

RESEARCH ARTICLE

An evaluation of assumptions underlying respondent-driven sampling and the social contexts of sexual and gender minority youth participating in HIV clinical trials in the United States

Andrea L. Wirtz¹, Jessica R. Iyer¹, Durryle Brooks¹, Kimberly Hailey-Fair^{1,2}, Noya Galai¹, Chris Beyrer¹, David Celentano¹, Renata Arrington-Sanders^{1,2} and PUSH Study Group

⁵**Corresponding author:** Andrea L Wirtz, Johns Hopkins Bloomberg School of Public Health, 615 N. Wolfe St., Office W3505, Baltimore, Maryland 21205, USA. Tel: +1 410 502 0800. (awirtz1@jhu.edu)

Abstract

Introduction: Respondent-driven sampling (RDS) has been an effective sampling strategy for HIV research in many settings, but has had limited success among some youth in the United States. We evaluated a modified RDS approach for sampling Black and Latinx sexual and gender minority youth (BLSGMY) and explored how lived experiences and social contexts of BLSGMY youth may impact traditional RDS assumptions.

Methods: RDS was implemented in three US cities, Baltimore, Philadelphia and Washington DC, to engage BLSGMY aged 15 to 24 years in HIV prevention or care intervention trials. RDS was modified to include targeted seed recruitment from venues, Internet and health clinics, and provided options for electronic or paper coupons. Qualitative interviews were conducted among a sub-sample of RDS participants to explore their experiences with RDS. Interviews were coded using RDS assumptions as an analytic framework.

Results: Between August 2017 and October 2019, 405 participants were enrolled, 1670 coupons were distributed, with 133 returned, yielding a 0.079 return rate. The maximum recruitment depth was four waves among seeds that propagated. Self-reported median network size was 5 (IQR 2 to 10) and reduced to 3 (IQR 1 to 5) when asked how many peers were seen in the past 30 days. Qualitative interviews (n = 27) revealed that small social networks, peer trust and targeted referral of peers with certain characteristics challenged network, random recruitment, and reciprocity assumptions of RDS. HIV stigma and research hesitancy were barriers to participation and peer referral. Other situational factors, such as phone ownership and access to reliable transportation, reportedly created challenges for referred peers to participate in research.

Conclusions: Small social networks and varying relationships with peers among BLSGMY challenge assumptions that underlie traditional RDS. Modified RDS approaches, including those that incorporate social media, may support recruitment for community-based research but may challenge assumptions of reciprocal relationships. Research hesitancy and situational barriers are relevant and must be addressed across any sampling method and study design that includes BLSGMY in the United States.

Keywords: HIV; respondent-driven sampling; African-American; Latinx; Gay; Bisexual; Transgender; Youth

Additional information may be found under the Supporting Information tab for this article.

Received 30 October 2020; Accepted 25 February 2021

Copyright © 2021 The Authors. *Journal of the International AIDS Society* published by John Wiley & Sons Ltd on behalf of the International AIDS Society.

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

1 | INTRODUCTION

Respondent-driven sampling (RDS) has gained popularity in HIV research over the last two decades as a means to sample populations for whom a sampling frame is unavailable [1,2]. This has typically included populations who are most affected by the HIV epidemic, such as sexual and gender minoritized (SGM) populations who have sex with men (transgender women and gay and bisexual men). The popularity of RDS is

owed to its dual benefits of enabling access to populations who may otherwise be challenging to recruit and the estimation of population prevalence (RDS inference) through the use of weighted estimates [2-4]. Several studies have demonstrated that lengthy referral chains enable recruitment to reach deep into social networks and engage individuals with greater risk behaviours and who may be less connected to services, leading to recent use of RDS for HIV intervention research [5-8].

Numerous international investigations demonstrate the effectiveness of RDS to engage adult SGM populations in HIV research [9,10] particularly where same-sex relationships or gender identities may be criminalized or stigmatized and communities have subsequently forged strong networks [11-13]. In these contexts, peer recruitment effectively builds on social networks among individuals with shared experiences. In the United States, increasing closure of SGM meeting places and physical venues that are critical to other sampling methods, increase the appeal of RDS methods [14]. Here, RDS has been effective for engaging racially diverse youth in research on sensitive topics of HIV, sexual exploitation and substance use [15-20]. Recently, however, studies in the United States have demonstrated challenges in achieving target sample size and recruitment depth for Black and Latinx SGM populations and SGM youth (SGMY), resulting in modifications of RDS methods to attain study recruitment targets [21-23].

There are assumptions that are inherent within RDS, which are critical to effective, representative recruitment and inference. These assumptions include: (1) the target population is well networked, (2) peer relationships are reciprocal (undirected network), (3) recruitment within one's social network is random and (4) the sample is selected with replacement [24]. The violation of these assumptions on population interferences has been explored in reviews and simulation studies [25], though less is known about individual participant experiences and behaviours that may challenge these assumptions, particularly how the social context of limited disclosure, stigma and racism experienced in Black and Latinx SGMY (BLSGMY) may impact RDS assumptions [26].

The US national strategy, *Ending the HIV Epidemic (EHE)* [27], is concentrated on the engagement of those most affected by the HIV epidemic – including BLSGMY – in HIV surveillance and prevention and care programmes [28]. Thoughtful analysis of effective sampling and recruitment methods is necessary to identify optimal ways to engage BLSGMY in HIV programming to achieve EHE goals. This mixed-methods analysis aimed to evaluate the use of a modified RDS approach and to explore the challenges of traditional RDS assumptions among BLSGMY.

2 | METHODS

This analysis was conducted using baseline RDS data that were drawn from the parent intervention study, *Providing Unique Support for Health (PUSH)*. The PUSH study utilized a modified recruitment approach which included RDS with targeted seed selection to recruit and enrol eligible BLSGMY who have sex with men to status-dependent randomized clinical trials (RCTs) that compared coach-based support to standard of care for HIV care or prevention [29]. The selection of modified RDS to identify engage BLSGMY in these interventions was based on the above premise that with sufficient recruitment depth, RDS could identify youth who were not adequately reached by services and who could benefit most from prevention and care interventions. After identifying the low propagation of RDS early in the study, we conducted qualitative interviews with a subsample of diverse PUSH participants across the study sites to explore their experiences and challenges in RDS recruitment.

