FINDING THE INVISIBLE
RESPONDENT DRIVEN SAMPLING
FINDING THE INVISIBLE
MICHELE R. DECKER, SCD
DAESHA RAMACHANDRAN, MHS
PRESENTATION OVERVIEW

• Background
• Strengths & limitations of RDS
• Overview of RDS study design
• Analytic considerations
• Reporting RDS research
• Moving beyond “typical” RDS populations
WHAT IS RDS?

• Introduced by Doug Heckathom, 1997
• Chain referral, link tracing network sampling method based on social network theory
• Offers rigor over other chain referral methods in recruitment and analysis
• Useful when traditional sampling methods are not possible
• Primary used for hard to reach (“invisible”) populations
  • men who have sex with men (MSM), female sex workers (FSWs), and injection drug users (IDUs)
WHAT IS RDS?

- Key assumptions: an underlying network AND sampling frame is unknown
- Good for
  - Migrants
  - MSM
  - FSWs
  - IDUs
  - Jazz musicians, artists
  - Child prostitutes
- Not good for
  - Broadly defined demographic groups for which the sampling frame is known (e.g. African Americans; households below the poverty line; students in school; people who get flu vaccines)
A METHOD UNDER REFINEMENT!

- Since its introduction in 1997, RDS is increasingly implemented for hard to reach populations
  - 1997-2005: 19 publications
  - 2006-2010: 212 publications
  - 2011-2012: 173 publications
- Researchers are increasingly
  - testing the core assumptions
  - evaluating applicability and cost-effectiveness,
  - optimizing implementation
- Not without controversy!
RDS IS THE METHOD OF CHOICE

• Effective at reaching more invisible, marginalized sub-populations
  • Johnston et al. found RDS to better capture hidden pockets of FSWs
• Can be a faster and, at times, more cost-effective method of sampling
  • Kendall et al. (2008) tested TLS, snowball and RDS methods and found RDS to be faster, more inclusive and most cost effective of methods
• Those recruited more likely to be eligible
  • Robinson et al. compared RDS and targeted sampling for IDUs in the U.S.
...BUT THERE ARE STILL ISSUES

- Even when population is hidden & hard to reach, RDS is not the best method
  - E.g. Simic et al. found weak social network among FSWs in Russia and former Soviet states
  - Platt et al. found RDS was safer but more costly and did not expose more marginalized groups
- Estimates can still be biased with unreliable CIs
  - McCreesh et al. found that RDS failed to reduce bias in sample estimates;
HOW RDS WORKS

1. Formative phase
2. Recruitment
   • Controlled
   • Tracked
   • Trained participants recruit
   • Incentivized
3. Analysis
   • Can be weighted
   • Can adjust for non-independence
FORMATIVE PHASE IS CRITICAL

• Is the population adequately networked?
• Identify seeds
• Identify subpopulations particularly those who are less “networked” with one another (eg by gender, drug use etc.)
FORMATIVE PHASE

• Logistical considerations
  • Determine optimal times and locations for recruitment
  • Explore incentivization structure including need for secondary incentives
• Material development
  • Survey questions
  • Coupons
• Methods of collecting formative information include focus groups, in-depth interviews, key-information interviews, mapping

• For further reading:
FORMATIVE RESEARCH WITH FSWS: KRASNOYARSK, RUSSIA

<table>
<thead>
<tr>
<th>SEEDS</th>
<th>Left bank</th>
<th>Right bank</th>
</tr>
</thead>
<tbody>
<tr>
<td>IDU</td>
<td></td>
<td>2-3</td>
</tr>
<tr>
<td>Non-IDU</td>
<td>2-3</td>
<td>2-3</td>
</tr>
</tbody>
</table>
RECRUITMENT

• Recruitment begins with seeds
  • Initial, purposive recruits, often facilitated by a partner NGO
  • Generally 3-10; FSWs found to require more seeds
  • Seeds should be well-known and respected by members of the target population
  • Capable of generating buy-in
  • Comprise key sub-populations identified in formative phase

• Recruitment is controlled
  • Fixed number of recruitment coupons to each participant (usually 3-5)
  • No participant can recruit more than 3-5 individuals!
RECRUITMENT GOALS

- Long referral chains
- Final sample is independent of “seeds”
- Sample reaches equilibrium and includes a variety of subgroups
- Final sample reflects population of recruiting network
THEORETICAL RECRUITMENT CHAIN

TRACKING RECRUITS: COUPONS

• Unique RDS (coupon) number is recorded and used to link participants
• Coupon also contains:
  • Organization name
  • Location, hours of data collection
  • Contact information
  • Expiration date
  • Sometimes designed in two parts: recruitment and payment
COUPON EXAMPLE

---

**Project co petro**

---

Singathande ukuba uluzwa kuphando lwethu
nceda une kukityelela o:

Youth Health Centre, Site B, Khayelitsha
(Fumana isiskhiso sandla ay khona angokuba umva
wali kuphiso). Sivance ukushabalule

