SURVEY FUTURES

SURVEY DATA COLLECTION METHODS COLLABORATION

Report 10: Respondent driven sampling: Evidence review

Mark Todd[1], Toby Li[1], Olga Maslovskaya[2], Carina Cornesse[3], Curtis Jessop[1]

[1]National Centre for Social Research; [2]University of Southampton; [3]GESIS

January 2026

Survey Futures is an Economic and Social Research Council (ESRC)-funded initiative (grant ES/X014150/1) aimed at bringing about a step change in survey research to ensure that high quality social survey research can continue in the UK. The initiative brings together social survey researchers, methodologists, commissioners, and other stakeholders from across academia, government, private and not-for-profit sectors. Activities include an extensive programme of research, a training and capacity-building (TCB) stream, and dissemination and promotion of good practice. The research programme aims to assess the quality implications of the most important design choices relevant to future UK surveys, with a focus on inclusivity and representativeness, while the TCB stream aims to provide understanding of capacity and skills needs in the survey sector (both interviewers and research professionals), to identify promising ways to improve both, and to take steps towards making those improvements. Survey Futures is directed by Professor Peter Lynn, University of Essex, and is a collaboration of twelve organisations, benefiting from additional support from the Office for National Statistics and the ESRC National Centre for Research Methods. Further information can be found at www.surveyfutures.net.

Research Strand 1 of Survey Futures ("Enhanced Sampling Frames and Procedures"), led by Professor Paul Smith (University of Southampton) aims to establish the feasibility and advantages of individual-based sampling and respondent-driven sampling (RDS) to recruit of hard-to-reach groups, and to identify good practice in the use of non-probability sampling. Research Strand 1 has three sub-projects:

(1) Uses of administrative and alternative data in sampling.

(2) Combining probability and non-probability-based samples.

(3) Use of RDS for boosting probability-based samples.

The suggested form of citation for this report is:

Todd M, Li T, Maslovskaya O, Cornesse C & Jessop C (2026) 'Respondent driven sampling: evidence review', *Survey Futures Report* no. 10. Colchester, UK: University of Essex. Available at https://surveyfutures.net/reports/.

Executive Summary

This Evidence Review examines how Respondent-Driven Sampling (RDS) has been used in practice, with a specific focus on sampling and recruitment and not on statistical estimation. RDS is increasingly applied in research on populations that lack a conventional sampling frame, yet the effectiveness of its implementation varies substantially across studies. The purpose of this review is to introduce researchers and survey practitioners to RDS and facilitate their design choices by clarifying how recruitment practices and fieldwork decisions shape the performance of RDS.

The review is guided by a set of research questions:

1. How does Respondent-Driven Sampling perform in practice across different populations, recruitment modes, and study designs?
2. How do recruitment structures (seeds, referral caps, incentives, chain length) shape recruitment trajectories and sample composition in RDS?
3. Which fieldwork, operational, and design strategies are associated with more successful RDS implementation?
4. How do different recruitment modes (face-to-face, online, and mixed-mode) affect the feasibility, integrity, and outcomes of RDS?
5. Where, why, and for whom does RDS break down as a sampling and recruitment method?
6. Where are the gaps in the existing evidence base regarding RDS?

While the review draws on a substantial body of peer-reviewed literature published between 1997 and 2024, it is not a systematic review or a meta-analysis. Key indicators are reported too inconsistently across studies to support formal aggregation, and the review instead adopts a rigorous narrative approach that synthesises commonly reported design features, patterns, strengths, and weaknesses.

The review is organised in two main sections. The first section documents how RDS recruitment is designed and implemented in published studies, focusing on features that are under researchers' control. This includes how target populations are defined, how initial participants are selected, how peer recruitment is structured, how incentives and peer-referral limits are used, and how recruitment patterns unfold in practice. The second main section provides an analytical evaluation of this evidence. By bringing the findings across studies together, the review assesses how recruitment practices interact with social network

structure and operational decisions to shape recruitment success, sample composition, and the credibility of inferences.

The review highlights both the flexibility of RDS and its recurring vulnerabilities. While RDS can be effective in accessing groups that lack conventional sampling frames, its performance is sensitive to weak connectivity within and between social networks, incomplete data, and violations of the core methodological assumptions. In particular, the evidence points to a tension between RDS's reliance on social networks and its tendency to under-represent individuals who are weakly connected or socially isolated.

The review concludes that RDS is a complex sampling method that is neither a universally reliable solution nor an inherently flawed method. Its credibility depends on careful formative research, strategic recruitment design, ongoing fieldwork monitoring, flexibility and adaptation to a specific context. For researchers concerned with representativeness and external validity, the review emphasises the importance of cautious implementation and, where possible, the use of complementary or hybrid sampling approaches. Greater transparency and standardisation in reporting, through guidelines such as STROBE-RDS, are encouraged to strengthen comparability across studies and further development of this method.

Content Page

Glossary of Terms

Bottleneck	A point where recruitment flow between sub-groups is restricted, limiting mixing and recruitment depth.
Bottleneck plots	Graphs used by researchers to see where recruitment slows or fails to mix between subgroups.
Chains	Recruitment sequences that begin with a seed and extend across multiple waves.
Convergence	The process by which sample characteristics stabilise across waves; see also Equilibrium.
Coupon	A paper or digital invitation that participants use to recruit peers into the study.
Eligibility criteria	Rules defining who can take part in the study.
Equilibrium	A state where sample composition stabilises across key variables across recruitment waves and becomes less dependent on initial seeds.
FSWs (Female Sex Workers)	Women who exchange sexual services for money or goods.
HIV (Human Immunodeficiency Virus)	A virus affecting the immune system; HIV referenced in relation to health-at-risk populations.
Homophily	The tendency for participants to recruit others similar to themselves, contributing to clustering and potential bias.
IDU (Intravenous Drug User)	A person who injects drugs.
Incentives (or Dual Incentives)	Compensation given for participating and for successfully recruiting peers.
LGBTQ+ (Lesbian, Gay, Bisexual, Transgender and Queer+)	A collective term for sexual and gender minority groups.
MSM (Men Who Have Sex with Men)	Men defined by sexual behaviour rather than identity.
Network	The social connections through which recruitment occurs.
Network size (or Degree)	The number of people in the target population that a participant knows and could potentially recruit.
NGO (Non-Governmental Organisation)	Independent organisations that often assist with identifying or recruiting seeds due to community access and trust.
RDS (Respondent-Driven Sampling)	A peer-recruitment sampling method using traceable referrals to reach populations without conventional sampling frames.
RDS-II	A commonly used RDS estimator that weights cases by reported network size.
Recruitment tree diagrams	Visual representations of recruitment chains showing who recruited whom.
Recruitment-transition matrix	A table showing which groups tend to recruit which other groups.
Referral Cap	The maximum number of peers a participant may recruit, used to control recruitment speed and concentration.
Replacement seed	A new seed added if recruitment slows or stops.
Screening	Procedures used to confirm participant eligibility, ranging from self-report to in-person verification.
Seed	An initial participant selected by researchers to start recruitment chains in RDS.
STBBIs (Sexually Transmitted and Blood-Borne Infections)	Infections transmitted through sexual contact or blood; referenced in relation to health-at-risk populations.
STI (Sexually Transmitted Infection)	An infection transmitted through sexual contact.
Waves	Rounds of recruitment, starting with the first participants (seeds) as wave 0, and continuing as participants recruit others through referral chains.

1. Introduction

1.1 Background & Objectives

The aims of the Survey Futures research programme are to provide understanding of the key properties of alternative approaches to data collection in areas where these are not yet fully understood, with a view to providing guidance and tools for the implementation of good practice in the short- to medium-term, and to explore possibilities for game-changing advances in other areas of survey data collection practice, recognising that these advances might only be realised in the medium- to long-term.

Research Strand 1: Enhanced sampling frames and procedures (led by Professor Paul Smith)

This Research Strand (RS1) aims to 1. establish the feasibility and advantages of individual-based sampling, 2. identify good practice in the combining of probability and non-probability samples and 3. assess effectiveness of the use of Respondent-Driven Sampling (RDS) to recruit hard-to-reach groups. This evidence review is produced for the third subproject.

Many surveys suffer from under-representation and biases within population sub-groups, but these groups are often of particular interest. RDS using a probability-based 'seeds' may provide a relatively robust but cost-effective way to reach and potentially 'boost' these under-represented population sub-groups. RS1 aims to review existing evidence on the use of RDS for recruiting representative samples of hard-to-survey groups, summarise current knowledge and practice, and finally produce guidelines for survey practitioners.

1.2 Definition and History of RDS

RDS is a chain-referral sampling method developed in the late 1990s by Douglas Heckathorn to access hidden or hard-to-reach populations lacking a defined sampling frame (Heckathorn, 1997). Building on earlier snowball sampling techniques, RDS introduced a structured recruitment process supported by mathematical modelling, theoretically enabling the generation of statistically robust population estimates. RDS begins with a set of purposefully selected initial participants, referred to as 'seeds', who usually (but not always) belong to the target population. These seeds are given a limited number of uniquely coded 'coupons' to recruit peers from within their social networks. Recruits then receive their own coupons to invite others and continue the recruitment process (through waves), resulting in successive, traceable 'chains' of participants. Commonly, incentives are provided both for participating in the study and for successfully recruiting peers, this is to increase participation rates and maintain momentum within recruitment chains (known as the "dual incentive scheme"). All main terms used in the review can be found in Glossary on page 7.

RDS population estimation is based on the idea that as people are recruited in traceable chains that disperse over time, the influence of the initial seeds fades and the sample begins to reflect characteristics of the broader population. When this happens (after 4-6 waves of

recruitment) and the characteristics of the sample remains stable from one wave to the next, the sample is said to have reached a steady state, or “equilibrium”. In theory, once this stability is reached, the data can be used to estimate characteristics of the population without systematic bias.

Statistical methods developed for RDS-generated data, such as the RDS-II estimator (Volz-Heckathorn 2008), are designed to correct for known problems in non-probability samples, which include differences in how people recruit others and differences in the size of social networks. Under certain conditions, these adjustments allow RDS estimates to approximate what would be obtained from a probability-based sample, though this depends on assumptions that are discussed later. Tracing links along referral chains is vital for the calculation of RDS estimators. These links provide the necessary data to account for the sampling structure in RDS, track the differential recruitment patterns, and to adjust for potential biases in sample recruitment. This link-tracing capability distinguishes RDS from other forms of non-probability sampling, such as snowball sampling and underpins its potential methodological advantage.

Since its development, RDS is being increasingly employed in public health, epidemiology, and social research to access a diverse array of “hard-to-reach” populations. These include marginalised or stigmatised groups such as the LGBTQ+ community, Intravenous Drug Users (IDU), sex workers, the homeless, undocumented migrants, people with medical conditions and individuals involved in informal or illicit economies. In addition to producing population estimates, RDS supports the investigation of social network structures, patterns of disease transmission, and the design and delivery of tailored health and social interventions in settings where traditional survey methods are impractical or unfeasible.

2. RDS Design Aspects and Common Features

This section documents the RDS design aspects that are reported in the reviewed literature and are under the researcher’s control, focusing specifically on sampling and recruitment features rather than data analysis or estimation procedures. It covers target populations, seed selection and screening, peer recruitment mechanisms, incentive structures, referral limits, and observed recruitment patterns as they are reported across studies.

2.1 Target Populations and Seed Selection in RDS

RDS is typically used to target “hard-to-reach” or “hidden populations,” groups characterised by their difficulty in sampling due to stigma, marginalisation, and a lack of conventional sampling frames. The RDS studies in this review can be grouped into four groups: Health-at-Risk Populations, Socially Marginalised Populations, Specialised Populations with some General Population Samples (see Table 1).

Table 1: Target Populations Groups

Target Population	Barriers to Conventional Sample Frames	Example Studies
Health-at-Risk Populations		
Intravenous Drug Users (IDU)	• Defined by high-risk injecting behaviour and criminalisation; their stigmatisation and legal risk make them difficult to sample.	(Lee et al., 2017 ; McFall et al., 2023 ; Johnston et al., 2013)
Adult smokers	• Behaviourally defined population (current smoking status); RDS is used where smokers are not easily identified via available sample frames.	(Rajani et al., 2013)
People infected with or highly at risk of sexually transmitted and blood-borne infections (STBBIs)	• High sensitivity, stigmatised population with privacy concerns which limit conventional sampling; RDS is used to gain access to discrete social networks for surveillance and risk-behaviour monitoring	(Thompson, 2013; Fordjuoh et al., 2024)
Women seeking or having had an abortion	• Highly sensitive, stigmatised population, criminalised in some contexts, RDS can reach women outside formal clinical pathways.	(Zan et al., 2024)
Socially Marginalised Populations		
Female Sex Workers (FSWs)	• Stigmatised population, criminalised in most contexts and face economic marginalisation. Social networks may be accessed by RDS via location, venue or online platform.	(Ouma et al., 2023; FISHER and MERLI, 2014; Pravosud and Andreeva, 2016)
Sexual Minorities (i.e., Lesbian, Gay, Bisexual, Transgender and Queer (LGBTQ+))	• Stigmatised population, criminalised in some contexts, RDS leverages social networks.	(Yauck et al., 2021; Welch, 1975; Middleton et al., 2021; Martin, Johnson and Hughes, 2015; Lu et al., 2011; Michaels et al., 2019; Inghels et al., 2021; Ong,

		2022; Diexer et al., 2023)
Migrant, Refugees, Ethnic Minorities, and Indigenous Populations	• Geographically mobile or dispersed. Language, legal status, and mistrust of institutions obstruct conventional sampling.	(Frere-Smith, Luthra and Platt, 2014) (Pham et al., 2023) (Leonard, 2025) (Snyder et al., 2024)
Rural Populations and Informal Settlements	• Low administrative visibility, weak sampling frames, physical access constraints and limited services, RDS leverages social networks.	(McCreesh et al., 2013 ; Truong et al., 2022 ; Jayaweera et al., 2023)
Specialised Populations		
Precarious Employment and Homecare Workers	• Fragmented employment settings (agencies, informal private homes) with limited centralised registries or union structures. RDS uses peer connectedness among workers.	(FerrerRosende et al., 2023, Jonsson et al., 2019)
University Students and Young People	• Low response rates for population from address-based sampling, strong social networks structured by courses, residence, and societies.	(Wejnert, 2010, Hildebrand et al., 2013)
General Population Samples		
Region-Specific	• RDS to access geographically defined groups, for example Guam or Albany county, USA.	(Badowski et al., 2017, Lee, 2024)
RAND American Life Panel- online panel	• Hybrid model combining online survey infrastructure with peer referral to boost representation of sub-populations with traditionally low response rates.	(Schonlau, Weidmer and Kapteyn, 2014)

The selection and number of seeds are crucial to the success of RDS. Most studies (61) used non-probability methods of recruiting initial seeds through NGOs, community groups, and digital platforms, while probability-based methods were rare. Tailored screening of seeds is used to ensure eligibility and varies by population type, such as Intravenous Drug Users (IDU) or LGBTQ+ individuals.

2.1.1 Number of Seeds

Table 2 show how many seeds were used in the studies reviewed, grouped into four main categories. It is important to note that in some studies seeds are added during fieldwork after recruitment stalled.

Table 2: Distribution of initial seed counts across studies

Initial seed count	Number of studies (%)	Typical characteristics
≤10	22 (32%)	Small or focused RDS studies
11–30	22 (32%)	Moderate-sized studies
31–100	11 (16%)	Larger studies
>100	13 (19%)	Large-scale, multi-site, panel, or multi-arm studies

Low seed counts (1–10) were used in 22 studies (32%), but this often causes problems. When recruitment slows or stops, researchers often must add more seeds later and sometimes dramatically , for example, by expanding from 8 to 43 seeds (Ngwakongnwi et al., 2014), from 7 to 8 (Inghels et al., 2021), or varying across sites (3–35 seeds; Wejnert et al., 2012; and 3–10 per region; Pham et al., 2023). Reviews and meta-analyses that report on papers with initial seed counts in this range reported a median of 7.5 initial seeds rising to 10 by study end (Johnston et al., 2016), typical ranges of 5–10 seeds (Wejnert and Heckathorn, 2008; Gile, Johnston and Salganik, 2015), and simulations comparing 1 versus 10 seeds (Lu et al., 2011). Use of a single seed was cited as ineffective (Kubal, Shvab and Wojtynska, 2014). Moderate seed counts (11–30) appeared in another 22 studies (32%), and most of these (19) did not add any new seeds later. But in Lachowsky et al. (2016), recruitment shortfalls prompted substantial seed additions, increasing from 30 to 119 seeds. Reviews that report on papers with initial seed counts in this range are Johnston et al., (2013) with up to 11 seeds and Kubal, Shvab and Wojtynska, (2014) with between 1 to over 20 initial seeds.