2.1 | Setting and participants

PUSH was a multi-site study conducted in Baltimore, MD (Johns Hopkins University), Philadelphia, PA (Children's Hospital of Philadelphia); and Washington, DC (Children's National Health System and Whitman-Walker Health). In these cities, Black and Latinx populations account for at least half of the population [30] and BLSGMY are priority populations for HIV prevention. All study sites have substantial clinical, research and outreach expertise among SGMY. Furthermore, study teams regularly engage with the communities in each city, the majority of staff who interact with participants are from the SGMY communities, and study teams regularly meet with youth advisory boards at each site.

To be eligible for enrolment, participants were as follows: aged 15 to 24, per the UN definition of youth [31]; from the three study sites; birth-assigned male member; self-identified Black and/or Latinx ethnicity and reported oral/anal sex with a cisgender male member in the prior 12 months. RDS recruits were required to present a valid RDS referral coupon to the study team. We focused on birth-assigned men to include male identified, trans feminine and gender variant youth given the sexual and gender diversity and dynamic sexual and gender identities of adolescents [32-34].

Qualitative participants were a subsample of PUSH participants, with an effort to obtain a maximum variation sample in terms of the number of successful peer referrals. Participants had the option to decline qualitative interviews without any impact on their participation in the parent study.

2.2 | Sampling and recruitment

The PUSH study used a modified RDS methodology. This included the use of RDS coupled with targeted recruitment from clinics, physical venues, online including social media and community outreach (e.g. word of mouth at community events, such as House Balls and through key informants). All eligible and participating youth who were direct recruits from these non-RDS sources were then eligible to become RDS seeds themselves and refer other participants. Similar modifications have been implemented in other studies among SGMY [35]. PUSH seeds and recruits were asked to complete an in-person screening and informed consent. Participants regardless of ultimate enrolment in an RCT were provided with RDS recruitment coupons and asked to return at a later date to obtain secondary recruitment incentives.

Consistent with RDS recommendations and best practices [36], RDS implementation was informed by formative research conducted among 18 key informants across the three cities [26]. Once PUSH launched, recruitment followed standard RDS procedures [2] with ongoing recruitment monitoring but added the following modifications for youth, based on prior studies [21]. Eligible and participating youth were offered electronic coupons (e-coupons) with which to recruit peers. Seeds and recruiters received a weblink by text message during the study visit. The link directed the participant to a page where they could manage and share e-coupons with peers from their social and sexual networks. Participants could continue to access the weblink after leaving the study visit. Paper coupons and study fliers were also available upon request. Finally, study staff worked closely with participants on-site to

support referral and conducted systematic follow-up among participants with outstanding coupons to provide additional guidance, answer questions and replace lost coupons.

E-coupons took the form of a text message sent to selected peers inviting them to the study and providing a unique numeric code, study telephone number, information on site operating hours and the e-coupon expiration date. Peers were asked to display the unique e-coupon code at screening. Text messages contained no information that the study was specific to BLSGMY populations nor focused on the topic of HIV, though participants were encouraged to discuss this verbally with their peers. Initially, participants were offered up to five coupons; however, this was expanded approximately six months after the launch of the study to 10 coupons, with up to five reimbursed. Participants were provided with a \$50 incentive for completing the initial study visit activities and \$5 to 15, depending on IRB restrictions at the site-level, for each eligible and participating peer recruit.

Qualitative participants were recruited from among the participants who completed the enrolment visit, regardless of whether they agreed to participate in one of the RCTs. Participants were verbally invited to participate in the qualitative interviews when they came to the study site for their second RDS visit or a subsequent study visit (if enrolled in one of the RCTs). We aimed to interview approximately 7 to 12 per city with a goal of interviewing participants with varying levels of success with peer referral.

2.3 | Data collection

Upon enrolment, participants were asked to complete a structured, self-administered survey. The survey included network size questions for RDS weighting. This included questions about the number of SGM who have sex with men that the participant knows, the number of these individuals who identify as Black and/or Latinx, and the number of these individuals who are aged 15 to 24 in their city. Of these, participants were then asked how many they have seen within the last 30 days [29].

Qualitative participants met separately with a trained qualitative researcher at their site. Qualitative interviews were conducted in private and followed semi-structured interview guides (Appendix S1). Interviews lasted an average of 39 minutes and explored overarching domains of attitudes, beliefs and experiences with the RDS recruitment strategy with the intent of understanding how to better address any potential challenges associated with inviting their peers/friends to the study.

2.4 | Data analysis

Quantitative analysis included descriptive statistics and recruitment diagnostics consistent with RDS, including analysis of RDS recruitment networks, coupon distribution and return rate and participant self-reported network size. RDSAT and Netdraw software programmes were used to create network graphs and Stata Statistical Software, version 15 (College Station, TX) was utilized for other descriptive statistics. Descriptive statistics with frequencies and Chi-square tests were calculated to compare participant characteristics by recruitment source including clinic, venue and Internet-based targeted seed recruitment and RDS recruitment.

Qualitative interviews were audio-recorded, transcribed and de-identified. Transcripts were entered into qualitative data analysis software, NVivo. All transcripts underwent an initial round of thematic analysis led by two trained qualitative analysts. Upon observation of emergent themes that related to RDS assumptions, we created an analytic framework based on the RDS assumptions for a second round of coding and analysis. Other emergent themes related to general barriers to HIV research among SGM youth were subsequently added. Coding was conducted in tandem for the first three interviews and then codes were reviewed and discussed for consistency across coders and to identify additional codes. Codes were modified until high agreement was achieved between coders ($Kappa > 0.80$). Memos were written for each code. Codes were refined and elaborated during the process of analysis through the constant comparison method. Transcripts were subsequently coded separately and 20% of the overall sample was checked to insure consistency of coding across interviews.

2.5 | Human subjects considerations

Ethical review was provided by the Johns Hopkins Bloomberg School of Public Health, the University of Pennsylvania and Children's National Medical Center Institutional Review Boards. This study received a waiver of parental consent for participants below the age of 18 years. The majority of staff members who interacted with participants were from the community. The study teams regularly engaged in community events and outreach, such as house balls and PRIDE events. Youth advisory boards were also convened regularly in each city for review and feedback on the parent study methods including RDS approach, study instruments and the intervention. Transportation via ride-sharing apps and bus tokens were provided to participants with limited transportation to minimize research disparities associated with transportation barriers.