9am to 4pm (Saturday & Sunday)

Ungama i appointment okanye ukuba umva ukwazi
sifumvelise ku

072 217 7630

---

This coupon expires:

---

Le kupon u gakuni ka isahlala yokuseumo sibhese lenimpilo
yokushaba simhal kuNempile wakho

Hamba yokuhlala sibhese o ab o:

Youth Health Centre, Site B, Khayelitsha

Phakathi ko 11am and 4pm Saturday

072 217 7630

*Noela ufiles okubhala ngeumva ukela coupon*
COUPON CONSIDERATIONS

- Coupons should be:
  - attractive yet discreet
  - durable to withstand traveling in pockets and wallets
  - difficult to reproduce for entrepreneurial participants
  - able to convey study information easily for limited literacy populations
COUPON CONSIDERATIONS

• Logistics: once coupons are in circulation, it is hard to adjust study location and hours
  • need to have someone on-site during the data collection hours posted!
• Alternates to a physical location are sometimes necessary
  • When participants lack time or transportation costs associated with going to a given site
  • Participants can call or “flash dial” the study and arrange a mutually agreeable time and location for participating
MONITORING RECRUITMENT

- Recruitment chains must be monitored each day in the early phases of data collection.
- Monitor recruitment patterns:
  - Do we need to launch a new seed?
  - Are there barriers to recruits getting enrolled?
- Monitor the **coupon procedures** that support our adjusted analyses:
  - Coupon numbers are the links in recruitment chains that allow us to adjust for the complex design of RDS.
  - Without proper coupon procedures, all we have is a convenience sample 😞
FSW RECRUITMENT: KAZAN
Training Recruits to Recruit

- Our seeds, and subsequently our participants, are our study ambassadors!
- We need to motivate them to invite their peers to participate
  - In part through incentives
  - In part through training
- Enable confidence in recruiting through training
  - When giving the coupons, ask them to think about who they might invite, what they might say
INCENTIVES

- Primary incentives
  - given upon survey completion
- Secondary incentives (optional)
  - given upon successful enrollment of individuals’ recruits
- Can be
  - monetary or non-monetary
  - stratified (to reach more hidden or high-risk groups)
- As with any study, possible pitfalls include inducing false reports of behavior to receive incentive (e.g. reporting MSM behavior when not MSM)
ANALYTIC CONSIDERATIONS

- Free analytic software for RDS
  - Netdraw to monitor recruitment and generate graphs
  - RDSAT for RDS-specific analyses
- RDS analysis is inconsistent across studies!
- Two primary issues:
  - Generalizability
  - Non-independence
ANALYSIS CONSIDERATIONS: WEIGHTS

- **PROBLEM:** Some participants will know more people than other participants!
  - Not everyone has equal chance of inclusion
- **ANSWER:** weighting by network size
  - During the survey, each participant is asked about their “network size”, i.e., how many other [insert eligibility criteria] they know
  - Weights calculated by RDSAT and can be exported to SAS or STATA for analysis
  - In practice: not always done
ANALYSIS CONSIDERATIONS: NON-INDEPENDENCE

• PROBLEM: aren’t people likely to recruit others like themselves?
  • Non-independence

• ANSWER(s):
  • Evaluate homophily, i.e., tendency for in-group recruitment, via RDSAT
  • Based on? HIV status, age, race, gender, sub-populations identified in formative phase
ANALYSIS CONSIDERATIONS: NON-INDEPENDENCE

- RDSAT can adjust for complex design in prevalence and cross-tab analyses; currently does not support complex analyses
- As a result, handling RDS samples like SRS is common
- Other have demonstrated use of complex survey analysis procedures (svy)
  - This offers the benefit of both weighting and adjusting for clustering by recruitment chain
- In theory, if chains are long enough, samples should be independent from seeds.
- But we need an approach for when they are not!
REPORTING ON RDS STUDIES

- Little consistency in reporting on RDS with regard to key parameters

- Recent calls for standardization of parameters
  - Eg, coupons distributed, coupons returned, measures to reduce potential repeat participation

For further reading:
BEYOND HIV RESEARCH

• RDS is increasingly applied for other types of high-risk populations

• Migrant workers at the Thai-Burma border (Trafficking Assessment Project; Robinson/Decker)

• Adolescents in urban slum communities (WAVE Study; Blum/Sonenstein)

• Undocumented immigrant populations
THE WORD ON THE STREET

- RDS is fast and effective at reaching hard-to-reach populations
- It can be challenging analytically
- Generalizability to the underlying population is hard to establish

- The bottom line:
- Right now, it is the best available method of conducting research with hard-to-reach populations
FOR FURTHER READING

- [www.respondentdrivensampling.org/main.htm](http://www.respondentdrivensampling.org/main.htm)
  Information, key references and analytic guidance from Doug Heckathom

  Research and recent RDS developments from Matthew Salganic

- [http://www.lisagjohnston.com/](http://www.lisagjohnston.com/)
  RDS how-to, research and recent RDS developments from Lisa Johnston