Larger seed counts (31–100) were used in 11 studies (16%), and most (8) did not increase from initial seeds. When new seeds were introduced as recruitment slowed, there were increases from 8 to 43 seeds (Ngwakongnwi et al., 2014) and 89 additions post-initial phase (Lachowsky et al., 2016). Wang et al. (2024b) is a review reporting on papers with ranges of 4 to 46 seeds. Very large seed counts (>100) featured in 13 studies (19%), usually in multi-site, multi-arm, or hybrid designs. Examples include 410 panel-derived seeds (Michaels et al., 2019), 143 seeds across RDS and WebRDS modes (Hildebrand et al., 2015), and geographically split samples exceeding 200 seeds across cities or regions (Lee et al., 2017; Lee et al., 2020; Lee, Ong and Elliott, 2020). In some cases, hundreds of seeds were introduced strategically as recruitment faltered (Martin, Johnson and Hughes, 2015). Reviews in this group report very wide aggregate ranges, from 1 to 358 (Leonard, 2025), 1 to 1,015 (Helms et al., 2021) and 549 across rounds/samples (Avery et al., 2021).

Across the evidence a significant proportion of seeds fail to recruit anyone, and early parts of recruitment chains fail quite often. Using a very small number of seeds increases vulnerability to stalled recruitment. Researchers should plan for this reality by being ready to add more seeds during fieldwork if necessary.

The ratio of seeds to total sample size was inconsistently reported in the studies reviewed to enable the analysis. Leonard (2025) suggests the ratio of seeds to total sample size is a critical metric in RDS studies due to its implications for the efficacy and representativeness of the sample. A high seed-to-sample ratio suggests a limited reach beyond the initial recruits, potentially undermining the study's ability to capture a diverse population. Better reporting of the ratio may help facilitate guidance about the number of seeds required to achieve a desired sample size. Leonard (2025) also proposes that the metric may reflect the effectiveness of recruitment design aspects such as the incentive structure and formative research. Appropriate incentives are crucial for motivating participants to recruit others, thereby enhancing the recruitment chains and reducing the seed-to-sample ratio. Additionally, the ratio could be indicative of the impact of formative research: if studies that engage in thorough preparatory work tend to achieve lower seed-to-sample ratios, it could be a measure of more effective and extensive recruitment. Future RDS research should prioritise explicitly reporting metrics such as the number of seeds and the achieved sample so that the seed-to-sample ratio can be assessed as an indicator of success and methodological robustness.

2.1.2 Origin and Recruitment of Seeds

Most RDS studies in the reviewed literature use purposive, non-probability seed selection, reflecting the method's employment in targeting populations who lack a defined sampling frame. Of the studies reviewed, 61 reported using non-probability-based seed selection, 9 studies employed random or probability-based approaches, and 7 studies involved simulations or theoretical modelling to assess how different seed characteristics or selection methods may impact recruitment dynamics and sample properties. Purposive seed selection methods are tailored to the local context and population which can limit methodological comparability and reproducibility between RDS studies.

Seed recruitment strategies often involve working with collaborators with knowledge of the community such as local charities, non-governmental organisations (NGOs), clinics in specific locations, and health outreach projects (Lee et al., 2020; Gile, Johnston and Salganik, 2015; Wirtz et al., 2021). Formative and participatory research involved the use of focus groups, in-depth interviews, and community advisory boards, especially in the Men Who Have Sex with Men (MSM) studies reviewed (Liu et al., 2012; Lachowsky et al., 2016) and the youth-focused recruitment from sports clubs and youth programs (Hildebrand et al., 2015; Hildebrand et al., 2013). Mixed-mode recruitment utilised both online and offline social networks, such as

Facebook advertising, online recruitment panels, distributing flyers, contacting church groups and word-of-mouth (Helms et al., 2021; Jonsson et al., 2019; Ngwakongnwi et al., 2014). Other seed recruitment included a research teams' personal connections, such as among migrant populations (Górny and Napierała, 2016) and contacting prior study participants. Venue-based seed selection was particularly utilised for sex workers and diaspora communities (Ouma et al., 2023; Inghels et al., 2021; Frere-Smith, Luthra and Platt, 2014; McGowan et al., 2023).

Although rare, probability-based seed selection (9 studies) included random selection of blocks from census tracts with a high percentage of Mexican Americans (Welch, 1975), random selection from the National Alcohol Survey (Middleton et al., 2021) and AmeriSpeak panel (Michaels et al., 2019; Xia et al., 2019), and sampling from representative cohorts, e.g., the American Community Survey (Diexer et al., 2023). 7 studies used simulations or theoretical modelling to assess how different seed characteristics or variations in peer-referral may impact recruitment dynamics and sample properties. For example, seeds were randomly drawn or selected in proportion to social network size from the 'Qruiser network' study sample (Lu et al., 2011; Goel and Salganik, 2010), or peer-referral was modelled with random selection (Luis et al., 2017).

2.1.3 Screening and Eligibility Procedures by Target Population

Seed screening is an essential step to verify that seeds belong to a target population. Screening methods varied in mode and rigor, some studies conducted eligibility assessments prior to participation via telephone, online screening forms or pre-survey questionnaires (Bostwick, Hughes and Everett, 2015; Górny and Napierała, 2016; Helms et al., 2021). Others required in-person verification at study sites, including document checks and visual confirmation before survey invitations were issued (Lee et al., 2020). Some studies rely solely on self-reporting, which may be less reliable, whilst others implement more stringent multi-step procedures to verify eligibility. For example, eligibility screening for Intravenous Drug Use (IDU) typically included self-reported recent injection behaviour or knowledge of drug preparation techniques, as well as physical evidence such as track marks (Heckathorn, 1997; Heckathorn et al., 2002) with varying time frames for recent use, including within 12 months (Wang et al., 2024a; Wejnert et al., 2012), 4 weeks (Mills et al., 2014) or weekly drug use (Liu et al., 2012). Sometimes additional demographic criteria were included, such as age (18+), local residency, and language proficiency (Wang et al., 2024a; Wang et al., 2024b; Wejnert et al., 2012). Some studies focused on subgroups or adjacent at-risk populations, not exclusively defined by injection status (Johnston et al., 2013; McFall et al., 2023; Young et al., 2014; Lee et al., 2017).

Screening of Female Sex Workers (FSWs) focused on occupational identity, with some studies targeting specific subgroups, such as FSWs with children (Ouma et al., 2023) or those in specific urban areas (Fisher and Merli, 2014; Pravosud and Andreeva, 2016). In other cases,

FSWs were included as part of broader at-risk groups (Johnston et al., 2013). Eligibility criteria for LGBTQ+ populations and specifically Men Who Have Sex with Men (MSM), included questions about sexual behaviour (e.g., recent sex with men), self-identified sexual orientation, gender identity and residency. Some studies focused on specific age groups such as 15–24 (Wirtz et al., 2021), or 18+ (Truong et al., 2013; Lachowsky et al., 2016; Inghels et al., 2021). Purposive selection was used to ensure diversity across attributes like ethnicity and HIV status (Lachowsky et al., 2016) within the target population. Eligibility for LGBTQ+ populations was also ascertained from prior questions in national surveys or panel membership (Middleton et al., 2021; Martin, Johnson and Hughes, 2015) or inferred based on membership of digital communities (Lu et al., 2011). Some studies employed mixed-seed strategies where both LGBTQ+ members of the target population and non-LGBTQ+ individuals were eligible for seed selection (Michaels et al., 2019).

Age was the primary eligibility criterion for Adolescents and Youth (typically 14–19) (Hildebrand et al., 2015; Hildebrand et al., 2013; Truong et al., 2022), with some studies also screening at youth groups or sports participation (Hildebrand et al., 2013). Recruitment for Students was typically limited to university-affiliated individuals and focused on residences (Wejnert, 2010), but other studies provided less detail about the process (Wejnert and Heckathorn, 2008). Simulated networks using student populations were used in experimental designs (Goel and Salganik, 2010). Other studies targeting young adults had eligibility criteria to select across a wide range of sociodemographic variables, such as age, gender, region, and education level (e.g., Bauermeister et al., 2012), social network size (Bostwick, Hughes and Everett, 2015), and employment status (Jonsson et al., 2019).

Studies with a reproductive health focus screened women based on recent reproductive events (e.g., abortion within the past three years) (Zan et al., 2024) or demographic factors such as age and residence (Jayaweera et al., 2023; Fordjuoh et al., 2024). A study of rural Ugandan tribes focused on male household heads, with eligibility determined by household status and village location (McCreesh et al., 2013; McCreesh et al., 2012). Screening for migrant populations incorporated nationality, ethnicity, or migration history. Examples included Korean Americans (Lee et al., 2020), Venezuelan migrants (Pham et al., 2023), and recent Polish and Pakistani immigrants in London (Frere-Smith, Luthra and Platt, 2014). The studies required participants to identify with specific cultural groups and practices or possess language proficiency relevant to the study context (e.g., Francophone or Yemeni diaspora) (Ngwakongnwi et al., 2014; McGowan et al., 2023; Leonard, 2025). Studies targeting workers in specific economic conditions used employment status as the key screening criteria (for example, the precariously employed (Jonsson et al., 2019) or homecare workers (FerrerRosende et al., 2023)).

2.1.4 Active Peer Recruitment

Active Peer-to-Peer recruitment is central to the logic of RDS: seeds recruit peers directly via incentivised systems (e.g., coupons, vouchers, digital codes) and in 64 studies this approach was explicitly detailed. 45 of the studies provided detail about peer-to-peer recruitment, whilst 19 studies were more ambiguous. Structured systems are common and are typically traceable via personalised links and unique codes. The most common recruitment systems were:

- Coupon-based systems: Widely used across studies (e.g., Young et al., 2014; Thompson, 2014; Salganik, 2006; Wylie and Jolly, 2013; Jayaweera et al., 2023), these are numbered coupons that resemble tickets and are mostly distributed physically, via SMS/email, or on social media platforms like KakaoTalk (Lee et al., 2020) and used to access the study and incentives.
- Digital-only referrals: some studies reported use of referral links (Jonsson et al., 2019; FerrerRosende et al., 2023) and invitation codes or PINs (Thompson, 2014; Michaels et al., 2019; Sosenko and Bramley, 2022).

A smaller group of studies (19) implied the language of active seed referral (e.g., “initiated recruitment”), without clarifying the specific mechanisms. Whilst the RDS framework implies active engagement, the lack of detail precludes further classification.

2.2 Incentive Structures and Their Impact on Recruitment

The use of dual incentives (rewards for both participation and successful peer recruitment) is a core feature of RDS methodology and has been shown to boost recruitment efficiency and improve sample diversity. As summarised in Table 3, monetary incentives remain the most common form of incentivisation. Within the evidence, incentives typically range from $5 to $60 for participation (median: $20), and $0.75 to $20 for referrals (median: $10). For example, one study offered $50 for participation and $10 for each successful referral (Ong, 2022), while others used more modest incentives, such as $6–$10 per task (Liu et al., 2012; Heckathorn, 1997).

Electronic vouchers ranging from $6 to $15 were also used in digital or web-based RDS contexts, offering flexibility in delivery (e.g., $15 store voucher for participation and $10 for referral (Hildebrand et al., 2013)).

In low-resource or culturally sensitive settings, small non-monetary items such as soap, notebooks, or transportation support were provided instead and valued between $1 to $25 (McCreesh et al., 2012; McCreesh et al., 2013). These token incentives often served ethical and logistical purposes, especially where direct payments may have posed challenges.

Several studies introduced lottery-based incentives, including raffles for high-value prizes (e.g., iPads, $500 draws per 50 referrals, or entries toward a $2,000 travel voucher), which

aimed to boost engagement while reducing coercive pressure (Truong et al., 2013; Lachowsky et al., 2016). Other studies adopted "Steering Incentives" which are similar to 'Conditional Incentives' in general survey methods. They are tiered and increase payments incrementally for reaching recruitment targets, or provide a bonus for enrolling specific subgroups (such as female Intravenous Drug Users (IDUs) (Heckathorn, 1997).

Three studies emphasised non-monetary or socially motivated incentives to participation, appealing to altruism, community solidarity, or the symbolic value of contributing to public health goals, such as support for MSM communities in Vietnam (Bengtsson et al., 2012), and further detailed in Table 3. While these approaches may reduce ethical risks, they often lead to slower recruitment or fewer waves.

Overall, while incentive strategies varied widely across the reviewed literature, their effectiveness depended heavily on context. Studies that aligned incentive types and values with local norms, ethical standards, and logistical constraints, such as using digital delivery in online studies or culturally appropriate tokens in community-based settings, tended to report more successful recruitment outcomes. Table 3 categorises the various incentive structures used across the reviewed literature on RDS studies. It includes monetary and non-monetary incentives, with typical value ranges and examples. The diversity of incentive strategies reflects efforts to balance recruitment efficiency with ethical, cultural, and logistical considerations.

Table 3: Types of incentives

Incentive Type	Description & Examples	Typical Value Range	Context of Use	References for examples
Monetary: Cash	Cash or direct monetary rewards for participation or recruitment.	Participation: $5–$60, median $20 Recruitment: $0.75–$20, median $10	Most common	
Monetary: Electronic Vouchers	Digital vouchers or electronic rewards (e.g., (1) shopping vouchers £5–£10, (2) mobile credits equivalent to lunch costs; (3) electronic store gift voucher:	$6- $15 for participation $6 -10 for recruitment	Online studies; contexts requiring digital delivery.	1. (Sosenko and Bramley, 2022) 2. (Pham et al., 2023) 3. (Hildebrand et al., 2013)
Monetary: Lottery	Chance-based rewards. Examples: iPad lottery, a $500	Range widely: $250–$2000 prizes	Increasing participant interest	(Bengtsson et al., 2012)

	raffle draw for every 50 referrals.		through large, chance-based incentives.	
Non-Monetary: Token	Small tangible items or gifts, such as soap, salt, school notebooks, or gasoline vouchers (e.g., (1) $1 equivalent items; (2) Bus tokens; (3) Free HIV screenings and counselling	Approx. $1–$25 equivalent	Low-resource settings; culturally sensitive contexts.	1. (McCreesh et al., 2012; McCreesh et al., 2013) 2. (Wirtz et al., 2021) 3. (Yauck et al., 2021; Gile, Johnston and Salganik, 2015; Heckathorn et al., 2002)
Steering, Tiered or Incremental Incentives	(1) initial £5 incentive increasing to £10 after poor referrals; (2) an extra $5 bonus for recruiting female IDUs to balance gender representation. (2 + 3) Bonus for recruiting within target population	$5-6 incremental increases) $5-$10 bonuses for recruiting 18-25 yo or Female IDUs	To motivate ongoing active recruitment (of a certain group). Boosting target population	1. (Frere-Smith, Luthra and Platt, 2014) 2. (Heckathorn , 1997) 3. (Heckathorn , 2002).
Non-Incentivised / Social Incentives	(1) Relies on altruism, (2) community motivation, (3) social recognition.	Non-monetary or service-based	Community-driven contexts; ethical considerations.	1. (Ong, 2022) 2. (Bengtsson et al., 2012; Heckathorn, 1997) 3. (Rajani et al., 2013)

Note: based on 68/99 reviewed studies, others did not provide relevant information

2.3 Achieved Sample Sizes and Sample Representativeness

2.3.1 Achieved Sample Size Variation in RDS Studies

The number of participants recruited in RDS studies vary widely, reflecting differences in study design, target populations, and logistical constraints. This variation illustrates the adaptability of RDS to different research aims and contexts.

I. Small samples (fewer than 300 participants):

Seven studies recruited relatively small numbers of participants, with a focus on hidden or narrow subgroups or specialised research questions. Martin et al. (2015) recruited 62 Sexual Minority Women (SMW) in Chicago, and Diexer et al. (2023) sampled 98 Men Who Have Sex with Men (MSM) in Amsterdam. Additional studies reported recruitment totals such as 93 (McGowan et al., 2023), 145 (Ngwakongnwi et al., 2014), 190 (Heckathorn, 2002), and 243 (Heckathorn, 2007). In one case, recruitment occurred across two separate time points, with 299 participants in 2006 and 292 in 2009 (Mills et al., 2014). While these smaller studies are often valuable for focused inquiry, they may offer limited generalisability of the results and statistical power.

II. Medium-sized samples (300–1,000 participants):

The majority of RDS studies reviewed (51) fall within this medium-sized range category. Examples include 503 people who use drugs in Appalachia, US (Young et al., 2014), 583 Female Sex Workers (FSW) in Liuzhou, China (Fisher and Merli, 2014), and 975 FSW in five Ukrainian cities (Pravosud and Andreeva, 2016). Other studies recruited 813 participants (Crawford, 2016), 638 across Los Angeles and Miami (Lee et al., 2020) and multiple cohorts between 394 and 524 participants (Wang et al., 2024a). One review noted a median sample size of 325 across studies, with a range from 100 to 1,056 (Johnston et al., 2016). This sample size category is common in RDS research, offering a balance between feasibility and analytical robustness.

III. Large samples (1,000–5,000 participants):

Five studies reported sample sizes exceeding 1,000 participants. Yauck et al. (2021) recruited 1,179 gay and bisexual men in Montréal, and Truong et al. (2022) included 3,061 adolescents in Kisumu, Kenya. Other studies included 1,203 participants in Ecuador and 1,231 in Peru (Zanoni et al., 2023), as well as 3,448 participants in a single large-scale study (Bauermeister et al., 2012). One study combined face-to-face RDS with web-based recruitment, reaching a total of 1,012 participants (Hildebrand et al., 2015). These larger studies enable more detailed sub-group analyses and increase the precision of statistical estimates.