3 | RESULTS

3.1 | Quantitative results

Between August 2017 and October 2019, a total of 442 participants were recruited and 405 enrolled in the study, including 305 (69.6%) seeds. A total of 1670 coupons were distributed with 133 returned, yielding a 0.079 return rate. RDS networks remained small, with four waves being the maximum recruitment depth among seeds that propagated (Figure 1). In terms of network size, masculine-identified participants reported knowing a median of 9 (IQR: 3 to 20) SGM individuals, knew a median of 5 (IQR: 2 to 16) Black or Latinx SGM, and remained at a median of 5 (IQR 2 to 15) when asked about those who were aged 15 to 24 and living in the three cities. Trans feminine and genderqueer youth reported knowing a median of 7 (IQR: 3 to 20) SGM individuals, knew a median of six (IQR: 2 to 15) Black or Latinx SGM, and a median of six (IQR 2 to 15) peers meeting the above criteria who were aged 15 to 24 living in the three cities. Ultimately, participants reported having seen a median of 3 (IQR 1 to 5) of these peers within the past 30 days, regardless of the participant's gender identity. Thirty-eight (9.5%) of participants, including 32 masculine (9%) and six transfeminine or

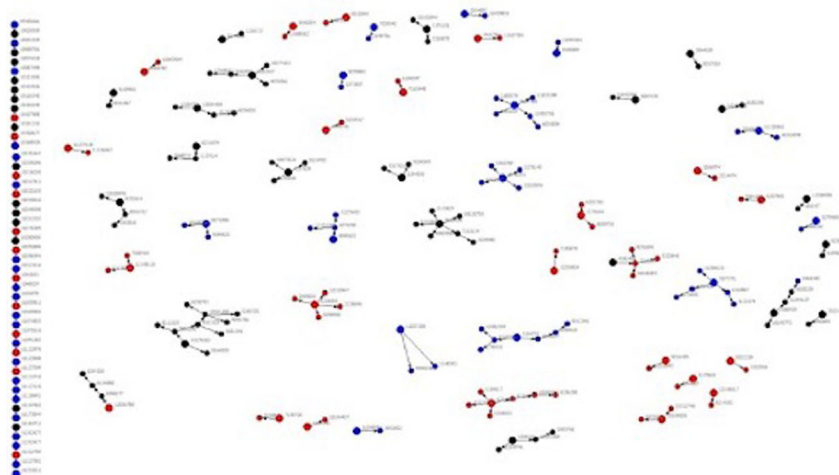


Figure 1. RDS network recruitment diagram: recruitment of Black and Latinx SGMY in Baltimore, Washington DC and Philadelphia. Blue, Washington, DC; Black, Philadelphia; Large nodes represent seeds; Red, Baltimore.

genderqueer participants (11%) reported knowing no other peer who identified within this population. Participants requested a median of five coupons (IQR: 3 to 5) for recruitment, 12% of which were requested as paper-based coupons rather than e-coupons.

Clinic-based targeted recruitment produced the largest subsample with 168 enrolled participants (41.5%), followed by RDS ($n = 123$, 30.4%) and substantially lower among venue ($n = 77$, 19.0%) and Internet-based targeted seed recruits ($n = 37$; 9.1%). There were few observable differences between participants recruited via these sources. Table 1 describes characteristics of study participants stratified by their recruitment source as RDS recruits or seed recruits. Targeted seed recruitment from clinic-based settings was more likely to result in enrolment of participants who were less likely to report unstable housing in the last year, more likely to report a prior positive HIV diagnosis, or more likely to report PrEP use compared to other sources of recruitment. Participants recruited through targeted seed recruitment via the Internet were more likely to report Latinx ethnicity compared to other sources, though the sample from this source was limited.

3.2 | Qualitative results

A total of 27 youth participated in in-depth interviews between May 2018 and December 2019, including gay or bisexual cisgender men ($n = 23$) and transgender or gender variant youth ($n = 4$). Appendix S2 displays the demographic characteristics of qualitative participants from the three sites. Despite a median networks size of six peers (IQR: 3 to 10), most qualitative participants reported difficulties in recruiting Black and/or Latinx SGMY peers, which challenged the success of RDS. Several of the challenges reported by qualitative participants directly affect the assumptions underlying RDS.

3.2.1 | RDS assumption 1: a networked population

RDS requires a population to be well-networked for the sampling process to function appropriately. One of the most

salient themes across all interviews was the reported low connectivity across the population of BLSGMY and very small, tightly knit networks (Table 2; Online Appendix S2). Reasons for a small number of peers outlined by participants included mistrust among peers and simply not knowing many peers that identified as gay or bisexual men, transfeminine or gender variant. Participants also struggled to identify peers that met the eligibility requirements for age and race. Peers were frequently described as older than the 15- to 24-year eligibility requirements and/or were not Black or Latinx race/ethnicity.

3.2.2 | RDS assumption 2: a relationship is reciprocal

RDS assumes that peer relationships are reciprocal (undirected). Frequently this is understood that *Peer A* knows *Peer B* sufficiently well to recruit *Peer B*, but also that *Peer B* knows *Peer A* sufficiently well to recruit *Peer A* – that is they are not strangers to each other. This assumption appeared to be less frequently violated (Table 2; Online Appendix S3). The majority of respondents expressed that they would be sceptical if approached by a stranger and reported preferentially referring peers they trusted. Participants also indicated that research study recruitment was not a priority when conversing with acquaintances or strangers.

Although participants were not asked to recruit via social media or dating apps, two participants reported using social media applications (Facebook, Jack'd and Grindr) to recruit individuals. One participant expressed more comfort recruiting strangers over social media platforms than in person. Another participant described building rapport with strangers via social media and then providing study information, which successfully supported peer referral.

