

# System Assisted Link Tracing (SALT): Cheap, Easy, Fast, Continuous and Statistically Valid BBS-like Sampling for Monitoring Key Populations

Ian E. Fellows

## Abstract

This article introduces a new sampling method, System Assisted Link Tracing (SALT), aimed at addressing some challenges of Bio-Behavioral Surveys (BBS) implemented using Respondent Driven Sampling (RDS) in monitoring key populations in the global HIV response. The SALT method, while sharing some characteristics with RDS, leverages programmatic staff and facilities, offers continuous monitoring, and provides instant analysis. It uses a software system to guide staff through the sampling process and manage recruitment speed, with flexibility to adjust settings based on the specific study. The software also generates real-time statistical analysis. This approach intends to reduce the cost and time often associated with BBS. Furthermore, it provides a framework for reducing impact on program staff time while ensuring accurate sampling and data collection. The system's configuration, monitoring, and analysis tools also promise efficient interaction for study administrators and policy stakeholders.

## Introduction

Monitoring key populations is a critical component of the global HIV response due to the outsized role these populations play in HIV transmission. Bio-Behavioral Surveys (BBS) implemented using Respondent Driven Sampling (RDS) are the primary sampling methodology used. Bio-Behavioral Surveys have been criticized as:

1. **Expensive:** Training study personnel, standing up study sites, monitoring sampling, and performing specialized statistical analyses all add to implementation cost.
2. **Slow:** The time from conception to actionable results can be considerable due to the time it takes to have a protocol approved, run an RDS with sufficient chains to reach convergence, and perform the statistical analysis.
3. **Out of Date:** Due to expense, BBS are run rarely and so their results are often out of date. Additionally, monitoring changes in a population requires multiple Bio-Behavioral Surveys over time.

Despite these limitations, BBS have been widely and successfully applied to a wide range of populations of interest. No alternative statistical sampling methods have been proposed that are superior to RDS on any of these three areas.

We propose a new sampling method called System Assisted Link Tracing (SALT), which borrows many characteristics from RDS, but aims to be superior on each of the limitations above. Like RDS, SALT samples the population using link tracing and uses coupon passing to identify and survey new individuals. Also like RDS, SALT generates a population representative sample. However, there are several differences that reduce expense and friction compared to a traditional RDS survey. They are:

1. **Leveraging programmatic staff and facilities:** SALT is implemented on the ground by programmatic staff at program facilities. These staff require limited to no training because SALT uses software to guide staff and participants through the sampling process, providing clear and easy to accomplish instructions.
2. **Continuous monitoring:** While RDS is a discrete survey performed at in a short time window, SALT recruits slowly and continuously. Seed selection and rate of recruitment are controlled by the management software to ensure that the sample is always at convergence. This allows population trends to be identified and results in a sample that is never out of date.
3. **Instant Analysis:** Appropriate statistical analyses are performed on a real-time basis by the software system and displayed in a dashboard. This results in zero lag time from subject participation to their data being analyzed and available for decision making.

## RDS and SALT Sampling

The most popular sampling method used for BBS studies is respondent driven sampling (RDS). RDS is link tracing design that traverses the underlying social networks of population members.

### RDS Sampling

1. Recruit a small convenience sample of individuals from the population (seeds).
2. Survey the recruited individuals including collecting information on the size of their social network (network size).
3. If the desired sample size has not been met, provide surveyed subjects with coupons that can be used to recruit members of their network who have not yet participated in the study.
4. GOTO step 2 for recruited contacts.

It is desirable for the RDS chains to be long to reduce or eliminate the bias inherent in the initial convenience sample of seeds. As chains become longer, the sampling process reaches an equilibrium, where the probability an individual is sampled is proportional to their network size. At this point the RDS sample is population representative.

SALT sampling is similar to RDS but is continuous and uses software to guide staff and determine recruitment speed.

### SALT Sampling

Set-up: Populate the recruitment pool with a small convenience sample.

Ongoing: If the SALT management software has signaled to sample an individual from the recruitment pool, contact the selected individual and recruit them. GOTO step 1.