IV. Very large or aggregated datasets (more than 5,000 participants):

Three studies involved large-scale recruitment or pooled data across multiple locations or rounds. McFall et al. (2023) reported 6,012 participants who engaged in Intravenous Drug Use (IDU) across six Indian cities. The National HIV Behavioral Surveillance System included 26,705

participants across two rounds (NHBS-IDU1 and NHBS-IDU2), with a final analytic sample of 21,686 (Wejnert et al., 2012). One meta-analysis aggregated recruitment data from 53 samples, totalling 36,547 participants (Avery et al., 2021).

Overall, smaller studies (<300 participants) are valuable for focused inquiries but may offer limited generalisability and statistical power. Medium-sized samples (300–1,000 participants) are most common in RDS research, balancing feasibility and analytical robustness.

2.3.2 Referral Caps

Fixed referral caps are a core feature of RDS methodology, aimed at balancing recruitment speed and dispersal by preventing a few highly connected people from dominating the sample. The majority of studies cap referrals at three coupons per participant, with over ten studies explicitly reporting this limit (for example, Bostwick et al., 2015; Crawford, 2016; Gile, 2011; Goel and Salganik, 2010; Liu et al., 2012; Ngwakongnwi et al., 2014). Four studies permitted 4 to 6 referrals, particularly when recruitment needed to be accelerated (Bengtsson et al., 2012; Heckathorn, 2007; Jonsson et al., 2019; Wang et al., 2024a; Lachowsky et al., 2016). Dynamic or adaptive referral limits start with lower caps, before increasing to maintain recruitment-chain progression. This can reduce the need for new seeds to be added and can help reach harder-to-access sub-groups if highly connected recruiters are present (Wirtz et al., 2021; Truong et al., 2013; Helms et al., 2021). Section 4.5 provides more details about recruitment monitoring and diagnostic tools used to help inform these decisions.

2.3.3 Recruitment Productivity Patterns

Recruitment patterns in RDS shape key methodological outcomes, including sample composition, chain length, and susceptibility to bias. Drawing on a typology informed by the reviewed evidence (*Table 4*), three general patterns emerge: low productivity, moderate productivity, and high productivity (super-seeds - individuals who recruit disproportionately large portions of the sample). These patterns are strongly influenced by referral caps, recruitment incentives, and characteristics of initial seeds and their social networks.

Table 4: Types of recruitment pattern

Recruitment Pattern	Description	Typical Range (throughout recruitment)	Notes	Number of studies with such pattern
Low Productivity	Many seeds fail to recruit any participants, short recruitment chains are common, limited coupon distribution.	0–3 recruits per seed	Many seeds are non-generative - seeds commonly failed to recruit any participants. -Administrative errors occasionally resulted in up to 4 recruits per seed.	4
Moderate Productivity	Most seeds recruit a modest number of peers; stable, moderate-length recruitment chains; moderate coupon limits usually enforced (3–5 coupons).	3–10 recruits per seed	The majority of studies limit recruits per seed (3–5), though occasionally relaxed to increase participation. - Seeds recruiting between 3–10 peers usually maintain sustainable chains.	53
High Productivity (Super-Seeds)	A few seeds recruit disproportionately high numbers, forming extensive recruitment chains; often facilitated by more relaxed coupon limits.	10–241 recruits per seed	Exceptional cases observed where individual seeds recruit large proportions of total sample, indicating strong	5

			network influence.	

Note: based on 62/99 studies, others did not provide relevant information

I. Low Productivity Patterns

Low productivity is typically defined by minimal recruitment success per seed, often in the range of 0–3 recruits, with many seeds failing to generate any referrals. Studies report that a significant proportion of seeds or participants do not recruit any peers. One study noted that over 50% of seeds failed to recruit even a single participant (Young et al., 2014). This pattern also has short recruitment chains, usually only 1–3 waves, which can impair the ability to reach equilibrium and increase seed-dependence bias. Recruitment limitations were often directly shaped by lower coupon caps (Lee et al., 2020; Górny and Napierała, 2016; Ngwakongnwi et al., 2014). In some cases, administrative or design constraints limited participants to only two referrals, especially in contexts seeking to tightly control chain expansion (Lee et al., 2020).

Low productivity may stem from multiple causes, including poor seed selection, insufficient peer connectivity in the target population, and low motivation to recruit or participate (e.g. due to inadequate incentives). It signals the need for better formative research, improved seed monitoring to reduce early-chain drop-off and adaptive recruitment strategies to enhance representativeness.

II. Moderate Productivity Patterns

Moderate productivity patterns were the most common in RDS studies. Typically, seeds recruit between 3 and 10 peers, with referral limits of three to five coupons per participant (Crawford, 2016; Goel and Salganik, 2010; Liu et al., 2012; Yan et al., 2020). This structure encourages moderate chain growth while minimising the risk of over-representation by a few highly connected individuals. In one example, participants were issued three recruitment coupons each, resulting in mean recruitment of approximately 1.69 peers per successful recruiter, with 55.3% of participants successfully recruiting at least one peer (Wejnert, 2010). Several studies initially provided three referrals per participant but later increased the cap to five or ten in response to recruitment slowdowns (Wirtz et al., 2021; Truong et al., 2013). Moderate productivity patterns often reflect deliberate methodological choices, balancing recruitment control and speed, ethical concerns, and sample representativeness. However, they still require active monitoring to ensure inclusion of socially isolated sub-groups, who may be less likely to be reached through standard recruitment chains.

III. High Productivity Patterns and Super-Seeds

In a smaller subset of studies, a few extremely productive seeds (often called "super-seeds") recruited a disproportionately large numbers of participants. These high-productivity scenarios typically involve more relaxed coupon limits, or seeds embedded in highly connected social networks. Recruitment figures in these cases ranged from 10 to over 240 participants per seed, and representing the majority of the final sample. In one example, a single seed recruited 241 individuals, comprising 77% of the total sample (McCreesh et al., 2012). Another study found that three seeds recruited 92.4% of the entire participant pool, with individual seeds recruiting up to 259 participants (Ferrer-Rosende et al., 2024). Super-seed effects were facilitated by high referral caps with up to ten referrals permitted in some contexts (Wirtz et al., 2021; Truong et al., 2013; Helms et al., 2021). While high productivity can enhance recruitment speed and help achieve equilibrium in some variables, it also poses risks to representativeness. Recruitment dominated by a small number of highly connected individuals can lead to network clustering, recruitment bias or 'Homophily' (a tendency of recruiters to recruit people like themselves) with over-representation of specific sub-groups and under-representation of others. Unless there is a strong quality control process, super-seeds could be a signal of fabricated data, fraud or respondent duplication.

Recruitment productivity in RDS studies falls along a spectrum from low to high, with each pattern carrying distinct methodological implications. Low productivity highlights challenges in network engagement and seed selection, often requiring mid-course adjustments. Moderate productivity (supported by typical referral limits of 3–5) strikes an operational balance, promoting stable recruitment while preserving representativeness. High productivity, often led by a few super-seeds, may speed recruitment but introduces heightened risk of bias. The risks within each pattern underscore the need for careful diagnostic tools, which are detailed in section 4.5, and to assess potential biases introduced by super-seeds.

2.3.4 Recruitment Chain Length and Representativeness

Table 5 shows the variation in recruitment chain length within the evidence.

Table 5: Chain length

Chain Length	Typical Characteristics	Examples / Notes
Short Chains (1–3 waves)	- Often results from unproductive seeds - High seed bias - Limited diversity - Insufficient for equilibrium	- Studies reported mean waves as low as 1.75; many chains died at seed stage or after one referral wave.

		- Web-RDS are often limited to 2–3 waves.
Moderate Chains (4–9 waves)	- Commonly observed range - Usually adequate for equilibrium - Improved diversity - Reduced seed bias, but still sensitive to homophily effects	Median recruitment chains frequently ranged from 4-9 waves; equilibrium often reached within 4-6 waves.
Long Chains (10–20 waves)	- Associated with equilibrium achievement - good representativeness and diversity - Reduced dependency on initial seeds - Risk of clustering or compounded homophily if too extensive	Longest chains up to 28 waves; effective at capturing diverse subpopulations, though sometimes geographically or demographically clustered. Equilibrium stable by wave 10 in many cases.

Note: based on 62/99 studies, others did not provide relevant information

Chain length varies greatly even within the same study, indicating variability in a seeds' effectiveness. Recruitment chains ranged from extremely short with a mean chain length of 1.75 waves, with a minimum of 0 waves due to unproductive seeds) and reaching up to 29 waves (see Table 5).

Short chains (1–3 waves) result from unproductive seeds and can lead to high seed dependence and limited demographic reach, failing to achieve equilibrium for key variables (Hildebrand et al., 2015; Lachowsky et al., 2016). For example, some studies reported that up to one-third of seeds were unproductive and failed to recruit any participants, necessitating the introduction of new seeds mid-study (Hildebrand et al., 2013; Avery et al., 2021; Helms et al., 2021).

Moderate-length chains (4–9 waves) were the most common and generally sufficient for achieving equilibrium with improved diversity and reduced influence of the characteristics of initial seeds (Heckathorn, 1997; Heckathorn et al., 2002; Liu et al., 2012; Wang et al., 2024a). This suggests that recruitment chains shorter than 4 waves may be insufficient to achieve independence from the initial seed bias.

Long recruitment chains (≥10 waves) were positively associated with representativeness and equilibrium. Studies documenting chains of 13, 16, or even 28 waves found this enabled deeper penetration into social networks, reducing initial seed influence and improving diversity of respondents (Wang et al., 2024b; Gile, Johnston and Salganik, 2015; Lachowsky et al., 2016). However, chains extending beyond 20 waves raised concerns about compounded homophily, geographic clustering and other structural biases as longer chains can become concentrated within specific sub-groups.

3. Analysis and Evaluation of RDS Recruitment

3.1 Recruitment Mode Effects

The evidence highlights how recruitment mode, such as face-to-face, online, phone and mixed-mode approaches change the RDS process itself in a way distinct to generic survey mode effects. Recruitment mode influences coupon and link control, seed productivity, recruitment chain growth, the ability to verify eligibility and 'who recruited whom', the risk of duplicate participation, and interacts with digital access and resource constraints. Table 6 summarises the recruitment modes used in the studies, where they were reported.

Table 6: Online and Offline Recruitment

Recruitment Mode	Number of studies
Face-to-Face Only	27
Online Only	18
Mixed-Mode	12
Unclear/Other	1

3.1.1 Offline (Face-to-Face) Recruitment

Face-to-face recruitment methods dominate the RDS studies analysed because of the mode's application to accessing and engaging hard-to-reach populations. Physical recruitment shapes who can enrol and how quickly recruitment proceeds. In-person engagement allows researchers to build trust and rapport with participants through outreach and locally established community organisations in a way that can be critical for stigmatised or hesitant groups. RDS explicitly relies on trust embedded in social networks, and offline settings allow staff to reinforce this through personal interaction (Heckathorn, 1997, Liu et al., 2012). Interviewer training, confidentiality and a welcoming environment can serve to increase comfort, engagement and coupon exchange. Offline RDS frequently bundles services such as HIV testing, counselling, or STI screening along with participation incentives. These services, such as those used in the US National HIV Behavioral Surveillance system on injection drug use (NHBS-IDU), allow for additional non-monetary participation incentives (e.g., free HIV tests) that affect recruitment success and coupon return, largely unavailable to fully online RDS designs (Wejnert et al., 2012). Górny and Napierała, (2016) report the positive effect of in-person peer-referral on study data quality, with RDS respondents more 'primed' to trust the interviewer and the associated institution through the mechanism of peer-referral. They found peers had conveyed to respondents a positive impression about the interview, its atmosphere, duration and the character of the questions, before answering survey questions.

Face-to-face interviewing also allows for rigorous eligibility screening procedures, because coupon redemption is monitored by the members of staff. In Intravenous Drug Use (IDU) RDS studies, eligibility confirmation involved including visual identification of IDU status through a 7-step process, which included physical evidence (recent track marks) and detailed knowledge assessment (e.g., the ability to describe injection techniques), ensuring that only eligible seeds were selected by trained staff (Heckathorn, 1997). This type of verification directly protects the integrity of the recruitment chain, something which is much harder to validate remotely. Offline RDS approaches conducted in homes, clinics, or community venues were advantageous for populations with limited access to or familiarity with technology. Studies conducted entirely in-person in rural or low-resource settings were able to reach individuals who may not otherwise be accessible via web-based platforms and included target populations in Uganda (McCreesh et al., 2013; Ouma et al., 2023), FSW in rural China (Fisher and Merli, 2014), and Pakistani migrants (Frere-Smith, Luthra and Platt, 2014).

However, these advantages came at the cost of significantly higher resource demands. Offline studies required investment in staff training, physical infrastructure and venues, and transportation logistical coordination (Heckathorn, 1997; Wang et al., 2024b; Górny and Napierała, 2016). Offline RDS is geographically anchored and if it is not located across multiple sites, recruitment clusters around a single location. This can affect participant burden in terms of travel, time costs, and any potential stigma associated with visiting survey sites. Dependence on geographic location can bias the achieved sample toward those with easier physical access. Offline studies can also lead to slower recruitment and more limited scalability, as the mode is dependent on staff presence and site capacity. Physical coupon exchange represents a physical bottleneck, which is more difficult to rapidly expand compared to online approaches. This limitation becomes evident when contrasted with web-based RDS, which can recruit hundreds or thousands of participants in weeks rather than months (Wejnert & Heckathorn, 2008).

3.1.2 Online Recruitment

Online RDS has the advantage of removing physical bottlenecks. Asynchronous and remote participation enables rapid recruitment chain progression when the target population is well-connected online (Wejnert & Heckathorn, 2008; Bengtsson et al., 2012). The mode can reduce participant burden and the versatility of platforms available can broaden network access opportunities, ranging from web-based questionnaires and email/SMS referrals to integrations with Zoom, WhatsApp, Open Data Kit (ODK) infrastructures to a variety of incentive delivery options (Bauermeister et al., 2012; Jonsson et al., 2019; McGowan et al., 2023). For example, Ong (2022) integrated Zoom interviews with web surveys, Wejnert (2010) used a fully email-based coupon referral system and Schonlau et al. (2014) relied on existing internet panels for both recruitment and follow-up. Digital systems can automatically log recruiter–recruit links, referral limits, and incentive eligibility, reducing manual errors that are possible with paper-based coupon systems (Wejnert & Heckathorn, 2008; Helms et al., 2021).

However, within the evidence there is a trade-off between these advantages and the fact that online RDS is vulnerable to duplicate participation, identity misrepresentation, and incentive fraud (Helms et al., 2021). Sosenko and Bramley (2022) note that it is often unable to screen vital eligibility criteria such as residence, age or ethnicity, and that online RDS opens a *"Pandora's box"* of fraud risks as recruitment is incentivised, which can challenge the credibility of RDS sample data. Web-RDS also requires secure platforms, device compatibility, IT expertise, and notable technical setup costs (McGowan et al., 2023). Technical failures such as mobile incompatibility, broken links, email spam filtering and server issues have been reported as recruitment barriers influencing referral chain progression in multiple web-RDS studies (Helms et al., 2021). Other online mode effects reported are not unique to RDS. Helms et al. (2021) note persistent concerns about survey content engagement in web-RDS relative to face-to-face designs.

3.1.3 Mixed-Mode Recruitment (Hybrid Designs)

The key strength of mixed-mode recruitment is that it allows researchers to tailor to participants' preferences and technological access, maximising reach for heterogeneous population sub-groups. However, the fact that different modes reach different subgroups violates RDS assumptions if not accounted for. Pham et al. (2023) compared telephone vs Internet RDS for Venezuelan refugees and migrants in Colombia and found that whilst the phone-based recruitment reached the target sample size in both locations, the Internet strategy succeeded in reaching their target sample size in Bogotá but not in Norte de Santander, a lower-connectivity remote border region. Furthermore, combining modes increases operational complexity and creates difficulties in harmonising consent, eligibility screening, incentive systems, referral, and link-tracing (Ngwakongnwi et al., 2014). Synchronising paper coupons with digital links and preventing cross-mode duplication requires bespoke systems and careful monitoring (Wejnert & Heckathorn, 2008; Helms et al., 2021). Hildebrand et al. (2015) compares face-to-face RDS against WebRDS in a study of adolescent alcohol norms and highlights how mode changes the mechanics of peer referral and recruitment management, not just response patterns. Helms et al., (2021) notes that many mixed-mode studies fail to report results separately due to analytical and operational complexity, undermining the credibility of those results.