3.2.3 | RDS assumption 3: recruitment of each individual is random

An inherent assumption underlying RDS is that individuals randomly recruit from within their network. Participants, however, frequently reported seeking to recruit peers who they anticipated would participate in research (Table 3; Online

Table 1. Demographic and other characteristics of Black and Latinx SGMY participants in Baltimore, Washington DC and Philadelphia, stratified by recruitment source

Characteristic	Recruitment source										p-value
	Total (N = 405)		RDS recruit (n = 123)		Clinic seed (n = 168)		Internet seed (n = 37)		Venue seed (n = 77)		
	n	Col %	n	Col %	n	Col %	n	Col %	n	Col %	
Median network size (IQR) ^a	5	(2 to 10)	5	(2 to 15)	4	(2 to 10)	5	(3 to 10)	5	(2 to 10)	
Gender identity											
Masculine	347	86.5	102	83.6	146	88.0	33	89.2	66	86.8	0.702
Trans feminine, gender queer	54	13.5	20	16.4	20	12.0	4	10.8	10	13.2	
Sexual orientation											0.567
Gay	255	63.7	79	65.3	109	65.7	24	64.9	43	56.6	
Bisexual	88	22	26	21.5	33	19.9	11	29.7	18	23.7	
Heterosexual	25	6.2	8	6.6	9	5.4	2	5.4	6	7.9	
Other	32	8.0	8	6.6	15	9.0	0	0.0	9	11.8	
Race and ethnicity*											0.007
African American	229	57.2	74	61.2	91	54.8	23	62.2	41	53.9	
Black other	115	28.7	35	28.9	49	29.5	3	8.1	28	36.8	
Black Latino/Hispanic	56	14.0	12	9.9	26	15.7	11	29.7	7	9.2	
Education (completed)											0.713
Less than high school	87	21.7	32	26.2	33	19.9	8	21.6	14	18.4	
Graduate equivalency	17	4.2	5	4.1	6	3.6	3	8.1	3	3.9	
High school graduate	152	37.9	51	41.8	62	37.3	11	29.7	28	36.8	
Technical school	6	1.5	0	0.0	3	1.8	1	2.7	2	2.6	
Some college	96	23.9	22	18	45	27.1	8	21.6	21	27.6	
College graduate	43	10.7	12	9.8	17	10.2	6	16.2	8	10.5	
Currently employed (reference: no)	228	56.9	60	49.2	102	61.4	20	54.1	46	60.5	0.180
Currently living at home with parents (reference: no)	290	73.0	85	70.2	121	73.3	30	81.1	54	73.0	0.637
Without regular place to stay in past 12mo (reference: no)*	96	24.1	33	27.0	26	16.0	10	27.0	27	35.5	0.007
Currently have a mobile phone (reference: no)	372	93.2	115	95.0	155	93.9	36	97.3	66	86.8	0.084
Mobile phone plan (among those with a phone, n = 369)											0.806
Unlimited Internet data text	303	82.1	92	79.3	130	85.0	28	77.8	53	82.8	
Limited Internet data text	49	13.3	16	13.8	18	11.8	6	16.7	9	14.1	
Text only	8	2.2	3	2.6	2	1.3	1	2.8	2	3.1	
Unsure	9	2.4	5	4.3	3	2.0	1	2.8	0	0.0	
Ever exchanged sex (reference: no)	92	23.1	23	18.9	44	26.7	10	27	15	20.0	0.369
Probability of Substance Abuse/Dependence Diagnosis (CRAFTT; reference: no)	243	60.4	73	59.8	103	62.0	24	64.9	43	55.8	0.757
Self-reported diagnosis at last HIV test											0.001
Positive	117	29.2	26	21.3	73	44.0	6	16.2	12	15.8	
Negative	203	50.6	64	52.5	69	41.6	24	64.9	46	60.5	
Unsure or never tested	81	20.2	32	26.2	24	14.5	7	18.9	18	23.7	
Currently taking PrEP (among those with a positive test, n = 271; reference: no)	52	19.2	9	9.6	29	33.7	3	10.0	11	18.0	0.001

^aNetwork size question based on number of known Black or Latinx SGMY living in Baltimore, Washington DC and Philadelphia metro area.; *Statistical difference by recruitment source at $p < 0.05$.

Table 2. Social contexts of Black and Latinx SGMY and relationship to RDS assumptions of networks and reciprocity: explanatory quotes from qualitative participants

Domain	Consideration	Explanatory quote
Network	General small social network	<p>"I would, like I think I said earlier, I don't really interact with too many people day-to-day, a lot of time I spend at work or with son- so If I did have paper ones [coupons], it would be a better option, I would prefer to do it that way but like I said my access to people is sort of limited." – <i>Baltimore 4</i></p> <p>"I think... well, for me, my biggest challenge was knowing people that I would want to send it to, let alone would actually do it, but I feel like, you know, other people have more friends than I do. So I don't think that would be a problem for everyone..." – <i>Washington, DC 4</i></p>
	Small number of peers that fit sexual orientation eligibility criteria	"I know millions of females that I could have gave this too, you know what I mean, instead of just men. I don't really know that boys. Gay boys don't hang with gay boys that much." – <i>Baltimore 3</i>
	Small number of peers that fit the age or race/ethnicity eligibility criteria	<p>"Yeah. People that was black and Latino. I don't have too much friends that are black and Latino. They're all white. And I have some black friends." – <i>Baltimore 6</i></p> <p>"Well, I couldn't bring any friends in. I tried. It was just there are-- I'm the youngest in the house. I'm 21. And all the other girls in the house is, like, 27, 26, and 30. So, they wasn't able to make it in." – <i>Philadelphia 4</i></p>
Reciprocity	Referring strangers to the study	"I don't know. If it was random, then I probably would be like "No." I don't know. I would just ask for proof first and making sure that it's not something out of the ordinary, something crazy or something like that. [Interviewer: So you feel like it works better if somebody from your circle asks you to do it.] Yeah, people that you know, it would get them to come in easier without the whole being scared to do it." – <i>Philadelphia 1</i>
	Use of social media	"Facebook. I asked people on the social apps I'd be on, whether if it's Jack'd or Grindr, asked them if they wanted to come in, or I sent them certain information. Some of them have said yes, that they would, but didn't work out too well. [Interviewer: So posting on social media, do you think that that has worked?] Even for a response, people have responded, but I would never- well, I would, but I can never just walk up to somebody and be like 'Oh, guess what? This and that,' because I don't know whether or not that would be appropriate or not." – <i>Baltimore 2</i>

Appendix S4). This was characterized as targeting peers for recruitment who they perceived could benefit from study participation or in need of material resources. Financial incentive was frequently reported as a driver for individual participation. Conversely, participants reported avoiding referring peers that may have difficulty completing surveys due to literacy constraints, low perceived likelihood of participating, or who were past sexual partners.

Trust was critical to feeling comfortable sharing study information with peers. Participants preferentially referred peers who they felt could be trusted to keep their involvement confidential. The idea of participating in the study with friends was also reassuring and appealing to some participants. Furthermore, participants described having a fear of unintentional peer disclosure associated with the recruitment process and preferentially disseminated coupons to close friends who already knew about their sexuality.