1. Survey recruited individuals including collecting information on the size of their social network (network size). Also collect contact information for willing subjects, who will be added to the recruitment pool by the SALT management software.
2. Provide surveyed subjects with a number of coupons determined by the SALT software that can be used to recruit members of their network who have not yet participated in the study within the last X months.
3. GOTO step 1 for recruited contacts.

### Definitions

*SALT management software:* This is the software system that helps control and manage the sampling process. It determines how many coupons to give to subject in order to hit monthly recruitment targets. If the link tracing is not proceeding fast enough to meet the monthly target, this system also selects individuals from the recruitment pool to increase the sampling rate.

*Recruitment pool:* This is a list of individuals who have previously participated in the study along with contact information. As the study progresses, this pool is filled by study participants recruited through link tracing and thus converges to an equilibrium sample.

SALT sampling begins as RDS does. A convenience sample of seeds are added to the recruitment pool and are then selected by the management system for sampling. These seeds are given coupons to recruit other members of the population, who in turn may be given seeds to recruit others.

There are significant differences though. The SALT management software will generally aim to recruit more slowly. An RDS study may aim to recruit 400 individuals in two months. SALT is a continuous monitoring system and thus may aim to recruit that figure over a year. Additionally, SALT allows for re-enrollment after a certain number of months.

A key component of SALT is the recruitment pool. Slower sampling may lead to recruitment chains drying up. The recruitment pool allows us to pick up sampling deep into the recruitment trees, far away from the initial seed convenience sample. The pool initially starts with a convenience sample, but rapidly fills with study participants that are increasingly many waves of recruitment away from this initial set. When the SALT management software selects an individual from the pool, this person is already at equilibrium and is not a convenience seed.

This mechanism allows SALT to be always at equilibrium, and hence population representative, even when the link trace sampling runs dry.

The probability of selection from the recruitment pool is a configurable setting in the SALT management software. Administrators may choose to select only individuals who have not participated recently, they may exclude the initial convenience sample from selection and/or may adjust the selection probabilities for non-response (i.e. individuals who did not wish to be contacted for re-recruitment).

## SALT Analysis

Existing RDS estimators can be leveraged in the analysis of SALT data. The first step of the analysis is to restrict the sample to a time window. The length of the time window (6, 12, 24 months, etc.) should be chosen such that the time specificity of the inference is acceptable, and the sample size is sufficient to generate relatively narrow confidence intervals. Standard RDS estimators, such as the Salganik-Heckathorn, Successive Sampling, Volz-Heckathorn, or Homophily Configuration Graph may be used to adjust for the link tracing process.

Time trends may be visualized by calculating the estimators over rolling windows. For example, one could construct estimates each month based on the trailing 12 months of SALT data. These estimates and their corresponding confidence intervals would then form a time series, which can be plotted to identify trends.

## SALT from the Staff Perspective

Because SALT relies on program staff to for sampling implementation, it is critical that the impact on their time be minimized and that the system streamlines the sampling process such that minimal training is required to successfully complete all tasks. Additionally, while SALT may minimize impact, it is still recommended that facilities be resourced to handle the additional workload that SALT requires.

The primary way that study staff interact with SALT is through a tablet located at the facility. Approved staff members will be able to log into the tablet. Prior to being able to do anything else, staff must complete a video training course that describes the SALT sampling process, what their role is, and what the do's/don't are.

If the management system signals to recruit from the recruitment pool, it pings staff through either e-mail or SMS. When they log into the tablet, they have access to the contact information of the person they are supposed to recruit. A script guides them through the recruitment phone call and scheduling.

When an individual comes in with a coupon or from the recruitment pool sampling, the study staff starts the tablet survey. It instructs them to take the subject into a private room and then guides both the staff and subject through the survey. The survey may be administered by the study staff using the tablet, or the participant may take the survey on the tablet without staff

assistance. Audio-Computer Assisted Self-Interview (ACASI) software may be built into the tablet to assist in administering the survey without staff intervention. The length of the survey is kept considerably shorter than a typical BBS, with only actionable information being collected.

If ACASI is used, behavioral counseling may also be provided in the form of pre-recorded videos covering a range of topics (e.g. alcohol misuse, lack of condom use, rape, depression, commercial sex, etc.). The computer plays these as needed, based on response values in the ACASI.