3.2 Link-Tracing Methods

RDS is fundamentally a link-tracing sampling method. Documenting recruiter–recruit relationships (along with each participant's network size) theoretically underpin the estimation of population parameters (Heckathorn, 1997; Wejnert & Heckathorn, 2008), enables reconstruction of recruitment chains (often visualised as tree diagrams) and the calculation of metrics like homophily or recruitment depth. Preserving the peer-referral structure in the data can help researchers assess whether the sample has reached equilibrium and whether recruitment was random within social ties (Helms et al., 2021). These methods are detailed in section 4.5.

49 studies explicitly document link-tracing mechanisms, with uniquely numbered coupons being the most frequently employed (25 studies). Each coupon carries a unique identifier that links the recruit to the recruiter's ID. Digital/electronic tracking was mentioned in 9 studies, and management systems such as RDS Coupon Manager (an SPSS package) was referenced by Koyuncu et al. (2023) and Badowski et al. (2017), which helps track coupon distribution, prevent duplicates, and records recruitment relationships. In a study on identifying IDUs in India, McFall and colleagues (2023) linked respondents' biometric data (e.g. blood) to their records at local health facilities to rule out any duplication (known as biometric matching).

Some studies detailed the implementation of electronic systems other than coupons for link tracing. In web-based RDS (WebRDS), participants may receive digital codes or URLs that serve the same role as physical coupons (McGowan et al., 2023). Wejnert & Heckathorn (2008) demonstrated an early example of online RDS where seeds are recruited by email and given unique invitation links (or codes) to forward via email to their peers. Michaels et al. (2019) tracked recruitment using email addresses and personal identification numbers (PINs), while Inghels and colleagues (2021) used SMS-based text messages with embedded referral codes. Pham et al., (2023) and FerrerRosende et al., (2023) tracked recruitment through unique digital links distributed via WhatsApp, phone, or internet-based platforms. Sosenko and Bramley (2022) reported the use of smartphone-based systems incorporating unique invitation codes and phone numbers. Similarly, other studies (Bauermeister et al., 2012; Jonsson et al., 2019; McGowan et al., 2023; Lee et al., 2020) described digital recruitment processes employing web-based RDS platforms with embedded tracking IDs or digital tokens.

Manual methods were used in 2 studies, especially in earlier RDS studies or resource-limited settings. Welch (1975) tracked referrals using participants' names and addresses. Another study (Ouma et al., 2023) reported maintaining manual coupon log notebooks to monitor recruitment activities. A smaller subset of studies (3) documented mixed-mode paper-based and electronic tracking systems (Truong et al., 2013; Lachowsky et al., 2016; Wirtz et al., 2021). In these studies, participants could choose between physical and electronic referral coupons. All recruitment instances were assigned unique identifiers across all modes, and in some cases systematic follow-up was conducted to ensure linkage.

While link tracing is clearly crucial in RDS, many published studies provide minimal technical detail on how it was implemented. For example, authors often state that coupons were used to track recruitment, but do not specify the coupon design, logging procedure, or any software used. The STROBE-RDS reporting guidelines (White et al., 2015) were introduced to improve this situation, but their uptake remains limited. Newer variants (like web-based RDS) sometimes imply link-tracing mechanisms without describing them. For instance, a webRDS study might mention that seeds were given referral links but not provide detail about duplicate or fraudulent enrolments. McGowan et al., 2023 found that webRDS implementations and designs varied considerably and *"none of the available sources provide*

a thorough description of the RDS system architecture," making it difficult to compare strengths and weaknesses of each link-tracing approach.

3.3 Technical Challenges in Link-Tracing

A frequent challenge in RDS is that a substantial fraction of initial seed participants do not recruit anyone at all, with one study reporting that over 50% failed to generate recruits due to invalid addresses and procedural errors (Young et al., 2014). It is, therefore, often necessary to substitute unproductive seeds with new ones to jump-start recruitment, particularly in response to early recruitment failures (Zan et al., 2024). This kind of ad-hoc seeding is common in RDS fieldwork (Wejnert et al., 2012), but it can complicate analysis if not properly documented. Another issue arises when participants are mistakenly treated as seeds due to missing link data. Ideally, every recruit should come with a coupon or referral code identifying their recruiter. However, if a recruit shows up without a valid coupon (for example, they lost it, or the code was illegible), the study might enrol them to not lose the data, effectively considering them a new "seed" in the dataset. This misclassification of recruits as seeds breaks the recruitment chain and can distort network measures and undermine the validity of the resulting recruitment chain analysis (Wang et al., 2024b).

When RDS relies on digital platforms, a host of technical issues can impact recruitment and link-tracing, and one issue is basic internet and communication infrastructure. If participants have unstable internet access, online surveys may be interrupted, or referral links may not send or receive properly (e.g., Pham et al., 2023). An extreme example is where email invitations were filtered as spam during an online RDS pilot in Thailand. Despite careful setup, many invite emails never reached participants because email providers blocked them, which *"severely affected"* recruitment chain progression (Helms et al., 2021). Similarly, web-based RDS studies have reported that a portion of seeds' recruitment emails ended up in junk folders or were ignored as possible scams, leading to slower recruitment that required follow-up or alternate contact modes (Helms et al., 2021). Platform policies and constraints can also limit RDS. WhatsApp, for instance, forbids bulk unsolicited messaging, and a web-RDS project noted that WhatsApp's automated detection of bulk messaging was a barrier preventing seeds from easily forwarding invites to many peers at once (FerrerRosende et al., 2023; McGowan et al., 2023). SMS referrals can also be constrained by character limits and some carriers filter bulk messages containing referral links after being flagged as potential spam. In a Yemeni diaspora web-RDS study, the integration with SMS/WhatsApp was *"relatively complex and limited the information we could send*" within a single message, restricting referral success (McGowan et al., 2023). Helms et al., (2021) found that unexpected incompatibility with individual mobile platforms and devices (e.g., Android, iPhone) was a barrier to referral-link progression in several studies.

Maintaining the integrity of Link-Tracing data goes beyond recruitment itself and extends to the quality and completeness of survey responses and the personal network data collected. Because RDS relies on peer recruitment, sometimes people will enrol (perhaps to get an incentive) but not fully complete the survey, resulting in partially filled questionnaires or missing key information. This creates a dilemma as excluding those cases may significantly truncate a recruitment chain but including them means having missing data. Wejnert et al. (2012) refer to situations where eligible participants started but did not complete the full survey. Although their substantive responses were excluded from analysis, their recruitment data was retained to preserve the integrity of the recruitment chains. Specifically, 64 individuals with incomplete surveys were included solely for their recruitment information, ensuring continuity in the peer-referral structure critical to RDS modelling.

3.4 Network Effects, Clustering and Population Parameters

The distinctive contribution of RDS to sampling is its attempt to use chain-referral recruitment with estimation to analyse and adjust for recruitment and network-based selection biases. Therefore, the effectiveness of RDS inference is contingent upon interrelated assumptions about the underlying social network and the recruitment process in the statistical estimators used. The evidence suggests that valid inference does not depend on any single recruitment factor but on the joint satisfaction of several interdependent conditions, including the underlying network structure, recruitment randomness and the accuracy of reporting personal network size (degree). Without capturing these indicators accurately RDS risks producing misleading inferences about a population.

3.4.1 Theoretical Assumptions Underpinning RDS Estimators

RDS estimation relies on a stochastic framework, typically modelled as a random walk or first-order Markov process on a social network (Handcock and Gile, 2010; Hartfiel and Seneta, 1994; Heckathorn, 2002). This model assumes that recruitment transitions are memoryless, that is, only influenced by the recruiter's characteristics and not by earlier steps in the chain (Heckathorn, 2002). Core statistical assumptions include the following: respondents accurately report their personal network size, network ties are reciprocal, the network of the target population forms a single inter-connected component, peer recruitment is random, and sampling occurs with replacement (Heckathorn, 2007; Heckathorn, 2011; Johnston et al., 2016). However, RDS is typically conducted without replacement, as individuals cannot be recruited more than once (Crawford, 2016; Gile and Handcock, 2010; Jonsson et al., 2019). While with-replacement models facilitate estimation, they can introduce bias, particularly when sample fractions are large or when the network has uneven connectivity (Gile and Handcock, 2010; Wang et al., 2024a; Gile, Johnston and Salganik, 2015).

3.4.2 Network Structure, Clustering, and Homophily

Network structure fundamentally shapes recruitment dynamics and estimator accuracy. Many studies assume or report a single, connected, sufficiently dense social network in which

members of the target population are connected to one another, allowing recruitment chains to reach the full target population (Heckathorn, 2007; Lee et al., 2020; Johnston et al., 2016; Górny and Napierała, 2016). In reality, populations often contain disconnected sub-groups, which lead to clustering and homophily (where individuals are more likely to recruit those similar to themselves). Bottlenecks and low conductance between sub-groups violate the assumption that the eligible population network is (effectively) one component and that the recruitment process mixes well enough between groups for the sample composition to become seed-independent. The implication of this assumption is that estimates for minority sub-groups become unstable and 'equilibrium' claims may be superficial (i.e., do not mean global mixing). Homophily statistics are commonly calculated to quantify within-group recruitment bias, which can reinforce clustering effects (Heckathorn, 1997; Heckathorn et al., 2002; Wang et al., 2024b; Wejnert et al., 2012). Within-group biases are reported across various dimensions, including gender, ethnicity, age, and drug use behaviour (Heckathorn, 1997; Heckathorn et al., 2002; Wang et al., 2024a; Wang et al., 2024b). National, linguistic, workplace, and status-based homophily can be strong enough that recruitment is quasi-segregated. Sub-group proportions and sub-group specific outcomes remain dependent on the initial seeds and their effect on recruitment dynamics. Such clustering and network fragmentation can reduce representativeness and increase design effects.

3.4.3 Non-Random Recruitment and Preferential Referral

Despite theoretical assumptions of random recruitment, numerous studies report evidence of non-random preferential referral. The evidence shows that people recruit "reachable" peers, friends, coworkers, and those within their own social or demographic group (Heckathorn, 2002; Jonsson et al., 2019; Wirtz et al., 2021; Ngwakongnwi et al., 2014). Recruitment patterns were influenced by observable characteristics such as age, gender and education, resulting in samples skewed toward cooperative and socially proximate individuals (Lee et al., 2020; Liu et al., 2012). As participants in RDS studies are rewarded for successful recruitments, rational and self-interested participants could be expected to ignore contacts who are considered less likely to accept invitations (Lu et al., 2011). Such a combination of a group having a high probability of rejecting invitations and a high probability of being ignored violates the assumption that given someone's network, recruits are selected randomly. The method of coupon distribution itself, whether through social media or face-to-face, also influenced who was ultimately selected for recruitment (Helms et al., 2021).

Some studies explicitly tested for reciprocity with follow-up questions on the recruiter-recruit relationship, while others inferred non-random recruitment from sample imbalances (Gile, Johnston and Salganik, 2015). Although estimation methods often model recruitment as random, this assumption is frequently unsupported by empirical data (Liu et al., 2012; Wang et al., 2024a). In RDS, it appears peer-recruitment is not making simple random draws from a social network, and some individuals will be selected with greater probability than others. Such non-random recruitment undermines the assumption of equal inclusion probabilities.

3.4.4 External Validation and Estimator Performance

Several studies validated RDS-derived estimates by comparing them to external benchmarks, including census data, national surveys, hospital records, and administrative statistics (Heckathorn, 1997; Heckathorn et al., 2002; Lee et al., 2020; Xia et al., 2019; Ngwakongnwi et al., 2014). These comparisons showed mixed results as some studies reported close alignment, while others identified substantial deviations in key demographic variables (Jonsson et al., 2019; Perera and Ramanayake, 2019). Real social networks and the recruiting behaviour of people in those networks can hardly meet all the assumptions underlying the RDS estimators, and there is no consensus on one universally correct estimator (White et al., 2015). The RDS-II estimator (Volz-Heckathorn, 2008) was the most widely used within the studies reviewed and can (in principle) adjust for selection probabilities resulting from differential recruitment patterns and network sizes. The RDS-II estimator treats personal social network size (degree) as inclusion probabilities to weight values (Crawford, 2016; Wang et al., 2024a; Mills et al., 2014). But it still assumes random referral, within a fully connected population network, with with-replacement sampling (Lachowsky et al., 2016). While RDS-II improves upon naive means by accounting for unequal inclusion probabilities, it remains sensitive to degree misreporting and violations of referral randomness (Heckathorn, 2007; Wang et al., 2024a).

3.4.5 Self-Reported Degree and Weighting

Many studies report substantial variation in degree reporting, including over- and under-estimation, rounding errors, and inconsistencies across study waves (Goel and Salganik, 2010; McCreesh et al., 2013; Lee et al., 2017; Badowski et al., 2017; Liu et al., 2012). These inaccuracies can distort both prevalence and regression estimates, and can increase variance (Avery et al., 2021; Snyder et al., 2024; Jayaweera et al., 2023). While some studies adjusted for degree (network size) in their analyses, only a minority described their methods in detail (Johnston et al., 2016). Consequently, degree-based weighting introduces sensitivity to degree measurement error and may amplify noise.

3.5 Recruitment Monitoring, Diagnostic Tools and Adaptive Design

RDS relies on peer recruitment processes that are only partially controlled by researchers. As a result, population estimates depend on a set of assumptions about recruitment behaviour, network structure, and mixing between subgroups discussed above. The STROBE-RDS guidelines (White et al., 2015) emphasises that these assumptions should be treated as "*necessary approximations rather than binary conditions that are either met or violated"*. Monitoring recruitment and reporting diagnostics therefore serves to support claims that estimator assumptions are reasonable in a study context and to justify any adaptations made during fieldwork.

3.5.1 Recruitment Diagnostic Tools

White et al. (2015) do not present diagnostic tools as a single formal instrument. Instead, diagnostics are framed as empirical checks and monitoring outputs allow investigators to assess whether recruitment processes and RDS assumptions remain plausible as data collection progresses. Reporting diagnostics are not intended to prove that RDS "worked", but to allow readers to judge the reasonableness of the data-generating process.

Studies with simulation-based diagnostics assess how robust RDS estimation is to violations of underlying assumptions. Some studies used simplified mathematical models (such as Poisson or Binomial) to explore how different patterns of uneven peer recruitment might affect estimator performance (Perera and Ramanayake, 2019; Yan et al., 2020). These studies do not model real field behaviour directly but help clarify how sensitive RDS estimates are to unequal recruitment productivity, short vs long chains, dominance by highly productive recruiters ("super-seeds") and delayed convergence (or equilibrium). In these **simulations**, researchers specify rules for how many peers each participant recruits. A Binomial distribution reflects a situation where each participant has a limited number of opportunities to recruit, and each opportunity may or may not succeed. A Poisson distribution reflects a situation where recruitment events occur irregularly, most people recruit none or just a few, whereas a small number of recruiters recruit more.

Diagnostics based on observed recruitment data are the most directly relevant to monitoring fieldwork (Gile et al., 2015). These approaches rely on the link-tracing and network data collected during the study. A commonly reported tool is the recruitment '**transition' matrix**, which summarises who recruits whom across key characteristics such as gender, ethnicity, age, location, or reported behaviour. In simple terms, this matrix shows how often participants from one group recruit participants from the same or different groups. These matrices are used to estimate referral probabilities and to identify patterns of homophily, segregation, or 'bottlenecks' in recruitment (Heckathorn, 1997; Heckathorn, 2002; Heckathorn, 2007; Gile, Johnston and Salganik, 2015; Wang et al., 2024a; Yan et al., 2020). **Homophily statistics** are derived from these recruitment patterns and quantify the tendency for participants to recruit others similar to themselves. High homophily indicates strong within-group recruitment and limited mixing between groups, which can reinforce clustering and reduce sample diversity. Within-group recruitment biases have been reported across dimensions including gender, ethnicity, age and behaviour (Heckathorn, 1997; Heckathorn et al., 2002; Wang et al., 2024a; Wang et al., 2024b).

Recruitment chains are often reconstructed and visualised as **'recruitment tree' diagrams**, which show the branching structure, depth, and productivity of recruitment. These visualisations allow investigators to identify short or stalled chains, dominant "super-seeds" and imbalances in recruitment depth. Preserving the peer-referral structure in the data enables assessment of whether recruitment has penetrated sufficiently beyond initial seeds,

if global mixing has occurred, and whether recruitment appears random within social ties (Helms et al., 2021). Several studies have also used **'bottleneck plots'**, which show how estimates behave across recruitment chains, and how estimates of key characteristics evolve separately for each seed. In simple terms, these plots show whether different recruitment chains mix and converge on similar estimates or remain isolated, indicating restricted pathways between subgroups (Gile, Johnston and Salganik, 2015). Bottlenecks, clustering, and network fragmentation have been shown to influence estimator convergence (or equilibrium), reduce sample diversity, and increase design effects (Heckathorn, 2002; Heckathorn et al., 2002; Yan et al., 2020).