3.2.4 | RDS assumption 4: sample is selected with replacement

RDS, particularly the analytic estimators, assumes that the sample is selected with replacement. This assumption is

frequently violated by study designs that prioritize sampling without replacement and exclude recruits who have previously participated in the research. The design of sampling without replacement also challenged recruitment for participants with small dense networks (Table 3; Online Appendix S4). Recruitment opportunities were limited for participants whose entire peer network had already participated in the study. Participants reported feeling that recruitment was a competition due to overlapping networks.

3.3 | Barriers to engaging in research

Challenges to RDS assumptions existed against a backdrop of competing priorities and situational barriers that broadly challenged the recruitment of BLSGMY to HIV-related research. Competing priorities for youth, such as work, school and their health were priority over peer referral to a study (Table 4; Online Appendix S5). Participants shared that concerns about drug testing, fear of needles and concerns of breached privacy associated with research participation posed challenges in peer referral and participation.

Situational barriers also posed challenges for peer recruitment. Participants described facing personal challenges in

Table 3. Social contexts of Black and Latinx SGMY and relationship to RDS assumptions of random recruitment and sampling with replacement: explanatory quotes from qualitative participants

Domain	Consideration	Explanatory quote
Random recruitment	Characteristics of peers targeted for recruitment	<p>“Sorry, I asked those five people just because I knew that they would be interested in giving their input and basically the research. I wouldn’t ask any other random people because they probably wouldn’t be as interested, but I knew people that I hang out with, people who I know who do outside work in the community would be interested in wanting to work with the research.” - <i>Philadelphia</i> 2</p> <p>“Because y’all got to understand, a lot of youth are homeless and what-you-call-them, so a lot of times food vouchers or food things and money is definitely going to-- will wheel a youth in. That’s how I started, struggling. ‘This is a little \$30-\$40 survey,’ boom. ‘They got food vouchers, too, and you bring this,’ dah, dah, dah. Yeah, youth struggle so you never know what the struggle might be” – <i>Baltimore</i> 5</p> <p>“[Interviewer: Are there certain kinds of people that you feel more comfortable for inviting?] Or somebody that already needs to get tested. Like you always need to get tested, so why not join the study where you can benefit from it and still get tested and still help other people?” – <i>Washington, DC</i> 5</p>
	Characteristics of peers avoided for recruitment	<p>“I feel like some of my friends don’t know how to read or spell. I don’t know... I don’t know if they would be able to get through it [the survey].” – <i>Baltimore</i> 6</p> <p>“That’s another thing that’s holding me back is that a lot of these people that I would send it to, past hookups, I do not... I don’t even want them in my phone really, so I don’t even want to look through them because all of them were unsafe sexual encounters and so they would not even, you know, look at something like this... I don’t think that they even would want to come into Whitman-Walker, you know, so. But they’re the ones who need it the most, so.” – <i>Washington, DC</i> 4</p>
	Comfort and benefits associated with recruiting close friends	<p>“Yeah. I know more so because I was limited to the number of people I could refer, I sought out my close friends more so than other people that I just knew that would’ve been qualified for the survey, because I wanted to let them know about the opportunity more so than someone that I barely knew.” — <i>Washington, DC</i> 1</p> <p>“Yeah, there’s still people that I could’ve invited that I’m like.. “Uh-uh, I don’t know,” just because like I was just saying, I don’t know what their situation is right now and I don’t want something to pop up on their phone that they don’t want. Also I just don’t, I don’t know, I still feel like it can get back to me for some reason. <laughs>That’s a hesitation that I had.” – <i>Washington, DC</i> 4</p>
Sample is selected with replacement	Challenges associated with research practice of sampling without replacement	<p>“A lot of names had came to mind, but then the person that recommended me also recommended them because our friend groups, they overlap. So then it was like, “Oh, they already did the study,” so then I couldn’t invite them.” — <i>Washington, DC</i> 3</p>

recruiting peers that prevented them from completely participating in RDS. For example some participants lacked or had inconsistent access to a cell phone, thus limiting the sharing of study information. Participants also reported lacking transportation that challenged study visits and some reported that peers did not possess a government-issued form of identification that was required for study enrolment.

Concerns related to privacy and confidentiality impacted not only willingness to recruit peers but also individual willingness to participate in the study. Participants who had not disclosed their sexual orientation or HIV status described a reluctance to participate in the study due to concerns of privacy and confidentiality (Table 4; Online Appendix S4). The focus of the study also mattered. Multiple participants expressed apprehension about being HIV tested or involved

with an HIV-focused study due to concerns of unintentional disclosure of one’s HIV status, behavioural risks for HIV, or perceptions of one’s HIV status.

4 | DISCUSSION

The use of RDS to sample BLSGMY faced challenges in this multisite study in the United States. RDS seeds and recruitment networks failed to propagate beyond a limited number of waves, despite participant reports of median network sizes of approximately five peers. While a network size of five peers seems feasible for peer referral, other research has shown that SGMY were 60% more likely to effectively refer at least one peer when they had a network size of 10 or more [21].

Table 4. Barriers to engaging in HIV research among Black and Latinx SGMY: explanatory quotes from qualitative participants