Laboratory panels may be taken either immediately after consent and eligibility are determined or at the end of the survey. An HIV rapid test may be taken at the start so that the result is available to the participant by the time they complete the survey. Negative results may be returned electronically on the tablet whereas positive results are returned by a staff member.

If the management system determines that coupons should be issued, these are printed out by the tablet and the staff is instructed to explain and provide them to the participant. Participant payment, both for their time and for any recruitments they do is handled flexibly by the management system depending on the population. One payment modality is automatic phone payments and another is in-person cash payments verified by facial photographs taken via the tablet.

## SALT Software

The software system is the critical component that makes SALT possible. It is designed to be deployable across populations and surveys with separate configuration options for each study customizing the implementation. The tablet provides strict guard rails for the survey staff, limiting necessary training and reducing workload. The management system controls the sampling process, ensuring that study goals are met. The analytics platform serves the role of continuously available statistician, bringing actionable information to policy makers on a real-time basis.

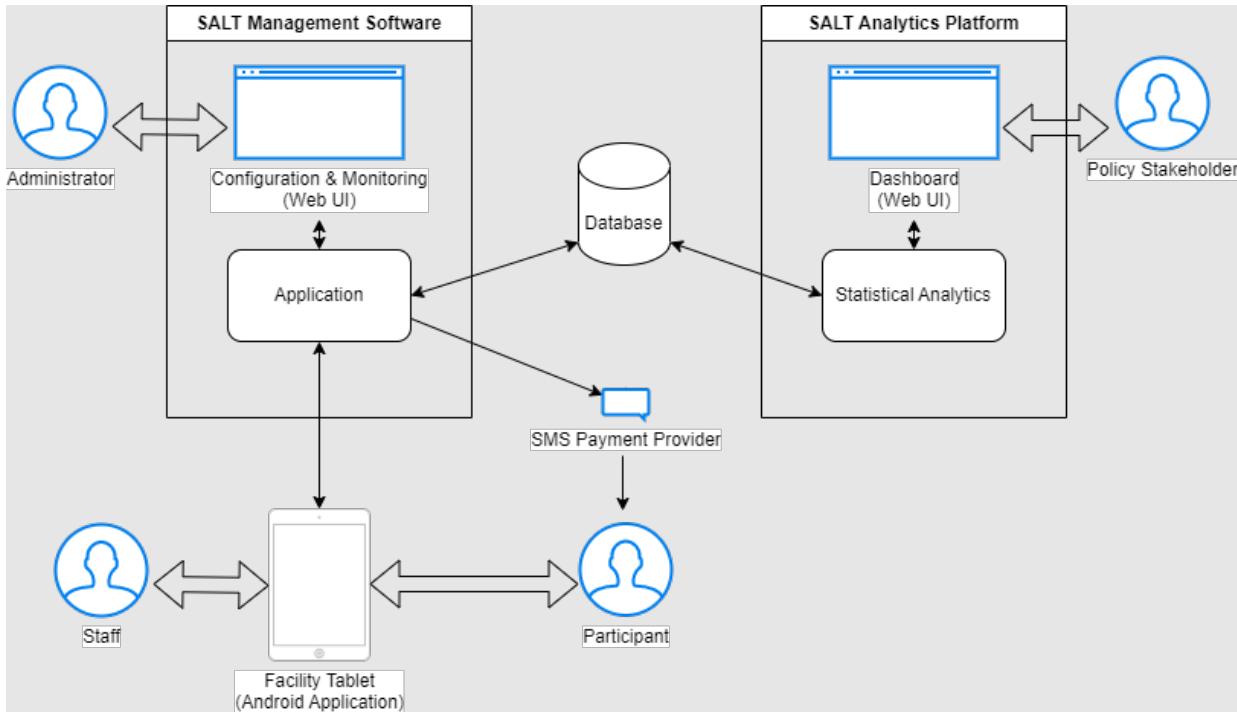


Figure 1: SALT Software Platform Structure

The study administrator interacts with the SALT Management Software through the Configuration and Monitoring UI. Here she is able to change options like the rate of data collection, maximum number of coupons, and recruitment pool criteria. She can also view recruitment diagnostics.