Seed-based diagnostics compare recruitment patterns and estimates across initial seeds. These include examination of chain length, recruitment speed, and whether recruitment equilibrates across chains. **'Convergence plots'** are descriptive and show how estimates of key characteristics change across successive recruitment waves to assess whether the overall sample composition stabilises as recruitment progresses (Gile, Johnston and Salganik, 2015). For example, the proportion of women by wave, the proportion reporting a behaviour by wave, or the mean network degree by wave. Equilibrium is generally defined as the stabilisation of key sample characteristics over successive recruitment waves, indicating reduced dependence on initial seeds (Jonsson et al., 2019; Liu et al., 2012). Evidence from multiple studies indicates that short chains often fail to reach demographic stability, leaving estimates vulnerable to seed-induced bias, whereas longer chains are associated with reduced seed dependence and improved diversity, with equilibrium commonly reached within four to six waves.

3.5.2 Adaptation and Variation during Data Collection

The STROBE-RDS guidelines recognise that modifications during RDS studies are common and may be appropriate, provided they are well documented and respect the functional and analytic assumptions of RDS. Changes to coupon distribution include altering the number of coupons per recruiter, reducing the probability that recruits receive coupons, or temporarily halting coupon distribution. Dynamic or adaptive referral limits have been used in some studies, starting with lower caps and increasing them based on recruitment diagnostics to maintain momentum and reduce the need for additional seeds (Wirtz et al., 2021; Truong et al., 2013; Helms et al., 2021). The addition or substitution of seeds is a common adaptive response to early recruitment failures, stalled chains or under-representation of key sub-groups. Substituting unproductive seeds or adding new ones can help jump-start recruitment and extend network penetration, particularly when diagnostics indicate insufficient recruitment depth or clustering (Zan et al., 2024). It may also be necessary to respond to a variety of recruitment problems which include addressing coupon selling, imposters, or duplicate recruits.

3.5.3 Diagnostics as Justification for Adaptation

Across the literature, adaptive recruitment strategies are sometimes linked to diagnostic evidence rather than treated as ad hoc protocol deviations. Continuous monitoring of recruitment speed, number of waves per seed, chain expansion or stalling, and convergence diagnostics have been used to inform decisions such as extending recruitment, increasing coupon numbers, or introducing new seeds. Monitoring recruitment progression is important because recruitment-chains are treated as key dimensions of RDS sample data, not merely a means to an end. Recent work has emphasised adaptive designs that monitor recruitment progression toward equilibrium using diagnostics. When equilibrium is not achieved, extending recruitment waves or introducing new seeds is warranted (Gile, Johnston and Salganik, 2015; Rudolph et al., 2024). Crawford (2016) demonstrates how the timing and modification of recruitment events and coupon use can affect whether recruitment becomes confined within subgroups or successfully traverses broader social networks.

3.6 RDS in Comparison to Alternative Sampling Methods

Respondent-driven sampling (RDS) was developed as a response to the inferential limitations of non-probability convenience approaches, such as snowball, quota or venue-based sampling (VBS), which were applied to study hidden or hard-to-reach populations (Heckathorn, 2011; Górny and Napierała, 2016). In venue-based quota sampling, researchers first identify specific physical or social locations (venues) where members of a target population are known to gather (e.g. workplaces, markets, churches, clubs). They then recruit respondents at those venues to fill predefined quotas (such as by nationality, age, or gender). Similarly, with snowball sampling, researchers begin with an initial convenience sample and ask participants to recruit others from their social networks.

RDS development attempts to address theoretical concerns about these convenience samples, by emphasising the role of the initial seeds in recruitment, gathering network size and link-tracing data to model the recruitment process itself. Explicit reporting on these distinctly RDS elements is essential to assessing the method versus snowball sampling (White et al., 2015). RDS introduces statistical estimators intended to adjust for network-based biases, with their validity depending on extent to which assumptions are met. Snowball sampling likewise relies on peer recruitment from an initial convenience sample but does not incorporate link-tracing or the formal estimation framework that RDS introduces to address seed dependence. Biases arising from network structure clustering and homophily in snowball sampling are therefore difficult to quantify (Heckathorn, 2011). Quota or venue-based approaches, whilst sometimes yielding high participant numbers, do not support chain-based recruitment. The method rarely uses “seeds”, instead it usually relies on the physical presence of participants at known locations.

In a study comparing venue-based quota sampling to RDS to recruit migrants, Górny and Napierała (2016) found that RDS produced higher-quality data (more complete, more precise), achieved higher sociometric diversity and better geographic dispersion. Venue-

based quota sampling's key limitation is that it systematically over-represents people who are present, visible and frequent at venues, while under-representing individuals who are more mobile, isolated, time-constrained, or deliberately avoiding public or institutional spaces. This makes 'coverage' depend on venue selection and prior assumptions about population composition (often missing for migrants) rather than true population structure (Górny and Napierała , 2016). *A priori* quota-based selection can appear balanced on paper but can be systematically misleading by hiding the mechanism of selection bias. In contrast, RDS can access previously hidden, unknown networks and provide data about the structure, the dependencies, and the sub-population bottlenecks. RDS's tendency to follow strong ties rather than administrative categories can make for a faithful reflection of where social density really is. Detailed comparison of RDS to other recruitment methods is presented in Table 7 below.

Table 7: Comparison of RDS to other recruitment approaches

Methodological Dimension	RDS	Alternative Sampling Methods	Comparative Robustness & Reliability
Risk of Duplicate Participation & Fraud	- Particularly online RDS due to coupon sharing, multiple enrolments. - Preventative measures include unique coupon codes, digital verification (IP, biometric), and careful manual tracking.	- Generally low in probability sampling, but higher in snowball or venue sampling due to weaker verification processes.	RDS has stronger fraud-prevention than typical snowball methods, yet weaker than traditional probability sampling.
Measurement & Reporting Errors	- High risk due to reliance on self-reported network sizes and behaviours.	- Lower risk in probability sampling methods using administrative verification, but similar high risk in snowball or convenience methods due to self-report reliance.	Similar to other referral-based methods but inferior to probability sampling.
Bias from Recruitment Patterns (Homophily, Seed Bias)	- Significant potential for bias due to network structures, homophily, seed selection, and recruitment stickiness. - Requires strategic seed selection, chain monitoring, and convergence checks.	- Snowball and venue-based methods have even higher bias potential due to less structured sampling. - Probability methods generally avoid these biases entirely due to randomisation.	Superior to snowball and venue methods but less robust compared to randomised probability sampling.

Technical & Administrative Issues	- High complexity (coupon management, chain tracking, online platforms). - Frequent administrative errors or coupon mismanagement. - Requires robust logistical and technical infrastructure.	- Probability sampling methods (particularly web-based survey) have simpler logistics and fewer technical complexities. - Snowball and venue methods have less formal management, resulting in different types of logistical errors.	May be more administratively challenging than random sampling methods.
Representativeness & Generalisability	- Representativeness can be achieved when properly managed. - Risks related to seed bias, recruitment chain limitations, and sample homogeneity. - Improved by formative research (for given target population), diverse seeds, and equilibrium monitoring.	- Probability sampling offers superior representativeness due to randomisation. - Snowball and venue-based methods typically show lower representativeness and higher biases.	RDS tends to be more representative than snowball sampling but less than probability sampling methods.
Statistical Estimation Reliability	- Dependent on validity of network assumptions, accurate reporting of degree, and unbiased recruitment. - Statistical estimators (**RDS-I, RDS-II, SS)** attempt to address biases.	- Reliability in probability sampling is higher with clear inclusion probabilities. - Snowball or convenience samples are not as reliable due to unclear inclusion probabilities and biases.	Proper RDS estimates are more likely to be reliable than those from snowball and convenient methods, yet less precise than those of probability sampling.
Cost and Feasibility	- Cost-effective for hard-to-reach populations due to peer-driven recruitment. - Feasibility challenged by network bottlenecks and seed productivity. - Online RDS can reduce costs but introduce fraud concerns.	- Probability methods are often expensive and logistically difficult for hidden populations. - Snowball sampling is less costly but introduces greater biases and lower reliability.	RDS balances cost-effectiveness and feasibility better than alternatives when studying hidden populations.

In summary, the main advantage of RDS relative to conventional snowball or venue-based sampling lies in its attempt to transform chain-referral recruitment into a framework with estimable properties. Despite these advantages, RDS remains less robust than traditional

probability-based sampling methods and its estimates are sensitive to violations of key assumptions regarding network connectivity, recruitment behaviour, and degree reporting. Its effectiveness is therefore conditional rather than assured and should be evaluated in light of how far these assumptions are plausibly met in specific study contexts.

4. Conclusion

Respondent-Driven Sampling (RDS) is a complex sampling method, the performance of which depends on interdependent recruitment practices, operational decisions, and analytical assumptions. Studies using RDS are frequently reported incompletely, limiting readers' ability to assess the extent to which core design, recruitment, and analytic features were implemented in ways consistent with the method's assumptions. Across the evidence, RDS demonstrates substantial flexibility and reach. When well-executed, it can penetrate the social networks of groups lacking coverage by traditional sample frames. However, the method has consistent vulnerabilities that may undermine its representativeness and inference when poorly executed and its assumptions are violated. There is also an internal contradiction within the evidence: because RDS relies on network structure to function, it may systematically under-represent very weakly networked individuals. Yet many of the target populations that the method is applied to are in fact defined by weak network embeddedness, marginality and social isolation.

4.1 Empirical Performance of RDS

The evidence shows that RDS has been effective in accessing hidden and diverse populations, including intravenous drug users (Heckathorn, 1997), precarious workers (Jonsson et al., 2019), and migrant populations across multiple international contexts (Górny and Napierała, 2016; Zanoni et al., 2023). The most common achieved sample size was 300–1,000 participants, reflecting a balance between feasibility and analytic robustness, while Web-RDS facilitated accelerated recruitment speed and scaling (Helms et al., 2021).

Recruitment productivity varies substantially across studies, with three dominant patterns observed: low productivity, moderate productivity (most common), and high productivity driven by "super-seeds." Moderate productivity patterns, typically characterised by referral caps of 3–5 coupons and recruitment chains of 4–9 waves, most often achieved equilibrium within 4–6 waves. Short chains (1–3 waves) were associated with high seed-dependence and a failure to reach equilibrium, while very long chains (10–20+ waves) allowed deeper network penetration but increased risks of clustering and compounded homophily (Heckathorn, 1997; Heckathorn et al., 2002; Wang et al., 2024a). Overall, recruitment patterns and chain length were strongly shaped by referral caps, incentive structures, seed characteristics and social network connectivity, which emphasises the need for ongoing monitoring using diagnostics such as recruitment trees, transition matrices and homophily indices.

4.2 Strategies Associated with Better Recruitment

Evidence consistently links stronger RDS performance to robust fieldwork and operational practices, including pre-survey outreach, formative research, and field site proximity to target populations (Abdul-Quader et al., 2006; Johnston et al., 2016). Formative assessment was central to evaluating whether key RDS assumptions were plausible, such as identifying population sub-groups, potential network bottlenecks, willing participants and logistical constraints. Most RDS studies rely on purposive, non-probability seeds selected through NGOs, community organisations, clinics and digital communities, often involving collaborators with specialised knowledge of a network (Gile, Johnston and Salganik, 2015; Lee et al., 2020; Wirtz et al., 2021). While fewer than 30 seeds were used in 64% of studies, reporting on the ratio of seeds to total sample size was inconsistent, limiting comparative assessment despite its suggested importance (Leonard, 2025).

Effective recruitment was associated with clear eligibility screening, prevention of repeat enrolment and the use of dual incentive schemes for participation and referral (Heckathorn, 1997; Shaghaghi, Bhopal and Sheikh, 2011; Lachowsky et al., 2016). Incentives aligned with local norms and delivery contexts (monetary or non-monetary) were linked to better recruitment outcomes, while fixed referral caps (most commonly three coupons) helped balance recruitment speed and dispersion. Adaptive strategies, including increasing referral caps or adding seeds in response to stalled recruitment, were also recommended.

4.3 Methodological Limitations: Where and Why Recruitment Breaks Down

Despite demonstrated strengths, the evidence documents recurring methodological limitations. Recruitment failure occurred in some populations due to poor seed selection, weak peer connectivity, inadequate incentives and inflexible designs that failed to adapt to stalled recruitment. Recruitment mode significantly shaped recruitment feasibility, data quality and representativeness, with face-to-face approaches offering stronger eligibility verification and rapport but at higher cost and with greater geographic clustering. Web-RDS increased scale and anonymity but introduced fraud risks and technical failures, which in some cases severely affected chain progression (Bauermeister et al., 2012; Helms et al., 2021; McGowan et al., 2023). The specific details of link-tracing systems are central to RDS validity but are frequently not reported. Technical failures, the misclassification of recruits as seeds, and incomplete or invalid link-tracing data can also undermine recruitment-chain reconstruction and valid network estimation (Wang et al., 2024b).

More fundamentally, the evidence highlights that structural limitations of RDS are related to network clustering and homophily. RDS systematically reproduces groups with larger, denser, and more reciprocal networks, while weakly networked, dispersed, or socially isolated sub-groups are less likely to be recruited. Multiple studies reported short chains and high seed

failure among highly marginalised populations (Wirtz et al., 2021; Inghels et al., 2021; Pham et al., 2023). These patterns raise unresolved questions about the extent to which RDS can capture socially isolated individuals within hidden populations.

Finally, estimator performance was found to be sensitive to violations of core assumptions. Evidence of degree misreporting, differential recruitment success, non-random recruitment, and homophily challenges the adequacy of RDS-II weighting, as inclusion probabilities may no longer be well approximated by degree alone (Heckathorn, 2002; Heckathorn, 2007; Mills et al., 2014; Wang et al., 2024a).

5. Recommendations for Practitioners

Seed Selection: Strategic and diverse seed selection is essential to minimise initial sampling bias and enhances population coverage. Studies have demonstrated that diverse seed selection supports equilibrium and reduces seed-related bias over time. Seeds should vary across socio-demographic and behavioural characteristics applicable to the study aims or target groups of interest. Utilisation of local recruitment collaborators such as NGOs, community organisations, peer networks and digital platforms is beneficial to seed selection and to accessing social networks. Formative research methods like ethnographic mapping and focus groups help identify well-positioned, willing seeds who are validated as eligible. Eligibility screening ensures that seeds belong to a target population and are commonly established via questionnaires or in-person checks (depending on criteria specific to target populations).

Incentive Design to Sustain Recruitment: Effective incentive structures in RDS studies are tailored to the local context. Dual incentives for participation and peer recruitment is a defining feature of RDS. Monetary incentives are common for online settings, while non-monetary items are used in low-resource environments. Targeted-incentives (conditional on sub-group membership) can increase recruitment among under-sampled populations, can target weakly networked individuals and help ensure participants achieve specific sampling aims, such as recruiting a minimum number of peers.

Sample Size: Smaller studies (<300 participants) are valuable for focused inquiries but may offer limited generalisability and statistical power. Medium-sized samples (300–1,000 participants) are most common in RDS research, balancing feasibility and analytical robustness.

Chain Length: Recruitment chains shorter than 4 waves may not mitigate seed bias. Moderate-length chains (4–9 waves) are sufficient with equilibrium most often achieved within 4–6 waves to reduce seed dependency and improve demographic diversity.

Manage Recruitment Per Seed: Optimal recruitment outcomes are achieved with referral caps of three to five coupons per participant, resulting in stable chains and increased demographic diversity. Recruitment below three peers often necessitates new seeds, while more than ten peers can cause "super-seed" effects, distorting the sample. Lower coupon limits produce longer, more evenly distributed recruitment chains, enhancing sample integrity.

Adaptive Recruitment and Diagnostics: Real-time monitoring and early replacement of non-productive seeds, or the truncation of dominant chains (from super seeds), can address homophily and network clustering. Adjusting referral limits to maintain chain progression can ensure broader sample representation and expand reach to sub-groups.

Link-Tracing: Comprehensive documentation of link-tracing methodology is crucial, specifying tools, platforms, and unique identifiers used. Quality control measures should be implemented to mitigate procedural errors and data quality risks.

Personal Network Questions: Personal network questions should be included in the survey to identify degree and non-random preferential recruitment. They should capture social network size, number of contacts in the target population and recruiter-recruit relationship details.

Standardised Reporting: RDS studies often fail to report recruitment processes, coupon details, seed handling, network size measurement, estimator choice and diagnostics. This makes it difficult to determine whether the study truly followed RDS principles, whether estimator assumptions are plausible and how recruitment dynamics shaped the sample. Standardised reporting of critical RDS operational indicators will improve the transparency, comparability and utility of RDS research. These reports should detail: the number and origin of seeds, the productivity of each seed (i.e., number of recruits per seed), achieved sample, incentive types, the length and sustainability of recruitment chains, equilibrium diagnostics, link-tracing systems, cost details and fraud control. The Strengthening the Reporting of Observational Studies in Epidemiology for RDS Studies (STROBE-RDS) guidelines are an excellent place to start and necessitate further adoption (White et al., 2015).

6. Evidence Gaps and Future Research

This review demonstrates that RDS has been widely implemented and extensively studied, yet several important gaps remain uncovered by our evidence.