Consideration	Explanatory quote
Competing priorities	"No, I didn't think about not inviting people, but it's like I didn't think about that, like thinking about, oh this- you know, like just going out here, like, 'You should come to PUSH.' I wasn't thinking about that. I don't know, that wasn't on my mind. I'm more thinking about what's going on with HIV and school and stuff like that. I wasn't thinking about coming right back." – <i>Baltimore 11</i>
Fear/Skepticism in research participation	"I mean, it was kind of, like I said before, a little nervous because I didn't know what exactly all the ins and outs of the research and what was it geared to. They just told me 'We're just trying to find information to better the community,' but I'm like 'Better the community how? There's so many things that can be worked on or can be addressed,' and then I was just a little bit nervous asking or giving my input on things that I've gone through in my life that I probably wouldn't share with any other body, but being as though it's research, you need to get all those variables of everything so you can have data or whatever. But you know, at first I was like I would like to do it just because I want to make a difference and I want my input to be in the research but having those feelings like what questions or what I have to answer or what you guys want to know was in the back of my mind." – <i>Philadelphia 2</i> "[Interviewer: Do you ever feel apprehensive about inviting people to join the study?] Sometimes, because I don't want them to question. Like, 'How the hell do you know about this?' Yeah, people are very nosy and they just. . . – <i>Washington, DC 5</i>
Situational barriers	"I told-- I brought in, I only brought in three [peers]. But I told, probably at the most, likely 10 people. I just couldn't make it. That's another concern. People can't make it so if they had transportation to get here it would work out. [Interviewer: Okay, so you think transportation was an issue?] With-- For most of them. [Interviewer: Okay, so the people that came into the study, those three people, what do you think made it easier for you to invite them and for them to come in?] They live close by." – <i>Baltimore 12</i> "[Interviewer:.. talk a little bit about how the text and your phone being broken was a challenge?] Yeah, so when I had lost my phone, it was hard to even remember about the coupons because I didn't have that reminder in my face. I forgot about the coupons." – <i>Washington, DC 5</i>
Sexual orientation	"Challenges? I'd say one challenge would be not being out of the closet but generally, generally speaking, if you go around an organization, if you participate in the study, you're most likely out of the closet. I mean, other cases, they're really not. So, I think in that case maybe people are scared that if they hand this out, then someone is going to know that they're gay or somebody down the line can tell someone that this person gave them this and they want to take their time to come out and make sure their parents or whoever are hearing from them and not someone else. . . ." – <i>Philadelphia 3</i> "I would say one main thing is like, say if you were to do the survey in secret and say you're not fully comfortable in your identity, then to pass out the coupons to whoever would in certain kind of imply something about yourself. So I feel like for those who aren't necessarily firm in who they are and their identity yet, then that would be one reason that they don't pass them out. . . ." – <i>Washington, DC 1</i>
HIV status	"Some people are because some people are actually scared to know their results. Like me for instance, I was young when I found out everything, so yeah. I can't lie. Now, I really would be scared to get my results because at this point in time, I'm a escort and everything, so me dealing with so many people and sexually wise and stuff, I would really be scared. I'd be like, 'Girl, I don't need this' or something like that." – <i>Baltimore 5</i> "[Interviewer: And I also wonder how that went, if you told anybody that there was HIV testing with the study?] That part I didn't mention. And, like, at the one place I didn't know that I had to in order to receive a gift card. And I don't know. Like, it's not a bad thing, it's just suspicious. For me, it's a little scary, because I'm private about it. . . Like, I don't like too many people knowing or-- because it's my business. It's like I'm a private person. I don't want everybody, like, to know, 'Oh, she has HIV.' Because people are so judgmental and they'll always say, 'Oh, well, you know, she does this, that, and a third. So, she has HIV.' People are rude. So, I just have to play my cards right." – <i>Philadelphia 4</i>

Coupling RDS with targeted recruitment from physical and online sites, however, supported access to and engagement of over 400 BLSGMY. A similar modified approach ("Starfish Sampling") was recently reported by Raymond and colleagues for sampling transgender men in San Francisco, which permitted participants who were recruited via venue-based sampling to then refer peers using standard RDS methods [35]. While

limited in its ability to produce population estimates, the authors noted that "starfish sampling could be considered for recruitment in populations when available tools are inadequate" [35].

Our qualitative data and that of others [21-23], however, suggested that BLSGMY infrequently have the well-developed networks that are observed among adults. Youth reported

knowing and associating with peers who identify with the same race/ethnicity or who identify as SGMY but may not have peers who identify across these characteristics. Underdeveloped peer networks may be due in part to emerging sexual orientation and gender identities, and limited disclosure to peers of the same age range or race/ethnicity. Furthermore, BLSGMY are more likely to experience marginalization, stigma and isolation, which may unequally compel some youth to engage in sampling strategies, while turning others away [26]. Targeting sampling to groups on the basis of race/ethnicity, age, gender identity and sexual orientation as well as common research practices of sampling without replacement may inherently break underlying social networks and violate the network assumption that is fundamental to RDS. RDS and related modified network-based approaches among BLSGMY that are less restrictive in eligibility criteria or permit the inclusion of other peer populations, such as female members of social networks, can potentially help to bridge networks and promote engagement in research [37].

Other RDS assumptions that are critical to implementation and analysis were also reportedly violated by youth participants. Youth frequently reported preferential recruitment of peers who either needed a financial incentive or would be interested in research. While well-intentioned, these practices may introduce bias and, in the case of recruiting to status-dependent RCTs, may challenge engagement of those who are not engaged in but could benefit from HIV prevention and care. This is also an important consideration for studies that select RDS for its reported ability to estimate population means, as preferential recruitment has been associated with biases in estimation [38].

The future of RDS among BLSGMY in the United States may rely on greater adaptation to technology but must do so with consideration for RDS assumptions, particularly if used to generate population-based estimates. Some participants described the use of social media and dating apps to distribute e-coupons, suggesting this is a viable option; however, it may also violate assumptions of reciprocity depending on the nature of the relationship within social media. Research has shown that youth view social media friends and followers as sources of social support [39,40]; thus, youth may experience similarly strong or stronger emotional ties with online peers as they do with peers they regularly see in-person. Social media has recently been integrated into RDS approaches among SGMY, improving enrolment despite that unique race and socio-economic differences were observed when compared to traditional RDS and nationally representative samples [22,41]. The authors of a webRDS study acknowledged racial disparities in consistency of computer and Internet access [41], drawing important consideration to the possibility for webRDS to potentiate disparities in health research. Taken together, these studies highlight the potential of social media to diversify samples recruited through RDS, but also suggest that sampling methods using social media alone may miss important populations who could benefit from public health interventions.

Study findings should be viewed in light of limitations. First, the proportion recruited via RDS may be underestimated, as anecdotal reports from staff suggest that coupons were provided to peers who participated in the study, but who forgot to display the study coupon and were possibly misclassified as

targeted seeds. Finally, the samples of transgender and gender-variant youth enrolled in the qualitative and quantitative components of this study are small, which limited our ability to detect unique differences across gender identity, and are unlikely to be representative. Other studies have recently faced similar challenges with the use of RDS to sample these populations, highlighting the importance of identifying a sampling method that is acceptable to transgender youth [22,35].

5 | CONCLUSIONS

Traditional RDS may have a limited role in sampling SGMY, particularly those who are racial or ethnic minoritized populations in the United States but may be improved through coupling with other sampling approaches and/or integration with social media platforms. Sexual and gender identity formation and peer disclosure are evolving processes among BLSGMY, potentially resulting in small social networks with varying degrees of trust and challenging traditional RDS assumptions and those of other peer-referral methods. Research hesitancy and situational barriers related to transportation and technology are relevant and must be addressed across any sampling method and study design that includes BLSGMY in the United States.