The policy stakeholder interacts with the Analytics Platform through the analysis dashboard. This dashboard displays rolling 6, 12 and 24 month estimates of population quantities of interest (Prevalence, viral suppression, etc.). The underlying statistical analytics software leverages estimators that adjust for the link-tracing design (e.g. Salganik-Heckathorn estimator, Homophily Configuration Graph estimator, etc.). However, the statistical sophistication is transparent to the stakeholder, who only views the estimates and confidence intervals.

The staff and participants interact with the system through the facility tablet. Additionally, the management system may interact with the participant to send them compensation mediated through a phone-based payment provider.

## Security and PII

All information stored in the SALT platform is secured using software architecture best practices. This includes encrypted communications, encrypted data storage, multifactor authentication, and user access controls.

Over and above these best practices, particular attention is paid to the personally identifiable information (PII) that is collected. The only PII that needs to be captured and stored by the system is the contact information for individuals in the recruitment pool. This may include their name and phone number. The degree of sensitivity of this information varies based on study populations, with some populations being relatively unstigmatized whereas in other populations individuals may encounter significant consequences if their membership is known.

Before storage, PII is encrypted on the tablet. Decryption may only be performed by authorized users of the tablet using a cryptographic secret that is only stored on the tablet and authorized users may only be created on the tablet itself by the on-site administrator of the tablet. Thus, even if the encrypted PII is removed from the facility, it is inaccessible without both the tablet and authorized user credentials for that tablet.

The encrypted PII may either be stored locally on the tablet, or centrally in the platform database. If it is stored locally, then no PII, even encrypted, ever leaves the facility. However, the tablet is generally in an active facility and is thus there may not be a completely secure environment to physically store it. That said, if the tablet is stolen, all information, including PII is encrypted and inaccessible to unauthorized users.

If the encrypted PII is stored centrally, then any stolen tablet would have no PII stored on it (even encrypted). However, the encrypted PII would all exist in one place, which may be more at risk of a governmental action. Even if the encrypted PII data were all confiscated by an undesirable actor, it would be inaccessible without the tablets and user credentials.

A consequence of this architecture is that the recruitment pool of a site is lost if the tablet is stolen or becomes corrupted. It is recommended that the site administrator keep an encrypted backup on site in a secure location.

## Conclusion

The proposed System Assisted Link Tracing (SALT) method presents a transformative approach to monitoring key populations in the context of global HIV response. The current standard, Bio-Behavioral Surveys (BBS) implemented using Respondent Driven Sampling (RDS), has had great success, but is not without its difficulties. These include high costs, time-consuming processes, and lagged results due to infrequent survey cycles. These challenges limit the effectiveness of public health initiatives as well as the timeliness and accuracy of data for policy decision-making.

SALT is a novel solution that tackles these hurdles head-on, leveraging technological systems and the human resources already embedded in the healthcare system. By utilizing programmatic staff and facilities, SALT removes the burden of extensive training and standing up independent study sites, thereby reducing the overall implementation cost. Additionally, SALT's design allows for continuous monitoring, which is a critical advancement over traditional

BBS surveys. This feature enables the collection of real-time, up-to-date data and helps in identifying emerging trends in population behavior more accurately and swiftly.

The real-time statistical analysis generated by the software results in zero lag time from subject participation to their data being available for interpretation and policy decision-making. This feature holds immense potential in changing HIV incidence environments, where immediate access to reliable data can significantly improve the response strategies.

The unique design of SALT also minimizes the impact on the time of program staff, ensuring that they can effectively fulfill their roles within the process. The system's user-friendly interface on tablets at the facilities, coupled with its robust recruitment management, allows for seamless integration of SALT into the daily operations of health facilities. This method maintains the integrity of the sampling process while not unduly burdening the healthcare personnel.

For study administrators and policy stakeholders, the SALT software offers efficient interaction points. Administrators can easily adjust and monitor study parameters, while policy stakeholders can access real-time data and analytics, all without needing extensive statistical knowledge.

Implementation trials of SALT are necessary to validate and refine its potential. The introduction of SALT could improve the HIV response in key populations by enhancing the accuracy, timeliness, and cost-effectiveness of population monitoring.