Limited Comparative Evaluation of Sampling Performance

The evidence reviewed has limited systematic comparison between RDS and alternative sampling approaches applied to the same target populations or compared to populations with known parameters. While studies differentially report recruitment reach, sample size, and

equilibrium diagnostics, these indicators were insufficiently evaluated relative to other feasible sampling strategies. The evidence lacked comparative studies that explicitly assess trade-offs between samples achieved by RDS and alternative methods in terms of coverage, bias, feasibility, costs and operational burden.

Incomplete Explanation of Recruitment Dynamics

The evidence documents recurring recruitment patterns such as low productivity chains, equilibrium-achieving recruitment, or dominant "super-seed" networks, but lacks empirical validation to determine whether the recruitment structures in RDS achieved samples correspond to plausible underlying social structures and known population segmentation. A focus on why certain sub-populations generate short or fragile chains in RDS (not just that they do) is lacking. Future RDS research would benefit from comparing achieved RDS chains to known network features and independent evidence about network connectivity, sub-group structure and social barriers. This would help address a core ambiguity about whether recruitment failure or clustering effects in RDS reflect poor design or genuine network structure.

Absence of Diagnostics for Exclusion and Weakly Networked Populations

An implicit concern in the evidence is the systematic under-representation of weakly networked or socially isolated individuals. Existing diagnostics focus on convergence and equilibrium among those recruited but offer no empirical strategy for identifying groups who remain unreached, or to estimate the size and characteristics of excluded subgroups. Future research may benefit from the development of diagnostics that address exclusion as well as inclusion. Hybrid designs that use RDS for well-networked groups and targeted outreach (or alternatives) for weakly networked groups may be necessary to evaluate how far RDS can adequately represent those characterised by marginality, dispersion, or limited peer ties. This would help determine clearer criteria regarding sampling method goodness-of-fit with a population, and define when RDS should not be used, not just how to optimise it.

Diagnostic and Estimator Performance Under Realistic Violations

The review focused on RDS recruitment and sampling, rather than estimators, but it became clear that the estimation and analysis of network structure are what distinguishes RDS as a method, not just the overall achieved sample. The evidence demonstrates that analysis, diagnostic monitoring and adaptation during fieldwork are beneficial to RDS sample outcomes. This, therefore, suggests that the separation of fieldwork and statistical analysis cannot be considered distinct within RDS, as intended at the beginning of this review. Future work should shift from assumption-based validation toward robustness-based evaluation, examining how diagnostic tools and estimators behave under realistic combinations of aspects such as degree misreporting, preferential recruitment, clustering and homophily. This would support clearer guidance on adaptive design choices, or when weighting improves inference and when it may exacerbate bias.

Recruitment Mode as a Methodological Modifier

The review shows that recruitment modes, such as face-to-face, online, or mixed-mode, substantially alters recruitment dynamics, verification processes, link-tracing integrity, fraud risk, and participant engagement. Future research should explicitly explore recruitment mode as part of the RDS mechanism itself, assessing whether digital adaptations preserve or transform the core assumptions (reciprocity, random referral, degree-based inclusion) underpinning RDS. This would clarify whether Web-RDS constitutes a methodological variant or a substantively distinct approach.

Improved Digital Infrastructure

The integration of digital tools into RDS, including electronic referrals, social media engagement, and Web-RDS platforms, remains under-researched and studies emphasise the need for the development of fraud prevention measures in online settings (Helms et al., 2021; Truong et al., 2013; McGowan et al., 2023; Bauermeister et al., 2012; Hildebrand et al., 2015). Additionally, more attention may be needed to address disparities in digital participation (such as socio-economic and cultural).

RDS as a Boost to Probability-based Surveys

Embedding RDS within existing probability-based survey designs by using seeds selected probabilistically from a survey sample, may allow RDS to function as a targeted boosting mechanism rather than an independent inferential framework. Because the initial sample is drawn from a known sampling frame, the probability survey provides a benchmark population structure against which RDS recruitment patterns and resulting estimates can be evaluated. Further exploration of this hybrid approach may allow expansion of sample sizes for networked subpopulations with declining response rates while preserving design-based comparability. This will facilitate the empirical assessment of bias, practical feasibility, and efficiency of RDS-derived estimates relative to established probability-based survey estimates.

Standardised Reporting

Finally, the evidence base is constrained by inconsistent reporting of key methodological features, including link-tracing systems, seed management, incentive design, recruitment diagnostics and chain structure. These omissions limit cumulative learning, hinder secondary analysis and improved transparency is essential for distinguishing between methodological and population-specific limitations. Full adoption of the STROBE-RDS (White et al., 2015) guidelines is an excellent place to start, and any further work on additional study-quality indicators would be beneficial.

Bibliography

Abdul-Quader, A.S., Heckathorn, D.D., Sabin, K. and Saidel, T. (2006). Implementation and Analysis of Respondent Driven Sampling: Lessons Learned from the Field. *Journal of Urban Health*, 83(S1), pp.1–5. doi:https://doi.org/10.1007/s11524-006-9108-8.

Avery, L., Rotondi, N., McKnight, C., Firestone, M., Smylie, J. and Rotondi, M. (2019). Unweighted regression models perform better than weighted regression techniques for respondent-driven sampling data: results from a simulation study. *BMC Medical Research Methodology*, 19(1). doi:https://doi.org/10.1186/s12874-019-0842-5.

Avery, L.M., Macpherson, A., Flicker, S. and Rotondi, M.A. (2021). A review of reported network degree and recruitment characteristics in respondent driven sampling implications for applied researchers and methodologists. *PLoS ONE*, 16(4), pp.e0249074–e0249074. doi:https://doi.org/10.1371/journal.pone.0249074.

Badowski, G., Somera, L.P., Simsiman, B., Lee, H.-R., Cassel, K., Yamanaka, A. and Ren, J. (2017). The efficacy of respondent-driven sampling for the health assessment of minority populations. *Cancer Epidemiology*, 50(Pt B), pp.214–220. doi:https://doi.org/10.1016/j.canep.2017.07.006.

Bauermeister, J.A., Zimmerman, M.A., Johns, M.M., Glowacki, P., Stoddard, S. and Volz, E. (2012). 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. *Journal of Studies on Alcohol and Drugs*, 73(5), pp.834–838. doi:https://doi.org/10.15288/jsad.2012.73.834.

Bengtsson, L., Lu, X., Nguyen, Q.C., Camitz, M., Hoang, N.L., Nguyen, T.A., Liljeros, F. and Thorson, A. (2012). Implementation of Web-Based Respondent-Driven Sampling among Men Who Have Sex with Men in Vietnam. *PLoS ONE*, 7(11), p.e49417. doi:https://doi.org/10.1371/journal.pone.0049417.

Bhatta, M., Majumdar, A., Banerjee, S., Ghosh, P., Biswas, S. and Dutta, S. (2023). Accumulation of Biological and Behavioral Data of Female Sex Workers Using Respondent-Driven Sampling: Protocol for a Systematic Review. *JMIR Research Protocols*, [online] 12, p.e43722. doi:https://doi.org/10.2196/43722.

Bostwick, W.B., Hughes, T.L. and Everett, B. (2015). Health Behavior, Status, and Outcomes Among a Community-Based Sample of Lesbian and Bisexual Women. *LGBT Health*, 2(2), pp.121–126. doi:https://doi.org/10.1089/lgbt.2014.0074.

Crawford, F.W. (2016). The Graphical Structure of Respondent-driven Sampling. *Sociological Methodology*, 46(1), pp.187–211. doi:https://doi.org/10.1177/0081175016641713.

Diexer, S., Teslya, A., Buskens, V., Matser, A., Stein, M. and Kretzschmar, M.E. (2023). Improving web-based respondent-driven sampling performance among men who have sex

with men in the Netherlands. *PLOS Digital Health*, 2(2), p.e0000192. doi:https://doi.org/10.1371/journal.pdig.0000192.

Erickson, B.H. (1979). Some Problems of Inference from Chain Data. *Sociological Methodology*, 10, p.276. doi:https://doi.org/10.2307/270774.

Ferrer-Rosende, P., Esteve-Matalí, L., Stuardo-Ávila, V., Fuentes-Alburquenque, M. and Navarro-Giné, A. (2024). Comparison between RDS and WebRDS for epidemiological and social studies. *Advance*. [online] doi:https://doi.org/10.31124/advance.172545789.92360927/v1.

FerrerRosende, P., FeijooCid, M., Isabel, F.M., SalasNicás, S., StuardoÁvila, V. and NavarroGiné, A. (2023). Implementation of webbased respondent driven sampling in epidemiological studies. *BMC Medical Research Methodology*, [online] 23(1), p.217. doi:https://doi.org/10.1186/s1287402302042z.

FISHER, J.C. and MERLI, M.G. (2014). Stickiness of respondent-driven sampling recruitment chains. *Network Science*, 2(2), pp.298–301. doi:https://doi.org/10.1017/nws.2014.16.

Fordjuoh, J., Dolezal, C., Bhengu, N., Harrison, A.D., Exner, T.M., HanassHancock, J. and Hoffman, S. (2024). PeertoPeer Chain Recruitment for Enrolling Young South African Women into an HIV Preexposure Prophylaxis (PrEP) Intervention Study: How Did It Perform? *AIDS and Behavior*, [online] 28(5), pp.1782–1794. doi:https://doi.org/10.1007/s10461023042560.

Frere-Smith, T., Luthra, R. and Platt, L. (2014). *Sampling recently arrived immigrants in the UK: Exploring the effectiveness of respondent driven sampling*. [online] University of Essex, Institute for Social and Economic Research (ISER). Available at: https://hdl.handle.net/10419/123811.

Gile, K.J. (2011). Improved Inference for Respondent-Driven Sampling Data With Application to HIV Prevalence Estimation. *Journal of the American Statistical Association*, 106(493), pp.135–146. doi:https://doi.org/10.1198/jasa.2011.ap09475.

Gile, K.J. and Handcock, M.S. (2010). 7. Respondent-Driven Sampling: An Assessment of Current Methodology. *Sociological Methodology*, 40(1), pp.285–327. doi:https://doi.org/10.1111/j.1467-9531.2010.01223.x.

Gile, K.J., Johnston, L.R. and Salganik, M.J. (2015). Diagnostics for Respondent-Driven Sampling. *Journal of the Royal Statistical Society*, 178(1), pp.241–269. doi:https://doi.org/10.1111/rssa.12059.

Goel, S. and Salganik, M.J. (2009). Respondent-driven sampling as Markov chain Monte Carlo. *Statistics in Medicine*, 28(17), pp.2202–2229. doi:https://doi.org/10.1002/sim.3613.

Goel, S. and Salganik, M.J. (2010). Assessing respondent-driven sampling. *Proceedings of the National Academy of Sciences*, 107(15), pp.6743–6747. doi:https://doi.org/10.1073/pnas.1000261107.

Górny, A. and Napierała, J. (2016). Comparing the effectiveness of respondent-driven sampling and quota sampling in migration research. *International Journal of Social Research Methodology*, 19(6), pp.645–661. doi:https://doi.org/10.1080/13645579.2015.1077614.

Handcock, M.S. and Gile, K.J. (2010). Modeling social networks from sampled data. *The Annals of Applied Statistics*, 4(1), pp.5–25. doi:https://doi.org/10.1214/08-aoas221.

Hartfiel, D.J. and E. Seneta (1994). On the theory of Markov set-chains. *Advances in Applied Probability*, 26(4), pp.947–964. doi:https://doi.org/10.2307/1427899.

Heckathorn, D.D. (1997). Respondent-Driven Sampling: A New Approach to the Study of Hidden Populations. *Social Problems*, 44(2), pp.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), pp.11–34. doi:https://doi.org/10.1525/sp.2002.49.1.11.

Heckathorn, D.D. (2007). 6. Extensions of Respondent-Driven Sampling: Analyzing Continuous Variables and Controlling for Differential Recruitment. *Sociological Methodology*, 37(1), pp.151–208. doi:https://doi.org/10.1111/j.1467-9531.2007.00188.x.

Heckathorn, D.D. (2011). Snowball versus Respondent-Driven Sampling. *Sociological Methodology*, 41(1), pp.355–366. doi:https://doi.org/10.1111/j.1467-9531.2011.01244.x.

Heckathorn, D.D., Semaan, S., Broadhead, R.S. and Hughes, J.J. (2002). Extensions of Respondent-Driven Sampling: A New Approach to the Study of Injection Drug Users Aged 18–25. *AIDS and Behavior*, 6(1), pp.55–67. doi:https://doi.org/10.1023/a:1014528612685.

Helms, Y.B., Hamdiui, N., Kretzschmar, M.E.E., Rocha, L.E.C., van Steenbergen, J.E., Bengtsson, L., Thorson, A., Timen, A. and Stein, M.L. (2021). Applications and Recruitment Performance of Web-Based Respondent-Driven Sampling: Scoping Review. *Journal of Medical Internet Research*, 23(1), p.e17564. doi:https://doi.org/10.2196/17564.

Hildebrand, J., Burns, S., Zhao, Y., Lobo, R., Howat, P., Allsop, S. and Maycock, B. (2015). Potential and Challenges in Collecting Social and Behavioral Data on Adolescent Alcohol Norms: Comparing Respondent-Driven Sampling and Web-Based Respondent-Driven Sampling. *Journal of Medical Internet Research*, 17(12), p.e285. doi:https://doi.org/10.2196/jmir.4762.

Hildebrand, J., Maycock, B., Howat, P., Burns, S., Allsop, S., Dhaliwal, S. and Lobo, R. (2013). Investigation of alcohol-related social norms among youth aged 14–17 years in Perth, Western Australia: protocol for a respondent-driven sampling study: Table 1. *BMJ Open*, [online] 3(10), p.e003870. doi:https://doi.org/10.1136/bmjopen-2013-003870.

Inghels, M., Kouassi, A.K., Niangoran, S., Bekelynck, A., Carillon, S., Sika, L., Koné, M., Danel, C., Annabel, Larmarange, J., ASSOUMOU, N., BEKELYNCK, A., DANEL, C., DOUMBIA, M., KONE, M., KOUADIO, A., Kouassi, A.K., NIANGORAN, S., OUANTCHI, H. and SIKA, L. (2021). Telephone peer recruitment and interviewing during a respondentdriven sampling (RDS) survey: feasibility and field experience from the first phonebased RDS survey among men who have sex with men in Côte d'Ivoire. *BMC Medical Research Methodology*, [online] 21(1), p.25. doi:https://doi.org/10.1186/s1287402101208x.

Jayaweera, R.T., Gerdts, C., Wesson, P., Motana, R., Muñoz, I., Bessenaar, T., McFarland, W. and Ahern, J. (2023). Respondent- Driven Sampling to Assess Cumulative Lifetime Incidence of Abortion in Soweto, South Africa: A Methodological Assessment. *American Journal of Epidemiology*, [online] 192(7), pp.1081–1092. doi:https://doi.org/10.1093/aje/kwad074.

Johnston, L.G., Chen, Y., SilvaSantisteban, A. and Fisher, R.H. (2013). An Empirical Examination of Respondent Driven Sampling Design Effects Among HIV Risk Groups from Studies Conducted Around the World. *AIDS and Behavior*, [online] 17(6), pp.2202–2210. doi:https://doi.org/10.1007/s1046101203948.

Johnston, L.G., Hakim, A.J., Dittrich, S., Burnett, J., Kim, E. and White, R.G. (2016). A Systematic Review of Published Respondent-Driven Sampling Surveys Collecting Behavioral and Biologic Data. *AIDS and Behavior*, 20(8), pp.1754–1776. doi:https://doi.org/10.1007/s10461-016-1346-5.

Jonsson, J., Stein, M., Johansson, G., Bodin, T. and Strömdahl, S. (2019). A performance assessment of web-based respondent driven sampling among workers with precarious employment in Sweden. *PLOS ONE*, 14(1), p.e0210183. doi:https://doi.org/10.1371/journal.pone.0210183.

Koyuncu, A., Ishizumi, A., Daniels, D., Jalloh, M.F., Wallace, A.S. and Prybylski, D. (2023). The Use of Adaptive Sampling to Reach Disadvantaged Populations for Immunization Programs and Assessments: A Systematic Review. *Vaccines*, 11(2). doi:https://doi.org/10.3390/vaccines11020424.

Kubal, A., Shvab, I. and Wojtynska, A. (2014). Initiation of the RDS Recruitment Process. *Applying Respondent Driven Sampling to Migrant Populations*, pp.37–48. doi:https://doi.org/10.1057/9781137363619.0011.

Lachowsky, N.J., Sorge, J.T., Raymond, H.F., Cui, Z., Sereda, P., Rich, A., Roth, E.A., Hogg, R.S. and 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). doi:https://doi.org/10.1186/s12874-016-0258-4.

Lee, R. (2024). *Respondent driven sampling : feasibility, validation and its potential use for estimating health indicators for large general populations*. [online] Scholars Archive.

Available at: https://scholarsarchive.library.albany.edu/legacy-etd/206/?utm_source=scholarsarchive.library.albany.edu%2Flegacy-etd%2F206&utm_medium=PDF&utm_campaign=PDFCoverPages [Accessed 18 Mar. 2025].