AUTHORS' AFFILIATIONS

¹Department of Epidemiology, Johns Hopkins Bloomberg School of Public Health, Baltimore, MD, USA; ²Division of Adolescent and Young Adult Medicine, Johns Hopkins School of Medicine, Baltimore, MD, USA

COMPETING INTEREST

The authors have declared no conflict of interest.

AUTHORS' CONTRIBUTIONS

RS, DC, CB, NG and ALW designed the parent study; RS and DC are the joint principal investigators. KHF and DB conducted the interviews, JI and a research assistant conducted the qualitative coding with oversight by RS. ALW and IJ jointly reviewed and analysed the data. ALW led the write-up of the findings with contribution from IJ. All authors have reviewed and provided input to the manuscript.

ACKNOWLEDGEMENTS

This study was supported by the National Institutes of Health (R01DA043089) through the NIDA and the Johns Hopkins University Center for AIDS Research (P30AI094189). The authors acknowledge the contribution of the investigators and staff of PUSH Study Group at participating research sites, including the Johns Hopkins University, Baltimore, MD (F Shorrock, J Conley and A Alvarenga); Children's Hospital of Philadelphia, PA (M Castillo, N Dowshen, A Schlupp, A Lopez and W Vickroy); Children's National Medical Center and Whitman Walker Health, Washington, DC (C Trexler, LJ D'Angelo, J Kwiat, R Carr, J Leslie and B Smith). The authors also thank the Johns Hopkins University and Children's Hospital of Philadelphia's Adolescent Medicine Teen Advisory Boards, who provided feedback on the study materials. The authors thank all the adolescents and young adults who participated in this project for their time and effort.

REFERENCES

1. Heckathorn D. Respondent-driven sampling: a new approach to the study of hidden populations. *Soc Probl.* 1997;44(2):174–99.
2. Salganik MJ, Heckathorn D. Sampling and estimation in hidden populations using respondent-driven sampling. In: Stolzenberg RM, editor. *Sociological Methodology*. Boston, MA: Blackwell Publishing; 2004. p. 193–238.

