | **95%CI** | **p-value** | **aOR** | | **95%CI** | | **p-value** |
| **Barranquilla and Soledad site** (Reference: Bogotá and Soacha) | 1.5 | (1.0-2.5) | 0.079 | | 2.0 | | 1.2-3.3 | 0.0064 | |
| **Gender** (Reference: woman) | |  |  | |  | |  |  | |
| Man | 3.0 | (1.9-5.0) | p<0.001 | 2.5 | | (1.4-4.3) | | 0.0011 |
| Transgender or nonbinary | 14.4 | (4.8-43.0) | p<0.001 | 5.4 | | (1.5-19.1) | | 0.0089 |
| **Key population** (Reference: general population) | 9.3 | (5.7-15.2) | p<0.001 | 4.7 | | (2.6-8.5) | | p<0.0001 |
| **Sexual exploitation** (Reference: no) | 8.4 | (3.9-18.1) | p<0.001 | 3.0 | | (1.2-7.5) | | 0.0163 |
| **Laboratory confirmed syphilis infection** (Reference: no) | 6.0 | (3.4-10.5) | p<0.001 | 3.6 | | (2.0-6.6) | | p<0.0001 |
| **Use of humanitarian services** (Reference: no use) | 1.8 | (1.1-2.9) | 0.030 | 2.0 | | (1.2-3.4) | | 0.0096 |
| Note: OR: unadjusted odds ratio; aOR: adjusted odds ratio calculated via a multivariable logistic regression model; Age, year of migration, any violence victimization in Colombia, symptoms of hazardous alcohol use, and psychologic distress were not associated with HIV infection in bivariate or multivariable models. Migration status was associated at the bivariate level but not in the multivariable model and was omitted for model fit. Final model is fit based on goodness of fit statistics and tested for collinearity; | | | | | | | | | |

**References**:

1. Wirtz AL, Page KR, Stevenson M, et al. HIV Surveillance and Research for Migrant Populations: Protocol Integrating Respondent-Driven Sampling, Case Finding, and Medicolegal Services for Venezuelans Living in Colombia. *JMIR Res Protoc* 2022; **11**(3): e36026.

2. Heckathorn DD. Respondent-Driven Sampling: A New Approach to the Study of Hidden Populations\*. *Social Problems* 2014; **44**(2): 174-99.

3. Gile KJ, Handcock MS. Respondent-Driven Sampling: An Assessment of Current Methodology. *Sociol Methodol* 2010; **40**(1): 285-327.

4. Schonlau M, Liebau E. Respondent-driven sampling. *The Stata Journal* 2012; **12**(1): 72-93.

5. Lansky A, Drake A, Wejnert C, Pham H, Cribbin M, Heckathorn DD. Assessing the assumptions of respondent-driven sampling in the national HIV Behavioral Surveillance System among injecting drug users. *Open AIDS J* 2012; **6**: 77-82.

6. MInisterio de Relaciones Exteriores. Distribucíon de Venezolanos en Colombia: Gobierno de Colombia, 2022.

7. Avery L, Rotondi N, McKnight C, Firestone M, Smylie J, Rotondi M. Unweighted regression models perform better than weighted regression techniques for respondent-driven sampling data: results from a simulation study. *BMC Med Res Methodol* 2019; **19**(1): 202.