Lee, S. (2009). Understanding Respondent Driven Sampling from a Total Survey Error Perspective. *Survey Practice*, 2(6), pp.1–6. doi:https://doi.org/10.29115/sp-2009-0029.

Lee, S., Ong, A.R., Chen, C. and Elliott, M. (2020). Respondent Driven Sampling for Immigrant Populations: A Health Survey of Foreign-Born Korean Americans. *Journal of Immigrant and Minority Health*, 23(4), pp.784–792. doi:https://doi.org/10.1007/s10903-020-01077-4.

Lee, S., Ong, A.R. and Elliott, M. (2020). Exploring Mechanisms of Recruitment and Recruitment Cooperation in Respondent Driven Sampling. *Journal of Official Statistics*, 36(2), pp.339–360. doi:https://doi.org/10.2478/jos-2020-0018.

Lee, S., Suzer-Gurtekin, T., Wagner, J. and Valliant, R. (2017). Total Survey Error and Respondent Driven Sampling: Focus on Nonresponse and Measurement Errors in the Recruitment Process and the Network Size Reports and Implications for Inferences. *Journal of Official Statistics*, 33(2), pp.335–366. doi:https://doi.org/10.1515/jos-2017-0017.

Leonard, M.M. (2025). Conducting Respondent-Driven Sampling with Ethnic Minority Populations: The State of the Field. *Survey Practice*, 19(S1), pp.1–19. doi:https://doi.org/10.29115/sp-2024-0020.

Liu, H., Li, J., Ha, T. and Li, J. (2012). Assessment of Random Recruitment Assumption in Respondent-Driven Sampling in Egocentric Network Data*. *Social Networking*, 01(02), pp.13–21. doi:https://doi.org/10.4236/sn.2012.12002.

Lu, X., Bengtsson, L., Britton, T., Camitz, M., Kim, B.J., Thorson, A. and Liljeros, F. (2011). The sensitivity of respondent-driven sampling. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 175(1), pp.191–216. doi:https://doi.org/10.1111/j.1467-985x.2011.00711.x.

Luis, Thorson, A.E., Lambiotte, R. and Liljeros, F. (2017). RespondentDriven Sampling Bias Induced by Community Structure and Response Rates in Social Networks. *J R Stat Soc Ser A Stat Soc*, [online] 180(1), pp.99–118. doi:https://doi.org/10.1111/rssa.12180.

Martin, K., Johnson, T.P. and Hughes, T.L. (2015). USING RESPONDENT DRIVEN SAMPLING TO RECRUIT SEXUAL MINORITY WOMEN. *Survey practice*, [online] 8(2), p.273. Available at: https://pmc.ncbi.nlm.nih.gov/articles/PMC5066809/ [Accessed 18 Mar. 2025].

McCreesh, N., Copas, A., Seeley, J., Johnston, L.G., Sonnenberg, P., Hayes, R.J., Frost, S.D.W. and White, R.G. (2013). Respondent driven sampling: determinants of recruitment and a method to improve point estimation. *PloS One*, [online] 8(10), p.e78402. doi:https://doi.org/10.1371/journal.pone.0078402.

McCreesh, N., Frost, S., Seeley, J., Katongole, J., Tarsh, M.N., Ndunguse, R., Jichi, F., Lunel, N.L., Maher, D., Johnston, L.G., Sonnenberg, P., Copas, A.J., Hayes, R.J. and White, R.G. (2012). Evaluation of Respondent-Driven Sampling. *Epidemiology (Cambridge, Mass.)*, [online] 23(1), pp.138–147. doi:https://doi.org/10.1097/EDE.0b013e31823ac17c.

McFall, A.M., Solomon, S.S., Lau, B., Latkin, C., Srikrishnan, Aylur K, Anand, S., Vasudevan, C.K., Kumar, M.S., Lucas, G.M. and Mehta, H. (2023). Respondentdriven sampling is more efficient than facilitybased strategies at identifying undiagnosed people who inject drugs living with HIV in India. *Drug and Alcohol Dependence*, [online] 249, p.110834. doi:https://doi.org/10.1016/j.drugalcdep.2023.110834.

McGowan, C.R., Ekoriko, P., Alhaffar, M., Cassidy-Seyoum, S., Whitbread, S., Rogers, P., Bell, L. and Checchi, F. (2023). Design and implementation of a web-based, respondent-driven sampling solution. *BMC Medical Informatics and Decision Making*, 23(1). doi:https://doi.org/10.1186/s12911-023-02217-0.

Michaels, S., Pineau, V., Reimer, B., Ganesh, N. and Dennis, J.M. (2019). Test of a Hybrid Method of Sampling the LGBT Population: Web Respondent Driven Sampling with Seeds from a Probability Sample. *Journal of Official Statistics*, 35(4), pp.731–752. doi:https://doi.org/10.2478/jos-2019-0031.

Middleton, D., Drabble, L.A., Krug, D., Karriker-Jaffe, K.J., Mericle, A.A., Hughes, T.L., Iachan, R. and Trocki, K.F. (2021). Challenges of virtual RDS for recruitment of sexual minority women for a behavioral health study. *Journal of Survey Statistics and Methodology*, [online] 10(2), pp.466–488. doi:https://doi.org/10.1093/jssam/smab039.

Mills, H.L., Johnson, S., Hickman, M., Jones, N.S. and Colijn, C. (2014). Errors in reported degrees and respondent driven sampling: Implications for bias. *Drug and Alcohol Dependence*, 142, pp.120–126. doi:https://doi.org/10.1016/j.drugalcdep.2014.06.015.

Ngwakongnwi, E., King-Shier, K.M., Hemmelgarn, B.R., Musto, R. and Quan, H. (2014). Comparison of sampling methods for hard-to-reach francophone populations: yield and adequacy of advertisement and respondent-driven sampling. *Open Medicine*, [online] 8(4), p.e120. Available at: https://pmc.ncbi.nlm.nih.gov/articles/PMC4242789/ [Accessed 28 Nov. 2024].

Ong, A. (2022). *Measuring Network Size and Recruitment Productivity in Respondent Driven Sampling*. [online] University of Michigan Library. Available at: https://deepblue.lib.umich.edu/bitstream/handle/2027.42/175650/aireneo_1.pdf?sequence=1&isAllowed=y [Accessed 18 Mar. 2025].

Ouma, S., Abbo, C., Natala, N., McCoy, M. and Kroupina, M. (2023). Victimization among adolescents of female sex workers: findings from the children of at-risk parents (CARP) study in Uganda. *BMC Pediatrics*, 23(1). doi:https://doi.org/10.1186/s12887-023-04131-w.

Perera, A.P.M. and Ramanayake, K.P.A. (2019). *ASSESSING THE EFFECTS OF RESPONDENT DRIVEN SAMPLING ESTIMATORS ON POPULATION CHARACTERISTICS*. [online] Asia International Conference on Multidisciplinary Research. Available at: https://www.academia.edu/40036649/ASSESSING_THE_EFFECTS_OF_RESPONDENT_DRIVEN_SAMPLING_ESTIMATORS_ON_POPULATION_CHARACTERISTICS [Accessed 5 Jun. 2025].

Pham, P.N., Johnston, L.G., Keegan, K., Wei, C. and Vinck, P. (2023). Innovative Strategies for Remotely Sampling HardtoReach Populations: Assessing Phone Versus Internet RespondentDriven Sampling Approaches Among Venezuelan Refugees and Migrants in Colombia. *Am J Epidemiol*, [online] 192(10), pp.1613–1623. doi:https://doi.org/10.1093/aje/kwad116.

Pravosud, V. and Andreeva, T. (2016). Respondent-driven sampling: characteristics associated with productive and non-productive recruitment. *Tobacco Control and Public Health in Eastern Europe*, [online] 6(1). doi:https://doi.org/10.6084/m9.figshare.4495910.v1.

Rajani, S., Cutrona, S., Erik, V., Rao, S. and Thomas, H. (2013). Web-based Peer-Driven Chain Referrals for Smoking Cessation. *Studies in health technology and informatics*, 192. doi:https://doi.org/10.3233/978-1-61499-289-9-357.

Rudolph, A.E., Fuller, C.M. and Latkin, C. (2013). The Importance of Measuring and Accounting for Potential Biases in Respondent-Driven Samples. *AIDS and Behavior*, 17(6), pp.2244–2252. doi:https://doi.org/10.1007/s10461-013-0451-y.

Rudolph, A. E., Nance, R. M., Bobashev, G., Brook, D., Akhtar, W., Cook, R., Cooper, H. L., Friedmann, P. D., Frost, S. D. W., Go, V. F., Jenkins, W. D., Korthuis, P. T., Miller, W. C., Pho, M. T., Ruderman, S. A., Seal, D. W., Stopka, T. J., Westergaard, R. P., Young, A. M., ... Delaney, J. A. C. (2024). Evaluation of respondent-driven sampling in seven studies of people who use drugs from rural populations: findings from the Rural Opioid Initiative. *BMC Medical Research Methodology*, 24(1), Article 94. https://doi.org/10.1186/s12874-024-02206-5

Salganik, M.J. (2006). Variance Estimation, Design Effects, and Sample Size Calculations for Respondent-Driven Sampling. *Journal of Urban Health*, [online] 83(S1), pp.98–112. doi:https://doi.org/10.1007/s11524-006-9106-x.

Salganik, M.J. and Heckathorn, D.D. (2004). 5. Sampling and Estimation in Hidden Populations Using Respondent-Driven Sampling. *Sociological Methodology*, 34(1), pp.193–240. doi:https://doi.org/10.1111/j.0081-1750.2004.00152.x.

Schonlau, M. and Liebau, E. (2012). Respondent-Driven Sampling. *The Stata Journal: Promoting communications on statistics and Stata*, 12(1), pp.72–93. doi:https://doi.org/10.1177/1536867x1201200106.

Schonlau, M., Weidmer, B. and Kapteyn, A. (2014). Recruiting an Internet Panel Using Respondent-Driven Sampling. *Journal of Official Statistics*, 30(2), pp.291–310. doi:https://doi.org/10.2478/jos-2014-0018.

Shaghaghi, A., Bhopal, R.S. and Sheikh, A. (2011). Approaches to Recruiting 'Hard-To-Reach' Populations into Re-search: A Review of the Literature. *Health promotion perspectives*, 1(2), pp.86–94. doi:https://doi.org/10.5681/hpp.2011.009.

Sirianni, A.D., Cameron, C.J., Shi, Y. and Heckathorn, D.D. (2021). Bias decomposition and estimator performance in respondent-driven sampling. *Social Networks*, 64, pp.109–121. doi:https://doi.org/10.1016/j.socnet.2020.08.002.

Snijders, T.A.B., van de Bunt, G.G. and Steglich, C.E.G. (2010). Introduction to stochastic actor-based models for network dynamics. *Social Networks*, 32(1), pp.44–60. doi:https://doi.org/10.1016/j.socnet.2009.02.004.

Snyder, M., McConkey, S., Brar, R., Anilniliak, J., Bourgeois, C., Dokis, B., Hardy, M., Joseph, S., Kilabuk, A., Mattina, J.-A., McKnight, C. and Smylie, J. (2024). Unmasking population undercounts, health inequities, and health service access barriers across Indigenous populations in urban Ontario. *Canadian Journal of Public Health*, [online] 115, pp.209–226. doi:https://doi.org/10.17269/s41997-024-00957-8.

Sosenko, F.L. and Bramley, G. (2022). Smartphone-based Respondent Driven Sampling (RDS): A methodological advance in surveying small or 'hard-to-reach' populations. *PLOS ONE*, 17(7), p.e0270673. doi:https://doi.org/10.1371/journal.pone.0270673.

Thompson, S. (2013). *Dynamic spatial and network sampling*.

Thompson, S. (2014). Link-tracing and respondent-driven sampling. In: R. Tourangeau, B. Edwards, T.P. Johnson, K.M. Wolter and N. Bates, eds., *Hard-to-Survey Populations*. [online] Cambridge University Press, pp.503–516. Available at: https://doi.org/10.1017/CBO9781139381635.

Thompson, S.K. (2017). Adaptive and Network Sampling for Inference and Interventions in Changing Populations. *J Surv Stat Methodol*, [online] 5(1), pp.1–21. doi:https://doi.org/10.1093/jssam/smw035.

Truong, H.-H.M., Fatch, R., Ogolla, D., Otieno, B., Amboka, S., Kadede, K., Cohen, C.R., Bukusi, E.A. and Guzé, M.A. (2022). Respondent-driven sampling to recruit adolescents in Kenya. *Annals of Epidemiology*, [online] 78, pp.68–73. doi:https://doi.org/10.1016/j.annepidem.2022.12.003.

Truong, H.M., Grasso, M., Chen, Y., Kellogg, T.A., Robertson, T., Curotto, A., Steward, W.T. and McFarland, W. (2013). Balancing Theory and Practice in Respondent-Driven Sampling: A Case Study of Innovations Developed to Overcome Recruitment Challenges. *PLOS ONE*, 8(8), pp.e70344–e70344. doi:https://doi.org/10.1371/journal.pone.0070344.

Tyldum, G. and Johnston, L.G. (2014). *Applying Respondent Driven Sampling to Migrant Populations*. *Palgrave Macmillan UK eBooks*. Palgrave Macmillan. doi:https://doi.org/10.1057/9781137363619.

Verdery, A.M., Mouw, T., Bauldry, S. and Mucha, P.J. (2015). Network Structure and Biased Variance Estimation in Respondent Driven Sampling. *PLOS ONE*, [online] 10(12), pp.e0145296-. Available at: https://doi.org/10.1371/journal.pone.0145296.

Vincent, K. and Thompson, S. (2017). Estimating Population Size With Link-Tracing Sampling. *Journal of the American Statistical Association*, 112(519), pp.1286–1295. doi:https://doi.org/10.1080/01621459.2016.1212712.

Volz, E. and Heckathorn, D.D. (2008). Probability Based Estimation Theory for Respondent Driven Sampling. *Journal of Official Statistics*, [online] 24(1), pp.79–97. Available at: https://www.researchgate.net/publication/283857994_Probability_Based_Estimation_Theory_for_Respondent_Driven_Sampling.

Wang, P., Afework Wogayehu, Bolden, B., Ibrahim, A.R. and Raymond, H.F. (2024a). Assessing Consistency of Respondent-driven Sampling Estimators by Using Repeated Surveys among People Who Inject Drugs (PWID) in New Jersey. *AIDS and Behavior*, 28(11), pp.3819–3835. doi:https://doi.org/10.1007/s10461-024-04461-5.

Wang, P., Wogayehu, A., Bolden, B., Ibrahim, A.R. and Raymond, H.F. (2024b). Assessing reliability of naïve respondent-driven sampling samples by using repeated surveys among people who inject drugs (PWID) in New Jersey. *Annals of Epidemiology*, [online] 94, pp.100–105. doi:https://doi.org/10.1016/j.annepidem.2024.05.001.

Wejnert, C. (2010). Social network analysis with respondent-driven sampling data: A study of racial integration on campus. *Social Networks*, 32(2), pp.112–124. doi:https://doi.org/10.1016/j.socnet.2009.09.002.

Wejnert, C. and Heckathorn, D.D. (2008). Web-Based Network Sampling. *Sociological Methods & Research*, 37(1), pp.105–134. doi:https://doi.org/10.1177/0049124108318333.

Wejnert, C., Pham, H., Krishna, N., Le, B. and DiNenno, E. (2012). Estimating Design Effect and Calculating Sample Size for Respondent-Driven Sampling Studies of Injection Drug Users in the United States. *AIDS and Behavior*, 16(4), pp.797–806. doi:https://doi.org/10.1007/s10461-012-0147-8.

Welch, S. (1975). Sampling by Referral in a Dispersed Population. *Public Opinion Quarterly*, 39(2), p.237. doi:https://doi.org/10.1086/268221.

White, R.G., Hakim, A.J., Salganik, M.J., Spiller, M.W., Johnston, L.G., Kerr, L., Kendall, C., Drake, A., Wilson, D., Orroth, K., Egger, M. and Hladik, W. (2015). Strengthening the Reporting of Observational Studies in Epidemiology for respondent-driven sampling studies:

'STROBE-RDS' statement. *Journal of Clinical Epidemiology*, [online] 68(12), pp.1463–1471. doi:https://doi.org/10.1016/j.jclinepi.2015.04.002.

Wirtz, A.L., Iyer, J.R., Brooks, D., Hailey-Fair, K., Galai, N., Beyrer, C., Celentano, D. and Arrington-Sanders, R. (2021). 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. *Journal of the International AIDS Society*, 24(5). doi:https://doi.org/10.1002/jia2.25694.

Wylie, J.L. and Jolly, A.M. (2013). Understanding recruitment: outcomes associated with alternate methods for seed selection in respondent driven sampling. *BMC Medical Research Methodology*, [online] 13(1), p.93. doi:https://doi.org/10.1186/147122881393.

Xia, K., Pineau, V.J., Ganesh, N., Michaels, S. and Reimer, B. (2019). *A Comparison of Estimation Methods for Web-Based Respondent Driven Sampling*. [online] Academia.edu. Available at: https://www.academia.edu/100494989/A_Comparison_of_Estimation_Methods_for_Web_Based_Respondent_Driven_Sampling [Accessed 5 Jun. 2025].