3. Heckathorn D. Respondent driven sampling II: deriving valid population estimates from chain-referral samples of hidden populations. *Soc Probl.* 2002;49(1):11–34.
4. Magnani R, Sabin K, Saidel T, Heckathorn D. Review of sampling hard-to-reach and hidden populations for HIV surveillance. *AIDS.* 2005;19 Suppl 2: S67–72.
5. Solomon SS, McFall AM, Lucas GM, Srikrishnan AK, Kumar MS, Anand S, et al. Respondent-driven sampling for identification of HIV- and HCV-infected people who inject drugs and men who have sex with men in India: a cross-sectional, community-based analysis. *PLoS Med.* 2017;14:e1002460.
6. Wirtz AL, Naing S, Clouse E, Thu KH, Mon SHH, Tun ZM, et al. The Parasol protocol: an implementation science study of HIV continuum of care interventions for gay men and transgender women in Burma/Myanmar. *JMIR Res Protoc.* 2017;6:e90.
7. Fogel JM, Sivay MV, Cummings V, Wilson EA, Hart S, Gamble T, et al. HIV drug resistance in a cohort of HIV-infected MSM in the United States. *AIDS.* 2020;34(1):91–101.
8. Wei C, McFarland W, Colfax GN, Fuqua V, Raymond HF. Reaching black men who have sex with men: a comparison between respondent-driven sampling and time-location sampling. *Sex Transm Infect.* 2012;88(8):622–6.
9. Malekinejad M, Johnston LG, Kendall C, Kerr LRF, Rifkin MR, Rutherford GW. Using respondent-driven sampling methodology for HIV biological and behavioral surveillance in international settings: a systematic review. *AIDS Behav.* 2008;12 4 Suppl:S105–30.
10. Johnston LG, Hakim AJ, Ditttrich S, Burnett J, Kim E, White RG. A systematic review of published respondent-driven sampling surveys collecting behavioral and biologic data. *AIDS Behav.* 2016;20(8):1754–76.
11. Wirtz AL, Zelaya CE, Latkin C, Peryshkina A, Galai N, Mogilnyi V, et al. The HIV care continuum among men who have sex with men in Moscow, Russia: a cross-sectional study of infection awareness and engagement in care. *Sex Transm Infect.* 2016;92(2):161–7.
12. Wirtz AL, Trapence G, Kamba D, Gama V, Chalera R, Jumbe V, et al. Geographical disparities in HIV prevalence and care among men who have sex with men in Malawi: results from a multisite cross-sectional survey. *Lancet HIV.* 2017;4(6):e260–9.
13. Hylton E, Wirtz AL, Zelaya CE, Latkin C, Peryshkina A, Mogilnyi V, et al. Sexual identity, stigma, and depression: the role of the "Anti-gay Propaganda Law" in mental health among men who have sex with men in Moscow, Russia. *J Urban Health.* 2017;94(3):319–29.
14. Mattson G. Are Gay Bars Closing? Using business listings to infer rates of gay bar closure in the United States, 1977–2019. *Socius.* 2019;5:2378023119894832.
15. Curtis R, Terry K, Dank M, Dombrowski K, Khan B, Muslim A, et al. The commercial sexual exploitation of children in New York City. New York: Center for Court Innovation; 2008.
16. Decker MR, Marshall BD, Emerson M, Kalamar A, Covarrubias L, Astone N, et al. Respondent-driven sampling for an adolescent health study in vulnerable urban settings: a multi-country study. *J Adolesc Health.* 2014;55 6 Suppl:S6–12.
17. Duncan DT, Hickson DA, Goedel WC, Callander D, Brooks B, Chen Y-T, et al. The social context of HIV prevention and care among black men who have sex with men in three U.S. cities: the Neighborhoods and Networks (N2) cohort study. *Int J Environ Res Public Health.* 2019;16(11):1922.
18. Franchino-Olsen H, Chesworth BR, Boyle C, Rizo CF, Martin SL, Jordan B, et al. The prevalence of sex trafficking of children and adolescents in the United States: a scoping review. *Trauma Violence Abuse.* 2020. <https://doi.org/10.1177/1524838020933873>.
19. Hotton A, Quinn K, Schneider J, Voisin D. Exposure to community violence and substance use among Black men who have sex with men: examining the role of psychological distress and criminal justice involvement. *AIDS Care.* 2019;31(3):370–8.
20. Quinn K, Fong C, Guarino H, Mateu-Gelabert P. Development, validation, and potential applications of the hepatitis C virus injection-risk knowledge scale (HCV-IRKS) among young opioid users in New York City. *Drug Alcohol Depend.* 2019;194:453–9.
21. Kuhns LM, Kwon S, Ryan DT, Garofalo R, Phillips G, Mustanski BS. Evaluation of respondent-driven sampling in a study of urban young men who have sex with men. *J Urban Health.* 2015;92(1):151–67.
22. Arayasirikul S, Chen Y-H, Jin H, Wilson E. A Web 2.0 and epidemiology mash-up: using respondent-driven sampling in combination with social network site recruitment to reach young transwomen. *AIDS Behav.* 2016;20(6): 1265–74.
23. Murrill CS, Bingham T, Lauby J, Liu K-I, Wheeler D, Carballo-Diéguez A, et al. Respondent-driven sampling in a multi-site study of black and latino men who have sex with men. *J Natl Med Assoc.* 2016;108(1):69–76.
24. Volz E, Heckathorn D. Probability based estimation theory for respondent driven sampling. *J Off Stat.* 2008;24(79).
25. Léon L, Des Jarlais D, Jauffret-Roustide M, Le Strat Y, et al. Update on respondent-driven sampling: Theory and practical considerations for studies of persons who inject drugs. *Methodological Innovations.* 2016;9: 2059799116672878.
26. Arrington-Sanders R, Hailey-Fair K, Wirtz AL, Morgan A, Brooks D, Castillo M, et al. Role of structural marginalization, HIV stigma, and mistrust on HIV prevention and treatment among young black latinx men who have sex with men and transgender women: perspectives from youth service providers. *AIDS Patient Care STDS.* 2020;34(1):7–15.
27. Fauci AS, Redfield RR, Sigounas G, Weakke MD, Giroir BP. Ending the HIV epidemic: a plan for the United States. *JAMA.* 2019;321(9):844–5.
28. Harper S. Race without racism: how higher education researchers minimize racist institutional norms. *Rev Higher Educ.* 2012;36(1):9–29.
29. Arrington-Sanders R, Hailey-Fair K, Wirtz AL, Brooks D, Castillo M, Dowshen N, et al. Providing Unique Support for Health (PUSH) Protocol: Mobile-enhanced Randomized Control Trial to identify, engage and improve outcomes among Young Black and Latinx men and transgender women who have sex with men living in three urban cities in the United States across the HIV care and prevention continuums. *JMIR Res Protoc.* 2020;9(9):e17269.
30. U.S. Census Bureau. QuickFacts: Philadelphia City, Pennsylvania; Washington City, District of Columbia; Baltimore City, Maryland. 2019 [cited 2020 Aug 15]. Available from: <https://www.census.gov/quickfacts/fact/table/philadelphiacitypennsylvania,washingtoncitydistrictofcolumbia,baltimorecitymaryland/PST045219>
31. United Nations. Youth. 2021. Available from: <https://www.un.org/en/sections/issues-depth/youth-0/#:~:text=For%20statistical%20purposes%2C%20however%2C%20the,of%2015%20and%2024%20years>. Accessed 2 October, 2020.
32. Phillips G, Feinstein BA, Levy M, Kuo I, Glick SN, Fields SD, et al. Changes in sexual and gender identity and their associations with internalized homophobia among black men who have sex with men in the HPTN 061 BROTHERS Cohort. *Arch Sex Behav.* 2020;49(7):2375–87.
33. Katz-Wise SL, Reischer SL, Hughto JW, Keo-Meier CL. Differences in sexual orientation diversity and sexual fluidity in attractions among gender minority adults in Massachusetts. *J Sex Res.* 2016;53(1):74–84.
34. Turban JL, Keuroghlian AS. Dynamic gender presentations: understanding transition and "De-Transition" among transgender youth. *J Am Acad Child Adolesc Psychiatry.* 2018;57(7):451–3.
35. Raymond HF, Chen YH, McFarland W. "Starfish Sampling": a novel, hybrid approach to recruiting hidden populations. *J Urban Health.* 2019;96(1):55–62.
36. Johnston LG, Whitehead S, Simic-Lawson M, Kendall C. Formative research to optimize respondent-driven sampling surveys among hard-to-reach populations in HIV behavioral and biological surveillance: lessons learned from four case studies. *AIDS Care.* 2010;22(6):784–92.
37. Schneider JA, Cornwell B, Ostrow D, Michaels S, Schumm P, Laumann EO, et al. Network mixing and network influences most linked to HIV infection and risk behavior in the HIV epidemic among black men who have sex with men. *Am J Public Health.* 2013;103(1):e28–36.
38. Gile KJ, Handcock MS. Respondent-driven sampling: an assessment of current methodology. *Sociol Methodol.* 2010;40(1):285–327.
39. Nabi RL, Prestin A, So J. Facebook friends with (health) benefits? Exploring social network site use and perceptions of social support, stress, and well-being. *Cyberpsychol Behav Social Netw.* 2013;16(10):721–727.
40. McInroy LB, Craig SL. Online fandom, identity milestones, and self-identification of sexual/gender minority youth. *J LGBT Youth.* 2018;15(3): 179–96.
41. Bauermeister JA, Zimmerman MA, Johns MM, Glowacki P, Stoddard S, Volz E. Innovative recruitment using online networks: lessons learned from an online study of alcohol and other drug use utilizing a web-based, respondent-driven sampling (webRDS) strategy. *J Stud Alcohol Drugs.* 2012;73(5):834–8.

SUPPORTING INFORMATION

Additional information may be found under the Supporting Information tab for this article.

Appendix S1. Qualitative interview guides.

Appendix S2. Description of Black and Latinx SGMY participants of qualitative interviews (N = 27).

Appendix S3. Social contexts of Black and Latinx SGMY and relationship to four RDS assumptions: explanatory quotes from qualitative participants.

Appendix S4. Social contexts of Black and Latinx SGMY and relationship to RDS assumptions of random recruitment and

sampling with replacement: explanatory quotes from qualitative participants.

Appendix S5. Barriers to engaging in HIV research among Black and Latinx SGMY: explanatory quotes from qualitative participants.