Yan, Y., Hanlon, B., Roch, S. and Rohe, K. (2020). Asymptotic seed bias in respondent-driven sampling. *Electronic Journal of Statistics*, 14(1). doi:https://doi.org/10.1214/20-ejs1698.

Yauck, M., Moodie, E., Apelian, H., Peet, M.-M., Lambert, G., Grace, D., Lachowsky, N.J., Hart, T.A. and Cox, J. (2021). Sampling from networks: respondent-driven sampling. *Epidemiologic methods*, 10(1). doi:https://doi.org/10.1515/em-2020-0033.

Young, A.M., Rudolph, A.E., Quillen, D. and Havens, J.R. (2014). Spatial, temporal and relational patterns in respondentdriven sampling: evidence from a social network study of rural drug users. *J Epidemiol Community Health*, [online] 68(8), p.792. doi:https://doi.org/10.1136/jech2014203935.

Zan, L.M., Rossier, C., Owolabi, O., Baguiya, A., Ouedraogo, R., Bangha, M. and Kim, C. (2024). Using respondent-driven sampling to measure abortion safety in restrictive contexts: Results from kaya (burkina faso) and nairobi (kenya). *Demographic Research*, [online] 50, pp.1387–1422. doi:https://doi.org/10.4054/DemRes.2024.50.47.

Zanoni, W., Romero, J., Chuquimarca, N. and Abuelafia, E. (2023). *Dealing with Hard-to-Reach Populations in Panel Data: Respondent-Driven Survey (RDS) and Attrition*. Inter-American Development Bank. doi:https://doi.org/10.18235/0005194.

Appendix A: The RDS Evidence Review: Methodology, Scope and Research Questions

Objectives and Research Questions

This Research Strand aims to evaluate the practical performance, implementation conditions, strengths and methodological limitations of Respondent-Driven Sampling (RDS) as applied across empirical studies. The review addresses the following inter-related research questions:

1. How does Respondent-Driven Sampling perform in practice across different populations, recruitment modes, and study designs?

2. How do recruitment structures (seeds, referral caps, incentives, chain length) shape recruitment trajectories and sample composition in RDS?

3. Which fieldwork, operational, and design strategies are associated with more successful RDS implementation?

4. How do different recruitment modes (face-to-face, online, and mixed-mode) affect the feasibility, integrity, and outcomes of RDS?

5. Where, why, and for whom does RDS break down as a sampling and recruitment method?

6. Where are the gaps in the existing evidence base regarding RDS?

Metrics and Coding Criteria

The following metrics were detailed in the specification and defined by the Survey Futures research team (Carina Cornesse, Olga Maslovskaya, Curtis Jessop) to be extracted and coded from each individual study:

- Study identifiers (authors, publication year, links)
- Number and characteristics of seeds
- Seed origin and restriction to target population
- Definition of target population
- Recruitment statistics: number of recruits, recruitment per seed, recruitment chain length
- Indicators of recruitment effectiveness
- Reported technical issues (e.g., link generation failures)
- Use of link-tracing methods

- Mode of survey administration (online/offline/mixed)
- Recruitment mechanics (active vs. passive seed involvement)
- Role of RDS in study design (standalone vs. supplement)
- Participant access and sharing protocols
- Use and value of incentives, including dual incentives for recruiters and recruits
- Analytical methods (descriptive vs. inferential)
- Quality assurance and fraud prevention procedures

Inclusion and Exclusion Criteria for the Evidence Review

Inclusion Criteria:

- Peer-reviewed academic papers, grey literature, and unpublished reports
- Studies published in English
- International and UK-based studies
- No temporal restrictions
- Studies that implement or test RDS designs or theoretical simulations
- Studies employing RDS either as a standalone recruitment method or to supplement a separate sample
- Studies where recruitment is respondent-driven or intermediated by field staff

Exclusion Criteria:

- Studies utilising RDS for small-scale qualitative samples not aimed at representativeness

There were no exclusions based on the specific RDS approach, survey mode, population focus, or study objectives.

Search Strategy and Source Identification

The Evidence Review was intended to be as rigorous as reasonably practicable, but not systematic. A list of 50 key papers and sources was provided by the Survey Futures research team (Carina Cornesse, Olga Maslovskaya, and Curtis Jessop) to guide the initial scoping and to validate search efficiency. Searches were conducted using Google Scholar and Scopus with multiple search iterations to maximise coverage.

Google Scholar

- Search unfiltered by publication year
- Grey literature included
- Six search iterations performed
- Final search successfully identified 35 of the 50 target papers
- Best-performing search string: “Respondent Driven Sampling, RDS”
- Total search hits: 7,600
- A technical limitation with the API request implemented in Python was that the list was restricted to extracting the first (most relevant) 980 results before request/response timeout
- Search was conducted on 08 January 2025

Scopus

- Sixteen search iterations performed
- Final search successfully identified 41 of the 50 target papers; 9 remaining were outside Scopus coverage (books, conferences)
- Best-performing search string: TITLE-ABS-KEY ("Respondent Driven Sampling" OR ("Network sampling" OR "Link Tracing Sampling" OR "link tracing") OR ("stochastic actor-based models" OR "sampling by referral"))
- Final search hit count: 2,859
- A total of 355 records were retained
- Final search was conducted on 20 February 2025

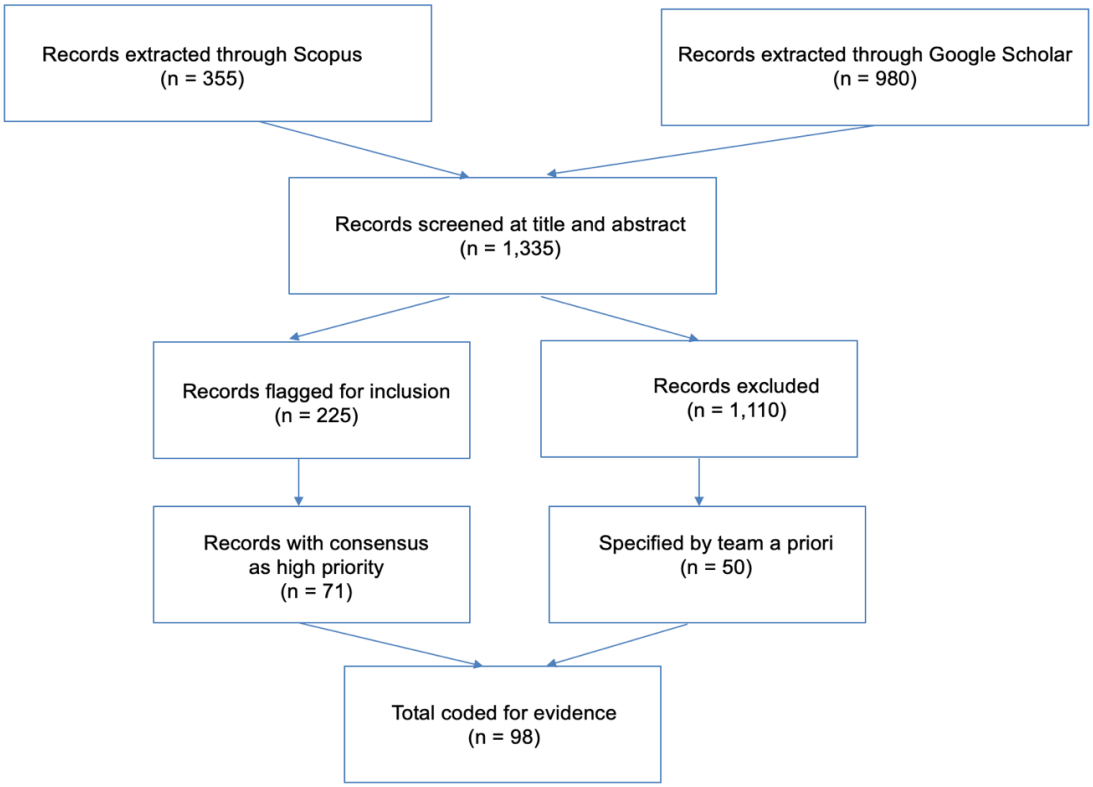
  The final dataset for screening combined 980 Google Scholar hits with 355 Scopus records, resulting in 1,335 studies, which were then screened by title and abstract.

Screening and Coding Process

From the 1,335 records, 225 were flagged for potential inclusion and prioritised based on methodological relevance, alignment with the inclusion criteria and according to how well each covered the metrics / coding criteria. Two reviewers (Toby Li and Mark Todd) conducted independent screening to prioritise the evidence, and any discrepancies were resolved through consultation with a wider research team (Olga Maslovskaya, Curtis Jessop, Carina Cornesse). This led to the identification of 71 high priority papers to include in the review. Of these, 48 were newly coded during the review process within the project time constraints.

Combined with the 50 pre-specified papers, 98 studies were coded overall and analysed for this evidence review (please see Figure 1 for details). Coding was initially conducted using AI (Google Gemini Flash 2.0 and ChatGPT 4.5; for the prompts see Appendix B), before papers were read and coding verified by the reviewers. The review is a narrative synthesis of the strengths and weaknesses of RDS sampling and recruitment, key metrics were inconsistently reported in different papers to enable meta-analyses.

Figure 1: PRISMA Flowchart



Appendix B: AI coding prompts

Item1 Google Gemini Flash 2.0 prompt used to help with coding: {

I am going to provide you a selected text from a relevant respondent-driven sampling academic literature. Can you extract the relevant information according to the metrics? The metrics are:

1. Study identifiers (authors, date, links, etc.)

2. Number of seeds?

3. Origin of seeds (e.g. a general population survey? community workers? click workers?)

4. Are the seeds restricted to the target population (E.g. if our target population is LGBT people, do our seeds all need to be LGBT people, or could it be gen pop)?

5. Definition of target population (e.g. a specific ethnic minority group? general population?)

6. Number of successfully recruited people?

7. Number of recruits per seed (mean/min/max)?

8. Length of recruitment chains (mean/min/max)

9. Quality indicators used on how effective RDS approach was (or if none used

10. Study's and/or our evaluation of effectiveness (or otherwise) of RDS approach

11. Were there any technical issues? If so, which (e.g., recruitment link generation failures)?

12. Was a link tracing procedure applied (i.e. can researchers trace who was recruited by whom)?

13. Was the recruitment procedure/survey conducted online or offline or both?

14. Did the seeds actively engage in the recruitment or did they only provide the contact details of the prospective recruits to the fieldwork agency?

15. Was the RDS a standalone approach or used to boost.in conjunction with another study?

16. Protocols for sharing/providing access to the survey for recruited participants

17. Was the data used in an exploratory / descriptive way or did researchers attempt inferential statistics? Which and how? Which modeling approaches were used?

18. Whether incentives were used. Value of incentives. Were incentives used for both recruiter and participant?

19. Potential issues: duplications, fraud/methods for prevention & quality assurance

Can you extract the relevant information from the following text according to the metrics listed above ? : [insert selected section of the academic paper] }

Item2 ChatGPT 4.5 prompt used to help with coding: *{*

Using only the data and explicit findings from the attached paper, provide detailed coding for the following components. Do not include assumptions, external knowledge, or generalized RDS principles beyond what is explicitly stated in the paper. Use only direct evidence, results, or statements from the paper. If any component is not addressed in the paper, state 'Not reported in the paper' rather than inferring or assuming information. Produce results in table format:

1. Study identifiers (authors, date, links, etc.)

2. Number of seeds?

3. Origin of seeds (e.g. a general population survey? community workers? click workers?)

4. Are the seeds restricted to the target population (E.g. if our target population is LGBT people, do our seeds all need to be LGBT people, or could it be gen pop)?

5. Definition of target population (e.g. a specific ethnic minority group? general population?)

6. Number of successfully recruited people?

7. Number of recruits per seed (mean/min/max)?

8. Length of recruitment chains (mean/min/max)

9. Quality indicators used on how effective RDS approach was (or if none used

10. Study's and/or our evaluation of effectiveness (or otherwise) of RDS approach

11. Were there any technical issues? If so, which (e.g., recruitment link generation failures)?

12. Was a link tracing procedure applied (i.e. can researchers trace who was recruited by whom)?

13. Was the recruitment procedure/survey conducted online or offline or both?

14. Did the seeds actively engage in the recruitment or did they only provide the contact details of the prospective recruits to the fieldwork agency?

15. Was the RDS a standalone approach or used to boost.in conjunction with another study?

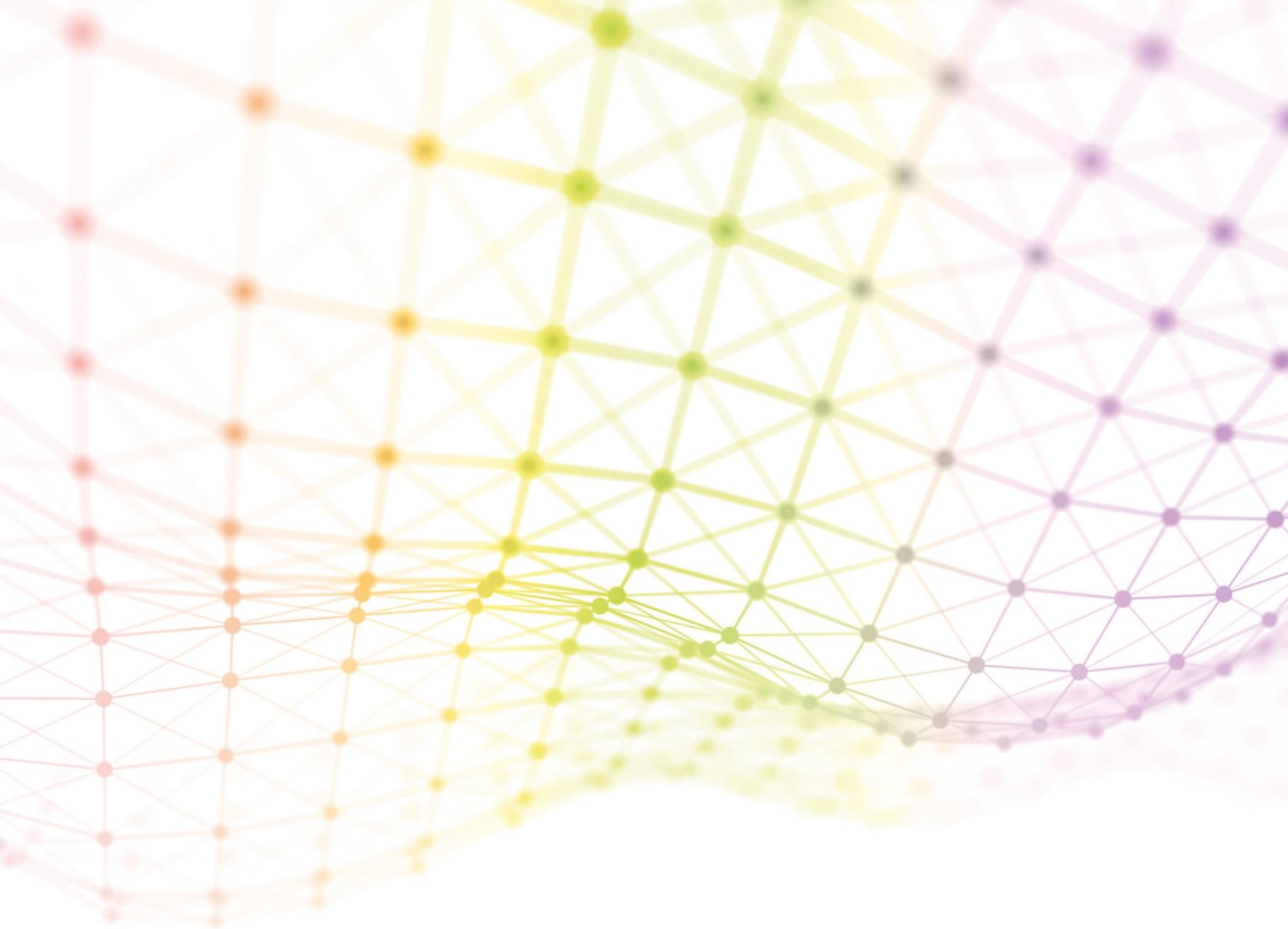
16. Protocols for sharing/providing access to the survey for recruited participants

17. Was the data used in an exploratory / descriptive way or did researchers attempt inferential statistics? Which and how? Which modeling approaches were used?

18. Whether incentives were used. Value of incentives. Were incentives used for both recruiter and participant?

19. Potential issues: duplications, fraud/methods for prevention & quality assurance

}



University of Essex

University of Southampton

UKRI Economic and Social Research Council

NCRM National Centre for Research Methods

Office for National Statistics

UCL

Warwick The University of Warwick

Manchester 1824 The University of Manchester

City University of London Est 1894

National Centre for Social Research

LSE The London School of Economics and Political Science

Ulster University

Unil UNIL | Université de Lausanne

Ipsos

verian Formerly Kantar Public

www.surveyfutures.net